Three-dimensional structure and receptor-recognition site of bombyxin-II, an insulin-related peptide of the silkmoth Bombyx mori

（カイコのインスリン族ペプチド、ボンビキシン-Ⅱの
立体構造と受容体認識部位）

永田宏次

Three-dimensional structure and receptor-recognition site of bombyxin-II, an insulin-related peptide of the

①

Three-dimensional structure and receptor-recognition site of bombyxin-II, an insulin-related peptide of the silkmoth *Bombyx mori*

（カイコのインスリン族ペプチド、ボンビキシン-II の立体構造と受容体認識部位）

永　田　　宏　次

Contents

Chapter 1
Introduction

1-1. Biochemical background of bombyxin research

(a) Hormonal control of insect metamorphosis

The metamorphosis of insects is regulated by effector hormones controlled by neurosecretory peptide hormones in the brain (Figure 1-1) (for reviews, see Granger and Bollenbacher, 1981; Gilbert and Goodman, 1981). The molting process is initiated in the brain, where neurosecretory cells release prothoracicotropic hormone (PTTH). PTTH stimulates the prothoracic glands to produce ecdysone, which is converted into 20-hydroxyecdysone (molting hormone). Each molt is occasioned by one or more pulses of 20-hydroxyecdysone. The second major effector hormone in insect development is juvenile hormone (JH), which is secreted by the corpora allata. This hormone is responsible for preventing metamorphosis. As long as JH is present, the 20-hydroxyecdysone-stimulated molts result in a new larval inster. In the last larval inster, JH levels drop below a critical threshold value. The subsequent molt, occuring in the relative scarcity of JH, shifts the organism from larva to pupa. During pupation, the corpora allata release no JH, and the 20-hydroxyecdysone-stimulated pupa will eventually metamorphose into the adult insect.

All these hormones, ecdysone (Butenandt and Karlson, 1954; Hüber and Hoppe, 1965), 20-hydroxyecdysone (Hoffmeister, 1966) and JH (Röller et al., 1967), were isolated and chemically characterized in 1960's, except for the brain-secretory peptide hormone, PTTH, which was isolated at length in 1987 (Kataoka et al., 1987, 1991; Kawakami et al., 1990).

(b) Purification of bombyxin as a "false PTTH" of the silkmoth Bombyx mori

PTTH was discovered as the first neurosecretory hormone throughout the animal kingdom, which Kopec (1922) described as an endocrine factor that induced pupation of the gypsy moth

Lymantria dispar. In spite of many researchers' efforts to purify PTTH, its chemical structure was not determined because of a formidably minute content of PTTH in an insect brain. In 1960, Ishizaki and Ichikawa began purifying the PTTH from brains of the silkmoth Bombyx mori. Since aqueous extract of Bombyx brains was able to induce adult development of brain-removed dormant pupae of both the silkmoth Bombyx mori and the saturniid moth Samia cynthia ricini when injected into pupae, they assumed that the biological activity of the Bombyx PTTH was species-nonspecific in activating pupae of the two species. Because of several technical advantages, they used Samia brain-removed pupae as an assay animal for the Bombyx PTTH purification (Figure 1-2). After extensive efforts towards the isolation for more than 20 years, Suzuki, Ishizaki and colleagues finally obtained the "PTTH" in a pure form (Suzuki et al., 1982). But this pure PTTH preparation was unexpectedly inactive to Bombyx brain-removed pupae. Soon after that, it was revealed that the Bombyx brain extract contained two distinct molecules with prothoracicotropic activity: (1) the genuine Bombyx PTTH (30 kDa) active to Bombyx but not to Samia and (2) a PTTH-like substance (5 kDa) heterologously active to Samia but not to Bombyx (Ishizaki et al., 1983b). Therefore, the crude extract of Bombyx brains was active to both Bombyx and Samia because it contained two types of molecules that were specifically active to the respective species. After that, they purified the genuine Bombyx PTTH using Bombyx brain-removed pupae as an assay animal (Kataoka et al., 1987). Despite the similar biological activity, the Bombyx PTTH (Kataoka et al., 1991; Ishibashi et al., 1993) and bombyxin (Nagasawa et al., 1986; Maruyama et al., 1988) have dissimilar structures (Figure 1-3). The Bombyx PTTH is a 30-kDa glycoprotein consisting of two identical peptide chains of 109 amino acid residues, seven disulfide bonds and glycosides. In contrast, bombyxin is a 5-kDa peptide consisting of two different peptide chains of 20 and 28 amino acid residues and three disulfide bonds. Recently, the Samia PTTH was identified (Adachi-Yamada et al., in preparation), which is a Bombyx PTTH-like glycoprotein consisting of two identical peptide chains of 125 amino acid residues, seven disulfide bonds and glycosides. Despite the apparently same biological activity, the Samia PTTH and bombyxin have no sequence similarity (Figure 1-3).

(c) Primary structure of bombyxin: homology with insulin

Although the activity of bombyxin was disappointing, determination of amino-acid sequence and disulfide-bond location of bombyxin-II, a representative molecular species of bombyxin, revealed that it was similar in primary structure to insulin, a vertebrate peptide hormone which lowers the blood sugar level (Nagasawa et al., 1984a). Both the molecules consist of two peptide chains (A- and B-chains; 20 and 28 residues in bombyxin-II; 21 and 30 residues in human insulin) which are 50% and 32% identical in sequence (Nagasawa et al., 1986) and three disulfide bonds linked in the same way (A6-A11, A7-B7 and A20-B19) (Nagasawa et al., 1988) (Figure 1-4). Bombyxin was thus identified as the first insulin-related peptide of invertebrate origin. Later, several insulin-related peptides have been identified from a wide range of invertebrate species (Figure 1-4), suggesting the insulin-related peptides are generally present in invertebrates. These peptides may play important roles in metabolism, growth and reproduction of invertebrates as do the vertebrate counterparts, insulin, insulin-like growth factors (IGF-I, -II) and relaxin. Three-dimensional structures of bombyxin-II and other members of insulin superfamily were proposed (Figure 1-5) using interactive computer graphics and energy minimization techniques, assuming homology with porcine insulin, the structure of which was determined by X-ray analysis (Jhoti et al., 1987; Murray-Rust et al., 1992). Recently, the intrinsic functions of bombyxin have been revealed. Bombyxin lowers the concentration of haemolymph trehalose, the major blood sugar of Bombyx mori, in a dose-dependent manner (Mizoguchi et al., in preparation), and it induces meiosis in the ovary (Orikasa et al., 1993). Thus, bombyxin is similar to insulin not only in primary structure but also in biological function (for the effect of insulin to induce meiotic division, see El-Etr et al., 1979).

(d) Chemical synthesis of bombyxin

Since the natural bombyxin is very scarce (50 μg from 650,000 Bombyx heads), efficient synthetic method should be required for further structural and functional studies of bombyxin

such as the three-dimensional structure determination and the production of anti-bombyxin antibody. The peptide chains were easily synthesized by the solid-phase peptide synthesis developed by Merrifield (1964). Because bombyxin is a heterodimeric peptide containing three intramolecular disulfide bonds, regioselective formation of these disulfide bonds was the key of the synthesis of bombyxin. Nagasawa et al. (1988) reported the first chemical synthesis of bombyxin-II, by the random combination of the A- and B-chains. The yield was low (4%), because this method also gave many kinds of by-products such as the oligomers of the A-chain and the B-chain and disulfide bond isomers of bombyxin-II. Maruyama et al. (1992) developed a regioselective disulfide-bond formation method for the synthesis of bombyxin-IV (Figure 1-6). This method utilizes the orthogonal thiol protecting groups (Trt, labile to TFA; MBzl and tBu, stable to TFA and labile to TFMSA; Acm, stable to TFA and TFMSA and labile to I_2), and the stepwise deprotection of these protecting groups forms the three disulfide bonds regioselectively. By this method, bombyxin-IV and -II were obtained in good yields of 48% and 69%, respectively (Maruyama et al., 1992; Nagata et al., 1992b). Owing to this efficient synthetic method, we can now obtain enough amount of bombyxin-II to the analysis of three-dimensional structure, production of antibody and preparation of affinity gel. I synthesized 90 mg of bombyxin-II and many kinds of related molecules, which were used for the above structural and functional studies and structure-activity relationship studies.

1-2. The purpose of the study

(a) The three-dimensional structure of bombyxin-II is essential for understanding its function

As described in 1-1 (b), bombyxin-II, although exogenous, is able to recognize the receptor on the Samia prothoracic glands (referred to as the Samia bombyxin receptor) with high affinity (though K_d is unknown, EC_{50} = 30 pM) and is able to activate the glands to synthesize and release ecdysone, as if it were a prothoracicotropic hormone of Samia cynthia ricini (Figure 1-3B). Despite the sequence similarity between bombyxin-II and human insulin, no cross-activity was observed between them (Nagasawa et al., 1884a; Fernandez-Almonacid and Rosen, 1987; Nagata et al., 1992b). In order to elucidate the structural basis of receptor specificity between bombyxin-II and human insulin and to investigate the molecular evolution of insulin-superfamily peptides, the three-dimensional structure of bombyxin-II must be elucidated and compared in detail with that of insulin. The first purpose of this thesis is therefore to determine the three-dimensional structure of bombyxin-II.

At present, two methods are available for structure determination of a peptide (or a protein) at atomic resolution: (1) X-ray crystallography and (2) nuclear magnetic resonance (NMR) (Wüthrich, 1986). Because the NMR analysis can provide high-resolution structural information in solution and does not need protein crystals, I chose the NMR method for the three-dimensional structure analysis of bombyxin-II. Steps involved in determination of the three-dimensional structure of a peptide by NMR are shown in Figure 1-7. NMR can yield the following structural constraints: (1) distance: NOE $\propto 1/r^6$, (2) dihedral angle: $^3J_{HN\alpha} \rightarrow \phi$, $^3J_{\alpha\beta} \rightarrow \chi 1$ and (3) hydrogen bond: amide exchange. A family of structures consistent with the NMR constraints are calculated.

(b) Identification of the receptor-recognition site of bombyxin-II is also essential for understanding its function

In order to elucidate the molecular basis of specific recognition between bombyxin-II and the Samia bombyxin receptor, identification of the receptor-recognition site on the bombyxin-II molecule is essential. The second purpose of this thesis is therefore to identify and characterize the receptor-recognition site of bombyxin-II.

I synthesized a large large numbers of (1) analogs of bombyxin-II with one or more amino-acid substitutions and (2) chimera molecules of bombyxin-II and human insulin using the regioselective disulfide-bond formation method (Figure 1-6) (Maruyama et al., 1992; Nagata et al., 1992b), and evaluated their potencies in the bombyxin-like activity (the PTTH-like activity to Samia cynthia ricini) using the in vivo Samia pupal assay (Figure 1-2) (Ishizaki and Ichikawa, 1967).

1-3. The arrangement of chapters in this thesis

This thesis consists of seven chapters, of which Chapters 2–5 are the main text. The contents of Chapters 2–5 are summarized in Table 1-1.

In Chapter 2, the three-dimensional structure of bombyxin-II is determined. The structure is compared with those of vertebrate counterparts, human insulin and human relaxin 2. Bombyxin-II is shown to lack the B-chain C-terminal turn and β-strand, which is of critical importance for activity of insulin.

In Chapter 3, the three-dimensional structures of the hybrid molecules of bombyxin-II and human insulin, bonsulin (bombyxin-II A-chain + human insulin B-chain) and imbyxin (human insulin A-chain + bombyxin-II B-chain), are determined. Bonsulin adopts an insulin-like main-chain fold as expected, while imbyxin takes a distorted main-chain fold which is different from those of bombyxin-II and human insulin. A solvent-dependent equilibrium in conformation is observed in imbyxin. Bioassays show that the B-chains rather than the A-chains determined the receptor specificity between bombyxin-II and human insulin.

Table 1-1
Contents

	Contents	
Chapter	Determination of the three-dimensional structure	Localization or identification of the receptor-recognition site of bombyxin-II
Chapter 2	bombyxin-II	
Chapter 3	bonsulin imbyxin	the B-chain (rather than the A-chain)
Chapter 4	bonsylin-(6-18) (bonsulin)	the middle part (B6 to B18) of the B-chain
Chapter 5		B6, B11, B12, B14, B15, B16, B17, B18 (A1, A3, A20-B19)

In Chapter 4, the three-dimensional structure of a chimera molecule of bombyxin-II and human insulin, bonsylin-(6-18), are determined and compared with that of bonsulin. Although they are sequentially different exclusively in the B-chain middle part (B6 to B18), where bonsylin-(6-18) has the bombyxin-type residues while bonsulin has the insulin-type residues. Unexpectedly, the largest conformational difference between them are observed in the sequentially identical B-chain C-terminal part but not in the sequentially different B-chain middle part. Bioassay for bombyxin-like activity showed bonsylin-(6-18) is fully active, while bonsulin is completely inactive. Therefore, the middle part of the bombyxin-II B-chain is of critical importance for bombyxin-like activity.

In Chapter 5, in order to identify the side chains important for bombyxin-like activity in the middle part (B6–B18) of the bombyxin-II B-chain, a series of Ala-scanning analogs of bombyxin-II (the native Ala residues were replaced by the insulin-type residues) are synthesized and their bombyxin-like activity is evaluated. The important residues for bombyxin-like activity are mapped on the three-dimensional structure of bombyxin-II, and the exposed side chains of the important residues are considered to be involved in the recognition of the Samia bombyxin receptor. The identified receptor-recognition site of bombyxin-II is compared with those of the vertebrate counterparts, human insulin and human relaxin 2.

In Chapter 6, "Concluding remarks" are described.

In Chapter 7, "Experimental procedures" are described in detail on some specified topics, although each of Chapters 2–5 include "Materials and methods" section.

Chapter 2
Three-dimensional structure of bombyxin-II, an insulin-related brain-secretory peptide of the silkmoth Bombyx mori: comparison with insulin and relaxin

Summary

Bombyxin, the insulin-related peptide of the silkmoth Bombyx mori, plays important roles in metabolism and reproduction of the moth. The solution structure of bombyxin-II was determined by two-dimensional ^{1}H nuclear magnetic resonance (NMR) spectroscopy and simulated annealing calculations. To our knowledge, this is the first three-dimensional structure determined for an invertebrate insulin-related peptide. The structure of bombyxin-II is similar to that of insulin. However, there are significant differences in the C-terminal region of the B-chain, where bombyxin-II, like relaxin, adopts an extension of a helix instead of a sharp turn followed by a strand as in insulin. The receptor-binding surface of bombyxin-II is proposed based on its structure-activity relationship which is structurally similar to that of relaxin but distinct from that of insulin. The different exposed patches on the surface should confer receptor-recognition specificity between bombyxin and insulin. The structure should provide novel views to the receptor recognition and molecular evolution of insulin-superfamily peptides.

1. Introduction

Several insulin-related peptides of invertebrate origin have so far been identified (Figure 2-1). Bombyxin, a brain-secretory peptide of the silkmoth Bombyx mori, was identified as the first insulin-related peptide of invertebrate origin (A-chain, 50%; B-chain, 32% identical to those of human insulin) (Nagasawa et al., 1984a, 1986). Bombyxin lowers the concentration of haemolymph trehalose, the major blood sugar of Bombyx mori, in a dose-dependent manner (Mizoguchi et al., unpublished data), and induces meiosis in the ovary (Orikasa et al., 1993). Hence, bombyxin is similar to insulin not only in primary structure but also in biological function (for the effect of insulin to induce meiotic division, see El-Etr et al., 1979). In addition, when administered to a brain-removed dormant pupa of the saturniid moth Samia cynthia ricini, bombyxin promotes adult development of the pupa by stimulating the prothoracic glands to synthesize and release ecdysteroid, the insect molting hormone (Nagasawa et al., 1984b). Five molecular species of bombyxin have so far been isolated from the heads of the silkmoth Bombyx mori using the Samia pupal assay (Ishizaki and Ichikawa, 1967); the primary structure is determined completely for bombyxin-II, -IV and partially for bombyxin-I, -III, -V (Nagasawa et al., 1986; Jhoti et al., 1987; Maruyama et al., 1988) (Figure 2-1). Despite the sequence similarity between bombyxin and insulin, no cross-activity was observed between them (Nagasawa et al., 1884a; Fernandez-Almonacid and Rosen, 1987; Nagata et al., 1992b). In order to elucidate the structural basis of receptor specificity between bombyxin and insulin and to investigate the molecular evolution of insulin-superfamily peptides, we have determined the three dimensional structure of bombyxin-II, a representative molecular species of bombyxin, by two-dimensional ^{1}H nuclear magnetic resonance (NMR) spectroscopy and simulated annealing calculations. The structure is compared with those of vertebrate insulin-superfamily peptides, insulin and relaxin, and the implications for receptor-binding determinants and divergent molecular evolution of insulin-superfamily peptides are discussed.

2. Materials and methods

(a) NMR measurements

Bombyxin-II was chemically synthesized by the combination of solid-phase peptide synthesis of the two peptide chains and regioselective formation of the three intramolecular disulfide bonds (Maruyama et al., 1992; Nagata et al., 1992a,b). The synthetic bombyxin-II was dissolved at a concentration of 3 mM in 70%/30% (v/v) $^2H_2O/C^2H_3CO_2{}^2H$ (pH*2.0, pH* indicates direct meter reading) or at 4 mM in 70%/30% (v/v) $H_2O/C^2H_3CO_2{}^2H$ (pH*2.0). ^1H NMR spectra were measured at 600 MHz on a JEOL JNM-α600 spectrometer at 28°C. DQF-COSY (Rance et al., 1983), PE-COSY (Müller, 1987), TOCSY (45 ms mixing time) with a modified DIPSI-2 pulse sequence (Cavanagh and Rance, 1992) and NOESY (75 or 150 ms mixing time) (Jeener et al., 1979; Macura et al., 1981) were recorded in the phase-sensitive mode (States et al., 1982). Water resonance was suppressed by DANTE pulse (Zuiderweg et al., 1986). Two-dimensional spectra were recorded using a data size of 512 (t_1) x 2048 (t_2) (512 x 4096 for PE-COSY) with a spectral width of 6500 Hz. After zero-filling once in the t_2 and twice in the t_1 dimension, 2048 x 2048 real data matrix were finally obtained and digital resolution was 3.2 Hz/point in both dimensions (512 x 4096 real data matrix and 1.6 Hz/point digital resolution in the F_2 dimension for primitive exclusive COSY).

(b) Structure calculations

Interproton distance constraints were derived from NOE crosspeak intensities (peak height) in the NOESY spectra (75 ms mixing time) according to the method of Hatanaka et al. (1994). Crosspeaks in the NOESY spectra were picked and edited with NMRZ (New Methods Research, Inc., Syracuse, NY). The peak intensities were translated into distances on the basis of the relation of NOE intensity $\propto$ (distance)$^{-6}$ and a standard distance of sequential d_{NN} in α helix = 2.8 Å (Wüthrich, 1986). The upper-bound distance constraints were the calculated distance plus 0.5 Å. The lower-bound constraints were set to 1.8 Å. The distances involving methylene and methyl protons and ring protons of tyrosine were referred to as single $(\langle r^{-6} \rangle)^-$

$1/6$ average distances so that no corrections for center averaging were made (Clore et al., 1986). Dihedral angle constraints were obtained based on the analysis of DQF-COSY, PE-COSY and/or NOESY spectra (Wagner et al., 1988). The three-dimensional structures were calculated by the simulated annealing method with X-PLOR (Molecular Simulations, Inc., Walthan, MA) using the distance and dihedral angle constraints. A final set of 10 converged structures was selected from 100 calculations on the basis of agreement with the experimental data and van der Waals energy. A mean structure was obtained by averaging the coordinates of the structures that were superimposed in advance to the best converged structure and then minimizing under the constraints (Clore et al., 1986).

3. Results

(a) Secondary structure

Bombyxin-II was synthesized by the solid-phase peptide-chain synthesis and regioselective disulfide-bond formation (Maruyama et al., 1992; Nagata et al., 1992a,b). It was difficult to dissolve bombyxin-II, like insulin, into H_2O at a concentration above 1 mM between pH 4 and 8 because of its self-association properties. At pH 2.0, bombyxin-II was apparently dissolved at 2 mM into H_2O, but was still aggregated, as evidenced by the line broadening of ^1H resonances (Figure 2-2A). In the case of insulin, addition of organic solvent (20% acetic acid or 35% acetonitrile) was successfully used to minimize the peptide aggregation without destroying the globular structure of the insulin molecule (Kline and Justice, 1990; Hua and Weiss, 1991; Hua et al., 1991). Similarly, the solvent 70%/30% (v/v) H_2O (or ^2H_2O)/$C^2H_3CO_2{}^2H$ reduced the self-association of bombyxin-II and allowed spectra of good-quality to be obtained for at least a few months. The addition of acetic acid did not cause a remarkable change in the pattern of ^1H NMR spectra (Figure 2-2B), indicative of overall conservation of native conformation of the peptide. Hence, the two-dimensional ^1H NMR spectra were measured in the presence of 30% $C^2H_3CO_2{}^2H$ at pH 2.0 (direct meter reading), peptide concentrations of 3–4 mM and 28°C.

The resonances were assigned to individual protons in a sequence-specific manner using the sequential assignment method (Table 2-1) (Wüthrich, 1986). The sequential assignment method consists of two stages. The first stage of assignment involves the identification of the systems of spin-spin coupled resonances which belong to a particular amino acid residue. This was achieved using DQF-COSY and TOCSY experiments (Figure 2-3). The second stage of assignment involves the assignment of an amino-acid spin systems identified in the first stage to a specific residue in the peptide. This sequence-specific assignment is achieved by correlating one amino acid spin system with the spin systems of its neighboring residues in the sequence. There is no resolvable spin-spin coupling between protons of adjacent residues and, therefore, COSY-type spectra can not be used to delineate the sequential connectivities. Instead, this stage

of assignment relied on the short-range through-space connectivities observed in NOESY spectra (Figure 2-4). The successive strong $\underline{d}_{NN}$, $\underline{d}\alpha N(i,i+3)$ and $\underline{d}\alpha\beta(i,i+3)$ NOE connectivities are characteristic of α-helix (Wüthrich, 1986). The analysis of these NOE connectivities revealed that bombyxin-II was composed of three α-helices in the A-chain N-terminal region (H_{AN}, residues IleA2 to LeuA8), in the A-chain C-terminal region (H_{AC}, residues ValA13 to TyrA19) and in the B-chain central region (H_B, residues ArgB9 to AlaB22) (Figure 2-4).

(b) Tertiary structure

A total of 535 distance constraints which included 229 intraresidue, 138 sequential ($|i - j| = 1$), 132 short-range ($2 \leq |i - j| \leq 5$) and 36 long-range ($|i - j| \geq 6$) constraints were derived from the assigned NOE crosspeaks measured with a mixing time of 75 ms. Dihedral angle constraints including 18 ϕ and 6 $\chi 1$ were obtained. The three-dimensional structures were calculated with X-PLOR (Brünger, 1990) using the simulated annealing protocol (YASAP) on the 563 above-mentioned experimental constraints and 3 distance constraints of the disulfide bonds (Nagasawa et al., 1988). A total of 100 calculations were carried out, and a final set of 10 structures was selected on the basis of agreement with the experimental constraints and van der Waals energy, with the cut-off taken at $F_{NOE} + F_{repel} < 197.00$ kcal/mol (Table 2-2). The number of inter-residue distance constraints and average root-mean-square deviations (RMSDs) around the mean structure for each residue (Figure 2-5A) and the Ramachandran plot for the 10 structures (Figure 2-5B) are shown. The structure was well-defined except for the N-terminal region of the B-chain (residues pGluB(-2) to HisB4) and the peptide-chain termini (residues GlyA1, CysA20 and GlyB23 to AspB25) (Figure 2-6A). The root-mean-square deviations (RMSDs) between the final 10 structures and the mean structure were 0.58 ± 0.15 Å for the backbone heavy atoms (N, C^{α}, C') and 1.03 ± 0.18 Å for all non-hydrogen atoms in the well-defined regions (residues IleA2 to TyrA19 and ThrB5 to AlaB22).

The A-chain of bombyxin-II consists of two antiparallel helices, H_{AN} (residues IleA2 to LeuA8) and H_{AC} (residues ValA13 to TyrA19), which are connected by a loop (residues

ArgA9 to SerA12) (Figure 2-6A). The B-chain contains an N-terminal less well-defined region (residues pGluB(-2) to HisB4), an extended arm (residues ThrB5 to GlyB8), a central helix (H_B, residues ArgB9 to AlaB22), and a C-terminal coil (residues GlyB23 to AspB25). The structure including the three helices is stabilized by the three disulfide bonds (CysA6-CysA11, CysA7-CysB7 and CysA20-CysB19) and a hydrophobic core (residues IleA2, LeuA16, TyrA19, TyrB6, LeuB11, LeuB15 and LeuB18) (Figure 2-6B).

It should be noted that a three-dimensional structure of bombyxin-II was proposed using interactive computer graphics and energy minimization techniques, assuming homology with porcine insulin, the structure of which was determined by X-ray analysis (Figure 1-5C) (Jhoti et al., 1987). In the modeled structure, the B-chain C-terminal region of bombyxin-II adopts a type III turn between CysB19 and AlaB22 and an extended C-terminal segment in a similar way to human insulin. However, in the solution structure of bombyxin-II, the B-chain central helix continues to AlaB22 and the C-terminus (residues GlyB23 to AspB25) adopts a coiled structure.

4. Discussion

(a) Structure comparison with insulin and relaxin

The overall main-chain fold of bombyxin-II in solution is similar to those of insulin in solution (Brookhaven Protein Data Bank entry 1HIU, Hua et al., 1991), insulin in the crystalline T-state (4INS, Baker et al., 1988) and relaxin in crystal (6RLX, Eigenbrot et al., 1991) (Figure 2-7); the root-mean-square deviations are 1.31 Å, 1.28 Å and 1.45 Å, respectively, for the main-chain atoms (N, C^{α}, C') within the common helical regions (residues A2 to A8, A13 to A19 and B9 to B19 or B22). The common structural features include: (1) an A-chain with two helices joined by an extended loop, (2) a B-chain with an extended N-terminus followed by a central helix, (3) three disulfide bonds and (4) a hydrophobic cluster.

Seven residues thoroughly conserved in the insulin-superfamily peptides (the six Cys residues and GlyB8) are essential to construct the characteristic backbone fold of the insulin-superfamily peptides. The linkages of the three disulfide bonds in bombyxin-II are identical to those in insulin, IGFs and relaxin (CysA6-CysA11, CysA7-CysB7, CysA20-CysB19) (Nagasawa et al., 1988). GlyB8, with a positive ϕ angle (Figure 2-5B), enables the main chain to turn sharply after CysB7 to form the helix from the residue at B9. The characteristic "insulin core" structure is stabilized by a hydrophobic cluster including the residues at A2 (Ile in bombyxin-II/Ile in human insulin/Leu in human relaxin 2), A16 (Leu/Leu/Leu), A19 (Tyr/Tyr/Phe), B6 (Tyr/Leu/Leu), B11 (Leu/Leu/Leu), B15 (Leu/Leu/Gln), B18 (Leu/Val/Ile), which are highly conserved as hydrophobic through the superfamily, and the A6-A11, A20-B19 disulfide bonds (Figures 2-1 and 2-6B).

Despite the overall structural similarity, the structure of the B-chain C-terminal region of bombyxin-II is different from that of insulin (Figure 2-7). Bombyxin-II, like relaxin, adopts a helix and a coiled structure, instead of a sharp turn and an extended β-strand as do insulin and IGFs (Cooke et al., 1991; Sato et al., 1993; Terasawa et al., 1994). Insulin has two Gly residues at B20 and B23, whose ϕ angles are positive in crystal (Blundell, 1972), whereas

bombyxin and relaxin have one (B23 in bombyxin-II, -V; B20 in relaxin) or no (in bombyxin-IV) Gly residue at these positions, indicating that both the Gly residues are required to form a sharp turn. Furthermore, the extended structure of the insulin B-chain C-terminus is stabilized by the intramolecular hydrophobic interactions between PheB24 and ValB12, LeuB15, TyrB16 and between TyrB26 and LeuB11, IleA2, ValA3 (Jørgensen et al., 1992), whereas in bombyxin-II and relaxin, which lack the residue at B26, such interactions should be appreciably weakened. A high-potency monomeric insulin analog, des-pentapeptide(B26-B30)-insulin, in which the B-chain terminates at B25 as in bombyxin-II and relaxin, still adopts a turn and extended structure in the B-chain C-terminal region (Bi et al., 1984; Hua and Weiss, 1990, 1991; Hua et al., 1992). Therefore, not due to the truncation of the B-chain C-terminus but due to lack of a Gly residue at either B20 or B23, the B-chain C-terminal region of bombyxin-II and relaxin takes an extension of the B-chain helix rather than an insulin-like turn. All of the invertebrate insulin-related peptides so far characterized except sponge insulin (Figure 2-1) (Robitzki et al., 1989) and hystricomorph (guinea pig, coypu, casiragua, cuis) insulins lack in Gly at B20 (Blundell and Wood, 1975), which suggests that they might take a helix extention as do bombyxin-II and relaxin.

(b) Receptor-binding determinants

The structure-activity relationship studies of bombyxin show that the residues at the A-chain N-terminus (GlyA1) (Maruyama, 1991) and on the B-chain central helix (HisB10, ArgB13 and/or AspB17) (Nagata et al., unpublished data) are important for receptor binding, while the residues in the B-chain N-terminus (pGluB(-2)-GlnB(-1)-ProB0-GlnB1) and the B-chain C-terminus (AlaB22-GlyB23-ValB24-AspB25) are not required for receptor binding (Minoru Tanaka, personal communication; Maruyama, 1991). Although there are only a small number of experimental data, mapping of these important residues on the structure of bombyxin-II localizes the putative receptor-binding surface to the A-chain N-terminus and the A-chain C-terminus and the B-chain central helix. In the case of insulin, extensive studies on the structure-function relationship have been made and the residues in the A-chain N-terminus

(GlyA1-IleA2-ValA3-AspA4-GluA5), the A-chain C-terminus (TyrA19 and AsnA21), the B-chain helix (ValB12 and TyrB16) and the B-chain C-terminal β-strand (PheB24-PheB25-TyrB26) have been shown to be important for the receptor recognition (Blundell et al., 1972; Pullen et al., 1976; Tager, 1987, 1990, Murray-Rust et al., 1992). In the case of relaxin, the residues ArgB9 and ArgB13 on the B-chain helix have been demonstrated to be involved in receptor binding (Büllesbach and Schwabe, 1988). In addition, the residues TyrA(-1), PheA19, ValB12, GlnB15 and IleB16 have been proposed as the receptor-recognition surface of relaxin (Eigenbrot et al., 1991). It is to be noted that these two Arg residues are also conserved in bombyxin-II, suggesting that they may be involved in the receptor binding of bombyxin-II. Therefore, the three molecules are considered to have common receptor-recognition sites which include the A-chain N-terminus, the A-chain C-terminus, the B-chain central helix and, in the case of insulin, the additional site of the B-chain C-terminal β-strand (Figure 2-8). The putative receptor-recognition surface of bombyxin-II and relaxin includes an exposed hydrophobic patch which is surrounded by polar and charged residues, and is distinct from that of insulin in not only the conformation of the B-chain C-terminal section but also in the distribution of the side-chain functional groups (Figure 2-8A). In contrast, the hydrophobic surface in insulin is covered by the B-chain C-terminal strand, which makes a characteristic patch including an aromatic cluster (residues PheB24-PheB25-TyrB26) on the proposed receptor-recognition surface of insulin (Figure 2-8B), and plays a pivotal role in the expression of insulin activity (Nakagawa and Tager, 1986, 1987, 1993; Mirmira and Tager, 1989, 1991; Derewanda et al., 1990; Hua et al., 1991; Mirmira et al., 1991). Hence, the B-chain C-terminal regions of bombyxin/relaxin and insulin are obviously different from each other both structurally and functionally. The different exposed patches on the surface should confer the specificity in receptor recognition on bombyxin, relaxin and insulin (Nagasawa et al., 1984a; Fernandez-Almonacid and Rosen, 1987; Nagata et al., 1992b).

Based on the solution structures of human insulin and its active mutant, [GlyB24]human insulin, Hua et al. (1991) proposed a model for the receptor recognition of insulin that when insulin binds to the receptor, the detachment of the B-chain C-terminal β-strand from the core

should occur which reorganizes the protein surface, from the locked (inactive) state to the unlocked (active) state, exposing side chains that are strictly conserved (IleA2-ValA3) in the N-terminal α-helix of the A-chain (Baker et al., 1988; Mirmira and Tager, 1989; Derewanda et al., 1990; Hua et al., 1991) (Figure 2-8B). In the unlocked state but not in the locked state, these hydrophobic side chains would be accessible for direct contact with the insulin receptor, which could afford insulin a high affinity for the receptor (Hua et al., 1991). In contrast to insulin, bombyxin-II and relaxin can take only the unlocked state, with the hydrophobic surface including the hydrophobic side chains at A2-A3 (IleA2-ValA3 in bombyxin-II; LeuA2-AlaA3 in human relaxin 2) exposed to solvent (Figure 2-8B); therefore, they need not reorganize the molecular surface when binding to respective receptors. Hence, we propose that bombyxin-II should recognize its receptor, in a similar manner to relaxin but in a distinct way to insulin, by the exposed hydrophobic patch (including residues IleA2, ValA3, TyrA19, AlaB12, LeuB15, AlaB16 and A20-B19 disulfide bond) and its surrounding polar and charged groups (including residues GlyA1, ArgB9, HisB10, ArgB13, AspB17) in the A-chain N- and C-termini and on the exposed side of the B-chain central helix, without involvement or conformational change of the B-chain C-terminal section.

(c) Phylogeny of insulin-superfamily peptides on the structural basis

The structural and functional differences of the B-chain C-terminal region of bombyxin-II/relaxin and insulin suggest that bombyxin/relaxin might have evolved a distinct mechanism of ligand-receptor recognition from that of insulin. Although the core structure of vertebrate insulin-superfamily peptides ("the insulin core") is conserved in bombyxin-II/relaxin, they have evolved a distinctive receptor-recognition patch from that of insulin, which confers the biological specificity between them. The receptor-recognition patch of bombyxin-II is similar to that of relaxin; examination of functional relationship between bombyxin and relaxin is a future topic for study. The receptor-binding site of the ancestral molecule probably involved the common framework (the A-chain N- and C-termini and the B-chain central helix) (Murray-Rust,J., et al., 1992), and the involvement (residues PheB24-PheB25-TyrB26) and associated

conformational changes in the B-chain C-terminus of insulin (Baker et al., 1988; Mirmira and Tager, 1989; Derewanda et al., 1990; Hua et al., 1991) may have appeared at a later stage in evolution to distinguish insulin from bombyxin/relaxin. The molecular phylogenetic tree of the insulin-superfamily peptides was constructed based on the sequence similarity which indicated that bombyxin was more closely related to insulin than to relaxin (Murray-Rust et al., 1992). However, our data on the three dimensional structure and proposition for the mechanisms of receptor recognition of bombyxin-II demonstrate that bombyxin-II is more closely related to relaxin rather than to insulin. The presence of more than thirty molecular species of bombyxin in the silkmoth Bombyx mori with various arrangements of polar or charged side chains on the B-chain helix (Kondo et al., unpublished data) suggests that bombyxins with various physiological functions might be specified by their receptors by the arrangements of the exposed side chains on the B-chain helix in the moth rather than by the involvement of B-chain C-terminus in receptor recognition as the cases of insulin and IGFs in vertebrates. The structure of bombyxin-II should provide novel views to the receptor recognition and divergent molecular evolution of insulin-superfamily peptides. The atomic coordinates of the 10 calculated structures and the averaged energy minimized structure will be deposited in the Brookhaven Protein Data Bank (1BON and 1BOM, respectively).

Chapter 3
Structure and activity of bonsulin and imbyxin, the hybrid molecules of bombyxin-II and human insulin

Summary

In spite of 40% sequence similarity, bombyxin-II, an insulin-like peptide of the silkmoth Bombyx mori, and human insulin are not cross-active to each other. To localize the receptor-specificity determinants between them, we synthesized their hybrid molecules: bonsulin (bombyxin-II A-chain–human insulin B-chain) and imbyxin (human insulin A-chain–bombyxin-II B-chain). Biological evaluation of the hybrid molecules has revealed that their B-chains determine the receptor specificity while their A-chains are partly interchangeable. We expected that these hybrid molecules would have a characteristic core structure shared by insulin-superfamily peptides. Bonsulin retains the core structure characteristic of insulin-superfamily peptides as expected. But, imbyxin unexpectedly have a distorted structure. The entire α-helix in the A-chain N-terminal part and the first α-helix turn in the B-chain middle part are lost and less well-defined in conformation in imbyxin. CD indicate imbyxin is in a TFE-dependent conformational equilibrium: without TFE, the distorted conformation is dominant, while with 30% TFE, imbyxin appears to take a bombyxin-like conformation. This conformational study demonstrate that an artificially designed molecule does not always fold as expected.

3-1. Introduction

Bombyxin is a brain-secretory peptide of the silkmoth Bombyx mori, which was identified as the first insulin-related peptide of invertebrate origin (Nagasawa et al., 1984a). Bombyxin-II, a representative molecular species of bombyxin, consists of two peptide chains (A-chain, 20 residues; B-chain, 28 residues) sequentially similar to those of insulin (50% and 32% identical, respectively, to those of human insulin) and three disulfide bonds linked in the same way as in insulin (A6-A11, A7-B7 and A20-B19) (Nagasawa et al., 1986; Nagasawa et al., 1988). Furthermore, bombyxin-II has a characteristic core structure similar to those of vertebrate insulin-superfamily peptides (Chapter 2). Despite the structural similarity between bombyxin-II and human insulin, they are not cross-active to each other. In order to localize the receptor-specificity determinants, we have synthesized the hybrid molecules of human insulin and bombyxin-II, and examined their bombyxin-like and insulin-like activities. The bioassays of the hybrid molecules demonstrated that the receptor-specificity determinants between insulin and bombyxin lie in their B-chain rather than their A-chains (Table 3-1). In this chapter, I describe the three-dimensional structure determination of the hybrid molecules, and discuss the relationships between their conformations and activity.

3-2. Experimental procedures

(a) NMR measurements

The synthetic bonsulin or imbyxin was dissolved at a concentration of 3 mM in 70%/30% (v/v) $^2H_2O/C^2H_3CO_2{}^2H$ (pH*2.0, pH* indicates direct meter reading) or at 4 mM in 70%/30% (v/v) $H_2O/C^2H_3CO_2{}^2H$ (pH*2.0). ^1H NMR spectra were measured at 600 MHz on a JEOL JNM-α600 spectrometer at 28°C. DQF-COSY (Rance et al., 1983), PE-COSY (Müller, 1987), TOCSY (45 ms mixing time) with a modified DIPSI-2 pulse sequence (Cavanagh and Rance, 1992) and NOESY (75 or 150 ms mixing time) (Jeener et al., 1979; Macura et al., 1981) were recorded in the phase-sensitive mode (States et al., 1982). Water resonance was suppressed by DANTE pulse (Zuiderweg et al., 1986). Two-dimensional spectra were recorded using a data size of 512 (t_1) x 2048 (t_2) (512 x 4096 for PE-COSY) with a spectral width of 6500 Hz. After zero-filling once in the t_2 and twice in the t_1 dimension, 2048 x 2048 real data matrix were finally obtained and digital resolution was 3.2 Hz/point in both dimensions (512 x 4096 real data matrix and 1.6 Hz/point digital resolution in the F_2 dimension for PE-COSY).

(b) Structure calculations

Interproton distance constraints were derived from NOE crosspeak intensities (peak height) in the NOESY spectra (75 ms mixing time) according to the method of Hatanaka et al. (Hatanaka et al., 1994). Crosspeaks in the NOESY spectra were picked using a homemade C program and edited with Felix (Biosym Technologies, Inc., San Diego, CA). The peak intensities were translated into distances on the basis of the relation of NOE intensity ∞ $(\text{distance})^{-6}$ and a standard distance of sequential d_{NN} in α helix = 2.8 Å (Wüthrich, 1986). The upper-bound distance constraints were the calculated distance plus 0.5 Å. The lower-bound constraints were set to 1.8 Å. The distances involving methylene and methyl protons and ring protons of tyrosine were referred to as single $(\langle r^{-6} \rangle)^{-1/6}$ average distances so that no corrections for center averaging were made (Clore et al., 1986). Dihedral angle constraints

were obtained based on the analysis of DQF-COSY, PE-COSY and/or NOESY spectra (Wagner et al., 1988). The three-dimensional structures were calculated by the simulated annealing method with X-PLOR (Molecular Simulations, Inc., Walthan, MA) using the distance and dihedral angle constraints. A final set of 10 converged structures was selected from 100 calculations on the basis of agreement with the experimental data and van der Waals energy. A mean structure was obtained by averaging the coordinates of the structures that were superimposed in advance to the best converged structure and then minimizing under the constraints (Clore et al., 1986).

(c) CD measurements

CD spectra were recorded on a JASCO J-600 spectropolarimeter in the wavelength range of 200–250 nm with 1-mm path length cells at room temperature. Four scans were accumulated. Samples were dissolved in 50 mM sodium phosphate buffer (pH 6.8)/50 mM NaCl at a residue concentration of 20 μM. In the TFE addition experiment, imbyxin was dissolved in 50 mM sodium phosphate buffer (pH 6.8) and 0, 10, 20, 30, 40, 50% (v/v) TFE at a residue concentration of 20 μM. The peptide concentration was determined using the sum of molar extinction coefficient at 280 nm of each residue: Trp (5400 l/cm/mol), Tyr (1100 l/cm/mol) and disulfide (200 l/cm/mol).

3-3. Results

(a) Three-dimensional structure of bonsulin and imbyxin

Bioassay revealed that bonsulin and imbyxin showed an weak insulin-like activity and an weak bombyxin-like activity, respectively (Table 3-1) (Nagata et al., 1992). These data indicate that the functional specificity is determined by their B-chains rather than their A-chains. In order to investigate the relationships between biological activity and three-dimensional structure of the hybrid molecules, three-dimensional structures of bonsulin and imbyxin were analyzed by NMR.

The two-dimensional ^1H NMR spectra of bonsulin and imbyxin were measured in the presence of 30% $C^2H_3CO_2{}^2H$ as in the case of bombyxin-II: 4 mM peptide dissolved in 70% H_2O/30% $C^2H_3CO_2{}^2H$ at pH* 2.0 (pH* indicates direct meter reading) and 3 mM peptide dissolved in 70% H_2O/30% $C^2H_3CO_2{}^2H$ at pH* 2.0. Two-dimensional ^1H NMR spectra, DQF-COSY, TOCSY (45 ms mixing time), NOESY (75-ms or 150-ms mixing time) and PE-COSY, were measured at 600 MHz on a JEOL JNM-α600 spectrometer. The resonances were assigned to individual protons in a sequence-specific manner using the sequential assignment method (Tables 3-2 and 3-3). The successive strong d_{NN}, $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ NOE connectivities were characteristic of α-helix (Wüthrich, 1986). The analysis of these NOE connectivities revealed that bonsulin was composed of three α-helices in the A-chain N-terminal region (H_{AN}, residues IleA2 to LeuA8), in the A-chain C-terminal region (H_{AC}, residues ValA13 to TyrA19) and in the B-chain central region (H_B, residues ArgB9 to CysB19), while imbyxin was composed of two α-helices in the A-chain C-terminal region (H_{AC}, residues ValA13 to TyrA19) and in the B-chain central region (H'_B, residues AlaB11 to AlaB22) (Figure 3-1). The peptide chains in bonsulin take a similar secondary structure to those in the original molecule, while the peptide chains in imbyxin take a different secondary structure to those in the original molecules. The N-terminal helix (IleA2 to ThrA8) in the human insulin A-chain and the first helix turn (ArgB9 to LeuB11) in the bombyxin-II B-chain was lost in

imbyxin, indicating that imbyxin should take a distinctly different main-chain conformation from bombyxin-II and human insulin.

In the case of bonsulin, a total of 623 distance constraints which included 271 intraresidue, 164 sequential ($|i - j| = 1$), 141 short-range ($2 \leq |i - j| \leq 5$) and 47 long-range ($|i - j| \geq 6$) constraints were derived from the assigned NOE crosspeaks measured with a mixing time of 75 ms, and dihedral angle constraints for 26 ϕ and 8 $\chi 1$ were obtained. In the case of imbyxin, a total of 509 distance constraints which included 228 intraresidue, 162 sequential, 83 short-range and 36 long-range constraints were derived from the assigned NOE crosspeaks measured with a mixing time of 75 ms, and dihedral angle constraints for 22 ϕ were obtained. The three-dimensional structures were calculated with X-PLOR (Brünger, 1990) using the simulated annealing protocol (YASAP) on the experimental constraints and 3 distance constraints of the disulfide bonds (CysA6-CysA11, CysA7-CysB7 and CysA20-CysB19). A total of 100 calculations were carried out, and a final set of 10 structures was selected on the basis of agreement with the experimental constraints and van der Waals energy (Tables 3-4 and 3-5). The number of inter-residue distance constraints and average root-mean-square deviations (RMSDs) around the mean structure for each residue (Figure 3-2) and the Ramachandran plot for the 10 structures (Figure 3-3) are shown. The structure of bonsulin was well-defined except for the peptide-chain termini (residues GlyA1, CysA20, PheB1 to AsnB3 and LysB29-ThrB30) (Figure 3-4); the RMSDs between the final 10 structures and the mean structure were 0.71 ± 0.11 Å for the backbone heavy atoms (N, C^{α}, C') and 1.59 ± 0.22 Å for all non-hydrogen atoms in the well-defined regions (residues IleA2 to TyrA19 and GlnB4 to ProB28) (Figure 3-4A). On the other hand, the structure of imbyxin was well-defined in the C-terminal halves of the A-chain (residues LeuA13 to TyrA19) and of the B-chain (residues AlaB12 to AlaB22), but was disordered in the N-terminal halves of the A-chain (residues GlyA1 to SerA12) and of the B-chain (residues pGluB(-2) to LeuB11) (Figure 3-4B). The average pairwise root-mean-square deviation (RMSD) was 0.82 Å for the backbone heavy atoms (N, C^{α}, C') and 1.58 Å for all non-hydrogen atoms in the well-defined regions (residues LeuA13 to TyrA19 and AlaB12 to AlaB22). Although the A- and B-chains of imbyxin were

sequentially identical to the A-chain of human insulin and the B-chain of bombyxin-II, respectively, these peptide chains took different conformations in imbyxin from what they took in human insulin and bombyxin-II (Figure 3-1B).

(b) CD spectra of bonsulin and imbyxin

Figure 3-5 shows the CD spectra of human insulin, bombyxin-II and their hybrid molecules, bonsulin and imbyxin dissolved in 50 mM sodium phosphate buffer (pH 6.8)/50 mM NaCl. The CD spectrum of bonsulin was similar to that of human insulin and bombyxin-II, while the CD spectrum of imbyxin was dissimilar to the CD spectra of the other molecules, indicating the helix content of imbyxin is lower than the others'. Thus the CD spectra of bonsulin and imbyxin dissolved in 50 mM sodium phosphate buffer (pH 6.8)/50 mM NaCl and the NMR structures of bonsulin and imbyxin dissolved in 70%/30% $C^2H_3CO_2{}^2H/H_2O$ (pH 2.0) are consistent. A serial addition of TFE (0, 10, 20, 30, 40 and 50%) to imbyxin dissolved in 50 mM sodium phosphate buffer (pH 6.8) changed the CD spectrum of imbyxin to a characteristic one of the insulin-superfamily peptides. The result indicates that imbyxin can take a bombyxin-II-like main-chain conformation in the presence of 30% TFE. After addition of 50% TFE, the TFE concentration was then lowered gradually by buffer exchange using a Centricon 3 microconcentrator (Amicon), and finally the TFE was almost entirely removed. The CD spectrum of imbyxin after the TFE removal was essentially the same as that of imbyxin before the TFE addition. Therefore, the conformation of imbyxin changed reversibly depending on the solvent conditions: in the absence of TFE, imbyxin takes distorted conformations in aqueous solution, while in the presence of 30% TFE, imbyxin can adopt a bombyxin-II-like main-chain conformation. The distorted main-chain conformation is more stable for imbyxin in aqueous solution without TFE than the bombyxin-II-like main-chain conformation. TFE is utilized as an α-helix stabilizer or inducer (Creighton, because it is poor in ability to form the hydrogen bond, and thus will strengthen the intramolecular local hydrogen bonds in peptides. In spite of the distorted conformations of imbyxin in aqueous solution, imbyxin retains a weak

bombyxin-like activity. Therefore, imbyxin should contain requirements for recognition of the Samia bombyxin receptor.

3-4. Discussion

(a) Conformational comparison of the hybrid molecules with bombyxin-II and human insulin

We have synthesized hybrid molecules of human insulin and bombyxin-II to localize the receptor-recognition specificity determinants (Nagata et al., 1992b). Bioassays revealed that bonsulin (bombyxin-II A-chain + human insulin B-chain) possessed insulin-like activity exclusively, while imbyxin (human insulin A-chain + bombyxin-II B-chain) possessed bombyxin-like activity exclusively, which indicated that the receptor-specifity determinants lie in the B-chains (Nagata et al., 1992b). Models of the three-dimensional structure of bonsulin and imbyxin was constructed on the basis of three-dimensional structures of human insulin and bombyxin-II. The model building suggested that both the hybrid molecules, bonsulin and imbyxin, could assume the core structure characteristic of the insulin-superfamily peptides, because human insulin and bombyxin-II have a similar core structure. But the three-dimensional structure determination of imbyxin demonstrates that it does not take the characteristic core structure but rather a distorted structure. The A-chain N-terminal helix (H_{AN}, residues IleA2 to ThrA8) and the first turn (ArgB9 to LeuB11) of the B-chain middle–C-terminal helix were lost and disordered in conformation.

(b) TFE-dependent equilibrium in conformation of imbyxin

The CD spectra of imbyxin changes reversibly depending on the concentration of TFE and becomes similar to those of bombyxin-II and human insulin in 30% of TFE. Therefore, the solvent-dependent equilibrium in the conformation of imbyxin is proposed (Figure 3-6): the conformer I, which has the distorted conformation, is dominant in aqueous solution without TFE, while the conformer II, which has the bombyxin-like main-chain conformation is dominant in aqueous solution containing 30% TFE. This result demonstrates that the heterologous A- and B-chains do not always match well, which is very instructive to me, for I will go on the structure-activity relationship studies of chimera molecules of bombyxin-II and

human insulin (Chapter 4) to localize the requirements for recognition of the Samia bombyxin receptor.

Although imbyxin takes a distorted conformation in the absence of TFE, it shows a weak bombyxin-like activity (0.002-fold that of bombyxin-II) and no insulin-like activity. The result can be interpreted that no conformers of imbyxin are capable of recognizing the human insulin receptor, while some conformers are capable of recognizing the Samia bombyxin receptor specifically. Therefore, despite the distorted conformation of imbyxin, it is confirmed that the A-chains of bombxyin-II and human insulin are interchangeable, while their B-chains are not, and the B-chains determine the receptor specificity between bombyxin-II and human insulin (Nagata et al., 1992b). In Chapter 4, I further localize the important site for the bombyxin-like activity using the chimera molecules of bombyxin-II and human insulin.

Chapter 4

Localization and characterization of the molecular surface of bombyxin-II required for recognition of the Samia cynthia ricini bombyxin receptor

Critical importance of the B-chain middle part

Summary

Bombyxin-II, a brain-secretory peptide of the silkmoth Bombyx mori, shares 40% sequence identity and the characteristic core structure with human insulin. In spite of the structural similarity, no cross-activity is seen between them. To localize the receptor-recognition region of bombyxin-II, chimera molecules of bombyxin-II and human insulin are synthesized, and their bombyxin-like activity was evaluated. Two chimera molecules which were sequentially identical except for the B-chain middle part possessed distinct potencies in bombyxin-like activity. Bonsylin-(6-18), which possessed the B-chain middle part of bombyxin-II, was fully active, whereas bonsulin, which possessed the B-chain middle part of human insulin, was completely inactive. The solution structure determination of bonsylin-(6-18) and bonsulin demonstrated that their B-chain middle parts took similar main-chain conformations but formed dissimilar patches. Therefore, the patch formed by the middle part of bombyxin-II B-chain is of critical importance for recognition of the bombyxin receptor.

1. Introduction

Bombyxin is a brain-secretory peptide of the silkmoth Bombyx mori, which was identified as the first insulin-related peptide of invertebrate origin (Nagasawa et al., 1984a). Bombyxin-II, a representative molecular species of bombyxin, consists of two peptide chains (A-chain, 20 residues; B-chain, 28 residues) sequentially similar to those of insulin (50% and 32% identical, respectively, to those of human insulin) and three disulfide bonds linked in the same way as in insulin (A6-A11, A7-B7 and A20-B19) (Figure 4-1) (Nagasawa et al., 1986; Nagasawa et al., 1988). Furthermore, bombyxin-II has a characteristic core structure similar to those of vertebrate insulin-superfamily peptides (Chapter 2). Bombyxin induces adult development of brain-removed dormant pupae of the saturiid moth Samia cynthia ricini, a relative species of Bombyx mori, when injected into the pupae, by binding to its "receptor" on the prothoracic glands to stimulate the synthesis and release of ecdysone, a steroid hormone required for insect molting and metamorphosis (Nagasawa et al., 1984a,b). Since the production of bombyxin-related peptides (referred to as the Samia bombyxin-related peptides or SBRPs) in the Samia brain was demonstrated (Kimura-Kawakami et al., 1992), the intrinsic role of the "bombyxin receptor" on the Samia prothoracic glands can be the receptor for the intrinsic SBRPs. Naturally, bombyxin and SBRPs should recognize the receptor on the Samia prothoracic glands in a similar way. In contrast, bombyxin is inactive to the intrinsic prothoracic glands of Bombyx mori (Ishizaki et al., 1983; Kiriishi et al., 1992). The functions of bombyxin in Bombyx mori so far demonstrated are to lower the concentration of haemolymph trehalose, the major blood sugar of the moth (Mizoguchi et al., unpublished data), and to induce meiosis in the ovary (Orikasa et al., 1993). Despite the structural similarity between bombyxin-II and human insulin, they have no cross-activity: bombyxin-II has no affinity to the insulin receptor (Fernandez-Almonacid and Rosen, 1987), whereas human insulin has no affinity to the Samia bombyxin receptor (Nagasawa et al., 1984a). In order to elucidate the structural basis of the receptor specificity, (1) the bombyxin-like activity of bombyxin-II–human insulin chimera molecules was evaluated to localize the region important for the Samia bombyxin receptor-

recognition, (2) the three-dimensional structures of the two chimera molecules with 84% sequence identity but distinct potencies were determined to identify the surface structure required for the Samia bombyxin receptor-recognition and (3) the receptor-recognition surfaces of bombyxin-II, human insulin and human relaxin were compared to characterize each surface structure.

2. Materials and methods

(a) Peptide synthesis

Bombyxin-II, human insulin and their hybrid molecules were synthesized by the combination of solid-phase peptide synthesis of the peptide chains and regioselective formation of the three intramolecular disulfide bonds according to the method (Maruyama et al., 1992; Nagata et al., 1992b). All the molecules were successfully obtained in high yields, for instance, 41% and 38% from the constituent peptide chains in the cases of bonsylin-(6-18) and bonsulin, respectively. Peptide identity was confirmed by FAB-MS or MALDI-TOF-MS measurement.

For the synthesis of des-octapeptide(B23-B30)-bonsulin, bonsulin was digested by trypsin. Lyophilized bonsulin (128 μg, 23 nmol) was dissolved in 370 μl of 0.1 M Tris-HCl buffer (pH 7.8) containing 0.01 M $CaCl_2$/CH_3CN (9:1, v/v). To the peptide solution, 92 μl of trypsin solution (1.0 mg dissolved in 10 ml of 0.003 M HCl) was added, and the reaction mixture was incubated at 37°C for 24 hr. After the reaction mixture was acidified to pH 2.0 by adding 10% aqueous TFA, the reaction product was purified by reverse-phase HPLC [HPLC system, JASCO Gulliver LC-900 system; column, SenshuPak Pegasil ODS (4.6 x 150 mm); eluent, 20%-40% (80-min linear gradient) CH_3CN in 0.1% TFA; column temperature, 40°C; detection, absorbance at 280 nm]. Under the elution conditions, bonsulin and des-octapeptide(B23-B30)-bonsulin were eluted at the retention times of 54.2 min and 49.0 min, respectively. HPLC analysis of the product revealed that the reaction gave des-octapeptide(B23-B30)-bonsulin as the major product (yield, 87 μg, 82%).

(b) Bioassay for bombyxin-like activity

The bioassay was done according to the method (Ishizaki and Ichikawa, 1967). The brains of pupae of the saturniid moth Samia cynthia ricini were surgically removed within 20 hr after pupation. This operation blocked adult development of the pupae, and most pupae should survive as pupae without initiating adult developement for months or even more than a year.

The test material dissolved in 0.1 M Tris-HCl (pH 7.8) containing 0.04% bovine serum albumin (BSA) was injected (10 µl per pupa) into the brain-removed pupae 2–5 months after the brain removal. During the period the sensitivity of brainless pupae to bombyxin-II, which was used as the positive control, did not change appreciably. If the material was active, adult development of the pupae started, and a complete adult body was formed about 20 days after injection. Judgement of the response was done 4–6 days after injection by checking the epidermal retraction. When the wing part of a pupa was moistened with water and viewed with intense illumination, tracheae of the wing epidermis was observed. The tracheae in the pupae that initiated development were deep, apart from the cuticle, and when a point of wing was gently pushed moved freely independent of the movement of the cuticle, in contrast to those of nondeveloping pupae, which were in firm contact with cuticle. The potency was evaluated by assaying a series of twofold diluted solutions of the test material. The response was dose-dependent. A range of the test material concentrations where a clearly positive response was seen indefinitely separated from that with negative response, sometimes accompanied with a boundary concentration where only some pupae developed. The relative potency of the test material was calculated as follows: (ED_{50} of bombyxin-II [mol/pupa])/(ED_{50} of the test material [mol/pupa]), where ED_{50} of bombyxin-II was typically 50 fmol/pupa.

(c) NMR measurements

The synthetic bonsylin-(6-18) or bonsulin was dissolved at a concentration of 3 mM in 70%/30% (v/v) $^2H_2O/C^2H_3CO_2^2H$ (pH*2.0, pH* indicates direct meter reading) or at 4 mM in 70%/30% (v/v) $H_2O/C^2H_3CO_2^2H$ (pH*2.0). ^1H NMR spectra were measured at 600 MHz on a JEOL JNM-α600 spectrometer at 28°C. DQF-COSY (Rance et al., 1983), PE-COSY (Müller, 1987), TOCSY (45 ms mixing time) with a modified DIPSI-2 pulse sequence (Cavanagh and Rance, 1992) and NOESY (75 or 150 ms mixing time) (Jeener et al., 1979; Macura et al., 1981) were recorded in the phase-sensitive mode (States et al., 1982). Water resonance was suppressed by DANTE pulse (Zuiderweg et al., 1986). Two-dimensional spectra were recorded using a data size of 512 (t_1) x 2048 (t_2) (512 x 4096 for PE-COSY) with

a spectral width of 6500 Hz. After zero-filling once in the t_2 and twice in the t_1 dimension, 2048 x 2048 real data matrix were finally obtained and digital resolution was 3.2 Hz/point in both dimensions (512 x 4096 real data matrix and 1.6 Hz/point digital resolution in the F_2 dimension for PE-COSY).

(d) Structure calculations

Interproton distance constraints were derived from NOE crosspeak intensities (peak height) in the NOESY spectra (75 ms mixing time) according to the method of Hatanaka et al. (1994). Crosspeaks in the NOESY spectra were picked using a homemade C program and edited with Felix (Biosym Technologies, Inc., San Diego, CA). The peak intensities were translated into distances on the basis of the relation of NOE intensity $\propto$ (distance)$^{-6}$ and a standard distance of sequential d_{NN} in α helix = 2.8 Å (Wüthrich, 1986). The upper-bound distance constraints were the calculated distance plus 0.5 Å. The lower-bound constraints were set to 1.8 Å. The distances involving methylene and methyl protons and ring protons of tyrosine were referred to as single $(\langle r^{-6} \rangle)^{-1/6}$ average distances so that no corrections for center averaging were made (Clore et al., 1986). Dihedral angle constraints were obtained based on the analysis of DQF-COSY, PE-COSY and/or NOESY spectra (Wagner et al., 1988). The three-dimensional structures were calculated by the simulated annealing method with X-PLOR (Molecular Simulations, Inc., Walthan, MA) using the distance and dihedral angle constraints. A final set of 10 converged structures was selected from 100 calculations on the basis of agreement with the experimental data and van der Waals energy. A mean structure was obtained by averaging the coordinates of the structures that were superimposed in advance to the best converged structure and then minimizing under the constraints (Clore et al., 1986).

3. Results

(a) Bombyxin-like activity of chimera molecules of bombyxin-II and human insulin

All the chimera molecules of human insulin and bombyxin-II used in this study were synthesized by the combination of solid-phase peptide chain synthesis and regioselective disulfide bond formation (Maruyama et al., 1992; Nagata et al., 1992b). Bombyxin-like activity of the chimera molecules assayed was the prothoracicotropic effect to the saturniid moth Samia cyntia ricini (Ishizaki et al., 1983; Nagasawa et al., 1984a). Injection of an effective dose of the test material with the bombyxin-like activity into Samia brain-removed pupae causes adult development of the pupae; otherwise the pupae remain dormant. The effect is dose-dependent: bombyxin-II induces adult development of the Samia pupae with the median effective dose (ED_{50}) of 0.25 ng per pupa. Since bombyxin-II effects directly on the Samia prothoracic glands to produce and secret ecdysteroid (Nagasawa et al., 1984b) and since no data have been obtained that suggest the presence of bombyxin binding proteins in the haemolymph of the moths, the potency is considered to be parallel with the affinity to the "bombyxin receptor" on the Samia prothoracic glands. The assay results for bombyxin-II–human insulin chimera molecules are summarized in Table 4-1. Of the two hybrid molecules of bombyxin-II and human insulin, imbyxin retains a weak bombyxin-like activity, whereas bonsulin lost the entire bombyxin-like activity, indicating that the B-chain of bombyxin-II rather than the A-chain contains structural element(s) required for the high-affinity recognition of the Samia bombyxin receptor. To localize the receptor recognition site on the bombyxin-II B-chain, bonsylins, which comprise the bombyxin-II A-chain and the chimera B-chain, were synthesized and bioassayed. Bonsylin-(6-18), in which the B-chain middle part of B6 to B18 is from the bombyxin-II B-chain while the N- and C-terminal parts of B(-2) to B5 and B19 to B30 are from the human insulin B-chain, possessed as high a potency as bombyxin-II, which showed that neither the N-terminal part (pGluB(-2) to ThrB5) nor the C-terminal part (TrpB20 to AspB25) of bombyxin-II B-chain did not play an important role in the Samia bombyxin

receptor recognition. Because the single residue replacement from bombyxin-II type to human insulin type at B6 [bonsylin-(7-18)] or at B18 [bonsylin-(6-17)] caused marked reduction in the bombyxin-like activity, the receptor recognition site in the bombyxin-II B-chain was localized to the middle part of TyrB6 to LeuB18.

(b) Three-dimensional structures of bonsylin-(6-18) and bonsulin

Because of the sequence identity except for the B-chain middle part, the distinct affinities for the Samia bombyxin receptor of bonsylin-(6-18) and bonsulin were considered to be due to the structural differences in the B-chain middle part of B6 to B18. To elucidate the structural basis of the distinct affinities to the Samia bombyxin receptor, their three-dimensional structures in solution were determined by two-dimensional ^1H NMR spectroscopy and simulated annealing calculations, and were compared. $H_2O/C^2H_3CO_2{}^2H$ (7:3, pH*2.0, pH* indicates direct meter reading) was used as the solvent for NMR measurements to avoid the peptide aggregation which occurred in H_2O in the same or a similar way as in the cases of bombyxin-II ($H_2O/C^2H_3CO_2{}^2H$ (7:3), pH*2.0) (Chapter 2) and human insulin ($H_2O/C^2H_3CO_2{}^2H$ (8:2, pH*1.9) (Hua et al., 1991), where the addition of organic solvent did not destroy the globular structures of bombyxin and insulin. The peptide concentration was 4 mM in $H_2O/C^2H_3CO_2{}^2H$ (7:3, pH*2.0) and 3 mM in $^2H_2O/C^2H_3CO_2{}^2H$ (7:3, pH*2.0) in both the cases of bonsylin-(6-18) and bonsulin. The resonances were assigned to individual protons in a sequence-specific manner using the sequential assignment method (Table 4-2 and Figure 4-2) (Wüthrich, 1986). The successive strong d_{NN}, $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ NOE connectivities were characteristic of α-helix. The analysis of these NOE connectivities revealed that both the molecules contain three α-helices in the A-chain N-terminal region (H_{AN}, residues IleA2 to LeuA8), in the A-chain C-terminal region (H_{AC}, residues ValA13 to TyrA19) and in the B-chain middle region (H_B, residues ArgB9 to AlaB22).

In the case of bonsylin-(6-18), a total of 523 distance constraints which included 254 intraresidue, 123 sequential ($|i - j| = 1$), 95 short-range ($2 \le |i - j| \le 5$) and 51 long-range ($|i - j| \ge 6$) constraints were derived from the assigned NOE crosspeaks measured with a mixing time of

75 ms, and dihedral angle constraints for 24 ϕ and 8 $\chi 1$ were obtained. In the case of bonsulin, a total of 623 distance constraints which included 271 intraresidue, 164 sequential, 141 short-range and 47 long-range constraints were derived from the assigned NOE crosspeaks measured with a mixing time of 75 ms, and dihedral angle constraints for 26 ϕ and 8 $\chi 1$ were obtained. The three-dimensional structures were calculated with X-PLOR (Brünger, 1990) using the simulated annealing protocol (YASAP) on the experimental constraints and 3 distance constraints of the disulfide bonds (CysA6-CysA11, CysA7-CysB7 and CysA20-CysB19) (Nagasawa et al., 1988). A total of 100 calculations were carried out, and a final set of 10 structures was selected on the basis of agreement with the experimental constraints and van der Waals energy (Table 4-3). The number of inter-residue distance constraints and average root-mean-square deviations (RMSDs) around the mean structure for each residue (Figure 4-3) and the Ramachandran plot for the 10 structures (Figure 4-4) are shown. The structure of bonsylin-(6-18) was well-defined except for the B-chain C-terminal part (GlyB23 to ThrB30) and the peptide-chain termini (GlyA1, CysA20 and PheB1 to AsnB3) (Figure 4-3A); the RMSDs between the final 10 structures and the mean structure were 0.46 ± 0.09 Å for the backbone heavy atoms (N, C^{α}, C') and 1.09 ± 0.13 Å for all non-hydrogen atoms in the well-defined regions (residues IleA2 to TyrA19 and GlnB4 to ArgB22). The structure of bonsulin was well-defined except for the peptide-chain termini (residues GlyA1, CysA20, PheB1 to AsnB3 and LysB29-ThrB30) (Figure 4-3B); the RMSDs between the final 10 structures and the mean structure were 0.71 ± 0.11 Å for the backbone heavy atoms (N, C^{α}, C') and 1.59 ± 0.22 Å for all non-hydrogen atoms in the well-defined regions (residues IleA2 to TyrA19 and GlnB4 to ProB28).

The overall main-chain folds of bonsylin-(6-18) and bonsulin were similar except for the B-chain C-terminal part; the RMSD was 1.98 Å for the main-chain atoms (N, C^{α}, C') within the well-defined core regions (residues IleA2 to TyrA19 and GlnB4 to CysB19) (Figure 4-5). The common fold of the A-chains consisted of two antiparallel helices, H_{AN} (residues IleA2 to LeuA8), H_{AC} (residues ValA13 to TyrA19) and a connecting loop (residues ArgA9 to SerA12). The common fold of the B-chains consisted less well-defined N-terminus (residues

PheB1 to AsnB3), an extended arm (residues GlnB4 to GlyB8) and a middle helix (H_B, residues ArgB9 or SerB9 to CysB19). The core structure, which was characteristic to the insulin-superfamily peptides, was stabilized by the three disulfide bonds (CysA6-CysA11, CysA7-CysB10 and CysA20-CysB19) and a hydrophobic cluster including the residues at A2 (Ile in bonsylin-Y/Ile in bonsulin), A16 (Leu/Leu), A19 (Tyr/Tyr), B6 (Tyr/Leu), B11 (Leu/Leu), B15 (Leu/Leu) and B18 (Leu/Val). Though not all the hydrophobic cluster-forming residues were the same in the B-chains of bombyxin-II and human insulin, the bombyxin-II A-chain was shown to be well-matched with the human insulin B-chain in forming the characteristic core structure, indicating that the formation of the hydrophobic cluster was adaptable to some extent.

Despite the sequence identity, the B-chain C-terminal parts (residues GlyB20 to ThrB30) of bonsylin-(6-18) and bonsulin were distinct in conformation (Figures 4-5 and 4-6). The B-chain C-terminal part of bonsulin comprised a turn (residues GlyB20 to GlyB23), a β-strand (residues PheB24 to ProB28) and a less well-defined terminus (residues LysB29-ThrB30) as that of human insulin, whereas that of bonsylin-(6-18) was poorly defined in conformation. In bonsulin, the hydrophobic interactions between ValB12 and TyrB26, ValB12 and ThrB27 and TyrB16 and PheB24 kept the B-chain C-terminal β-strand on the B-chain middle helix, whereas in bonsylin-(6-18), where ValB12 and TyrB16 in bonsulin were replaced by AlaB12 and AlaB16, such hydrophobic interactions were weakened, and the B-chain C-terminal part no longer stays on the B-chain helix (Figure 4-6). Therefore, the conformational differences in the B-chain C-terminal part were brought about by the differences in hydrophobic interactions between the B-chain middle part and the B-chain C-terminal part. Similar hydrophobic interactions were observed in human insulin (Hua et al., 1991; Jørgensen et al., 1992). Thus, ValB12 and TyrB16 contribute not only to the recognition of the insulin receptor (Hu et al., 1993) but also to the stabilization of the B-chain C-terminal strand, which should be stabilized further in the insulin dimer by intermolecular antiparallel β-sheet formation (Jørgensen et al., 1992). On the other hand, bombyxin-II does not require the β-strand stabilizers, ValB12 and TyrB16, since the B-chain C-terminal part of bombyxin-II does not take an insulin-like turn and

a β-strand (Baker et al., 1988; Hua et al., 1991) but rather a relaxin-like helix extension (Eigenbrot et al., 1991). From the different main-chain conformations in the B-chain C-terminal parts, a possibility arised that bonsulin could not bind to the Samia bombyxin receptor due to a steric hindrance caused by the B-chain C-terminal β-strand or by the insulin-like dimer formation through the C-terminal β-strand and middle helix of the human insulin B-chain. But the possibility was denied by the fact that des-octapeptide(B23-B30)-bonsulin, which lacked the B-chain C-terminal β-strand and thus could not form the insulin-like dimer, had no detectable bombyxin-like activity (relative potency, < 0.0001) as bonsulin. Therefore, it is concluded that the distinct affinities of bonsylin-(6-18) and bonsulin to the Samia bombyxin receptor were derived from the different side chains in the B-chain middle part but not from the diffrent main-chain conformation in the B-chain C-terminal part.

4. Discussion

(a) Critical importance of the middle part of the bombyxin-II B-chain in the recognition of the Samia cynthia ricini bombyxin receptor

In spite of the 84% sequence identity, bonsylin-(6-18) and bonsulin possess disparate potencies in the bombyxin-like activity. Bonsylin-(6-18) shows as high a potency as bombyxin-II, whereas bonsulin has no detectable activity (Table 4-1). The distinct affinities of bonsylin-(6-18) and bonsulin to the Samia bombyxin receptor are derived from the different side chains in the B-chain middle part of B6 to B18. Of the eight residues uniquely found in bonsylin-(6-18), the six residues (TyrB6, ArgB9, AlaB12, ArgB13, AlaA16 and AspA17) are exposed to the solvent (Figure 4-6) and some of them should play critical roles in recognizing the Samia bombyxin receptor. The two other residues, ThrB14 and LeuB18, are buried inside the molecule, thus might be less important in receptor recognition. The exposed six residues made up a characteristic patch on the molecular surface of bonsylin-(6-18), where the side chains of ArgB9-ArgB13-AspB17 and AlaB12-AlaB16 line in parallel. In contrast, the corresponding surface was replaced by a structurally and chemically different patch in bonsulin which includes LeuB6, SerB9, ValB12, GluB13, TyrB16 and LeuB17. Thus, the characteristic patch of bonsylin-(6-18) (referred to as the Y/RRD/AA patch) in the B-chain middle part is essential to high affinity binding to the Samia bombyxin receptor, and the different patch of bonsulin (referred to as the L/SEL/VY patch) cannot replace it. The critical importance of the Y/RRD/AA patch in the Samia bombyxin receptor recognition is confirmed by the facts that the patch is also found in the molecular surface of bombyxin-II and that it is altered to the L/SEL/VY patch in human insulin (Figure 4-7).

(b) Receptor-recognition surface of bombyxin-II

The structure-activity relationships of bombyxin analogs confirm the importance of the B-chain middle part (HisB10, ArgB13 and/or AspB17), and demonstrates, in addition, the importance of the A-chain N- and C-termini (GlyA1, ValA3 and A20-B19 disulfide)

(Maruyama, 1991; Nagata et al., 1992, unpublished data). Since these parts for recognition of the Samia bombyxin receptor are located in a molecular surface of bombyxin-II shown in Figure 4-7, the surface can be the receptor-recognition surface of bombyxin-II.

The receptor-recognition surface of bombyxin-II shown in Figure 4-7 (referred to as the surface A) partly overlaps the surface (refered to as the surface B) which was proposed previously (Figure 2-8) based on much smaller number of experimental data and on the classical receptor-binding region of insulin. Though the surface B contains most residues involved in the Samia bombyxin receptor recognition, it lacks several important residues such as TyrB6, HisB10 and AspB17. The surface A contains almost all the identified residues important in receptor recognition. Interestingly, the surface A corresponds to the recently proposed receptor-recognition surface of human insulin, which contains two receptor-recognition sites (Schäffer, 1994).

(c) Comparison of the receptor-recognition surface of bombyxin-II with those of insulin and relaxin

Since the receptor recognition surface of bombyxin-II was localized, it was compared with the receptor-recognition surfaces of human insulin and human relaxin. In the case of insulin, the A-chain N-terminus (GlyA1-IleA2-ValA3) (Nakagawa and Tager, 1992), the A-chain C-terminus (TyrA19 and AsnA21) (Kitagawa et al., 1984; Carpenter, 1966), the B-chain middle helix (ValB12 and TyrB16) (Hu et al., 1993) and the B-chain C-terminal β-strand (PheB24-PheB25) (Mirmira and Tager, 1989; Nakagawa and Tager, 1986; Mirmira et al., 1991) were shown to form a receptor-recognition site (Blundell et al., 1972; Pullen et al., 1976; Murray-Rust et al., 1992), and recently LeuA13 and LeuB17 have been demonstrated to form the other receptor-recognition site (Schäffer, 1994). In the case of relaxin, ArgB9 and ArgB13 on the B-chain middle helix were demonstrated to be involved in receptor binding (Büllesbach and Schwabe, 1988), and in addition, TyrA(-1), PheA19, ValB12, GlnB15 and IleB16 were proposed to be involved in the receptor-recognition surface of relaxin (Eigenbrot et al., 1991). Mapping of these residues on the three-dimensional molecular structure reveals that the these

molecules have a common receptor-recognition surface shown in Figure 4-8, which includes the A-chain N-terminus, the A-chain C-terminus, the B-chain middle part in common, and in addition the B-chain C-terminal β-strand in the case of insulin.

In the surface, the A-chains of bombyxin-II and human insulin contain several common residues (GlyA1, IleA2, ValA3, TyrA19 and CysA20-CysB19 disulfide) shown or considered to be important for recognition of both the Samia bombyxin receptor (Nagata et al., 1992; Nagata et al., unpublished data) and the human insulin receptor (Nakagawa and Tager, 1992; Kitagawa et al., 1984; Sieber et al., 1978). In constrast, most of the important residues in the B-chains of bombyxin-II and human insulin are different, for instance, the Y/RRD/AA patch and the L/SEL/VY patch (Nakagawa and Tager, 1991; Hu et al., 1993; Schäffer, 1994) in the middle part and the distinct functions of the C-terminal part. The B-chain C-terminal part of bombyxin-II, AlaB22 to AspB25, is not required for Samia bombyxin receptor recognition (Nagata et al., in preparation), whereas that of human insulin, in particular the aromatic triplet PheB24-PheB25-TyrB26, plays a pivotal role in insulin receptor recognition (Mirmira and Tager, 1989; Nakagawa and Tager, 1986; Hu et al., 1993; Mirmira et al., 1991). Thus, the human insulin A-chain can partly take the place of the bombyxin-II A-chain, while the human insulin can not replace the bombyxin-II B-chain, as demonstrated by the bioassays of imbyxin and bonsulin (Table 4-1).

The A-chain of human relaxin makes a unique patch around TyrA(-1) in its extended N-terminus, which is dissimilar to the A-chains of bombyxin-II and human insulin (Figure 4-8). In contrast, the B-chains of bombyxin-II and human relaxin have several common characters which are not found in the human insulin B-chain. First, two Arg residues at B9 and B13, which are important for relaxin receptor recognition (Büllesbach and Schwabe, 1988), are conserved in the B-chains of bombyxin-II and human relaxin, suggesting that they might also be involved in bombyxin receptor recognition. Second, the B-chain C-terminal parts of bombyxin-II and human relaxin take a similar helix extension. These facts suggest that the B-chains of bombyxin-II and human relaxin could be interchangeable, but their A-chains could not.

In conclusion, in the divergent evolution from a common ancestral molecule, an invertebrate insulin-superfamily peptide, bombyxin-II, and its vertebrate counterparts, human insulin and human relaxin, have retained the common surface for receptor recognition, while they have evolved peculiar surfaces, which are characterized by the variety of involved side chains and by the inclusion of additional parts, for instance, the B-chain C-terminal part of insulin and the extended A-chain N-terminal helix of relaxin.

Chapter 5
Identification of the receptor-recognition site of bombyxin-II, a silkmoth insulin-related peptide, and its comparison with those of insulin and relaxin

Ala-scanning mutagenesis of the functionally important B-chain middle part

Summary

Bombyxin is an insulin-like peptide produced in the cerebral neurosecretory cells of the silkmoth Bombyx mori. Bombyxin consists of two peptide chains with 40% sequence identity to those of human insulin and three disulfide bonds formed in the same way as in insulin. Despite the structural similarity between bombyxin and insulin, they are not cross-active. We have previously shown that (1) the receptor-specificity determinants lie in their B-chains rather than their A-chains and that (2) the middle part (B6 to B18) contains all the requirements in the bombyxin-II B-chain for recognition of the Samia cynthia ricini bombyxin receptor. In order to identify the active site of bombyxin-II, the importance of each side chain in the middle part of bombyxin-II B-chain was evaluated systematically by the Ala-scanning mutagenesis (the native Ala residues are replaced by insulin-type residues). The important residues for activity are mapped on the three-dimensional structure of bombyxin-II, and thus the receptor-recognition site is identified. The receptor-recognition site of bombyxin-II consists of two regions, and is mostly overlapped with that of insulin, but may be different from that of relaxin. Detailed comparison of these receptor-recognition sites are described.

1. Introduction

Bombyxin is a brain-secretory peptide of the silkmoth Bombyx mori, which was identified as the first insulin-related peptide of invertebrate origin (Nagasawa et al., 1984a). Bombyxin-II, a representative molecular species of bombyxin, consists of two peptide chains (A-chain, 20 residues; B-chain, 28 residues) sequentially similar to those of insulin (50% and 32% identical, respectively, to those of human insulin) and three disulfide bonds linked in the same way as in insulin (A6-A11, A7-B7 and A20-B19) (Figure 5-1) (Nagasawa et al., 1986; Nagasawa et al., 1988). Furthermore, bombyxin-II has a similar core structure to those of vertebrate insulin-superfamily peptides (Chapter 2). Bombyxin induces adult development of brain-removed dormant pupae of the saturiid moth Samia cynthia ricini, a relative species of Bombyx mori, when injected into the pupae, by binding to its "receptor" on the prothoracic glands to stimulate the synthesis and release of ecdysone, a steroid hormone required for insect molting and metamorphosis (Nagasawa et al., 1984a,b). Since the production of bombyxin-related peptides (referred to as the Samia cynthia ricini bombyxin-related peptides or SBRPs) in the Samia brain was demonstrated (Kimura-Kawakami et al., 1992), the intrinsic role of the "bombyxin receptor" on the Samia prothoracic glands can be the receptor for intrinsic SBRPs. Thus, bombyxin and SBRPs are considered to recognize the receptor in a similar way. In contrast, bombyxin is inactive to the intrinsic prothoracic glands of Bombyx mori (Ishizaki et al., 1983; Kiriishi et al., 1992). The functions of bombyxin in Bombyx mori so far demonstrated are to lower the concentration of haemolymph trehalose, the major blood sugar of the moth (M. Masumura et al., in preparation), and to induce meiosis in the ovary (Orikasa et al., 1993). Despite the structural similarity between bombyxin-II and human insulin, they have no cross-activity: bombyxin-II has no affinity to the insulin receptor (Fernandez-Almonacid and Rosen, 1987), whereas human insulin has no affinity to the Samia bombyxin receptor (Nagasawa et al., 1984a).

In Chapter 4, the requirements for recognition of the Samia bombyxin receptor was located to the middle part of the bombyxin-II B-chain (TyrB6 to LeuB18). To identify the important

side chains for recognition of the Samia bombyxin receptor, a series of Ala-scanning analogs (if Ala was the native residue, it was changed to the insulin-type amino acid residue) were synthesized and their potencies in bombyxin-like activity were evaluated. These and previously shown important residues are mapped on the three-dimensional structure of bombyxin-II, and thus the receptor recognition site of bombyxin-II is identified. The receptor-recognition site of bombyxin-II is compared with those of vertebrate counterparts, human insulin and human relaxin 2.

2. Materials and methods

(a) Solid-phase peptide chain synthesis

The A-chain of bombyxin-II was synthesized on an Applied Biosystems Inc. 430A peptide synthesizer by the solid-phase peptide synthesis based on the HOBt-NMP/Fmoc chemistry. A series of the Ala-scanning mutants and other analogs of the bombyxin-II B-chain were synthesized as follows. A larger batch of the C-terminal segment (CysB19 to AspB25) of bombyxin-II B-chain was synthesized on the ABI 430A peptide synthesizer with the HOBt-NMP/Fmoc protocol (a single 3-hr coupling with 4-fold excess of Fmoc-amino acid HOBt ester). The bombyxin-II B-chain-(B19–B25)-peptide-resin was distributed in distinct vessels (0.03 mmol each) of an Advanced ChemTech 396 multiple peptide synthesizer. A double-coupling protocol with DIC/HOBt activation, 6-fold excess and a coupling time of 40 min was used. After the peptide-chain assembly was completed, each peptide-resin was washed thoroughly with CH_3OH and CH_2Cl_2 and dried up. The peptide were deprotected and cleaved from the resin by the treatment with phenol/1,2-ethanedithiol/thioanisole/H_2O/triisopropylsilane/TFA (2.8:1.0:2.0:2.0:1.0:40, v/v/v/v/v/v, 3.0 ml) at room temperature for 1.5 hr, precipitated by adding cold diethyl ether, collected on the PTFE membrane filter and dissolved into 50% CH_3CN/0.1% TFA. After dilution with H_2O to a CH_3CN concentration of < 10%, the peptide solutions were alkalified to pH 8.3–8.5 by adding 0.5 M Tris-HCl buffer (pH 8.5). The peptide was reduced with DTT (3.0 mg/ml) under an atomosphere of N_2 at 37°C for 2 hr. After acidification to pH 3 with 10% TFA, the peptide solution was stored under an atomosphere of N_2 at -20°C till purification. The peptides were purified to homogeneity by preparative RP-HPLC on a JASCO LC-900 system under the following conditions: column, SenshuPak ODS-H (20 x 250 mm); eluent, 10-25-35% CH_3CN (10 min/40 min, linear gradient); flow rate, 8.0 ml/min; column temperature, 40°C; detection, absorbance at 300 nm. The purity of the fractionated peptides was checked by analytical RP-HPLC on a Shimadzu LC-9 system under the following conditions: column, SenshuPak Pegasil ODS (4.6 x 150 mm); eluent, 10-40% CH_3CN (40 min, linear gradient); flow rate, 1.0

ml/mil; column temperature, 40°C; detection, absorbance at 280 nm. The peptide identity was confirmed by FAB-MS analysis. The fractions containing purified B-chains were stored at -20°C.

(b) Regioselective formation of three-disulfide bonds

Regioselective formation of three-disulfide bonds was performed essentially according to the method of Maruyama et al. (1992) shown in Figure 5-2. The preparation of [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain containing the intra-A-chain disulfide bond CysA6-CysA11 was described previously. The second disulfide bond was formed between CysA20 and CysB19 by thiol-disulfide exchange: [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain (100 µg dissolved in 100 µl of 0.5 M Tris-HCl buffer (pH 7.5) was coupled with [Cys(Acm)B7,CysB19]B-chain (105 µg, four fifth molar equivalents to the A-chain, dissolved in 240 µl of 10% CH_3CN/0.1% TFA) at 45°C for 1 hr to yield [Cys(Acm)A7,B7,A6-A11,A20-B19-cystine]heterodimer as the major product. The heterodimer was purified to heterogeneity by the analytical RP-HPLC. The third disulfide bond was formed between CysA7 and CysB7 by the iodine oxidation of [Cys(Acm)A7,B7, A6-A11,A20-B19-cystine]heterodimer: the peptide (80 µg, 15 nmol) was incubated with I_2 (400 nmol) in 50 µl of 95% CH_3CO_2H in the presence of HCl (200 nmol) to yield the target molecule [A6-A11,A7-B7,A20-B19-cystine]heterodimer as the major product. The product was purified to homogeneity by the analytical RP-HPLC. The peptide identity was confirmed by FAB-MS analysis.

(c) Circular dichroism spectra measurement

CD spectra were recorded on a JASCO J-600 spectropolarimeter in the wavelength range of 200–250 nm with a 1-mm path length cell at room temperature. Four scans were accumulated. Samples were dissolved in 50 mM sodium phosphate buffer (pH 6.8)/50 mM NaCl at a residue concentration of 20 µM. The peptide concentration was determined using the sum of molar

extinction coefficient at 280 nm of each residue: Trp (5400 l/cm/mol), Tyr (1100 l/cm/mol) and disulfide (200 l/cm/mol).

(d) Bioassay for bombyxin-like activity

Bombyxin-like prothoracicotropic activity to the saturniid moth Samia cynthia ricini of the analogs was evaluated by the in vivo Samia pupal assay (Ishizaki and Ichikawa, 1967). The brains of pupae of Samia cynthia ricini were surgically removed within 20 hr after pupation. This operation blocked adult development of the pupae, and most pupae should survive as pupae without initiating adult developement for months or even more than a year. The test material dissolved in 0.1 M Tris-HCl (pH 7.8) containing 0.04% bovine serum albumin (BSA) was injected (10 µl per pupa) into the brain-removed pupae 2–5 months after the brain removal. During the period the sensitivity of brainless pupae to bombyxin-II, which was used as the positive control, did not change appreciably. If the material was active, adult development of the pupae started, and a complete adult body was formed about 20 days after injection. Judgement of the response was done 4–6 days after the injection by checking epidermal retraction. When the wing part of a pupa was moistened with water and viewed with intense illumination, tracheae of wing epidermis were observed. The tracheae in the pupae that initiated development were deep, apart from the cuticle, and when a point of wing was gently pushed, moved freely independent of the movement of the cuticle, in contrast to those of nondeveloping pupae, which were in firm contact with cuticle. The potency was evaluated by assaying a series of twofold diluted solutions of the test material. The response was dose-dependent. A range of the test material concentrations where a clearly positive response was seen indefinitely separated from that with negative response, sometimes accompanied with a boundary concentration where only some pupae developed. The relative potency of the test material was defined as follows: (ED_{50} of bombyxin-II [mol/pupa])/(ED_{50} of the test material [mol/pupa]), where ED_{50} of bombyxin-II was typically 25 fmol/pupa.

3. Results

(a) Synthesis of Ala-scanning analogs of bombyxin-II

In order to evaluate the importance of each side chain in the B-chain middle part of the bombyxin-II (TyrB6 to LeuB18) for recognition of the Samia bombyxin receptor, a series of Ala-scanning analogs, in which each amino acid residue is individually replaced by Ala, were synthesized. Since CysB7 and GlyB8 were important to stabilize the insulin fold, they were not replaced. The native Ala residues at B12 and B16 were replaced by the insulin-type residues, Val and Tyr, respectively.

Since the target molecules consisted of two peptide chains and three disulfide bonds, first the peptide chains were synthesized by the solid-phase peptide synthesis based on the Fmoc chemistry and then the disulfide bonds were formed in a stepwise and regioselective manner essentially according to the method of Maruyama et al. (1992). The A-chain, [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]bombyxin-II A-chain, was common to all the target peptides and was synthesized on an Applied Biosystems Inc. 430A peptide synthesizer based on the HOBt-NMP Fmoc chemistry. On the other hand, the mutated [Cys(Acm)B7,CysB19]bombyxin-II B-chains were synthesized in two steps. The C-terminal segment (CysB19 to AspB25) common to all the B-chains were assembled on HMP-resin support on the ABI 430A peptide synthesizer based on the HOBt-NMP Fmoc chemistry. The middle and N-terminal parts (pGluB(-2) to Leu/ValB18) was elongated on the bombyxin-II B-chain-(B19-B25)-peptide-resin on an Advanced ChemTech 396 multiple peptide synthesizer based on the DIC/HOBt chemistry. For regioselective disulfide-bond formation, multiple orthogional Cys side-chain protecting groups were used: Trt (labile to TFA) for CysA6, CysA11 and CysB19, tBu (stable to TFA and labile to TFMSA) for CysA20 and Acm (stable to TFA and TFMSA and labile to I_2) for CysA7 and CysB7. After the TFA treatment, the released B-chains were reduced entirely with DTT, purified by RP-HPLC at pH 2 and stored at -20°C. The yield of the [Cys(Acm)B7,CysB19]B-chains was 8–14%, if both the total coupling yield and the recovery during the peptide purification were taken into account (Table 5-1).

The three intramolecular disulfide bonds were formed in a regioselective and stepwise manner (Figure 5-2). The intra-A-chain disulfide CysA6-CysA11 was formed first by oxidation in air of the [CysA6,A11,Cys(Acm)A7,Cys(tBu)A20]A-chain obtained by the TFA cleavage. After the conversion of CysA20 side-chain protection from tBu to Pys, the [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain and four fifth molar equivalents of [Cys(Acm)B7,CysB19]B-chain were coincubated at pH 7.5 at 45°C for 1 hr. Thus the second disulfide CysA20-CysB19 was formed during the coincubation by thiol-disulfide exchange. [Cys(Acm)A7,B7,A6-A11,A20-B19-cystine]heterodimer was obtained in yields of 32–46% (Table 5-2). The third disulfide bond CysA7-CysB7 was formed by iodine oxidation of [Cys(Acm)A7,B7,A6-A11,A20-B19-cystine]heterodimer. For the iodine oxidation, 95% CH_3CO_2H was used as the solvent and 12-fold molar excess of HCl to the peptide was added to prevent His, Tyr and Trp side chains from oxidative degradation (Nagata et al., 1992a). The target peptide, [A6-A11,A7-B7,A20-B19-cystine]heterodimer, was obtained in yields of 54–66% (Table 5-3). The peptide identity was confirmed by FAB-MS analysis. Thus, all the analogs were successfully obtained in yields of 18–29% from the A- and B-chains by the regioselective method of disulfide bond formation.

(b) Effects of the mutations on the conformations of mutants

The secondary structure of bombyxin-II and its analogs was compared using CD. The peptides were dissolved at a residue concentration at 20 μM in 50 mM sodium phosphate buffer (pH 6.8)/50 mM NaCl. All the Ala-scanning analogs (including the native bombyxin-II and the AlaB12Val and AlaB16Tyr analogs) except the LeuB15Ala analog showed very similar CD spectra in the wavelength range of 200–250 nm. Therefore, all analogs but the LeuB15Ala analog should have similar main-chain fold characteristic of the insulin-superfamily peptides. By contrast, the LeuB15Ala analog must have a peculiar main-chain conformation.

(c) Bombyxin-like prothoracicotropic activity of the bombyxin-II analogs

In order to identify important side chains for receptor recognition of bombyxin-II, bombyxin-like activity of the Ala-scanning analogs was evaluated. Bombyxin-II is able to induce adult development of brain-removed dormant pupae of the saturniid moth Samia cynthia ricini, a relative species of the silkmoth Bombyx mori, in a dose-dependent manner (Ishizaki et al., 1983; Nagasawa et al., 1984a), with a median effective dose (ED_{50}) of 25 fmol/pupa. This is because bombyxin-II stimulates synthesis of ecdysteroid, an insect molting hormone, in the prothoracic glands of Samia cynthia ricini. Direct activation of the Samia prothoracic glands by bombyxin-II was demonstrated in vitro (Nagasawa et al., 1984b). Therefore, the primary action of bombyxin-II is to bind to its receptor (referred to as the Samia bombyxin recetor) on the Samia prothoracic glands. Since no data suggest the presence of bombyxin binding proteins in the Samia haemolymph, the potency of bombyxin-II and its analogs in the bombyxin-like prothorcicotropic activity to Samia cynthia ricini is considered to be parallel with the affinity to the Samia bombyxin receptor.

The assay results for bombyxin-II–human insulin chimera molecules are summarized in Table 5-4. The TyrB6Ala (in which the Tyr residue at B6 was replaced by an Ala residue), LeuB11Ala, AlaB12Val, ThrB14Ala, LeuB15Ala, AlaB16Tyr, AspB17Ala and LeuB18Ala mutants possessed reduced potencies than bombyxin-II, while ArgB9Ala, HisB10Ala, ArgB13Ala, and LeuB18Ala mutants retained as high a potency as bombyxin-II. The largest decrease in potency (40-fold) was observed when TyrB6 was replaced by Ala. The second largest decrease in potency (20-fold) was observed when AlaB12 was replaced by Val. The replacement of the following residues led to moderate decreases in potency such as LeuB15 by Ala (8-fold), LeuB11 by Ala (4-fold), ThrB14 by Ala (2-fold), AlaB16 by Tyr (4-fold), AspB17 by Ala (4-fold) and LeuB18 by Ala (2-fold). In contrast, the replacement of ArgB9, HisB10 and ArgB13 did not affect the potency. The results indicate that the residues TyrB6, LeuB11, AlaB12, ThrB14, LeuB15, AlaB16, AspB17 and LeuB18 should be directly involved in the receptor recognition and/or contribute to the maintenance of active conformations of receptor-recognition elements, while ArgB9, HisB10 and ArgB13 side chains are not important for receptor recognition.

(d) Mapping of the important residues for bombyxin-like activity on the three-dimensional structure of bombyxin-II

The important residues for bombyxin-like activity identified above were mapped on the three-dimensional structure (Figure 5-4) determined by two-dimensional ^{1}H NMR and simulated annealing calculations (Chapter 2). LeuB11, AlaB12, LeuB15 and AlaB16 form a hydrophobic patch (referred to as the patch I) on the molecular surface together with highly conserved hydrophobic residues, IleA2, ValA3, TyrA19 and CysA20-CysB19 disulfide (Figure 5-4A), while TyrB6, AspB17 and LeuB18 form another mostly hydrophobic patch (referred to as the patch II) on a distinct molecular surface together with ValA13 and LeuA16 (Figure 5-4B). Important residues for bombyxin-like activity so far identified (GlyA1, ValA3 and CysA20-CysB19 disulfide bond) (Maruyama, 1991; Nagata et al., 1992a, unpublished data) are involved in the patch I. The patches I and II are separated by the two Arg at B9 and B13, which are not important for bombyxin-like activity (Figure 5-4C).

4. Discussion

(a) Evaluation of importance of each side chain in the B-chain middle part of bombyxin-II by Ala-scanning mutagenesis

In spite of 40% sequence identity, no cross-activity is seen between bombyxin-II and human insulin (Nagasawa et al., 1884a; Fernandez-Almonacid and Rosen, 1987; Nagata et al., 1993). We have previously shown that (1) the receptor specificity determinants lie in their B-chains rather than their A-chains (Nagata et al., 1993) and that (2) the middle part (TyrB6 to LeuB18) contains all the requirements in the bombyxin-II B-chain for recognition of the Samia bombyxin receptor, while the C-terminal part (PheB24-PheB25) in addition to the middle part (LeuB6 to ValB18) of the human insulin B-chain contributes to recognition of the human insulin receptor (Chapter 4). In order to identify the receptor-recognition site of bombyxin-II, the importance of each side chain in the middle part of bombyxin-II B-chain was evaluated systematically by the Ala-scanning mutagenesis (the native Ala residues are replaced by the insulin-type residues).

Ala-scanning mutagenesis is a systematic method for the evaluation of the significance of each side chain (Beck-Sickinger et al., 1994). Each amino acid residue of the native molecule is individually replaced by Ala. It is assumed that single substitutions by Ala do not disturb the secondary structure or change the hydrophobicity. Therefore it is possible to study the role of the side chain functional groups for biological activity. In contrast, D-enantiomer scan, in which each amino acid residue is individually replaced by the corresponding D-enantiomer, can be utilized to study the significance of the orientation of each side chain. Until recently such scans were difficult to realize because of the great number of peptide analogs required. The situation has become easier after the introduction of simultaneous multiple peptide synthesis.

All the Ala-scanning analogs were obtained in good yields (18–29%) from the constituent peptide chains by the synthetic strategy of bombyxin comprising solid-phase synthesis of the peptide chains and regioselecitve formation of the disulfide bonds (Figure 5-2) (Maruyama et al., 1992; Nagata et al., 1992a,b). Since the synthetic strategy can be applied to the synthesis of any insulin analog having any amino acid sequence, it is invaluable to the structure-activity

relationship studies of insulin and its related peptides. The bombyxin-like prothoracicotropic activity to the saturniid moth Samia cynthia ricini of the analogs was evaluated by the in vivo Samia pupal assay (Ishizaki and Ichikawa, 1967). Bombyxin, when injected into brain-removed dormant pupae of Samia cynthia ricini, a relative species of Bombyx mori, induces adult development of the pupae by activating the prothoracic glands to stimulate the synthesis and release of ecdysone, a steroid hormone required for insect molting and metamorphosis, in a dose-dependent manner (Nagasawa et al., 1984a,b). Since bombyxin-II effects directly on the Samia cynthia ricini prothoracic glands (Nagasawa et al., 1984b) and since no data suggest the presence of bombyxin binding proteins in the Samia haemolymph, the potency is considered to be parallel with the affinity to the Samia bombyxin receptor on the prothoracic glands.

The individual replacement of ArgB9, HisB10 and ArgB13 by Ala does not decrease the potency, indicating that these side chains are important neither for the Samia bombyxin receptor recognition nor for the maintenance of active conformation of the molecule. Whereas the individual replacement of TyrB6, LeuB11, LeuB15, AspB17, LeuB18 by Ala, AlaB12 by Val and AlaB16 by Tyr decreases potency. Therefore, these side chains should be involved directly in the Samia bombyxin receptor recognition and/or contribute to maintain the active conformation of the molecule. Since the LeuB15Ala analog showed a peculiar CD spectrum at 200–250 nm, the side-chain of LeuB15 was demonstrated to be involved in the maintenance of the peptide conformation.

(b) Receptor-recognition site of bombyxin-II

The important side chains for bombyxin-like activity form two patches I and II on the bombyxin-II molecules, which are separated by the two Arg at B9 and B13 not important for bombyxin-like activity (Figure 5-4). The patch I contains LeuB11, AlaB12, LeuB15 and AlaB16 side chains. It also contains so far identified important residues for bombyxin-like activity (GlyA1, ValA3 and CysA20-CysB19 disulfide bond) (Maruyama et al., 1991; Nagata et al., 1992a, unpublished data) and highly conserved hydrophobic residues (IleA2 and TyrA19). The patch II contains TyrB6, AspB17 and LeuB18 side chains. It also contains

highly conserved hydrophobic residues (ValA13, LeuA16 and LeuA17). Of the important side chains, the side chains of AlaB12, AlaB16 in the patch I and of TyrB6 and AspB17 in the patch II should be directly involved in the recognition of the Samia bombyxin receptor, because these side chains are fully or mostly exposed to the solvent. Other important side chains in these patches, which are partly exposed, may also be involved in the receptor recognition and/or may contribute to maintain the active conformation. Since both the patches contain receptor-recognition residues, the patches I and II can be referred to as the receptor-recogntion regions I and II (rrrBBX-I and II), respectively. The rrrBBX-I and -II are separated by the two Arg residues at B9 and B13, which are not important for the receptor recognition (Figure 5-4C).

(c) Comparison of the receptor-recognition site of bombyxin-II, human insulin and human relaxin 2

In the case of insulin (Figure 5-5), the A-chain N-terminus (GlyA1-IleA2-ValA3), the A-chain C-terminus (TyrA19 and AsnA21), the B-chain helix (ValB12 and TyrB16) and the B-chain C-terminal β-strand (PheB24-PheB25) form a classical receptor-recognition region (the receptor-recognition region rrrISL-I) (Blundell et al., 1972; Pullen et al., 1976; Murray-Rust et al., 1992). Recently, another patch involving LeuA13 and LeuB17 has been proposed to be a second site for receptor recognition (Schäffer, 1994). Since LeuA13 and LeuB17 lie adjacent to the important LeuB6 (Nakagawa and Tager, 1991), the three Leu residues form the second receptor-recognition region, rrrISL-II. It is interesting the receptor-recognition regions of bombyxin-II, rrrBBX-I and -II, and those of human insulin, rrrISL-I and -II, are well-overlapped (Figure 5-5). Both the rrrBBX-I and the rrrISL-I involve the residues at A1 (Gly/Gly), A2 (Ile/Ile), A3 (Val/Val), A19 (Tyr/Tyr), A20 (Cys/Cys), A21 (none/Asn), B11 (Leu/Leu), B12 (Ala/Val), B15 (Leu/Leu), B16 (Ala/Tyr), and B19 (Cys/Cys). Both the rrrBBX-II and the rrrISL-II involve the residues at A13 (Val/Leu), B6 (Tyr/Leu) and B17 (Asp/Leu). By contrast, the B-chain C-terminal part is functionally distinct between bombyxin-II and human insulin. The B-chain C-terminal part of bombyxin-II is not required for the

Samia bombyxin receptor recognition, while the corresponding part of human insulin, particularly PheB24-PheB25, is of critical importance for the insulin receptor recognition.

The rrrBBX-I and the rrrISL-I involve many common residues such as GlyA1, IleA2, ValA3, TyrA19, CysA20, LeuB11, LeuB15 and CysB19. Despite the highly conservation, most of these hydrophobic residues in rrrBBX-I and those in the rrrISL-I lie in different environments: the hydrophobic patch of bombyxin-II is fully exposed to solvent, while that of human insulin is mostly covered by the B-chain C-terminal β-strand (Figure 5-5). Based on the solution structures of human insulin and its active mutant, [GlyB24]human insulin, Hua et al. (1991) proposed that detachment of the B-chain C-terminal β-strand from the core should occur in insulin on binding to its receptor, exposing the hydrophobic patch, particularly the IleA2 and ValA3 side chains (Figure 5-5). Thus, the common hydrophobic patches in the rrpBBX-I and the rrpISL-I could be accessible for direct contact with the respective receptors in a similar way.

The different residues between the rrrBBX-I and the rrrISL-I are (lack of A21)-AsnA21, AlaB12-ValB12 and AlaB16-TyrB16 (Figure 5-5). The individual replacement of AlaB12 and AlaB16 in bombyxin-II to the insulin type residues, Val and Tyr, decreases the bombyxin-like activity to 0.05 and 0.25, respectively (Table 5-4). On the other hand, replacement of ValB12 in insulin to the bombyxin-II type residue, Ala, decreases the insulin-like activity to 0.013 (Nakagawa and Tager, 1992). Although replacement of TyrB16 in insulin to the bombyxin-II type residue, Ala, has not been reported, mutagenic study at B16 of insulin indicates that bulk and/or aromaticity at this site is important for insulin-like activity (Hu et al., 1993), suggesting the replacement of TyrB16 by Ala also decreases the insulin-like activity. Since AlaB12 and AlaB16 side chains of bombyxin-II and ValB12 and TyrB16 side chains in human insulin are fully exposed to solvent, these side chains should be involved in recognizing the respective receptors, and moreover, determine the receptor specificity between bombyxin-II and human insulin.

Although bombyxin-II does not need the amino acid residue at A21, insulin needs it: the deletion of AsnA21 of insulin decreases the insulin-like activity to 0.02 (Carpenter, 1966; Yu

and Kitabchi, 1973). Three-dimensional structure analysis of chimera molecules of bombyxin-II and human insulin demonstrates that the deletion of AsnA21 increases conformational flexibility in the human insulin B-chain C-terminal part around PheB24, and further replacement of ValB12 and TyrB16 to the bombyxin-II type, Ala and Ala, results in detachment of the insulin-type B-chain C-terminal part from the core due to decrease in hydrophobic interactions between B12/B16 and the aromatic triplet PheB24-PheB25-TyrB26 (Chapter 4). Therefore, the residue at AsnA21 and side chains of ValB12 and TyrB16 peculiar to human insulin all contribute to the stabilization of the B-chain C-terminal β-strand in insulin besides the recognition of the human insulin receptor. In constrast, bombyxin-II, which lacks an insulin-like B-chain C-terminal β-strand, does not need the ValB12, TyrB16 side chains or the AsnA21 residue neither for the B-chain C-terminal β-strand stabilization nor for the Samia bombyxin receptor recognition. It is interesting that human relaxin 2, which does not have an insulin-like B-chain C-terminal β-strand, also lacks the residue at A21 as does bombyxin-II. The turn and β-strand in the B-chain C-terminal part is a common feature to the ligands for the insulin receptor, insulin, IGF-I and -II (Baker et al., 1988; Hua et al., 1991; Cooke et al., 1991; Terasawa et al., 1994). Whereas bombyxin-II and relaxin, which do not bind to the insulin receptor, do not possess the B-chain C-terminal β-strand (Chapter 2; Eigenbrot et al., 1991). The B-chain C-terminal β-strand characteristic to insulin contributes not only to the insulin receptor recognition but also to the biosynthesis of insulin by stimulating formation of the insulin dimer (Blundell, 1972), which further aggregates to form the insulin hexamer in the presence of Zn. The aggregation protects newly generated insulin from proinsulin converting enzyme during storage in the secretory granules in the β cells of the Langerhans islet (Blundell, 1972).

The rrrBBX-II and the rrrISL-II mostly involve different residues (ValA13-LeuA13, TyrB6-LeuB6, AspB17-LeuB17). Individual replacement of TyrB6 and AspB17 in bombyxin-II analogs to the insulin-type residue, Leu and Leu, decreases bombyxin-like activity to 0.05 and 0.05, respectively (Nagata et al., in preparation). On the other hand, individual replacement of LeuA13, LeuB6 and LeuB17 in insulin to other amino acid residues lower the insulin-like

activity (Nakagawa and Tager, 1991; Schäffer, 1994). Therefore, the amino acid side chains involved in the rrrBBX-II and the rrrISL-II should determine the receptor specificity between bombyxin-II and human insulin together with the side chains at B12 and B16 in the rrrBBX-I and the rrrISL-I.

In the case of relaxin, the middle part of the B-chain (ArgB9 and ArgB13) are essential for biological activity (Büllesbach and Schwabe, 1991). In addition, GluB10-LeuB11-ValB12 between the two essential Arg residues (Büllesbach and Schwabe, 1991) and TyrA(-1), PheA19, ValB12, GlnB15 and IleB16 near the Arg residues (Eigenbrot et al., 1991) are suggested to be important for activity. Although the two Arg residues at B9 and B13 are apparently conserved in bombyxin-II and human relaxin 2, they are functionally distinct in bombyxin-II and human relaxin 2. Those in bombyxin-II are not important for bombyxin activity ([AlaB9]bombyxin-II, [AlaB13]bombyxin-II and [CitB9,CitB13]bombyxin-II are all fully active (Nagata et al., unpublished data), while those in relaxin are essential for relaxin activity ([CitB9,CitB13]human relaxin 2 is inactive (Büllesbach and Schwabe, 1991)). Thus, human relaxin 2 has a slightly different receptor-recognition region (referred to as the rrrRLX) from that of bombyxin-II and human insulin, whch locates between the corresponding regions of the rrrBBX/ISL-I and the rrrBBX/ISL-II, and may partly overlap with the rrrBBX/ISL-I (Figure 5-5). Since the other residues proposed to be involved in their receptor recognition are mostly different, bombyxin-II and human relaxin 2 are probably incompatible in function. It is interesting that the relative orientation of the three molecules to the respective receptor could be the same, though the site for receptor recognition of the three molecules are not entirely identical (Figure 5-5). The conservation of the "face" for receptor recognition may be an indication of co-evolution of these ligands and their receptors from an ancestral molecule. Thus, the bombyxin receptor on the Samia cynthia ricini prothoracic glands might have a similar structure and function to the insulin receptor, i.e. an $\alpha_2\beta_2$ form and tyrosine kinase activity.

In conclusion, the receptor-recognition site of bombyxin-II, and the receptor specificity determinants between bombyxin-II and human insulin are identified. Comparison of receptor-recognition site of bombyxin-II, human insulin and human relaxin reveals that these insulin-

superfamily peptides have a common "face" for receptor recognition as well as the common core structure, while they have developed characteristic surfaces in the common "face" to recognize its receptor in a specific manner.

Chapter 6
Concluding remarks

6-1. Summary and general discussion

(a) Three-dimensional structure of bombyxin-II and its comparison with those of vertebrate counterparts, human insulin and human relaxin 2

In order to compare the three-dimensional structures of invertebrate and vertebrate insulin-superfamily peptides and to elucidate the receptor specificity between bombyxin-II and human insulin on the structural basis, the three-dimensional structure of bombyxin-II was determined by two-dimensional ^{1}H NMR spectroscopy and simulated annealing calculations (Chapter 2). The structure of bombyxin-II was determined at high resolution based on 535 distance and 24 dihedral constraints (Figure 6-1). The root-mean-square deviations (RMSDs) between the final 10 structures and the mean structure were 0.58 ± 0.15 Å for the backbone heavy atoms (N, C^{α}, C') and 1.03 ± 0.18 Å for all non-hydrogen atoms in the well-defined regions (residues IleA2 to TyrA19 and ThrB5 to AlaB22). (Figure 2-6). It is the first three-dimensional structure determined for an insulin-superfamily peptide of invertebrate origin. The determined solution structure of bombyxin-II is markedly different from its modeled structure (Jhoti et al., 1987) in the B-chain C-terminal part (B20 to B25). The solution structure takes an extention of the helix (B9 to B22) and coiled structure (B23 to B25) there, while the modeled structure adopts a type III turn (B19 to B22) and an extended structure (B23 to B25) in a similar way to human insulin.

The structure of bombyxin-II (determined by NMR) and those of vertebrate counterparts, human insulin (NMR) (Hua et al., 1991) and human relaxin 2 (X-ray crystallography) (Eigenbrot et al., 1991), are shown in Figure 6-1. Relaxin, called "the hormone of pregnancy", widens the birth canal in mammals prior to parturition (Schwäbe and Wüllesbach,

1990). The overall main-chain fold of bombyxin-II is similar to those of insulin (Hua et al., 1991; Jørgensen et al., 1992), and relaxin (Eigenbrot et al., 1991) (Figure 5). The characteristic "insulin core" structure is stabilized by a cluster of highly conserved hydrophobic residues at A2 (Ile in bombyxin-II/Ile in human insulin/Leu in human relaxin 2), A16 (Leu/Leu/Leu), A19 (Tyr/Tyr/Phe), B6 (Tyr/Leu/Leu), B11 (Leu/Leu/Leu), B15 (Leu/Leu/Gln), B18 (Leu/Val/Ile), which are highly conserved as hydrophobic through the superfamily, and three disulfide bonds.

The largest difference between the structures of bombyxin-II and human insulin lies in the B-chain C-terminal part (B20 to B25). Bombyxin-II, like relaxin, adopts a helix and a coiled structure there, instead of a sharp turn and an extended β-strand as do insulin and IGFs (Cooke et al., 1991; Sato et al., 1993; Terasawa et al., 1994). The turn and β-strand in insulin are stabilized by the two Gly residues at B20 and B23, whose ϕ angles are positive in crystal (Blundell et al., 1972), and the intramolecular hydrophobic interactions between PheB24 and ValB12, LeuB15, TyrB16 and between TyrB26 and LeuB11, IleA2, ValA3 (Jørgensen et al., 1992). Whereas, most of these turn-and-β-strand stabilizers are lost in bombyxin-II and human relaxin 2: bombyxin-II lacks GlyB20, human relaxin 2 lacks GlyB23, both molecules lack the residues at A21 and B26. Similarly, all the invertebrate insulin-related peptides so far characterized except sponge insulin (Robitzki et al., 1989) may take a helix extention in the B-chain C-terminal part. Thus the turn and β-strand found in insulin and IGFs are not common to all the members of the insulin superfamily but rather a characteristic property of the ligands for the insulin receptor and the type I IGF receptor. Actually, insulin requires the B-chain C-terminal part for recognition of the insulin receptor (Mirmira et al., 1991), while bombyxin-II does not require the corresponding part for recognition of the Samia bombyxin receptor (Maruyama, 1991; Chapter 4).

The B-chain C-terminal tail of insulin also contributes to formation of the insulin dimer, which in the presence of Zn^{2+} forms the insulin hexamer (Blundell, 1972). The aggregation properties of proinsulin and insulin seems to be advantageous in the biosynthesis of insulin,

because their zinc-containing hexamers are resistant to enzymatic proteolysis in the storage granules (Blundell, 1972).

(b) Three-dimensional structures of hybrid molecules of bombyxin-II and human insulin

Since both bombyxin-II and human insulin had a common core structure, it was expected that their hybrid molecules would also have the characteristic core structure. In Chapter 3, the three-dimensional structures of the hybrid molecules, bonsulin (bombyxin-II A-chain + human insulin B-chain) and imbyxin (human insulin A-chain + bombyxin-II B-chain) are determined. Bonsulin actually has the bombyxin/insulin-like core structure as expected, while imbyxin does not (Figure 6-2). Imbyxin takes a distorted conformation, in which the A-chain N-terminal α-helix (IleA2 to ThrA8) and the first turn (ArgB9 to LeuB11) of the B-chain α-helix are lost, in 70%/30% (v/v) water/acetic acid (pH 2.0) and in 50 mM sodium phosphate buffer (pH 6.8)/50 mM NaCl as demonstrated by NMR and indicated by CD, respectively. The CD spectra of imbyxin changes reversibly depending on the concentration of TFE and becomes similar to those of bombyxin-II and human insulin in 30% TFE. Therefore, the solvent-dependent equilibrium in the conformation of imbyxin is proposed (Figure 3-5): the conformer I, which has the distorted conformation, is dominant in aqueous solution without TFE, while the conformer II, which has the bombyxin-like main-chain conformation is dominant in aqueous solution containing 30% TFE. Because bonsulin lacks AsnA21, which stablizes the B-chain C-terminal β-strand by interacting with PheB24 in insulin, the B-chain C-terminal tail of bonsulin is less well-defined around PheB24.

In Chapter 4, the three-dimensional structures of bonsylin-(6-18), a chimera molecule of bombxyin-II and human insulin, is determined and is compared with that of bonsulin. They differ in sequence only in the B-chain middle part (B6 to B18), where bonsylin-(6-18) has the bombyxin-type residues while bonsulin has the insulin-type residues. Structure comparison demonstrates that they take similar main-chain conformations except for the sequentially identical B-chain C-terminal part (GlyB20 to ThrB30), where bonsulin takes an insulin-like

turn and β-strand, while bonsylin-(6-18) takes less-well defined conformations. The turn and β-strand is lost in bonsylin-(6-18), because it lacks the β-strand stabilizers, the ValB12, TyrB16 side chains and AsnA21 residue.

(c) Localization and identification of the receptor-recognition site of bombyxin-II

Identification of the receptor-recognition site is essential to elucidate the molecular basis of receptor specificity. Many kinds of bombyxin-II analogs, chimera molecules of bombyxin-II and human insulin (Nagata et al., 1992b) and disulfide-bond isomers of bombyxin-II (Nagata et al., 1992a) were synthesized and their bombyxin-like activity was evaluated (Table 6-1) to identify the important residues for bombyxin-like activity.

Because bombyxin-II and human insulin are not cross-active (Nagata et al., 1992b), the requirements for recognition of the Samia bombyxin receptor can be localized by evaluating the bombyxin-like activity of the hybrid and chimera molecules of bombyxin-II and human insulin.

In Chapter 3, in order to determine which peptide chain is more important than the other for receptor recognition, the hybrid molecules of bombyxin-II and human insulin were bioassayed. Bonsulin (bombyxin-II A-chain + human insulin B-chain) shows only the insulin-like activity but no bombyxin-like activity, while imbyxin (human insulin A-chain + bombyxin-II B-chain) shows only the bombyxin-like activity but no insulin-like activity. Since imbyxin takes multiple conformations, the result can be interpreted that no conformers of imbyxin are capable of recognizing the human insulin receptor, while some conformers are capable of recognizing the Samia bombyxin receptor. Therefore, it is concluded that while the A-chains of bombxyin-II and human insulin are interchangeable, their B-chains are not, and the B-chains determine the receptor specificity between bombyxin-II and human insulin (Nagata et al., 1992b).

In Chapter 4, in order to localize the requirements for recognition of the Samia bombyxin receptor in the bombyxin-II B-chain, the bombyxin-like activity of several chimera molecules of bombyxin-II and human insulin was evaluated. When the residues in the B-chain middle part (B6 to B18) of bonsulin is converted (from the insulin-type) to the bombyxin-type [bonsylin-

(6-18)], the bombyxin-like activity is fully restored. Therefore, the middle part (B6 to B18) exclusively is required for recognition of the Samia bombyxin receptor in the bombyxin-II B-chain. Comparison of the solution structures of bonsylin-(6-18) and bonsulin shows the peculiar patch formed in the middle part of bombyxin-II B-chain is of critical importance for recognition of the Samia bombyxin receptor.

In Chapter 5, in order to evaluate the significance of each side chain in the middle part of bombyxin-II B-chain (TyrB6 to LeuB18) systematically, a series of Ala-scanning analogs (the native Ala residues are replaced by insulin-type residues) were synthesized and bioassayed. The bioassay revealed that the side chains of TyrB6, LeuB11, AlaB12, ThrB14, LeuB15, AlaB16, AspB17 and LeuB18 play roles, directly or indirectly, for recognition of the Samia bombyxin receptor, while those of ArgB9, HisB10 and ArgB13 are not important for recognition of the Samia bombyxin receptor. Thus, all the residues important for recognition of the Samia bombyxin receptor in the bombxyin-II B-chain are identified. On the other hand, the residues in the N- and C-termini, GlyA1 (Maruyama et al., 1991), ValA3 (Nagata et al., in preparation) and A20-B19 disulfide (Nagata et al., 1992a), of the bombyxin-II A-chain are also shown to be important for the activity.

The residues important for activity are mapped on the three-dimensional structure of bombyxin-II, and the exposed side chains of the important residues are considered to be involved in the recognition of the Samia bombyxin receptor (Figure 6-3). The important residues forms two patches on the molecular surface: (1) The receptor-recognition region I of bombyxin-II (rrrBBX-I) involving the following important residues, GlyA1, ValA3, CysA20, LeuB11, AlaB12, LeuB15, AlaB16 and CysB19. The highly conserved residues, IleA2 and TyrA19, may also be involved in receptor recognition. (2) rrrBBX-II involving the following important residues, TyrB6, ThrB14, AspB17 and LeuB18. The highly conserved residues, ValA13, may also be involved in the Samia bombyxin receptor recognition. Therefore, the site of bombyxin-II for recognition of the Samia bombyxin receptor consists of two regions, rrrBBX-I and rrrBBX-II.

(d) Comparison of the receptor recognition site of bombyxin-II with those of vertebrate counterparts, human insulin and human relaxin 2.

In order to elucidate the structural basis of receptor specificity between bombyxin-II and human insulin and to compare the mechanisms of receptor recognition between invertebrate and vertebrate insulin-superfamily peptides, the recepetor-recognition site of bombyxin-II is compared with those of vertebrate counterparts, human insulin and human relaxin 2 (Figure 6-4).

The receptor-recognition site of bombyxin-II mostly overlaps with that of insulin, except for B-chain C-terminal part (Schäffer, 1994). Bombyxin-II does not require the B-chain C-terminal part for its activity, while human insulin requires the aromatic residues (PheB24-PheB25-TyrB26) in the B-chain C-terminal β-strand for the insulin receptor recognition (Mirmira et al., 1991). The difference in function of their B-chain C-terminal parts reflects the difference in their main-chain conformations. In bombyxin-II, the B-chain C-terminal part lies out of the receptor-recognition site. Whereas in insulin, the B-chain C-terminal β-strand covers the hydrophobic surface which corresponds to the rrrBBX-I of bombyxin-II, and the aromatic triplet makes a peculiar patch in the receptor-recognition site of insulin (rrrISL-I). Despite the overlap of the receptor-recognition sites of bombyxin-II and human insulin, their surfaces are not similar. The different residues at B12 (Ala in bombyxin-II/Val in human insulin), B16 (Ala/Tyr), B6 (Tyr/Leu), B17 (Asp/Leu) and B18 (Leu/Val) are shown to contribute to the determination of receptor specificity between bombyxin-II and human insulin. Other different residues in their receptor-recogniton sites may also contribute the receptor specificity.

Human relaxin 2 has a slightly different receptor-recognition region (referred to as the rrrRLX) (Eigenbrot et al., 1991; Büllesbach and Schwabe, 1991) from those of bombyxin-II and human insulin. The rrrRLX locates between the corresponding regions of the rrrBBX/ISL-I and the rrrBBX/ISL-II, and may partly overlap with the rrrBBX/ISL-I (Figure 6-4). Although the two Arg residues at B9 and B13 are apparently conserved in bombyxin-II and human relaxin 2, they are functionally distinct. The side chains of Arg at B9 and B13 of bombyxin-II are not required for recognition of the Samia bombyxin receptor (Chaper 5), while

these side chains in human relaxin 2 are essential for recogniton of the relaxin receptor (Büllesbach and Schwabe, 1988, 1991). The relative orientations of these ligands to their receptors are almost the same, which may be an indication of co-evolution of these ligands and their receptors from an ancestral molecule. Thus, the Samia bombyxin receptor may have a similar structure and function to the insulin receptor, an $\alpha_2\beta_2$ form and tyrosine kinase activity, as does the intrinsic bombyxin receptor of Bombyx mori (Minoru Tanaka, personal communication).

In the case of insulin, a model for the binding to the insulin receptor is proposed based on the structure-activity relationships of a number of insulin analogs (Figure 6-5) (Schäffer, 1994). In the model, the insulin molecule bridges the two α subunits through the two different intermolecular interactions, rrrISL-I—IrsIRα-I and rrrISL-II—IrsIRα-II (the IrsIRα-I and -II represent the ligand-recognition sites in the α subunits of the insulin receptor). The bridging step gives rise to a conformational change of the insulin receptor which is transmitted through the cell membrane, activating the tyrosine kinase. A similar mechanism of ligand—receptor recognition can be assumed for the bombyxin-II—the Samia bombyxin receptor interaction because of (1) the mostly overlapped receptor-recognition sites of bombyxin-II and human insulin and (2) the suggested $\alpha_2\beta_2$ heterotetramer structure and tyrosine kinase activity of the Samia bombyxin receptor. Thus, the two-site receptor-recognition of bombyxin-II would be significant in that it could induce asymmetry in the conformation of the Samia bombyxin receptor (an putative $\alpha_2\beta_2$ heterotetramer), leading to the pupative tyrosine kinase activation.

6-2. Perspectives on bombyxin research

In this thesis, I investigated the structural basis of the specific, high-affinity recognition between bombyxin-II and the Samia bombyxin receptor. The structure of bombyxin-II was analysized extensively, while the structure of the Samia bombyxin receptor is unknown. Therefore, chemical characterization of the Samia bombyxin receptor should be a future topic for study. For this purpose, a bombyxin-II derivative, [(6-biotinylamidocaproyl)GlnB(-2),Phe(pN3)B6]-bombyxin-II, is designed. Since TyrB6 in bombyxin-II lies in the rrrBBX-II and is important for recognition of the Samia bombyxin receptor, the azido group incorporated into PheB6 can be used for photoaffinity labeling of the Samia bombyxin receptor, while the biotin moiety at the N-terminus of the B-chain can serve as the probe for highly sensitive detection and as a ligand for affinity purification and it would not adversely affect receptor-binding affinity. The derivative can be applied to the following studies: (1) investigation of the receptor distribution in an individual worm or pupa, (2) affinity purification of the cross-linked ligand—receptor complexes and (3) identification of the bombyxin-II (particularly, the rrrBBX-II)-recognition site in the Samia bombyxin receptor (see Kurose et al., 1994). The final goal in this line would be the determination of the three-dimentional structure of the bombyxin-II—the Samia bombyxin receptor complex, by which the mechanism of the specific recognition could be solved at atomic resolution. The detailed analysis and comparison of the signaling pathways of bombyxin-II and the Samia PTTH would reveal the reason why bombyxin-II can behave as if it were a PTTH to Samia cynthia ricini.

The elucidation of physiological function(s) of bombyxin in Bombyx mori is the most significant biological topic for study. This requires the characterization and identification of the Bombyx bombyxin receptor, which is now in progress by Minoru Tanaka in our laboratory. Once the receptor is identified, the distribution of the receptor in an individual at verious stages of development would suggest the physiological function(s) of bombyxin.

I hope that not only structural and chemical but also biological studies of bombyxin and related peptides will progress using the results described in this thesis.

Chapter 7
Experimental procedures

7-1. Structural analysis

(a) Preparative HPLC

Preparative RP-HPLC was performed with the following conditions. Two representative elution programs were used.

program 1

HPLC system:	Shimadzu LC-6A
column:	SenshuPak ODS-H (20 x 250 mm) (Senshu Kagaku)
eluent:	10-50% CH_3CN (80 min)/0.10% TFA
	solvent A: 10% CH_3CN/0.10% TFA
	solvent B: 50% CH_3CN/0.09% TFA
flow rate:	8.0 ml/min
column temperature:	40°C
detection:	absorbance at 280 nm.

program 2

HPLC system:	Shimadzu LC-6A
column:	Asahipak ODP-90 (21.5 x 300 mm)
eluent:	10-50% CH_3CN (80 min)/10 mM $CH_3CO_2NH_4$ (pH 6.0)
	solvent A: 10% CH_3CN/10 mM $CH_3CO_2NH_4$ (pH 6.0)
	solvent B: 50% CH_3CN/9.5 mM $CH_3CO_2NH_4$ (pH 6.0)
flow rate:	8.0 ml/min
column temperature:	room temperature
detection:	absorbance at 280 nm.

If these elution programs did not give satisfactory results, the programs were modified to get a better resolution.

(b) Analytical HPLC

Analytical RP-HPLC was performed with the following conditions. Two representative elution programs were used.

program 1

HPLC system:	JASCO Gulliver
column:	SenshuPak Pegasil ODS (4.6 x 150 mm) (Senshu Kagaku)
eluent:	10-50% CH_3CN (40 min)/0.10% TFA
	solvent A: 10% CH_3CN/0.10% TFA
	solvent B: 50% CH_3CN/0.09% TFA
flow rate:	1.0 ml/min
column temperature:	40°C
detection:	absorbance at 280 nm.

program 2

HPLC system:	JASCO Gulliver
column:	SenshuPak Pegasil ODS (4.6 x 150 mm) (Senshu Kagaku)
eluent:	10-50% CH_3CN (40 min)/10 mM $CH_3CO_2NH_4$ (pH 6.0)
	solvent A: 10% CH3CN/10 mM $CH_3CO_2NH_4$ (pH 6.0)
	solvent B: 50% CH3CN/9.5 mM $CH_3CO_2NH_4$ (pH 6.0)
flow rate:	1.0 ml/min
column temperature:	room temperature
detection:	absorbance at 280 nm.

(c) FAB mass spectrometry

FAB mass spectrometry was carried out in the positive mode. Peptides were ionized by xenon atoms or cesium ions. The aquisition systems were used in the PROFILE mode. The peptides were dissolved in 20% CH_3CN containing 0.1% TFA at an average concentration of 1 μg/μl for measurements at a resolution of 1000 and 5 μg/μl for measurements at higher resolutions. Glycerol or 3-nitrobenzyl alcohol was used as the matrix.

(d) Amino acid sequence analysis

Peptides were sequenated on an Applied Biosystems Inc. 477A protein sequencer with an on-lined ABI 120A phenylthiohydantoyl amino acid analyzer.

(e) Circular dichroism spectra measurements

CD spectra were recorded on a JASCO J-600 spectropolarimeter with a 1-mm path length cell at room temperature. Four scans were accumulated. A peptide was dissolved in 50 mM sodium phosphate buffer at a residue concentration of 20 μM for the wavelength range of 200-250 nm. The peptide concentration was determined using the sum of molar

extinction coefficient at 280 nm of each residue: Trp (5400 l/cm/mol), Tyr (1100 l/cm/mol) and disulfide (200 l/cm/mol).

7-2. Solid-phase synthesis of peptide chains

The A- and B-chains with protected side chains were elongated on an Applied Biosystems Inc. 430A peptide synthesizer using the NMP-HOBt Fmoc cycle. The Fmoc-amino acids with protected side chains were Cys(Acm), Cys(tBu), Cys(MBzl), Cys(Trt), Asp(OtBu), Glu(OtBu), His(Trt), Lys(Boc), Asn(Trt), Gln(Trt), Arg(Pmc), Ser(tBu), Thr(tBu), Trp(Boc) and Tyr(tBu). In odrer to form the three disulfide bonds in the regioselective and stepwise fashion, orthogonal Cys protecting groups were used: Trt (TFA labile) for CysA6, CysA11, CysB19; tBu and MBzl (TFA stable and TFMSA labile) for CysA20 and CysB19, respectively; Acm (TFA- and TFMSA stable and I_2 labile) for CysA7 and CysB7. HMP (4-hydroxymethylphenoxymethyl)-resin was used as the solid support. The C-terminal amino acid was attached to the resin through an ester bond using the cycle "RFMC L1". In this cycle, 1.0 mmol of the Fmoc amino acid was activated with DCC as the asymmetric anhydride and subsequently coupled to 0.25 mmol of HMP-resin in the presence of 0.06 equivalent of DMAP as an esterification catalyst for 68 min. Addition of each successive amino acid residue was done using the cycle "RFMC3CL". This cycle involved a reaction and washing cycle containing the following steps: (1) preliminary washing of the polymer support bearing the protected amino acid or peptide chain; (2) cleavage of the terminal Fmoc-protecting group by treatment with 20% piperidine; (3) washing of the protecting group reaction product from the resin together with excess cleavege reagent; (4) addition of 4 equivalents of the Fmoc-amino acid HOBt ester to the washed polymer support; (5) gentle agitation of the resin slurry during the reaction period of 194 min; (6) removal of co-products and excess reagents by thorough washing.

(a) Preparation of Fmoc-amino acid cartridges

1. The caps were removed from the used cartridges.
2. The inside of the cartridges was washed thoroughly with CH3OH and DCM.
3. The washed cartridges were dried up.
4. 1.0 mol of Fmoc-amino acid was put into each cartidge.
5. The cartridge was capped.

Caution

The Fmoc-amino acid should be kept dried and cooled (at 4°C or -20°C) till use.

(b) Peptide-chain assembly on an Applied Biosystems Inc. 430A peptide synthesizer

The peptide-chain assembly was done using an ABI 430A peptide synthesizer essentially according to the users mannual.

1. The synthesis protocol "NMP-HOBt Fmoc" was used.
2. The static run file "OPT 31" was used.
3. The reaction vessel cycles "rbeg31", "RFMC L1", "RFMC3CL" and "rdefmoc" were used for the begin cycle, loading cycle, single coupling cycles and end cycle, respectively.
4. The default settings were used for concentrator vessel cycles and the activator vessel cycles.
5. The dynamic run files and substitutions for the A- and B-chains of bombyxin-II were to be print out.
6. After the peptide chain assembly was completed, the dried peptide resin was weighed and stored at 4°C or -20°C.

Caution

The reaction vessel end cycle "rdefmoc" performed the N-terminal Fmoc-group deprotection; otherwise, the N-terminal Fmoc-group should be removed using piperidine.

(c) TFA cleavage and deprotection

The simultaneous detachment of the peptide from the resin support and removal of all side-chain protecting groups of the amino acid residues except Acm of Cys, tBu of Cys and MBzl of Cys were done by TFA cleavage and deprotection. Scavengers were reagents that trapped TFA-liberated carbonium ions and thereby prevented them from undergoing deleterious side-reactions with sensitive amino acid residues. EDT was an efficient scavenger for tBu, Boc and Trt side-chain protecting groups, but it alone did not prevent Trp alkylation. Thioanisole accelerated Arg(Pmc) removal in TFA. Phenol offered some Tyr and Trp protection. TIPS was added to ensure complete quenching of Trt cations.

1. The cleavage mixture was prepared.
   the cleavage mixture

phenol	0.70 ml
EDT	0.25 ml
thioanisole	0.50 ml
H2O	0.50 ml
TIPS	0.25 ml
TFA	10 ml

2. The dried peptide-resin (0.1-1.5 g) was placed in a round-bottom flask that contained a micro stir bar, and the flask was cooled in an ice bath.

3. The cleavage mixture was cooled in an ice bath, then added to the cooled peptide-resin to give a total volume of 10 ml. After the cleavage mixture had been added, the flask was removed from the ice bath and allowed to warm to room temperature. The flask was stoppered, and the reaction mixture was stirred at room temperature for 1.5 hr.
4. After the reaction time had elapsed, 100 ml of cold Et_2O was added to the reaction mixture to precipitate the peptide.
5. Collect the precipitated peptide and the resin by filtering the mixture through a PTFE membrane filter (pore size 3.0 μm). The protecting group reaction product and excess cleavage reagent was washed with cold Et_2O from the peptide and the resin.
6. The peptide on the filter was dissolved in an aqueous solvent, and the peptide solution was filtered and collected in a beaker. Cold 50% CH_3CN/0.1 M Tris-HCl (pH 8.5) was used as the solvent for the A-chains of bombyxin-II, human insulin and their analogs, and the A-chain solution was subjected to air oxidation. Cold 50% CH_3CN/0.1% TFA was used for the B-chains of bombyxin-II, human insulin and their analogs, and the B-chain solution was stored at -20°C till the dasalting and purification by RP-HPLC.

Caution

The peptide solution should not be left at room temperature.

(d) Air oxidation of the A-chains

The intrachain disulfide bond between CysA6 and CysA11 was formed by air oxidation.

1. The A-chain solution obtained in TFA cleavage and deprotection was diluted with H2O to a peptide concentration of approximately 0.1 mg/ml.
2. The pH of the solution was adjusted to 8.5 with 0.5 M Tris, and the solution was stirred at 4°C for 5 days.
3. The oxidation process was monitored by the analytical RP-HPLC. In general, the oxidazed form ([Cys(Acm)A7,Cys(tBu)A20,A6-A11-cystine]A-chain) was eluted earlier than the reduced form ([CysA6,A11,Cys(Acm)A7,Cys(tBu)A20]A-chain) in the analysis system 1.
4. After the completion of the reaction, the peptide solution was subjected to purification.

(e) Purification of the A-chains

The air-oxidized A-chain ([Cys(Acm)A7,Cys(tBu)A20,A6-A11-cystine]A-chain) was purified by preparative RP-HPLC with the following conditions:

HPLC system:	Shimadzu LC-6A
column:	Asahipak ODP-90 (21.5 x 300 mm)
eluent:	20–50% CH_3CN (60 min)/10 mM $CH_3CO_2NH_4$ (pH 6.0)
	solvent A: 20% CH_3CN/10 mM $CH_3CO_2NH_4$ (pH 6.0)
	solvent B: 50% CH_3CN/9.5 mM $CH_3CO_2NH_4$ (pH 6.0)
flow rate:	8.0 ml/min
column temperature:	room temperature
detection:	absorbance at 280 nm.

1. The pH of the solution of air-oxidized A-chain was adjusted to 6.0 with 1 M CH_3CO_2H.
2. Filter the solution with the PTFE membrane filter (pore size 0.45 μm).
3. The solution was degassed with a sonicator.
4. The filtered, degassed peptide solution was loaded into the HPLC line from the inlet of solvent A. After the loading, the inlet was washed thoroughly with solvent B, followed with solvent A, and then the inlet was put in the solvent A.
5. After the run at the initial conditions for 20 min, the elution program was started.
6. Each peak component was fractionated separately.
7. After the identification of the A-chain by FAB-MS analysis, the fraction containing the purified A-chain was concentrated with a rotary evaporator.
8. The concentrated peptide solution was transferred to a 50-ml polypropylene tube, which had been weighed previously, and then subjected to lyophilization.
9. After the lyophilization, the tube containing the peptide was weighed, and thus the weight of peptide was calculated.
10. The lyophilized peptide was kept dried at -20°C.

(f) Purification of the B-chains

The B-chain ([Cys(Acm)B7,Cys(MBzl)B19]B-chain) was purified by preparative RP-HPLC with the following conditions:

HPLC system:	Shimadzu LC-6A
column:	ODS-H (20 x 250 mm)
eluent:	10–25–40% CH_3CN (10min/30 min)/0.1% TFA
	solvent A: 10% CH_3CN/0.10% TFA
	solvent B: 50% CH_3CN/0.09% TFA
flow rate:	8.0 ml/min
column temperature:	40°C
detection:	absorbance at 280 nm.

1. Dilute the B-chain solution with H_2O to lower the CH_3CN concentration to approximately 10%, and the pH of the solution of B-chain was adjusted to 3.0 with 10% TFA.
2. Filter the solution with the PTFE membrane filter (pore size 0.45 μm).
3. The solution was degassed with a sonicator.
4. The filtered, degassed peptide solution was loaded into the HPLC line from the inlet of solvent A. After the loading, the inlet was washed thoroughly with solvent B, followed with solvent A, and then the inlet was put in the solvent A.
5. After the run at the initial conditions for 20 min, the elution program was started.
6. Each peak component was fractionated separately.
7. After the identification of the B-chain by FAB-MS analysis, the fraction containing the purified B-chain was concentrated with a rotary evaporator.
8. The concentrated peptide solution was transferred to a 50-ml polypropylene tube, which had been weighed previously, and then subjected to lyophilization.
9. After the lyophilization, the tube containing the peptide was weighed, and thus the weight of peptide was calculated.
10. The lyophilized peptide was kept dried at -20°C.

(g) Deprotection of Cys(tBu) at B19

The tBu protection of Cys at B19 was deprotected by the standard TFMSA cleavage procedure.

1. The lyphilized [Cys(Acm)B7,Cys(tBu)B19]B-chain was placed in a round-bottom flask that contains a micro stir bar.
2. thioanisole 1.0 ml
   1,2-ethanedithiol 0.5 ml
   was added, and the mixture stirred at room temperature for 10 min.
2. TFA 10 ml
   was added, and the mixture stirred for 10 min.
3. TFMSA 1.0 ml
   was added dropwise with stirring at room temperature for 30 min.
4. When the reaction is complete,
   cold Et_2O 100 ml
   was added, and the peptide was precipitated. The mixture was stirred for 1 min more.
5. Collect the precipitated peptide by filtering the mixture through a PTFE membrane filter (pore size 3.0 μm). The protecting group reaction product and excess cleavage reagent was washed with cold Et_2O from the peptide and the resin.
6. The peptide on the filter was dissolved in cold 50% CH_3CN/0.1% TFA, and the peptide solution was filtered and collected in a beaker.

Caution

The peptide solution should not be left at room temperature.

(h) Purification of the B-chains

The B-chain ([Cys(Acm)B7,CysB19]B-chain) was purified by preparative RP-HPLC with the following conditions:

HPLC system:	Shimadzu LC-6A
column:	ODS-H (20 x 250 mm)
eluent:	10–25–40% CH_3CN (10min/30 min)/0.1% TFA
	solvent A: 10% CH_3CN/0.10% TFA
	solvent B: 50% CH_3CN/0.09% TFA
flow rate:	8.0 ml/min
column temperature:	40°C
detection:	absorbance at 280 nm.

1. Dilute the B-chain solution with cold H_2O to lower the CH_3CN concentration to approximately 10%, and the pH of the solution of B-chain was adjusted to 3.0 with 10% TFA.
2. Filter the solution with the PTFE membrane filter (pore size 0.45 μm).
3. The solution was degassed with a sonicator.
4. The filtered, degassed peptide solution was loaded into the HPLC line from the inlet of solvent A. After the loading, the inlet was washed thoroughly with solvent B, followed with solvent A, and then the inlet was put in the solvent A.
5. After the run at the initial conditions for 20 min, the elution program was started.
6. Each peak component was fractionated separately.
7. After the identification of the B-chain by FAB-MS analysis, the fraction containing the purified B-chain was concentrated with a rotary evaporator.
8. The concentrated peptide solution was transferred to a 50-ml polypropylene tube, which had been weighed previously, and then subjected to lyophilization.
9. After the lyophilization, the tube containing the peptide was weighed, and thus the weight of peptide was calculated.
10. The lyophilized peptide was kept dried at -20°C.

Caution

The peptide solution should be kept cold to prevent intermolecular disulfide bond formation.

7-3. Regioselective disulfide bond formation

(a) Conversion of the protection of CysA20 from tBu to Pys

[Cys(Acm)A7,Cys(tBu)A20,A6-A11-cystine]A-chain was converted to [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain in one step by the DPDS/TFMSA treatment.

1. Reagents A, B, C were prepared and ice-cooled.
   reagent A
   TFA 1.0 ml
   thioanisole 0.2 ml
   reagent B
   DPDS 6.6 mg dissolved in
   TFA 1.0 ml
   reagent C
   TFA 2.0 ml
   TFMSA 0.4 ml
2. [Cys(Acm)A7,Cys(tBu)A20,A6-A11-cystine]A-chain (2 μmol) was dissolved in reagent A 1.2 ml in an ice bath.
3. To the peptide solution, reagent B 1.0 ml was added dropwise and then reagent C 2.0 ml with stirring and cooling.
4. After the mixture was stirred for 20 min in an ice bath, 50 ml of cold Et_2O was added to the mixture to precipitate the peptide.
5. Collect the precipitated peptide by filtering the mixture through a PTFE membrane filter (pore size 3.0 μm). The protecting group reaction product and excess cleavage reagent was washed with cold Et_2O from the peptide.
6. The peptide on the filter was dissolved in cold 50% CH_3CN/0.1 M Tris-HCl (pH 7.5), and the A-chain solution was stored at -20°C till the dasalting.
7. The peptide solution was desalted by gel filtration with the following conditions:
   column: Sephadex LH-20 (25 x 660 mm) (Pharmacia)
   eluent: 20% CH_3CN/0.1 M $CH_3CO_2NH_4$ (pH 6.0)
   flow rate: approximately 0.75 ml/min
   column temperature: 4°C
   detection: absorbance at 280 nm
   fractionation: 6 ml each.
8. The peptide-containing fractions were collected and filtered with a PTFE membrane filter (pore size 3.0 μm).
9. The A-chain was purified by preparative RP-HPLC with the following conditions:
   HPLC system: Shimadzu LC-6A
   column: ODS-H (20 x 250 mm)
   eluent: 20-50% CH_3CN (30 min)/10 mM $CH_3CO_2NH_4$ (pH 6.0)
   solvent A: 20% CH_3CN/10 mM $CH_3CO_2NH_4$ (pH 6.0)
   solvent B: 50% CH_3CN/9.5 mM $CH_3CO_2NH_4$ (pH 6.0)
   flow rate: 8.0 ml/min
   column temperature: 40°C
   detection: absorbance at 280 nm.
10. Each peak component was fractionated separately.
11. After the identification of the A-chain by FAB-MS analysis, the purified A-chain solution was kept at -20°C or -80°C.

Caution

1. [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain was readily degraded to [Cys(Acm)A7,CysA20,A6-A11-cystine]A-chain and then [Cys(Acm)A7,A7',A6-A11,A6'-A11',A20-A20'-cystine]A-chain dimer under alkaline pH, especially above pH 8.5, and thus the pH of the A-chain solution should be below 7.5.
2. Once the [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain was lyophilized, it was difficult to redissolve the peptide to an aqueous solvent, so the peptide should not be lyophilized but rather kept in the freezer at -20°C or -80°C to avoid degradation.
3. It was difficult to dissolve the [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain into an acidic solvent such as 50% CH3CN/0.1% TFA.
4. Once [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain was treated with DTT or other thiol-containing materials, it was degraded to [Cys(Acm)A7,CysA20,A6-A11-cystine]A-chain and then [Cys(Acm)A7,A7',A6-A11,A6'-A11',A20-A20'-cystine]A-chain dimer.
5. [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain was not suitable for long term storage, so a certain amount of [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain should be left for future needs.

(b) Chain combination

To the [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain dissolved in approximately 30% CH_3CN/10 mM $CH_3CO_2NH_4$ (pH 6.0), four-fifth molar equivalents of the [Cys(Acm)B7,CysB19]B-chain was added dropwise

1. The peptide concentration of the [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain solution was adjusted to 280 μM.
2. The lyophilized B-chain was dissolved in 10% CH_3CN/0.1% TFA to a peptide concentration of 160 μM.
3. The retention times of the A- and B-chains on the analytical RP-HPLC were recorded.
3. To the [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain solution, four-fifth molar equivalents of the [Cys(Acm)B7,CysB19]B-chain solution was added dropwise with stirring at room temperature for 30 min.
4. The reaction process was monitored by the analytical RP-HPLC. In general, the heterodimer ([Cys(Acm)A7,B7,A6-A11,A20-B19-cystine]heterodimer) was eluted between the [Cys(Acm)B7,CysB19]B-chain and the [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain in the analysis system 1.
5. After the completion of the reaction, the pH of the solution was adjusted to 3.0 and, the solution was subjected to purification by RP-HPLC with the following conditions:

HPLC system:	Shimadzu LC-6A
column:	ODS-H (20 x 250 mm)
eluent:	10-50% CH_3CN (80 min)/0.10% TFA
	solvent A: 10% CH_3CN/0.10% TFA
	solvent B: 50% CH_3CN/0.09% TFA
flow rate:	8.0 ml/min
column temperature:	40°C
detection:	absorbance at 280 nm.

6. Each peak component was fractionated separately.
7. After the identification by FAB-MS analysis, the purified heterodimer solution was subjected to lyophilization.

Caution

1. [Cys(Acm)A7,Cys(Pys)A20,A6-A11-cystine]A-chain was readily degraded under the alkaline pH, especially at pH above 8.5.
2. [Cys(Acm)B7,CysB19]B-chain was readily form interchain disulfide bond to give the B-chain homodimer at neutral pH and high temperature, thus the lyophilized [Cys(Acm)B7,CysB19]B-chain was dissolved to the solvent just before use and the solution was kept cold on ice.
3. If the molar amount of the B-chain is larger than that of the A-chain, a side product such as [Cys(Acm)B7,B7',B19-B19'-cystine]B-chain homodimer would be also obtained.

(c) Iodine oxidation

1. To the lyophilized [Cys(Acm)A7,B7,A6-A11,A20-B19-cystine]heterodimer (1.7 μmol),

H2O	300 μl
CH3CO2H	2100 μl
1.0 M HCl	24 μl

were added with stirring.

2. To this solution,
13 mM I2 (3.3 mg/CH_3CO_2H 1.0 ml) 3600 μl was added dropwise with stirring.
3. After incubation at 45°C for 20 min, 0.25 M L-ascorbate 1200 μl was added.
4. The reaction mixture was diluted with H2O to the CH3CO2H concentration of 10%, and subjected to preparative RP-HPLC.

7-4. NMR spectra measurements

^1H NMR spectra were measured at 600 MHz on a JEOL JNM-α600 spectrometer at 28°C. DQF-COSY (Rance et al., 1983), PE-COSY (Müller, 1987), TOCSY (45 ms mixing time) with a modified DIPSI-2 pulse sequence (Cavanagh and Rance, 1992) and NOESY (75 or 150 ms mixing time) (Jeener et al., 1979; Macura et al., 1981) were recorded in the phase-sensitive mode (States et al., 1982). Water resonance was suppressed by DANTE pulse (Zuiderweg et al., 1986). Two-dimensional spectra were recorded using a data size of 512 (t_1) x 2048 (t_2) (512 x 4096 for PE-COSY) with a spectral width of 6500 Hz. After zero-filling once in the t_2 and twice in the t_1 dimension, 2048 x 2048 real data matrix were finally obtained and digital resolution was 3.2 Hz/point in both dimensions (512 x 4096 real data matrix and 1.6 Hz/point digital resolution in the F_2 dimension for primitive exclusive COSY).

1. Insert sample.
2. Lock (Sawtooth ---> Lock on).
3. Shim.
4. Tune.
5. Transmitter offset.
6. Calibrate pulse.

(a)Locking the sample

This consisted of manually adjusting the magnetic field to bring the deuterium signal of the sample into resonance with the lock frequency of the spectrometer. Once the lock was found it was important that the lock transmitter was not run at too high a power otherwise the field stabilization would be impaired and shimming would be unresponsive. The lock phase should also be adjusted to give the highest lock signal: this was essential for satisfactory shimming and also ensures that the stabilization circuit was working with a signal in pure dispersion.

(b) Shimming

The object of shimming was to improve the homogeneity of the magnetic field experienced by the sample. This was done by altering the currents in a set of correction coils, each of which was designed to counteract a particular type of field gradient. What was often more significant was the lineshape, particularly at the base of a solvent peak, since this controls how many solute signals were obscured by solvent. The height of the lock signal was used for monitoring the field homogeneity while shimming was in progress. As the experiment was carried out non-spinning, all shimming was also be carried out non-spinning. Because there were significant interactions among the various gradients produced by the axial shim system, the shims were adjusted interactively as a set to ensure that a global optimum was found.

(c) Pulse calibration

Multiple-pulse experiments all depended on applying pulses of particular flip-angles to the spins, and pulse calibrations were therefore an essential stage in setting-up an experiment. A proton observe pulse was calibrated by varying the width of the pulse in a simple pulse acquire sequence until the first and second nulls were found, which corresponded to flip-angles of 180° and 360°, respectively. Assuming that 90°, 180°, and 360° pulse-widths was 1:2:4, the 90° pulse-width was derived.

(d) Acquisition parameters

(a) The transmitter offset. Because each spectrum measured contained H_2O (or HDO) peak that was markedly more intense than the rest of the spectrum it was preferable to set the transmitter was set to be coincident with the H_2O (HDO) peak. This minimized spectral artifacts which arised from errors in digitizing the strong signal.

(b) The spectral width. The spectral widths were chosen to be sufficient to accommodate the spectrum of interest and to avoid adverse effects on the baseline.

(c) The receiver gain. The receiver gain was set as high as possible subject to the absolute requirement that neither the receiver front-end nor the ADC be overloaded. ADC overload if any was detected by looking for clipping of a single scan FID.

(e) Baseline optimization

Unless special care was taken spectra was obtained which have some degree of baseline distortion, typically a DC offset, a slope, a curvature, or some combination of the three. There were two main causes of baseline distortion:

(a) The time-domain signal was delayed in its passage through the receiver, primarily by the audiofrequency filters. This would create a frequency-dependent phase-shift in the spectrum unless sampling started immediately after the delay. Frequency-dependent phase-shifts gave rise to baseline roll.

(b) The receiver, and in particular the audio filters, were not able to respond faithfully to the beginning of the FID since by definition this was a transient which must have frequency components outside the bandwidth of the filter. The resulting distortion of the time-domain data inevitably caused a baseline distortion.

(b) gave rise mainly to a DC offset, leaving only (a) to be corrected at source. Good baselines could therefore be obtained in this case simply by adjusting the delay between the observe pulse and the start of acquisition to give a spectrum needing no first-order phase correction.

(f) Solvent irradiation

Because the absolute coherence between the saturation and the pulse was not able to be obtained with JEOL JNM-α600 spectrometer, a DANTE sequence—a train of pulses of intermediate strength was used in place of the CW irradiation for presaturation. This obviated the need for the very low power level and so also avoided the problem of coupling the low level to the probe. The pulse repetition rate was chosen so that it was at least as great as the offset between the solvent signal and the edge of the spectrum.

(g) General considerations in setting-up 2D experiments

The most important factors were the F_1 spectral width ω_1 and the number of acquisitions n. The two were related and in setting them it was helpful to appreciate that they determine the maximum value of the incrementable delay, since $t_{1max} = n/2\omega_1$. The significance of t_{1max} was that it controled the trade-off between the sensitivity and the attainable F_1 resolution: if t_{1max} was too short the resolution will be impaired and if it was too long the sensitivity will suffer. In biological work t_{1max} values of between 20 and 80 msec were generally used, depending on the application.

(a) The F_1 spectral width. Many of the comments made about the real-time spectral width also applied to ω_1. It was commonplace to use the minimum possible value, in the belief that this gave the best sensitivity and in the interests of attaining the desired t_{1max} with the minimum possible number of acquisitions.

(b) The number of acquisitions. If the above procedure was followed this was simply set to give the desired t_{1max} with the chosen F_1 spectral width. Between 500 and 1000 acquisitions were usually used.

(c) The repetition rate. If the delay between scans was too short artifacts were likely to arise as a result of the spin system not being in the same state at the start of each scan. Conversely, if the delay was too long the sensitivity would

be poor. A good compromise was to allow a delay between scans which was around one to one and a half times the T_1 of the protons of interest, and to precede the delay with a purging sequence which dephased all transverse and z magnetization, thus eliminating the variations with scans. Suitable purging sequences include:

(i) a pair of field-gradient pulses separated by a 90° rf pulse, and
(ii) a pair of spin-lock pulses with a 90° phase-shift between them.

(d) The number of scans per acquisition. This must be set taking the number of acquisitions into account so that the total number of scans in the experiment would give an adequate signal-to-noise ratio. The only constraint was that since 2D pulse sequences invariably used phase-cycling the number of scans must be set in multiples of the cycle size (or possibly in multiples of the size of a subcycle, depending on the sequence).

(e) F_1 baseline optimization. Two of the three causes of baseline distortion also applied to the F_1 dimension, the exception being that signal distortions produced by the receiver were of direct significant for F_2 but not for F_1. The F_1 baselines of hypercomplex experiments could be improved by acquiring the data such that the resulting 2D spectrum needed no first-order phase corrections in F_1.

(h) DQF-COSY

DQF-COSY was used to identify pairs of protons that have a mutual scalar coupling. The DQF-COSY sequence could give more favorable lineshapes—all the diagonal peaks and the cross-peaks in the same antiphase absorption-mode fine-structure than COSY, which gave the diagonal peaks and the cross-peaks which were 90° out of phase in both dimensions. DQF-COSY suffers from the generic COSY problem of giving poor sensitivity unless the relevant splitting is resolvable.

(i) NOESY

The NOESY experiment was usually used to identify pairs of protons that are undergoing cross-relaxation. The characteristic feature of the NOESY pulse sequence was the mixing time T. The cross-peaks were generated by magnetization transfer that took place during the mixing time so the length of this delay must be chosen according to the rate of transfer process. In contrast to the COSY experiment, both cross-peaks and diagonal peaks in NOESY spectra was obtained with absorption-mode lineshapes, and if any fine structure was resolvable in the NOESY cross-peaks all the multiplet components will also be in phase. If presaturation is used in a NOESY experiment, it was usually also necessary to irradiate the solvent during the mixing time. A complicating factor was that the NOESY pulse sequence can not distinguish between magnetization transfer caused by cross-relaxation and magnetization transfer caused by chemical exchange. In macromolecules both give cross-peaks of the same sign as the diagonal.

(j) TOCSY

The idea underlying the TOCSY experiment was that coupled spins could be made to exchange magnetization by applying a pulse train which eliminated the effects of the chemical shift but retains the scalar coupling. This phenoinenon was known as isotropic mixing and the advantage of it was that to a good approximation all the multiplet components in the resulting cross-peaks had the same phase; as a consequence, the cross-peaks did not suffer 'self cancellation' if the multiplet structure was poorly resolved. A second advantage was that the cross-peaks and diagonal peaks all have absorption-mode line-shapes, again to a good approximation. In practice the mixing time is usually chosen on a fairly rough-and-ready basis, with values of 25-50 msec being used to favour large couplings and single-step transfers, and values from 50 to 100 msec for smaller couplings and multi-step transfers.

7-5. Sequence-specific ^{1}H NMR assignment

(a) Stage 1: spin system identification

The first stage of assignment involves the identification of systems of spin-spin coupled resonances. Because there is no spin-spin coupling across the peptide bond each of these spin systems corresponds to the resonances of an individual amino acid residue. The random coil chemical shift values for resonances in the common amino acids are shown in Figure. Spin system information can be obtained from two types of experiment: DQF-COSY and TOCSY. Many of the amino acids have unique spin system topologies and will give rise to unique patterns of cross-peaks in the COSY and TOCSY spectra. The cross-peaks in the DQF-COSY spectra are characterized by a pattern of antiphase fine-structure which reflects the active and passive coupling constants of spin system. By adjusting the isotropic mixing period used in the TOCSY experiment long-range connectivities in the spin system can be observed; in favorable cases complete spin systems can be identified even for Arg and Lys residues.

(b) Stage 2: sequence-specific assignment

The second stage of assignment involves the assignment of an amino acid spin system identified in stage 1 to apecific residue in the protein sequence. This sequence-specific assignment is achieved by correlating one amino acid spin system with the spin systems of its neighboring residues in the sequence. There is no resolvable spin-spin coupling between protons of adjacent residues and, therefore, COSY-type spectra can not be used to delineate the sequential connectivities. Instead, this stage of assignment relies on the short-range through-space connectivities observed in NOESY spectra. The upper limit for the observation of an NOE effect is now around 5.0 Å. Wüthrich and co-workers showed that, for all sterically allowed values of Φ, Ψ, and χ_1 at least one of the distances between HN, Hα, and Hb protons of adjacent residues was short enough to give rise to an observable NOE effect. The identification of two of the three NOE effects, dNN, daN and dbN, is a more reliable criterion for sequential assignment.

Thus, the second stage of assignment would simply involve the identification of sequential NOE effects beginning at the N-terminal residue and continuing to the C-terminal residue. In practice a full set of NOE connectivities from the N- to C-termini is usually not observed. Breaks in the sequential assignment occur for several reasons. First, a sequential dNN NOE will not be resolved from the diagonal if the two HN resonances involved have very similar HN chemical shifts. Secondly, a break may occur at each proline residue. For these reasons the sequential assignment process is carried out in shorter peptide segments within the protein sequence.

7-6. Structure determination from NMR data I. Analysis of NMR data.

(a) Torsion angle constraints for ϕ from DQF-COSY

The $^3J_{HN\alpha}$ values were estimated on the DQF-COSY spectrum (Marion and Wuthrich, 1983), which was operated with zero filling two times in the t_2 dimension without application of any window function, but the apparent values J_{app} were somewhat larger because of the antiphase character of the cross-peaks (Neubaus et al., 1985), so the ϕ angle constraints of $-90° < \phi < -40°$ and $-160° < \psi < -80°$ were made to correspond to the apparent values $J_{app} < 5.5$ Hz and $J_{app} > 8.0$ Hz, respectively (Kline et al., 1988).

(b) Stereospecific assignments and torsion angle constraints for $\chi 1$ from PE-COSY and NOESY

Stereospecific assignments and the dihedral angle constraints from passive couplings revealed on the fine structures of cross-peaks ($C^{\alpha}H$, $C^{\beta}H$) of the PE-COSY spectrum in ^2H_2O solution, $^3J_{\alpha\beta 2}$ and $^3J_{\alpha\beta 3}$ were estimated (Milller, 1987). These coupling constants were combined with the intraresidual NH-$C^{\beta}H$ and $C^{\alpha}H$-$C^{\beta}H$ NOEs in order to achieve the stereospecific assignments of the prochiral $C^{\beta}Hs$ and to estimate the range of $\chi 1$ angles (Wagner et al., 1987; Hyberts et al., 1987).

1. Information on torsion angle constraints and its conversion to XPLOR format

```
% jot bbx01.dih
C6   ch1       t2g3
I2   ph1       <5.5
% dih2xpl bbx01.dih bbxdih01.xpl
% jot bbxdih01.xpl
***BBXA***
***BBXB***
assign
(segid 1BBA and resid   6 and name N  )
(segid 1BBA and resid   6 and name CA )
(segid 1BBA and resid   6 and name CB )
(segid 1BBA and resid   6 and name SG ) 1.0  -60.0   60.0 2
assign
(segid 1BBA and resid   1 and name C  )
(segid 1BBA and resid   2 and name N  )
(segid 1BBA and resid   2 and name CA )
(segid 1BBA and resid   2 and name C  ) 1.0  -65.0   75.0 2
Normal end. How lucky you are !
```

2. Stereospecific assignments of the prochiral $C^{\beta}Hs$ were specified in the assignment list (—.rec).

(c) NOESY cross-peak picking

Crosspeaks in the NOESY spectra recorded in ^1H_2O and ^2H_2O with a mixing time of 75 ms were picked and edited with NMRZ in the case of bombyxin-II, and were picked using a homemade C program and edited with Felix (Biosym Technologies, Inc., San Diego, CA) in the other cases.

1. Peak picking of NOESY cross-peaks and editing of the peak table (—.vxl)
2. Making of the noise list (—.noi)
3. Chemical shift adjustment and making of assignment list (—.rec)

(d) NOESY peak assignments

NOESY peak assignments were automatically made using the following procedure. First, we selected four structures out of ten that satisfied well the NOE constraints and van der Waals repulsion terms. We calculated the distance of each proton pair of all candidates using the four selected structures. Then, these distances were compared with those estimated from the peak intensities. If the differences between the calculated and the observed for a certain pair of protons were always smaller than 1.0 Å, this specific proton pair was selected out of other candidates. When the differences were always larger than 10.0 Å, this proton pair was discarded. In other cases, all the candidates of the proton pairs were reserved for consideration in the next calculations. This assignment process was very negative, but efficient enough to select unambiguously assigned NOESY cross-peaks and exclude incorrect candidates. After several

cycles of iterative NOESY peak assignment and structural calculation process, almost all the NOESY peaks were assigned.

1. Making of candidate list (—pdg)

```
% anlyptab bbxh.rec bbxnnv.txt 0.008 0.008 bbxnn01.pdg
 Number of NOE cross Peaks :   750
% anlyptab bbxh.rec bbxnav.txt 0.008 0.015 bbxna01.pdg
 Number of NOE cross Peaks :   264
% anlyptab bbxd.rec bbxaav.txt 0.015 0.015 bbxaa01.pdg
 Number of NOE cross Peaks :   390
```

2. Removing noises from the candidate list

```
% noise bbxnn01.pdg bbxnnn01.pdg bbxnn01.noi
                    :    0     1    >1  : Total
-------------------------------------------------
 0 pairs            :  179                :  179
 Only 1 pair        :   19    26          :   45
 More than 1 pairs  :   16     0    24    :   40
-------------------------------------------------
 Total              :  214    26    24    :  264
Uniquely assigned =   26
 tertiary     NOE =    7
  (longrange  NOE =    2)
 sequential   NOE =    9
 intraresidue NOE =   10
% noise bbxna01.pdg bbxnna01.pdg bbxna01.noi
                    :    0     1    >1  : Total
-------------------------------------------------
 0 pairs            :  247                :  247
 Only 1 pair        :   40   105          :  145
 More than 1 pairs  :   71     0   287    :  358
-------------------------------------------------
 Total              :  358   105   287    :  750
Uniquely assigned =  105
 tertiary     NOE =   38
  (longrange  NOE =   14)
 sequential   NOE =   27
 intraresidue NOE =   40
% noise bbxaa01.pdg bbxnaa01.pdg bbxaa01.noi
                    :    0     1    >1  : Total
-------------------------------------------------
 0 pairs            :   97                :   97
 Only 1 pair        :   29    29          :   58
 More than 1 pairs  :   74     0   161    :  235
-------------------------------------------------
 Total              :  200    29   161    :  390
Uniquely assigned =   29
 tertiary     NOE =    9
  (longrange  NOE =    3)
 sequential   NOE =    2
 intraresidue NOE =   18
```

(e) Reassignment of ambiguous NOEs on the basis of intermediate structures

```
% pdgscreendo2 bbxnnn01.pdg 5 556888.625 2.8 1.0
-----------------------------------
 NOE peak table : bbxnnn01.pdg
 Number of Peaks :   264
-----------------------------------
input pdb basename : bbx01s
 input pdb index number 1 : 3
 input pdb index number 2 : 8
 input pdb index number 3 : 2
 input pdb index number 4 : 4
 input pdb index number 5 : 5
                    :    0     1    >1  : Total
-------------------------------------------------
 0 pairs            :  214                :  214
 Only 1 pair        :    0    26          :   26
 More than 1 pairs  :    0    19     5    :   24
```

```
-----------------------------------------
 Total              :  214    45     5  :  264
Uniquely assigned =   45
 tertiary      NOE =    9
  (longrange   NOE =    2)
 sequential    NOE =   21
 intraresidue NOE =   15
% pdg2noenn bbxnnn01.pdgs n
-----------------------------------------
 NOE peak table : bbxnnn01.pdgs
 Number of Peaks :   264
-----------------------------------------
Distance Standard (Most intense 5 peaks of seq.HN-HN)
 1. Y19  HN    C20  HN    676353.000
 2. A125 HN    G126 HN    599739.000
 3. G126 HN    V127 HN    527816.000
 4. L16  HN    V15  HN    502562.000
 5. L8   HN    R9   HN    477973.000
Suppose intens. = 556888.625 when distance = 2.8 A.
% pdgscreendx2 bbxnnn01.pdgs 5 556888.625 2.8 10.0
-----------------------------------------
 NOE peak table : bbxnnn01.pdgs
 Number of Peaks :   264
-----------------------------------------
Input pdb basename : bbx01s
 input pdb index number 1 : 3
 input pdb index number 2 : 8
 input pdb index number 3 : 2
 input pdb index number 4 : 4
 input pdb index number 5 : 5
                    :    0     1    >1  : Total
-----------------------------------------
 0 pairs            :  214                :  214
 Only 1 pair        :    0    45         :   45
 More than 1 pairs  :    0     1     4   :    5
-----------------------------------------
 Total              :  214    46     4  :  264
Uniquely assigned =   46
 tertiary      NOE =    9
  (longrange   NOE =    2)
 sequential    NOE =   22
 intraresidue NOE =   15
% pdg2noenn bbxnnn01.pdgss nn
-----------------------------------------
 NOE peak table : bbxnnn01.pdgss
 Number of Peaks :   264
-----------------------------------------
Distance Standard (Most intense 5 peaks of seq.HN-HN)
 1. Y19  HN    C20  HN    676353.000
 2. A125 HN    G126 HN    599739.000
 3. D14  HN    V15  HN    529289.000
 4. G126 HN    V127 HN    527816.000
 5. L16  HN    V15  HN    502562.000
Suppose intens. = 567151.812 when distance = 2.8 A.
% pdgscreendo2 bbxnnn01.pdgss 5 567151.812 2.8 1.0
-----------------------------------------
 NOE peak table : bbxnnn01.pdgss
 Number of Peaks :   264
-----------------------------------------
Input pdb basename : bbx01s
 input pdb index number 1 : 3
 input pdb index number 2 : 8
 input pdb index number 3 : 2
 input pdb index number 4 : 4
 input pdb index number 5 : 5
                    :    0     1    >1  : Total
-----------------------------------------
```

```
 0 pairs          :  214             :  214
 Only 1 pair      :    0    46       :   46
 More than 1 pairs :    0     0     4 :    4
-----------------------------------------------
 Total            :  214    46     4  :  264
Uniquely assigned =   46
 tertiary     NOE =    9
  (longrange  NOE =    2)
 sequential   NOE =   22
 intraresidue NOE =   15
% pdg2noenn bbxnnn01.pdgsss nnn
-------------------------------------
 NOE peak table : bbxnnn01.pdgsss
 Number of Peaks :   264
-------------------------------------
Distance Standard (Most intense 5 peaks of seq.HN-HN)
 1. Y19  HN   C20  HN   676353.000
 2. A125 HN   G126 HN   599739.000
 3. D14  HN   V15  HN   529289.000
 4. G126 HN   V127 HN   527816.000
 5. L16  HN   V15  HN   502562.000
Suppose intens. = 567151.812 when distance = 2.8 A.
% pdgscreendx2 bbxnnn01.pdgsss 5 567151.812 2.8 10.0
-------------------------------------
 NOE peak table : bbxnnn01.pdgsss
 Number of Peaks :   264
-------------------------------------
input pdb basename : bbx01s
 input pdb index number 1 : 3
 input pdb index number 2 : 8
 input pdb index number 3 : 2
 input pdb index number 4 : 4
 input pdb index number 5 : 5
                    :    0     1    >1  : Total
-----------------------------------------------
 0 pairs          :  214             :  214
 Only 1 pair      :    0    46       :   46
 More than 1 pairs :    0     0     4 :    4
-----------------------------------------------
 Total            :  214    46     4  :  264
Uniquely assigned =   46
 tertiary     NOE =    9
  (longrange  NOE =    2)
 sequential   NOE =   22
 intraresidue NOE =   15
% pdg2noenn bbxnnn01.pdgssss nnnn
-------------------------------------
 NOE peak table : bbxnnn01.pdgssss
 Number of Peaks :   264
-------------------------------------
Distance Standard (Most intense 5 peaks of seq.HN-HN)
 1. Y19  HN   C20  HN   676353.000
 2. A125 HN   G126 HN   599739.000
 3. D14  HN   V15  HN   529289.000
 4. G126 HN   V127 HN   527816.000
 5. L16  HN   V15  HN   502562.000
Suppose intens. = 567151.812 when distance = 2.8 A.
```

Proceed to "Converting uniquely assigned NOEs into inter-proton distances"

(f) Manual assignment of NOESY cross-peaks for final structure calculations

A total of 100 calculations were carried out with XPLOR using the YASAP protocol on the NMR-derived distance information. We selected 10 out of 100 converged structures based on the criteria of the smallest residual energy values of distance constraints, dihedral angle constraints and van der Waals repulsion. After fitting to the best converged structure, the 10 structures were averaged and restrained-minimized to give a mean structure.

1. Manual assignment of the residual ambiguously assigned NOEs

```
% pdgscreen bbxnnn30.pdg bbx30Sa1.pdb 1.0 > nn1
% pdgscreen bbxnnn30.pdg bbx30Sa2.pdb 1.0 > nn2
```

```
% pdqscreen bbxnnn30.pdq bbx30Ss3.pdb 1.0 > nn3
% pdqscreen bbxnnn30.pdq bbx30Ss4.pdb 1.0 > nn4
% pdqscreen bbxnnn30.pdq bbx30Ss5.pdb 1.0 > nn5
% cp bbxnnn30.pdq bbxnnn30.pdqc
% jot bbxnnn30.pdqc
NO.   19     8.839   7.935   361940   2
 C122 HN     8.84    7.93    D120 HN
o C122 HN     8.84    7.93    W123 HN
NO.   71     8.601   7.030    76365   2
o Y109 HN     8.61    7.03    Y109 QD
 C110 HN     8.60    7.03    Y109 QD
NO.   74     8.585   7.961   234291   2
 L118 HN     8.59    7.97    R116 HN
o L118 HN     8.59    7.97    T117 HN
```

correct assignment was marked with o.
incorrect assignment was marked with x.
uncertain assignment was left as it was.

```
% pdqrenew bbxnnn30.pdqc bbxnnn30.pdqcn
                    :    0     1     >1  : Total
.....................................................
 0 pairs            :  214                :  214
 Only 1 pair        :    0    44         :   44
 More than 1 pairs  :    0     6     0   :    6
.....................................................
 Total              :  214    50     0   :  264
Uniquely assigned =  50
 tertiary      NOE =  10
  (longrange   NOE =   2)
 sequential    NOE =  25
 intraresidue NOE =  15
% pdqscreen bbxnan30.pdq bbx30Ss1.pdb 1.0 > na1
% pdqscreen bbxnan30.pdq bbx30Ss2.pdb 1.0 > na2
% pdqscreen bbxnan30.pdq bbx30Ss3.pdb 1.0 > na3
% pdqscreen bbxnan30.pdq bbx30Ss4.pdb 1.0 > na4
% pdqscreen bbxnan30.pdq bbx30Ss5.pdb 1.0 > na5
% cp bbxnan30.pdq bbxnan30.pdqc
% jot bbxnan30.pdqc
NO.   16     8.839   4.068   215790   2
o C122 HN     8.84    4.06    L118 HA
 C122 HN     8.84    4.07    A119 HA
NO.   41     8.733   4.628   788136   2
o S12  HN     8.73    4.63    S12  HA
 V13  HN     8.73    4.63    S12  HA
NO.   42     8.733   4.364   329200   2
o S12  HN     8.73    4.36    S12  HBL
 V13  HN     8.73    4.36    S12  HBL
```

correct assignment was marked with o.
incorrect assignment was marked with x.
uncertain assignment was left as it was.

```
% pdqrenew bbxnan30.pdqc bbxnan30.pdqcn
                    :    0     1     >1  : Total
.....................................................
 0 pairs            :  357                :  357
 Only 1 pair        :    0   351         :  351
 More than 1 pairs  :    0    41     1   :   42
.....................................................
 Total              :  357   392     1   :  750
Uniquely assigned =  392
 tertiary      NOE =  122
  (longrange   NOE =   38)
 sequential    NOE =  114
 intraresidue NOE =  156
% pdqscreen bbxaan30.pdq bbx30Ss1.pdb 1.0 > aa1
% pdqscreen bbxaan30.pdq bbx30Ss2.pdb 1.0 > aa2
% pdqscreen bbxaan30.pdq bbx30Ss3.pdb 1.0 > aa3
% pdqscreen bbxaan30.pdq bbx30Ss4.pdb 1.0 > aa4
% pdqscreen bbxaan30.pdq bbx30Ss5.pdb 1.0 > aa5
```

```
% cp bbxaan30.pdg bbxaan30.pdgc
% jot bbxaan30.pdgc
NO.    77     4.433    3.750    320329    2
 D14  HA      4.42     3.74     V15  HA
o P103 HA      4.43     3.75     P103 HDR
NO.    86     4.427    0.873   2663575    3
 D14  HA      4.42     0.88     V106 QG
o P103 HA      4.43     0.88     V106 QG
 C122 HA      4.43     0.88     V127 MG2
NO.   108     4.375    0.841   4398638    3
o L17  HA      4.38     0.84     L17  MDH
 L17  HA      4.38     0.84     L114 QD
 L17  HA      4.38     0.84     L118 QD
```

correct assignment was marked with o.
incorrect assignment was marked with x.
uncertain assignment was left as it was.

```
% pdgrenew bbxaan30.pdgc bbxaan30.pdgcn
                    :    0      1     >1   : Total
-------------------------------------------------
 0 pairs            :  199                  :  199
 Only 1 pair        :    0    171           :  171
 More than 1 pairs  :    0     20      0    :   20
-------------------------------------------------
 Total              :  199    191      0    :  390
Uniquely assigned =  191
 tertiary      NOE =   46
  (longrange   NOE =   13)
 sequential    NOE =   17
 intraresidue  NOE =  128
```

Proceed to "Converting uniquely assigned NOEs into inter-proton distances"

(g) Distance constraints from NOESY

Interproton distance constraints were derived from NOE cross-peak intensities in the NOESY spectra recorded in H_2O and ^2H_2O with a mixing time of 75 ms. The peak intensities in ^2H_2O NOESY were adjusted to those in H_2O using common cross-peaks observed in both NOESYs. The peak intensities were translated into distances on the basis of the relation of NOE intensity ∞ (distance)$^{-6}$ and a standard distance of sequential d_{NN} in α helix = 2.8 Å (Wüthrich, 1986). The upper-bound distance constraints were the calculated distance plus 0.5 Å. The lower-bound constraints were set to 1.8 Å. The distances involving methylene and methyl protons and ring protons of tyrosine were referred to as single $(<r^{-6}>)^{-1/6}$ average distances so that no corrections for center averaging were made (Clore et al., 1986).

1. Converting uniquely assigned NOEs into inter-proton distances

```
% pdg2noenn bbxnnn01.pdg bbxnn01.noe
-------------------------------------
 NOE peak table : bbxnnn01.pdg
 Number of Peaks :   264
-------------------------------------
Distance Standard (Most intense 5 peaks of seq.HN-HN)
 1. Y19  HN    C20  HN    676353.000
 2. A125 HN    G126 HN    599739.000
 3. G126 HN    V127 HN    527816.000
 4. L16  HN    V15  HN    502562.000
 5. L8   HN    R9   HN    477973.000
Suppose intens. = 556888.625 when distance = 2.8 A.
% pdg2noe2 bbxnan01.pdg bbxna01.noe 556888.625 2.8
-------------------------------------
 NOE peak table : bbxnan01.pdg
 Number of Peaks :   750
-------------------------------------
% pdg2noe2 bbxaan01.pdg bbxaa01.noe 1810339.349 2.8
-------------------------------------
 NOE peak table : bbxaan01.pdg
 Number of Peaks :   390
-------------------------------------
```

2. Conversion of distance constraints into XPLOR format

```
% cat bbxnn01.noe bbxna01.noe > bbxnnna01.noe
% cat bbxnnna01.noe bbxaa01.noe > bbxnnnaaa01.noe
% noe2xpl++e bbxnnnaaa01.noe bbx01.xpl
```

```
***BBXA***
***BBXB***
Normal end. How lucky you are !
```

(b) Other constraints

Additional constraints were obtained from disulfide bonds. All if the bombyxin-II, bonsulin, imbyxin and bonsylin-(6-18) contain three disulfide bonds CysA6-CysA11, CysA7-CysB7 and CysA20-CysB19. A disulfide bond was tracted as one distance constraint bctween the two sulfur atoms of which target values are set to 2.02 ± 0.02 Å. No hydrogen-bond constraints were used for structure calculation.

1. Information on disulfide bonds and hydrogen bonds

```
% cp ~hatanaka/WorkSpace/remsshb.ctl .
% jot remsshb.ctl
REM  bombyxin ver.1 (75ms)
SEG  BBXA 1 20
SEG  BBXB 101 128
SS   C6   C11
SS   C7   C110
SS   C20  C122
```

7-7. Structure determination from NMR data II. Computation.

(a) Simulated annealing calculations

Energy minimization and molecular dynamics calculations were made using the program XPLOR (version 2.1, MSI, Waltham, MA) on a Personal Iris 35, an Indigo or an Indigo2 xl (Silicon Graphics Inc., Mountain View, CA). Determination of the three-dimensional structures of proteins from inter-proton distance constraints and torsion angle constraints is to find a global minimum of a nonlinear target function. XPLOR makes use of Newton's equation of motion for this purpose. The target function has the form

$$F_{total} = F_{bond} + F_{angle} + F_{impr} + F_{repel} + F_{NOE} + F_{tor}$$

where F_{bond} and F_{angle} describe the covalent energy terms for maintaining correct bond lengths and bond angles, respectively. Fimpr describes energy involving chirality and planarity. Frepel describes a repulsion term for preventing unduly close contacts of atoms. F_{NOE} and F_{tor} describe square-well potentials for introducing penalties when teh interproton distances or torsion angles deviate from the acceptable value ranges. Detailed definitions are given by Clore et al. (1986) and Brünger (1990).

XPLOR is not a simple program but rather a interpreter of a program language adapted for restrained molecular dynamics. We used some tutorial programs in the XPLOR manual (Brünger, 1990) with slight modifications. Initial structures were generated by the program randomphipsi.inp using random f and y angles with peptide bonds and side chains in an extended conformation and perfect covalent geometry. The program sa.inp, alias YASAP, was used to carry out simulated annealing calculation. YASAP uses a similar but more robust simulated annealing protocol than that used in the hybrid distance geometry-simulated annealing protocol. In addition to the tutorial programs, we made some new XPLOR programs for analyses of the calculatod structures and UNIX shell programs. We call them XPLOR-UP (XPLOR utility programs). These utility programs greatly increase the efficiency of the calculations and analyses. For instance, a C-shell program run_yasap2c automatically generates parameter files containing sequentially changing random seeds and invokes sequentially RANDOMPHIPSI and YASAP for each parameter file to produce, for example, 100 structures. In the present version of XPLOR (version 2.1), the nbfix command for adjusting the van der Waals parameters for the sulfur sulfur interaction in a disulfide bridge does not function. Hence, after the restrained dynamics at 1000 K and just before cooling, a disulfide bridge is connected with a covalent bond with an appropriate patch command.

1. Making structure file and internal coordinates

```
% cp ~kohda/XP/XPLOR/generate.par .
% jot generate.par
<<segname>>    BBXA
<<sequence1>>  gly ile val asp glu cys cys leu arg pro
<<sequence2>>  cys ser val asp val leu leu ser tyr cys
<<patch1>>     patch hcys refe=nil=(segid BBXA and resid  6) end
<<patch2>>     patch hcys refe=nil=(segid BBXA and resid  7) end
<<patch3>>     patch hcys refe=nil=(segid BBXA and resid 11) end
<<patch4>>     patch hcys refe=nil=(segid BBXA and resid 20) end
<<psffile>>    bbxa.psf
<<icfile>>     bbxa.ic
! for multichain   generate psf & ic files for each chain  ! by GENERATE, and use
run_yasap2c
% GENERATE
generate psf & ic files
parameters from generate.par
***remark   log file is generate.oxp.
```

```
% jot generate.par
<<segname>>   BBXB
<<sequence1>>  pyg gln pro gln ala val his thr tyr cys
<<sequence2>>  gly arg his leu ala arg thr leu ala asp
<<sequence3>>  leu cys trp glu ala gly val asp
<<sequence4>>   blank
<<sequence5>>   blank
<<sequence6>>   blank
<<sequence7>>   blank
<<sequence8>>   blank
<<sequence9>>   blank
<<sequence10>>  blank
<<patch1>>    patch hcys refe=n11=(segid BBXB and resid 10) end
<<patch2>>    patch hcys refe=n11=(segid BBXB and resid 22) end
<<psffile>>    bbxb.psf
<<icfile>>     bbxb.ic
! for multichain   generate psf & ic files for each chain  ! by GENERATE, and use
run_yasap2c
% GENERATE
generate psf & ic files
parameters from generate.par
***remark   log file is generate.oxp.
```

2. Performing simulated annealing calculations

```
% cp ~kohda/XP/XPLOR/run_yasap2c .
% jot run_yasap2c
# C-shell
     @ loop = 10        #  Num of run
     set f  = 'bbx01'   #  Name of output file
     @ seed1 = 111111       # initial seed in randomphipsi.inp
     @ incre1 = 173           # increment of seed1
     @ ext  = 1           # start of index
# prepare randomphipsi.par
     if (-e randomphipsi.par) then
     rm randomphipsi.par
     endif
# A chain generation
     ex - randomphipsi.par << EOF
     a
<<parameter>> /usr/people/kohda/XP/XPLOR/parallhsa.pro_kohda
<<psffile>>  bbxa.psf
<<icfile>>   bbxa.ic
<<seed>>     $seed1
<<inipdb>>   ($f)iniA$ext.pdb
<<segment>>  BBXA
*
     wq
EOF
#
     RANDOM
     report_randomphipsi randomphipsi.oxp $f$ext.rpt
# B chain generation
     @ seed2 = $seed1 + 117
     rm randomphipsi.par
     ex - randomphipsi.par << EOF
     a
<<parameter>> /usr/people/kohda/XP/XPLOR/parallhsa.pro_kohda
<<psffile>>  bbxb.psf
<<icfile>>   bbxb.ic
<<seed>>     $seed2
<<inipdb>>   ($f)iniB$ext.pdb
<<segment>>  BBXB
*
     wq
EOF
#
     RANDOM
     report_randomphipsi randomphipsi.oxp temp.$$
```

```
     cat temp.$$ >> $f$ext.rpt
     rm temp.$$
# remove END and concat two pdb files
    grep -v END ($f)iniA$ext.pdb > ($f)ini$ext.pdb
    grep -v REMARK ($f)iniB$ext.pdb >> ($f)ini$ext.pdb
    rm ($f)iniA$ext.pdb ($f)iniB$ext.pdb
# prepare yasap.par
     if (-e yasap.par) then
     rm yasap.par
     endif
     ex - yasap.par << EOF
     a
<<parameter>> /usr/people/kohda/XP/XPLOR/parallhsa.pro_kohda
<<psffile>>  bbxa.psf @bbxb.psf
<<inipdb>>  ($f)ini$ext.pdb
<<noefile>>  bbx01.xpl
<<dihedfile>> bbxdih01.xpl
<<SSbond1>>     patch dis2 refe-1=(segid BBXA and resid 6) refe-2=(segid BBXA and
resid 11) end
<<SSbond2>>     patch dis2 refe-1=(segid BBXA and resid 7) refe-2=(segid BBXB and
resid 10) end
<<SSbond3>>     patch dis2 refe-1=(segid BBXA and resid 20) refe-2=(segid BBXB and
resid 22) end
<<SSbond4>>     blank
<<SSbond5>>     blank
<<SSbond6>>     blank
<<seed>>        1234567       ! seed for dynamics
<<finpdb>>      ($f)$ext.pdb
!<<stg2pdb>>    ($f)stg2$ext.pdb
!<<stg3pdb>>    ($f)stg3$ext.pdb
!<<stg4pdb>>    ($f)stg4$ext.pdb
.
     wq
EOF
#cat yasap.par
      YASAP
      report_yasap yasap.oxp $f$ext.rpt
# remove the initial structure
      rm ($f)ini$ext.pdb
       @ loop--
       @ ext++
       @ seed1 += $increm1
     end
exit
% run_yasap2c
 (New file)
generate a random-phi-psi structure
parameters from randomphipsi.par
***remark   log file is randomphipsi.oxp.
 (New file)
generate a random-phi-psi structure
parameters from randomphipsi.par
***remark   log file is randomphipsi.oxp.
 (New file)
yasap: yet another simulated annealing
parameters from yasap.par
***remark   log file is yasap.oxp.
% grep ERR *oxp
```

(b) Analyses of the calculated structures

Analyses of structures were carried out with XPLOR-UP, and displaying and plotting of structures were carried out with QUANTA (version 3.2, MSI, Waltham, MA) on a Personal Iris. Ribbon drawings of protein structures were generated by the program MolScript (Kraulis, 1991). For quantitative assessment of the convergence of the calculations, a partial sum

$$F_3 = F_{repel} + F_{NOE} + F_{tor}$$

was computed for each structure. The structures with the smallest F_3 values were selected as computationally converged structures. Percentage of retained structures against total calculations was high (> 90%) in all the final calculations. For quantitative comparisons of different structures, minimum RMS deviations (RMSDs) were calculated for the backbone atoms (N, C^{α}, C') or the heavy atoms of a specified range of residues. A mean structure was obtained by averaging the coordinates of the structures that are superimposed in advance with respect to the best converged structure. We excluded residues in poorly defined regions during the superposition due to the absence of NOEs other than intraresidue and sequential ones. Because such an averaged structure was poor in geometry, it was subjected to a restrained Powell minimization: this comprises 200 cycles with the soft van der Waals radii reduced by a factor of s = 0.25, 200 cycles with s = 0.5, and 800 cycles with s = 0.8 (Clore et al., 1986). The final coordinates were obtained by fitting to the mean structure for the backbone atoms in the well-defined regions. RMSDs per residue were calculated between the individual structures and the mean structures.

1. Analyses of the calculated structures

```
% sumreport bbx01
    bbx013   105.29   0.060 51 0
    bbx018   112.22   0.061 49 0
    bbx012   114.27   0.065 52 1
    bbx014   150.92   0.078 53 1
    bbx015   156.06   0.079 48 3
    bbx019   178.28   0.086 51 3
    bbx011   211.88   0.091 55 2
    bbx011   238.51   0.090 55 3
    bbx0110  235.04   0.093 59 3
    bbx017   238.51   0.090 55 3
    bbx016   254.88   0.098 65 3
% cp ~kohda/XP/XPLOR/violation.par .
% cp ~kohda/XP/XPLOR/repeat_violation .
% jot violation.par
<<psffile>>      bbxa.psf @bbxb.psf
<<SSbond1>>      patch dis2 refe=1=(segid BBXA and resid 6) refe=2=(segid BBXA and
resid 11) end
<<SSbond2>>      patch dis2 refe=1=(segid BBXA and resid 7) refe=2=(segid BBXB and
resid 10) end
<<SSbond3>>      patch dis2 refe=1=(segid BBXA and resid 20) refe=2=(segid BBXB and
resid 22) end
<<noefile>>          bbx01.xpl
<<dihedfile>>        bbxdih01.xpl
% jot repeat_violation
set f = 'bbx01'
set list_str = (3 8 2 4 5 9 1 10 7 6)
% repeat_violation bbx01.vio
repeat violation (foreach statement)
output to bbx01.vio
%
awk'$11>.5||$12>.5||$13>.5||$14>.5||$15>.5||$16>.5||$17>.5||$18>.5||$19>.5||$20>.5'
bbx01.vio
     04_4                                3    8    2    4    5    9    1   10    7    6
12    : BBXA    2   HD# - BBXA   19   HB# 5.4    0.035 0.000 0.056 0.070 0.000 0.011
0.000 0.000 0.504 0.055
55    : BBXA    7    HA - BBXB    8    HB 4.3    0.000 0.000 0.000 0.000 0.000 0.046
0.195 0.526 0.000 0.306
130   : BBXA   17    HN - BBXA   19    HN 4.2    0.129 0.135 0.269 0.666 0.135 0.333
0.000 0.138 0.361 0.516
148   : BBXA   19    HN - BBXB   15    HN 3.6    0.281 0.451 0.576 0.111 0.644 0.503
0.425 0.290 0.157 0.496
149   : BBXA   19   HE# - BBXB   15    HA 4.5    0.000 0.000 0.000 0.000 0.000 0.000
0.000 0.000 0.509 0.150
193   : BBXB    7    HA - BBXB    7   HB# 3.9    0.310 0.236 0.241 0.406 0.155 0.236
0.508 0.172 0.222 0.278
213   : BBXB    9   HB# - BBXB    9   HD# 4.0    0.062 0.086 0.000 0.109 0.077 0.758
0.115 0.221 0.230 0.210
280   : BBXB   20    HA - BBXB   22    HN 4.8    0.000 0.000 0.308 0.000 0.000 0.493
0.000 0.000 0.542 0.000
282   : BBXB   20    HA - BBXB   23   HE3 4.3    0.120 0.153 0.326 0.000 0.636 0.000
0.220 0.711 0.000 0.636
293   : BBXB   21   HD* - BBXB   23   HD1 5.8    0.401 0.489 0.420 0.335 0.605 0.241
0.572 0.647 0.283 0.771
```

```
300   : BBXB   22    HN - BBXB   23   HB# 5.2    0.161 0.197 0.000 0.056 0.080 0.531
0.288 0.000 0.368 0.131
% cp ~kohda/XP/XPLOR/chkdih.par .
% jot chkdih.par
<<psffile>>      bbxa.psf @bbxb.psf
<<SSbond1>>      patch dis2 refe=1=(segid BBXA and resid 6) refe=2=(segid BBXA and
resid 11) end
<<SSbond2>>      patch dis2 refe=1=(segid BBXA and resid 7) refe=2=(segid BBXB and
resid 10) end
<<SSbond3>>      patch dis2 refe=1=(segid BBXA and resid 20) refe=2=(segid BBXB and
resid 22) end
% CHKDIH bbx011.pdb bbx01.xpl
analyse a coordinate about dihedral angles
1BBA   6 N    6 CA   6 CB   6 SG  -60.0   60.0  -43.8 ok
1BBA   1 C    2 N    2 CA   2 C   -65.0   25.0  -43.9 ok
```

(c) Analysis of the finally obtained structure

```
% sumreport bbx38.xpl
in sorting
   pdbfile  En+j+r   rmsv  >0 >.5
   bbx3837   172.78   0.066 95 0
   bbx3865   176.17   0.065 93 0
   bbx3823   176.37   0.067 93 0
   bbx3849   178.09   0.067 90 0
   bbx3857   178.82   0.068 84 0
   bbx3858   185.00   0.066 104 0
   bbx3817   188.36   0.067 101 0
   bbx3878   188.40   0.068 99 1
   bbx3827   189.26   0.064 95 0
   bbx3877   190.03   0.067 103 1
% jot fselect
    set f = 'bbx38' #Name of files
    set w = 'bbx385' #Name of selected files
    foreach ext (37 65 23 49 57 58 17 78 27 77)
% fselect
% jot violation.par
<<psffile>>      bbxa.psf @bbxb.psf
<<SSbond1>>      patch dis2 refe=1=(segid BBXA and resid 6) refe=2=(segid BBXA and
resid 11) end
<<SSbond2>>      patch dis2 refe=1=(segid BBXA and resid 7) refe=2=(segid BBXB and
resid 10) end
<<SSbond3>>      patch dis2 refe=1=(segid BBXA and resid 20) refe=2=(segid BBXB and
resid 22) end
<<noefile>>          bbx38.xpl
<<dihedfile>>        bbxdih38.xpl

% jot repeat_violation
    set f = 'bbx385'
    foreach ext (1 2 3 4 5 6 7 8 9 10)
% repeat_violation bbx38.vio
repeat violation (foreach statement)
output to bbx38.vio
%                                                                    a     w     k
'$11>.5||$12>.5||$13>.5||$14>.5||$15>.5||$16>.5||$17>.5||$18>.5||$19>.5||$20>.5'
2bbx38.vio
193   : 1BBA   13    HA - 1BBA   17   HB# 4.1    0.487 0.000 0.445 0.000 0.000 0.487
0.488 0.504 0.031 0.505
% CHKDIH 2bbx380fit1.pdb 04_38dih.xpl
analyse a coordinate about dihedral angles
1BBA   6 N    6 CA   6 CB   6 SG  -60.0   60.0  -91.6 ok
1BBA   8 N    8 CA   8 CB   8 CG  -60.0   60.0 -103.8 ok
1BBA   9 N    9 CA   9 CB   9 CG   60.0   40.0   18.9 violation =    1.1
1BBA  14 N   14 CA  14 CB  14 CG  -60.0   60.0 -103.9 ok
1BBA  20 N   20 CA  20 CB  20 SG  -60.0   60.0  -94.6 ok
1BBA  22 N   22 CA  22 CB  22 SG  -60.0   60.0  -78.3 ok
1BBA   1 C    2 N    2 CA   2 C   -65.0   25.0  -76.8 ok
1BBA   2 C    3 N    3 CA   3 C   -65.0   25.0  -64.2 ok
```

```
1BBA  13 C  14 N  14 CA  14 C  -65.0   25.0  -62.9 ok
1BBA  16 C  17 N  17 CA  17 C  -65.0   25.0  -59.7 ok
1BBA  19 C  20 N  20 CA  20 C  -65.0   25.0  -86.4 ok
1BBA  16 C  17 N  17 CA  17 C  -65.0   25.0  -71.0 ok
1BBB  18 C  19 N  19 CA  19 C  -65.0   25.0  -97.5 violation =    7.5
1BBB  24 C  25 N  25 CA  25 C  -65.0   25.0  -87.6 ok
1BBB  26 C  27 N  27 CA  27 C  -65.0   25.0  -45.4 ok
1BBB  27 C  28 N  28 CA  28 C  -65.0   25.0  -43.2 ok
1BBA   3 C   4 N   4 CA   4 C  -65.0   25.0  -76.5 ok
1BBA   5 C   6 N   6 CA   6 C  -65.0   25.0  -93.4 violation =    3.4
1BBA  17 C  18 N  18 CA  18 C  -65.0   25.0  -95.2 violation =    5.2
1BBB   6 C   7 N   7 CA   7 C  -65.0   25.0  -70.4 ok
1BBB  12 C  13 N  13 CA  13 C  -65.0   25.0  -55.0 ok
1BBB  21 C  22 N  22 CA  22 C  -65.0   25.0  -54.8 ok
1BBB  22 C  23 N  23 CA  23 C  -65.0   25.0  -72.1 ok
1BBB  23 C  24 N  24 CA  24 C  -65.0   25.0  -70.7 ok
% jot adjust.par
<<psffile>>      bbxa.psf @bbxb.psf
<<refpdb>>       bbx381.pdb
<<range>>        segid BBXA and resid 2:8 or segid BBXA and resid 13:19 or segid
BBXB and resid 12:25
<<selection>>    name ca or name n or name c
<<resnum>>       48
% jot repeat_adjust2
     set f = bbx38      #  name of coor to be adjusted w/o num and .pdb
     set g = bbx38ad    #  name of adjusted coor       w/O num and .pdb
     foreach ext (1 2 3 4 5 6 7 8 9 10)
% repeat_adjust2
repeat adjust (foreach statement)
adjust a coordinate
adjusted coor is bbx38ad1.pdb
***remark   log file is adjust.oxp.
adjust a coordinate
adjusted coor is bbx38ad2.pdb
***remark   log file is adjust.oxp.
1BBA ASP 14 2
1BBA ILE 2 3- cd1
1BBA ILE 2 3- cg2
1BBA LEU 17 3+ cd1
1BBA LEU 17 3- cd2
1BBA TYR 19 2
1BBA VAL 15 3+ cg1
1BBA VAL 3 3+ cg1
1BBB ALA 15 3+
1BBB ALA 25 3-
1BBB ALA 5 3+
1BBB ARG 12 2
1BBB ASP 20 2
1BBB GLU 24 2
1BBB LEU 14 3+ cd2
1BBB LEU 18 3+ cd1
1BBB LEU 18 3- cd2
1BBB LEU 21 3+ cd1
1BBB LEU 21 3+ cd2
1BBB THR 17 3-
1BBB THR 8 3-
1BBB VAL 27 3+ cg1
1BBB VAL 27 3+ cg2
1BBB VAL 6 3- cg1
% jot average+.par
<<psffile>>      bbxa.psf @bbxb.psf
<<inpcoor>>      bbx38ad
<<refpdb>>       bbx38ad1.pdb
<<startindex>>   1
<<endindex>>     10
<<coornum>>      10
```

```
<<range>>          segid BBXA and resid 2:8 or segid BBXA and resid 13:19 or segid
BBXB and resid 12:25
<<selection>>      name ca or name n or name c
<<avepdb>>         bbx38ave.pdb
% AVERAGE+
compute averaged coordinates v2.11
parameters from average+.par
***remark   log file is average+.oxp
% jot refine.par
<<psffile>>        bbxa.psf @bbxb.psf
<<avepdb>>         bbx38ave.pdb
<<noefile>>        bbx38.xpl
<<dihedfile>>      bbxdih38.xpl
<<SSbond1>>     patch dis2 refe=1=(segid BBXA and resid 6) refe=2=(segid BBXA and
resid 11) end
<<SSbond2>>     patch dis2 refe=1=(segid BBXA and resid 7) refe=2=(segid BBXB and
resid 10) end
<<SSbond3>>     patch dis2 refe=1=(segid BBXA and resid 20) refe=2=(segid BBXB and
resid 22) end
<<avempdb>>        bbx38aven.pdb
% REFINE
refine an averaged structure
parameters from refine.par
***remark   log file is refine.oxp.
 X-PLOR>energy end
 --------------- cycle=     1 ----------------------------------------------
 | Etotal =1257.607   grad(E)=0.960      E(BOND)=41.548      E(ANGL)=986.191    |
 | E(IMPR)=59.221     E(VDW )=48.565     E(CDIH)=5.570       E(NOE )=116.512    |
 ---------------------------------------------------------------------------
 NOEPRI: RMS diff. =   0.066,  #(viol.> 0.0)=   105 of    535 NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.0)=   105 of    535 NOEs
 RMS deviation=   0.008
 RMS deviation=   2.227
 RMS deviation=   1.142
 RMS deviation=  31.082
% jot translate.par
! translate.par
<<psffile>>        bbxa.psf @bbxb.psf
<<inputpdb>>       bbx38aven.pdb
<<outputpdb>>      bbx38fit0.pdb

TRANSLATE
translate a set of coordinates
parameters from translate.par
***remark   log file is translate.oxp.

% jot fit+.par
<<psffile>>        bbxa.psf @bbxb.psf
<<meanpdb>>        bbx38fit0.pdb
<<inpcoor>>        bbx38ad             ! name w/o '.pdb' or full file name
<<startindex>>     1
<<endindex>>       10
<<range>>          segid BBXA and resid 2:8 or segid BBXB and resid 13:19 or segid
BBXB and resid 12:25
<<selection>>      name ca or name n or name c
<<outcoor>>        2bbx38fit           ! name w/o '.pdb' or full file name
% FIT+
fit coors to mean  v2.11
parameters from fit+.par
***remark   log file is fit+.oxp
% jot superpose
  @ loop = 10                  # number of coor
  set f = '2bbx38fit'          #  name of coor
  set h = '2bbx38sup'   # name of superpose coors
% jot rmsd+2c.par
<<psffile1>>       bbxa.psf
<<segid1>>         BBXA
```

```
<<psffile2>>          bbxb.psf
<<segid2>>            BBXB
<<1coor>>             bbx38fit
<<2coor>>             bbx38fit
<<fit_range1>>        resid 2:8 or resid 13:19
<<fit_selection1>>    name ca or name n or name c
<<fit_range2>>        resid 12:25
<<fit_selection2>>    name ca or name n or name c
<<rms_range1>>        resid 2:8 or resid 13:19
<<rms_selection1>>    name ca or name n or name c
<<rms_range2>>        resid 12:25
<<rms_selection2>>    name ca or name n or name c
% jot rmsd2c_res
      set psffile1 = bbxa.psf
      set psffile2 = bbxb.psf
      set segid1   = BBXA
      set segid2   = BBXB
      @ resnum   = 20       # A-chain length or B-chain length
      set aorb     =  1     # If A-chain, =1 or if B-chain, =2
      set mean = bbx38fit0.pdb     # Do not start with a digit
      set coor = bbx38fit           # $coor$extstart.pdb .... $coor$extend.pdb
      @ extstart = 1
      @ extend   = 10
<<fit_range1>>        resid 2:8 or resid 13:19
<<fit_selection1>>    name ca or name n or name c
<<fit_range2>>        resid 12:25
<<fit_selection2>>    name ca or name n or name c
<<rms_range1>>        resid 0
<<rms_selection1>>    name ca or name n or name c
<<rms_range2>>        resid 0
<<rms_selection2>>    name ca or name n or name c
% rmsd2c_res bbx38rmsd.resA
compute RMSDs per residue
% jot 2bbx38rmsd.resA
1       1.33438         0.495236
2       0.41839         0.254019
3       0.307839        0.171009
4       0.439927        0.211644
5       0.40061         0.187008
6       0.356931        0.218334
7       0.408437        0.197846
8       0.531024        0.25641
9       0.596197        0.231419
10      0.679866        0.314589
11      0.560068        0.304589
12      0.552146        0.205441
13      0.565983        0.254324
14      0.671118        0.236694
15      0.495066        0.228355
16      0.386212        0.16875
17      0.427876        0.283249
18      0.465905        0.284902
19      0.358653        0.188092
20      0.684244        0.216785
% jot rmsd2c_res
      set psffile1 = bbxa.psf
      set psffile2 = bbxb.psf
      set segid1   = BBXA
      set segid2   = BBXB
      @ resnum   = 20       # A-chain length or B-chain length
      set aorb     =  2     # If A-chain, =1 or if B-chain, =2
      set mean = bbx38fit0.pdb     # Do not start with a digit
      set coor = bbx38fit           # $coor$extstart.pdb .... $coor$extend.pdb
      @ extstart = 1
      @ extend   = 10
<<fit_range1>>        resid 2:8 or resid 13:19
<<fit_selection1>>    name ca or name n or name c
```

```
<<fit_range2>>       resid 12:25
<<fit_selection2>>   name ca or name n or name c
<<rms_range1>>       resid 0
<<rms_selection1>>   name ca or name n or name c
<<rms_range2>>       resid 0
<<rms_selection2>>   name ca or name n or name c
% rmsd2c_res bbx38rmsd.resB
compute RMSDs per residue
% jot 2bbx38rmsd.resB
1       11.0034         4.3248
2       9.06256         3.51225
3       7.23653         3.77764
4       6.11486         4.54026
5       3.98118         3.29107
6       2.5259 1.35299
7       1.81898         0.552834
8       0.800619        0.261316
9       0.760128        0.300883
10      0.753721        0.418554
11      0.614088        0.366372
12      0.586715        0.363154
13      0.614232        0.327427
14      0.473105        0.256066
15      0.506728        0.278276
16      0.47035         0.224147
17      0.581704        0.242166
18      0.611808        0.348524
19      0.485647        0.296741
20      0.420538        0.200943
21      0.333848        0.196702
22      0.469334        0.267563
23      0.505595        0.247
24      0.672197        0.295785
25      1.01983         0.394729
26      1.45361         0.278249
27      2.00768         0.950698
28      3.45744         2.35505
% jot rmsd+2c.par
<<psffile1>>         bbxa.psf
<<segid1>>           BBXA
<<psffile2>>         bbxb.psf
<<segid2>>           BBXB
<<1coor>>            bbx38fit
<<2coor>>            bbx38fit
<<fit_range1>>       resid 2:8 or resid 13:19
<<fit_selection1>>   name ca or name n or name c
<<fit_range2>>       resid 12:25
<<fit_selection2>>   name ca or name n or name c
<<rms_range1>>       resid 2:8 or resid 13:19
<<rms_selection1>>   name ca or name n or name c
<<rms_range2>>       resid 12:25
<<rms_selection2>>   name ca or name n or name c
% RMSD+2c -s 0 1-10
***remark   log file is rmsd+2c.oxp
bbx38fit0.pdb  bbx38fit1.pdb   0.429892
bbx38fit0.pdb  bbx38fit2.pdb   0.513556
bbx38fit0.pdb  bbx38fit3.pdb   0.460997
bbx38fit0.pdb  bbx38fit4.pdb   0.620042
bbx38fit0.pdb  bbx38fit5.pdb   0.661511
bbx38fit0.pdb  bbx38fit6.pdb   0.726766
bbx38fit0.pdb  bbx38fit7.pdb   0.694949
bbx38fit0.pdb  bbx38fit8.pdb   0.699238
bbx38fit0.pdb  bbx38fit9.pdb   0.488023
bbx38fit0.pdb  bbx38fit10.pdb   0.263027
RMSD (A)  Ave= 0.5558 Sig. 0.149593
% RMSD+2c -g 1 2 3 4 5 6 7 8 9 10 # 1 2 3 4 5 6 7 8 9 10
***remark   log file is rmsd+2c.oxp
```

```
RMSD (A)  Ave= 0.652739 Sig. 0.2749
% jot rmsd+2c.par
<<psffile1>>          bbxa.psf
<<segid1>>            BBXA
<<psffile2>>          bbxb.psf
<<segid2>>            BBXB
<<lcoor>>             bbx38fit
<<2coor>>             bbx38fit
<<fit_range1>>        resid 2:8 or resid 13:19
<<fit_selection1>>    not hydrogen
<<fit_range2>>        resid 12:25
<<fit_selection2>>    not hydrogen
<<rms_range1>>        resid 2:8 or resid 13:19
<<rms_selection1>>    not hydrogen
<<rms_range2>>        resid 12:25
<<rms_selection2>>    not hydrogen
% RMSD+2c -s 0 1-10
***remark   log file is rmsd+2c.oxp
bbx38fit0.pdb  bbx38fit1.pdb    0.772128
bbx38fit0.pdb  bbx38fit2.pdb    0.996368
bbx38fit0.pdb  bbx38fit3.pdb    1.02559
bbx38fit0.pdb  bbx38fit4.pdb    1.23568
bbx38fit0.pdb  bbx38fit5.pdb    1.09514
bbx38fit0.pdb  bbx38fit6.pdb    1.23453
bbx38fit0.pdb  bbx38fit7.pdb    0.987794
bbx38fit0.pdb  bbx38fit8.pdb    1.09183
bbx38fit0.pdb  bbx38fit9.pdb    1.0859
bbx38fit0.pdb  bbx38fit10.pdb   0.749834
RMSD (A)  Ave= 1.02748 Sig. 0.16438
% RMSD+2c -g 1 2 3 4 5 6 7 8 9 10 # 1 2 3 4 5 6 7 8 9 10
***remark   log file is rmsd+2c.oxp
RMSD (A)  Ave= 1.18019 Sig. 0.426229
% jot ramachandran.par
! ramachandran.par
<<psffile>>   bbxa.psf @bbxb.psf
% jot repeat_rama
    set f = 'bbx38'          # input file name w/o num & .pdb
    set g = 'bbx38_res'      # phi-psi list for each residue +num.phipsi
    set h = 'bbx38_res.Splus'  # to send to mac, cat file of $g$ext
    @ ext = 1
    @ loop = 10
    @ resnum = 48    # residue number
% repeat_rama
analyze phi and psi
% jot bbx38_res.Splus
-76.8  -6.2          2       1.0
-87.9  -3.6          2       1.0
-80.4  -5.0          2       1.0
-86.2  4.6           2       1.0
-47.2  -70.2  2      1.0
-79.0  2.6           2       1.0
-87.6  -22.3  2      1.0
-93.4  4.5           2       1.0
-86.9  -15.4  2      1.0
-90.5  -20.4  2      1.0
% jot elj.par
<<psffile>>         bbxa.psf @bbxb.psf
<<SSbond1>>    patch dis2 refe=1=(segid BBXA and resid 6) refe=2=(segid BBXA and
resid 11) end
<<SSbond2>>    patch dis2 refe=1=(segid BBXA and resid 7) refe=2=(segid BBXB and
resid 10) end
<<SSbond3>>    patch dis2 refe=1=(segid BBXA and resid 20) refe=2=(segid BBXB and
resid 22) end
% jot repeat_elj
    set f = 'bbx38'
    @ ext = 1
    @ loop = 10
```

```
% repeat_elj
EL-J (kcal/mol)       Ave= -154.771 Sig. 11.0654
% ELJ bbx38fit0.pdb
El-j -162.220 Eelec -1181.440
% jot repeat_analysis
    set f = 'bbx38fit'
    foreach ext (1 2 3 4 5 6 7 8 9 10)
% repeat_analysis bbx38fit.ana
repeat analysis (foreach statement)
output to bbx38fit.ana
% jot bbxfit38.ana
ANALYSIS output
 X-PLOR> coor @bbx38fit1.pdb
      ******** RMS viol in NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.5)=    0 of    535 NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.4)=    2 of    535 NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.3)=    5 of    535 NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.2)=   18 of    535 NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.1)=   47 of    535 NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.0)=   95 of    535 NOEs
       ******* RMS viol in bonds, angles, impropers
 RMS deviation=   0.008
 RMS deviation=   2.251
 RMS deviation=   1.183
 ------------- cycle=     1 ----------------------------------------------------
 | Etotal =1286.544   grad(E)=3.463      E(BOND)=43.116      E(ANGL)=1006.994    |
 | E(IMPR)=63.489     E(VDW )=55.539     E(CDIH)=1.451       E(NOE )=115.956     |
 -------------------------------------------------------------------------------
% ANALYSIS bbx38fit0.pdb
analyse a coordinate
report file is bbx38fit0.ana
% jot bbx38avem.ana
 X-PLOR> coor @bbx38fit0.pdb
      ******** RMS viol in NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.5)=    0 of    535 NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.4)=    1 of    535 NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.3)=    5 of    535 NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.2)=   20 of    535 NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.1)=   48 of    535 NOEs
 NOEPRI: RMS diff. class FINA =   0.066,  #(viol.> 0.0)=  105 of    535 NOEs
       ******* RMS viol in bonds, angles, impropers
 RMS deviation=   0.008
 RMS deviation=   2.228
 RMS deviation=   1.143
 RMS deviation=  31.082
 ------------- cycle=     1 ----------------------------------------------------
 | Etotal =1257.942   grad(E)=1.961      E(BOND)=41.725      E(ANGL)=986.295     |
 | E(IMPR)=59.366     E(VDW )=48.579     E(CDIH)=5.580       E(NOE )=116.416     |
 -------------------------------------------------------------------------------
% statistics bbx38.ana bbx38.sta
filename num>.5 num>.0 rms.viol E(noe) E(cdih) E(vdw) E(total) E(bond) E(angle)
E(impr) category Enoe+cdih+vdw
% jot 2bbx38.sta
bbx381.pdb     0      95      0.066  115.864       1.457  55.553 1286.225
       42.984 1006.946       63.421 0.065-0.07    172.874       150-175
bbx382.pdb     0      93      0.065  113.519       2.196  60.500 1286.735
       48.912  996.403       65.206 0.065-0.07    176.215       175-200
bbx383.pdb     0      93      0.067  120.844       4.016  51.423 1310.977
       45.132 1020.906       68.656 0.065-0.07    176.283       175-200
bbx384.pdb     0      90      0.067  118.591       2.267  57.197 1302.844
       44.448 1013.977       66.364 0.065-0.07    178.055       175-200
bbx385.pdb     0      84      0.068  124.008       4.697  50.060 1303.494
       43.500 1017.520       63.710 0.065-0.07    178.765       175-200
bbx386.pdb     0     104      0.066  118.099       5.101  61.735 1375.369
       43.819 1069.194       77.421 0.065-0.07    184.935       175-200
bbx387.pdb     0     101      0.067  120.644       5.098  62.522 1329.638
       45.105 1036.380       59.889 0.065-0.07    188.264       175-200
```

```
bbx388.pdb    1      99      0.068  122.021       4.938  61.306 1335.560
       46.589 1044.520       56.186 0.065-0.07    188.265        175-200
bbx389.pdb    0      95      0.064  109.493       9.621  70.041 1334.268
       46.411 1031.476       67.226 0.06-0.065    189.155        175-200
bbx3810.pdb   1     103      0.067  118.845       8.300  63.028 1317.145
       45.023 1019.198       62.751 0.065-0.07    190.173        175-200
% report_statistics bbx38.sta 24
rms noe   (A)        Ave= 0.0665 Sig. 0.0012693
rms cdih (deg)       Ave= 3.48779 Sig. 0.98995 +++++ num of cdih lx 24
E(noe)   (kcal/mol)  Ave= 118.193 Sig. 4.28247
E(cdih)  (kcal/mol)  Ave= 4.7691 Sig. 2.59655
E(vdw)   (kcal/mol)  Ave= 59.3365 Sig. 5.93404
Etotal   (kcal/mol)  Ave= 1318.23 Sig. 26.8255
E(bond)  (kcal/mol)  Ave= 45.1923 Sig. 1.74778
E(angle) (kcal/mol)  Ave= 1025.65 Sig. 20.7443
E(impr)  (kcal/mol)  Ave= 65.083 Sig. 5.65156
% report_bai bbx38.ana
rms bond:  0.008 +/- NaN
rms angle:   2.2718 +/- 0.0230786
rms impr:   1.1983 +/- 0.0519274
% statistics bbx38fit0.ana bbx38fit0.sta
filename num>,5 num>,0 rms,viol E(noe) E(cdih) E(vdw) E(total) E(bond) E(angle)
E(impr) category Enoe+cdih+vdw
% jot 2bbx38fit0.sta
bbx38fit0.pdb   0          105       0.066    116.416 5.580     48.579   1257.942
41.725  986.295 59.346  0.065-0.07      170.575 150-175
% jot deposit2PDB
set psf = 'bbxa.psf @bbxb.psf'            # psf file
@ nres  = 48                    # number of residues
set avf = 'bbx38fit0'         # mean structure w/o .pdb
@ loop  = 10                    # Number of individual structures
set fil = 'bbx38fit'            # Name of individual struct w/o num & .pdb
set segid_flag = 'y'              # y -> insert lsegid after aa name (2chain)
                                  # lsegid is the last char of SEGID
set deposit = 'bbx38_deposit.pdb'  # Name of file for deposition
```

Acknowledgments

I am greatly indebted to Professor Akinori Suzuki, Laboratory of Bioorganic Chemistry, The University of Tokyo, for the guidance, encourgement and erudite suggestions.

I wish to express my deep appreciation to Dr. Fuyuhiko Inagaki, Head of Department of Molecular Physiology, The Tokyo Metropolitan Institute of Medical Science, for his direct guidance and valuable comments throughout this work.

It would be sincere to express my gratitude to Emeritus Professor Hironori Ishizaki, The Second Laboratory of Zoology, Department of Biology, School of Science, Nagoya University, for the guidance, encourgement, erudite suggestions and the provision of invaluable brain-removed pupae of Samia cynthia ricini.

I like to express my gratefulness to Professor Hiromichi Nagasawa, Marine Research Institute, The University of Tokyo, Associate Professor Hiroshi Kataoka, Dr. Jiro Nakayama, Laboratory of Bioorganic Chemistry, The University of Tokyo, Dr. Daisuke Kohda, Dr. Hideki Hatanaka, Mr. Kenji Ogura, Department of Molecular Physiology, The Tokyo Metropolitan Institute of Medical Science.

Thanks are also due to Dr. Kazunori Maruyama, Mr. Minoru Tanaka, Mr. Hideki Nakase, Mr. Kuniaki Kojima, Mr. Mamoru Yamamoto and all members of the Laboratory of Bioorganic Chemistry, The University of Tokyo and of the Department of Molecular Physiology, Tokyo Metropolitan Institute of Medical Science for their help and friendship.

I acknowledge Fellowships of the Japan Society for the Promotion of Science for Japanese Junior Scientists.

This thesis was completed with whole heartful support from my family.

References

Adachi, T., Takiya, S., Suzuki, Y., Iwami, M., Kawakami, A., Takahashi, S. Y., Ishizaki, H., Nagasawa, H. and Suzuki, A. (1989). cDNA structure and expression of bombyxin, an insulin-like brain secretory peptide of the silkmoth Bombyx mori. J. Biol. Chem., **264**, 7681-7685.

Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M. C., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N. and Vijayan, N. M. (1988). The structure of 2Zn pig insulin crystals at 1.5 Å resolution. Phil. Trans. Roy. Soc. ser. B, **319**, 369-456.

Bi, R. C., Dauter, Z., Dodson, E., Dodson, G., Giordano, F. and Reynolds, C. (1984). Insulin structure as a modified and monomeric molecule. Biopolymers, **23**, 391-395.

Blundell, T. L. and Wood, S. P. (1975). In the evolution of insulin Darwinian or due to selectively neutral mutation? Nature, **257**, 197-203.

Blundell, T., Dodson, G., Hodgkin, D. and Mercola, D. (1972). Insulin: the structure in the crystal and its reflection in chemistry and biology. Adv. Protein Chem., **26**, 279-402.

Brünger, A. T. (1990). X-PLOR software manual version 2.1. Yale University Press, New Haven.

Büllesbach, E. E. and Schwabe, C. (1988). On the receptor binding site of relaxins. Int. J. Peptide Protein Res., **32**, 361-367.

Carpenter, F. H. (1966). Relationship of structure to biological activity of insulin as revealed by degradative studies. Amer. J. Med., **40**, 750.

Cavanagh, J. and Rance, M. (1992). Suppression of cross-relaxation effects in TOCSY spectra via a modified DIPSI-2 mixing sequence. J. Magn. Reson., **96**, 670-678.

Clore, G. M., Brünger, A. T., Karplus, M. and Gronenborn, A. M. (1986). Application of molecular dynamics with interproton distance restraints to three-dimensional protein structure determination. J. Mol. Biol., **191**, 523-551.

Cooke, R. M., Harvey, T. S. and Campbell, I. D. (1991). Solution structure of human insulin-like growth factor 1: a nuclear magnetic resonance and restrained molecular dynamics study. Biochemistry, **30**, 5484-5491.

Derewenda, U., Derewenda, Z. S., Dodson, G. G. and Hubbard, R. E. (1990). In Handbook of Experimental Pharmacology (Cuatrecasas,P. and Jacobs,S., eds.), Vol. 92, pp. 23-39, Springer-Verlag, Berlin.

Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C. and Swenson, D. (1989). Phenol stabilizes more helix in a new symmetrical zinc insulin hexamer. Nature, **338**, 594-596.

Dodson, E. J., Dodson, G. G., Hodgkin, D. C. and Reynolds, C. D. (1979). Structural relationships in the 2Zn insulin hexamer. Can. J. Biochem., **57**, 469-479.

Eigenbrot, C., Randal, M., Quan, C., Burnier, J., O'Connell, L., Rinderknecht, E. and Kossiakoff, A. A. (1991). X-ray structure of human relaxin at 1.5 Å. Comparison to insulin and implications for receptor binding determinants. J. Mol. Biol., **221**, 15-21.

El-Etr, M., Schorderet-Slatkine, S. and Baulieu, E-E. (1979). Meiotic maturation in Xenopus laevis oocytes initiated by insulin. Science, **205**, 1397-1399.

Fernandez-Almonacid, R. and Rosen, O. M. (1987). Structure and ligand specificity of the Drosophila melanogaster insulin receptor. Mol. Cell. Biol., **7**, 2718-2727.

Hatanaka, H., Oka, M., Kohda, D., Tate, S., Suda, A., Tamiya, N. and Inagaki, F. (1994). Tertiary structure of erabutoxin b in aqueous solution as elucidated by two-dimensional nuclear magnetic resonance. J. Mol. Biol., **240**, 155-166.

Hu, S., Burke, T. and Katsoyannis, P. G. (1993). Contribution of the B16 and B26 tyrosine residues to the biological activity of insulin. J. Protein Chem., **12**, 741-747.

Hua, Q. -X., Kochoyan, M. and Weiss, M. A. (1992). Paradoxical structure and function in a mutant human insulin associated with diabetes mellitus. Proc. Natl. Acad.Sci. USA, **89**, 2379-2383.

Hua, Q. X. and Weiss, M. A. (1990). Toward the solution structure of human insulin: sequential 2D ^{1}H NMR assignment of a des-pentapeptide analogue and comparison with crystal structure. Biochemistry, **29**, 10545-10555.

Hua, Q. X. and Weiss, M. A. (1991). Comparative 2D NMR studies of human insulin: sequential resonance assignment and implications for protein dynamics and receptor recognition. Biochemistry, **30**, 5505-5515.

Hua, Q. X., Shoelson, S. E., Kochoyan, M. and Weiss, M. A. (1991). Receptor binding redefined by a structural switch in a mutant human insulin. Nature, **354**, 238-241.

Ishibashi, J., Kataoka, H., Isogai, A., Kawakami, A., Saegusa, H., Yagi, Y., Mizoguchi, A., Ishizaki, H. and Suzuki, A. (1994). Assignment of disulfide bond location in prothoracicotropic hormone of the silkworm, Bombyx mori: a homodimeric peptide. Biochemistry, **33**, 5912-5919.

Ishizaki, H. and Ichikawa, M. (1967). Purification of the brain hormone of the silkworm Bombyx mori. Biol. Bull., **133**, 355-368.

Ishizaki, H., Mizoguchi, A., Fujishita, M., Suzuki, A., Moriya, I., O'oka, H., Kataoka, H., Isogai, A., Nagasawa, H., Tamura, S. and Suzuki, A. (1983). Species specificity of the insect prothoracicotropic hormone (PTTH): the presence of Bombyx- and Samia-specific PTTHs in the brain of Bombyx mori. Dev. Growth Diff., 25, 593-600.

Iwami, M., Adachi, T., Kondo, H., Kawakami, A., Suzuki, Y., Nagasawa, H., Suzuki, A. and Ishizaki, H. (1990). A novel family C of the genes that encode bombyxin, an insulin-related brain secretory peptide of the silkmoth Bombyx mori: isolation and characterization of gene C-1. Insect Biochem., **20**, 295-303.

Iwami, M., Kawakami, A., Ishizaki, H., Takahashi, S. Y., Adachi, T., Suzuki, Y., Nagasawa, H. and Suzuki, A. (1989). Cloning of a gene encoding bombyxin, an insulin-like brain secretory peptide of the silkmoth Bombyx mori with prothoracicotropic activity. Develop. Growth & Differ., **31**, 31-37.

Jeener, J., Meier, B. H., Bachmann, P. and Ernst, R. R. (1979). Investigation of exchange processes by two-demensional NMR spectroscopy. J. Chem. Phys., **71**, 4546-4553.

Jhoti, H., McLeod, A. N., Blundell, T. L., Ishizaki, H., Nagasawa, H. and Suzuki, A. (1987). Prothoracicotropic hormone has an insulin-like tertiary structure. FEBS Lett., **219**, 419-425.

Jørgensen, A. M. M., Kristensen, S. M., Led, J. J. and Balschmidt, P. (1992). Three-dimensional solution structure of an insulin dimer. A study of the B9(Asp) mutant of human insulin using nuclear magnetic resonance, distance geometry and restrained molecular dynamics. J. Mol. Biol., **227**, 1146-1163.

Kataoka, H., Nagasawa, H., Isogai, A., Tamura, S., Mizoguchi, A., Fujiwra, Y., Suzuki, C., Ishizaki, H. and Suzuki, A. (1987). Isolation and partial characterization of a prothoracicotropic hormone of the silkworm, Bombyx mori. Agric. Biol. Chem., **51**, 1067-1076.

Kawakami, A., Iwami, M., Nagasawa, H., Suzuki, A. and Suzuki, A. (1989). Structure and organization of four clustered genes that encode bombyxin, an insulin-related brain secretory peptide of the silkmoth Bombyx mori. Proc. Natl. Acad. Sci. USA, **86**, 6843-6847.

Kimura-Kawakami, M., Iwami, M., Kawakami, A., Nagasawa, H., Suzuki, A. and Ishizaki, H. (1992). Structure and expression of bombyxin-related peptide genes of the moth Samia cynthia ricini. Gen. Comp. Endocrinol., **86**, 257-268.

Kiriishi, S., Nagasawa, H., Kataoka, H., Suzuki, A. and Sakurai, S. (1992). Comparison of the in vivo and in vitro effects of bombyxin and prothoracicotropic hormone on the prothoracic glands of the silkworm, Bombyx mori. Zool. Sci., **9**, 149-155.

Kitagawa, K., Ogawa, H., Burke, G. T., Chanley, J. D. and Katsoyannis, P. G. (1984). Interaction between the A^2 and A^{19} amino acid residues is of critical importance for high biological activity in insulin. Biochemistry, **23**, 1405-1413.

Kline, A. D. and Justice, R. M. Jr (1990). Complete sequence-specific ^{1}H NMR assignments for human insulin. Biochemistry, **29**, 2906-2913.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr., **24**, 946-950.

Kurose, T., Pashmforoush, M., Yoshimasa, Y., Carroll, R., Schwartz, G. P., Burke, G. T., Katsoyannis, P. G. and Steiner, D. F. (1994). Cross-linking of a B25 azidophenylalanine insulin derivative to the carboxyl-terminal region of the a-subunit of the insulin receptor. J. Biol. Chem., **269**, 29190-29197.

Lagueux, M., Lwoff, M., Meister, M., Goltszené, F. and Hoffmann, J. A. (1990). cDNAs from neurosecretory cells of brains of Locusta migratoria (Insecta, Orthoptera) encoding a novel member of the superfamily of insulins. Eur. J. Biochem., **187**, 249-254.

Macura, S., Huang, Y., Suter, D. and Ernst, R. R. (1981). Two-dimensional chemical exchange and cross-relaxation spectroscopy of coupled nuclear spins. J. Magn. Reson., **43**, 259-281.

Maruyama, K. (1991) Ph.D. Thesis, The University of Tokyo.

Maruyama, K., Hietter, K., Nagasawa, H., Isogai, A., Tamura, S., Suzuki, A. and Ishizaki, H. (1988). Isolation and primary structure of bombyxin-IV, a novel molecular species of bombyxin from the silkworm, Bombyx mori. Agric. Biol. Chem., **52**, 3035-3041.

Maruyama, K., Nagata, K., Tanaka, M., Nagasawa, H., Isogai, A., Ishizaki, H. and Suzuki, A. (1992). Synthesis of bombyxin-IV, an insulin superfamily peptide from the silkworm, Bombyx mori, by stepwise and selective formation of three disulfide bonds. J. Protein Chem., **11**, 1-12.

Mirmira, R. G. and Tager, H. S. (1989). Role of the phenylalanine B24 sidechain in directing insulin interaction with its receptor. J. Biol. Chem., **264**, 6349-6354.

Mirmira, R. G. and Tager, H. S. (1991). Disposition of the pheylalanine B25 side chain during insulin-receptor and insulin-insulin interactions. Biochemistry, **30**, 8222-8229.

Mirmira, R. G., Nakagawa, S. H. and Tager, H. S. (1991). Importance of the character and configuration of residues B24, B25 and B26 in insulin-receptor interactions. J. Biol. Chem., **266**, 1428-1436.

Mizoguchi, A., Hatta, M., Sato, S., Nagasawa, H., Suzuki, A. and Ishizaki, H. (1990). Developmental change of bombyxin content in the brain of the silkmoth Bombyx mori. J. Insect Physiol., **36**, 655-664.

Mizoguchi, A., Ishizaki, H., Nagasawa, H., Kataoka, H., Isogai, A., Tamura, S., Suzuki, A., Fujino, M. and Kitada, C. (1987). A monoclonal antibody against a synthetic fragment of bombyxin (4K-prothoracicotropic hormone) from the silkmoth, Bombyx mori: characterization and immunohistochemistry. Mol. Cell. Endocrinol., **51**, 227-235.

Murray-Rust, J., McLeod, A. N., Blundell, T. L. and Wood, S. P. (1992). Structure and evolution of insulins: implications for receptor binding. BioEssays, **14**, 325-331.

Müller, L. (1987). P. E. COSY, a simple alternative to E. COSY. J. Magn. Reson., **72**, 191-196.

Nagasawa, H., Kataoka, H., Hori, Y., Isogai, A., Tamura, S., Suzuki, A., Guo, F., Zhong, X., Mizoguchi, A., Fujishita, M., Takahashi, S. Y., Ohnishi, E. and Suzuki, A. (1984b). Isolation and some characterization of the prothoracicotropic hormone from Bombyx mori. Gen. Comp. Endocrinol., **53**, 143-152.

Nagasawa, H., Kataoka, H., Isogai, A., Tamura, S., Suzuki, A., Ishizaki, H., Mizoguchi, A., Fujiwara, Y. and Suzuki, A. (1984a). Amino-terminal amino acid sequence of the silkworm prothoracicotropic hormone: homology with insulin. Science, **226**, 1344-1345.

Nagasawa, H., Kataoka, H., Isogai, A., Tamura, S., Suzuki, A., Ishizaki, H., Mizoguchi, A., Fujiwara, Y. and Suzuki, A. (1984a). Amino-terminal amino acid sequence of the silkworm prothoracicotropic hormone: homology with insulin. Science, **226**, 1344-1345.

Nagasawa, H., Kataoka, H., Isogai, A., Tamura, S., Suzuki, A., Mizoguchi, A., Fujiwara, Y., Suzuki, A., Takahashi, S. Y. and Ishizaki, H. (1986). Amino acid sequence of a prothoracicotropic hormone of the silkworm Bombyx mori. Proc. Natl. Acad. Sci. USA, **83**, 5840-5843.

Nagasawa, H., Maruyama, K., Sato, B., Hietter, H., Kataoka, H., Isogai, A., Tamura, S., Ishizaki, H., Semba, T. and Suzuki, A. (1988). Structure and synthesis of bombyxin from the silkworm, Bombyx mori. In Peptide Chemistry 1987 (Shiba,T. and Sakakibara,S., eds.), pp. 123-126, Protein Research Foundation, Osaka.

Nagasawa, H., Maruyama, K., Sato, B., Hietter, H., Kataoka, H., Isogai, A., Tamura, S., Ishizaki, H., Semba, T. and Suzuki, A. (1988). Structure and synthesis of bombyxin from the silkworm, Bombyx mori. In Peptide Chemistry 1987 (Shiba,T. and Sakakibara,S., eds.), pp. 123-126, Protein Research Foundation, Osaka.

Nagata, K., Maruyama, K., Nagasawa, H., Urushibata, I., Isogai, A., Ishizaki, H. and Suzuki, A. (1992a). Bombyxin-II and its disulfide bond isomers: synthesis and activity. Peptides, **13**, 653-662.

Nagata, K., Maruyama, K., Nagasawa, H., Tanaka, M., Isogai, A., Ishizaki, H. and Suzuki, A. (1992b). Synthesis of bombyxin-II, an insulin-like heterodimeric peptide of the silkmoth Bombyx mori, by regiospecific disulfide bond formation. In Peptide Chemistry 1991 (Suzuki,A., ed.), pp. 1-6, Protein Research Foundation, Osaka.

Nagata, K., Momomura, K., Tamori, K., Kadowaki, T., Maruyama, K., Tanaka, M., Kojima, K., Nagasawa, H., Kataoka, H., Isogai, A. and Suzuki, A. (1993). Insulin, bombyxin and their hybrids: synthesis and activity. In Peptide Chemistry 1992 (Yanaihara.N., ed.), pp. 416-419, ESCOM Science Publishers B. V., Leiden.

Nakagawa, S. H. and Tager, H. S. (1986). Role of the phenylalanine B25 side chain in directing insulin interaction with its receptor. Steric and conformational effects. J. Biol. Chem., **261**, 7332-7341.

Nakagawa, S. H. and Tager, H. S. (1986). Role of the phenylalanine B25 side chain in directing insulin interaction with its receptor. Steric and conformational effects. J. Biol. Chem., **261**, 7332-7341.

Nakagawa, S. H. and Tager, H. S. (1987). Role of hte COOH-terminal B-chain domain in insulin-receptor interactions. Identification of perturbations involving the insulin mainchain. J. Biol. Chem., **262**,12054-12058.

Nakagawa, S. H. and Tager, H. S. (1991). Implications of invariant residue Leu^{B6} in insulin-receptor interactions. J. Biol. Chem., **266**, 11502-11509.

Nakagawa, S. H. and Tager, H. S. (1992). Importance of aliphatic side-chain structure at positions 2 and 3 of the insulin A-chain in insulin-receptor interactions. Biochemistry, **31**, 3204-3214.

Nakagawa, S. H. and Tager, H. S. (1993). Importance of main-chain flexibility and the insulin fold in insulin-receptor interactions. Biochemistry, **32**, 7237-7243.

Orikasa, C., Yamauchi, H., Nagasawa, H., Suzuki, A. and Nagata, M. (1993). Induction of oocyte-nurse cell differentiation in the ovary by the brain during the initial stage of oogenesis in the silkworm, Bombyx mori (Lepidoptera: Bombycidae). Appl. Entomol. Zool., **28**, 303-311.

Pullen, R. A., Lindsay, D. G., Wood, S. P., Tickle, I. J., Blundell, T. L., Wollmer, A., Krail, G., Brandenburg, D., Zahn, H., Gliemann, J. and Gammeltoft, S. (1976). Receptor-binding region of insulin. Nature, **259**, 369-371.

Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. and Wüthrich, K. (1983). Improved spectral resolution in COSY proton NMR spectra of proteins via double quantum filtering. Biochem. Biophys. Res. Commun. **117**, 479-485.

Robitzki, A., Schröder, H. C., Ugarkovic, D., Pfeifer, K., Uhrenbruck, G. and Müller, W. E. G. (1989). Demonstration of an endocrine signaling circuit for insulin in the sponge Geodia cydonium. EMBO J., **8**, 2905-2909.

Sato, A., Nishimura, S., Ohkubo, T., Kyogoku, Y., Koyama, S., Kobayashi, M., Yasuda, T. and Kobayashi, Y. (1993). Three-dimensional structure of human insulin-like growth factor-I (IGF-I) determined by ^{1}H-NMR and distance geometry. Int. J. Peptide Protein Res., **41**, 433-440.

Sauber, F., Reuland, M., Ries, E., Holder, F. and Charlet, M. (1990). Invert. Reprod. Dev., **17**, 123-126.

Schäffer, L. (1994). A model for insulin binding to the insulin receptor. Eur. J. Biochem., **221**, 1127-1132.

Schwabe C. and Büllesbach, E. E. (1990). Relaxin. Comp. Biochem. Physiol., **96B**, 15-21.

Sieber, P., Eisler, K., Kamber, B., Riniker, B., Rittel, W., Märki, F. and de Gasparo, M. (1978). Synthesis and biological activity of two disulphide bond isomers of human insulin: [A7-A11,A6-B7-cystine]- and [A6-A7,A11-B7-cystine]insulin (human). Hoppe-Seyler's Z. Physiol. Chem., **359**, 113-123.

Smit, A. B., Geraerts, W. P. M., Meester, I., Heerikhuizen, H. v. and Joosse, J. (1991). Characterization of a cDNA clone encoding molluscan insulin-related peptide II of Lymnaea stagnalis. Eur. J. Biochem., **199**, 699-703.

Smit, A. B., Vreugdenhil, E., Ebberink, R. H. M., Geraerts, W. P. M., Klootwijk, J. and Joosse, J. (1988). Growth-controlling molluscan neurons produce the precursor of an insulin-related peptide. Nature, **331**, 535-538.

States, D. J., Haberkorn, R. A. and Ruben, D. J. (1982). A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. J. Magn. Reson., **48**, 286-292.

Tager, H. S. (1987). in Cohen,M.P. and Foa,P.P. (eds.), Hormone Resistance and Other Endocrine Paradoxes, Springer Verlag New York Inc., New York, pp. 35-61.

Tager, H. S. (1990). in Cuatrecasas,P. and Jacobs,S. (eds.), Handbook of Experimental Pharmacology, Springer-Verlag, Berlin, Vol. 92, pp. 41-64.

Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N. and Wüthrich, K. (1987). Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. J. Mol. Biol., **196**, 611-639.

Wüthrich, K. (1986). NMR of Proteins and Nucleic Acids. John Wiley & Sons, Inc., New York.

Zachary, D., Goltzené, F., Holder, F. C., Berchtold, J. P., Nagasawa, H., Suzuki, A., Mizoguchi, A., Ishizaki, H. and Hoffmann, J. A. (1988). Presence of bombyxin (4K-PTTH)-like molecules in neurosecretory granules of brain-corpora cardiaca complexes of Locusta migratoria developmental aspects. Int. J. Invert. Reprod. Dev., **14**, 1-10.

Zitnan, D., Sehnal, F. and Bryant, P. J. (1993). Neurons producing specific neuropeptides in the central nervous system of normal and pupariation-delayed Drosophila. Dev. Biol., **156**, 117-135.

Zitnan, D., Sehnal, F., Mizoguchi, A., Ishizaki, H., Nagasawa, H. and Suzuki, A. (1990). Developmental changes in the bombyxin- and insulin-like immunoreactive neurosecretory system in the wax moth, Galleria mellonella. Dev. Growth & Differ., **32**, 637-645.

Zuiderweg, E. R. P., Hallenga, K. and Olejniczak, E. T. (1986). Improvement of 2D NOE spectra of biomacromolecules in H_2O solution by coherent suppression of the solvent resonance. J. Magn. Reson., **70**, 336-343.

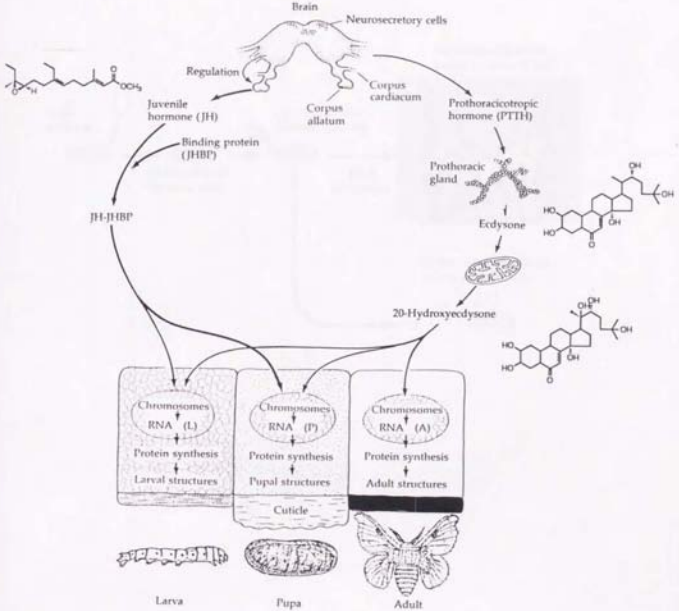


Figure 1-1. Schematic diagram illustrating the control of molting and metamorphosis in the silkmoth *Bombyx mori* (Adapted from Gilbert and Goodman, 1981).

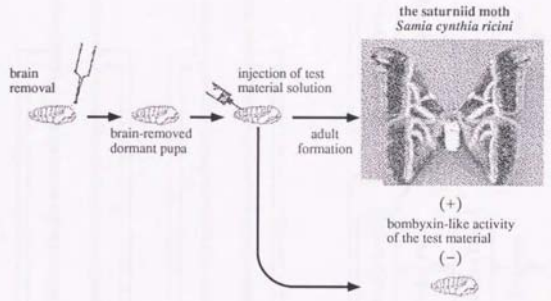


Figure 1-2. *In vivo Samia cynthia ricini* pupal assay for bombyxin-like prothoracicotropic activity.

A

Bombyx PTTH

GNIQVE--NQAIPDPPCTCKYKKEIEDLGENSVPRFIETRNCNKTQQPTCRPPYICKESLYSITILKRRETKSQESLEIPNELKYRWVAESHPVSVACLCTRDYQLRYNNN

GNIQVE--NQAIPDPPCTCKYKKEIEDLGENSVPRFIETRNCNKTQQPTCRPPYICKESLYSITILKRRETKSQESLEIPNELKYRWVAESHPVSVACLCTRDYQLRYNNN

Samia PTTH

GDLRREKHNQAIQDPPCSCGYTQTLLDFGKNAFPRHVVTRNCS-DQQQSCLFPYVCKETLYDVNILKRRETSTQISEEVPRELKFRWIGEKWQISVGCMCTRDYRNSTEDYQPRLLTKIIQQRDLS

GDLRREKHNQAIQDPPCSCGYTQTLLDFGKNAFPRHVVTRNCS-DQQQSCLFPYVCKETLYDVNILKRRETSTQISEEVPRELKFRWIGEKWQISVGCMCTRDYRNSTEDYQPRLLTKIIQQRDLS

bombyxin-II

GIVDECCLRPCSVDVLLSYC

<QQFQAVHTYCGRHLARTLADLCWEAGVD

B

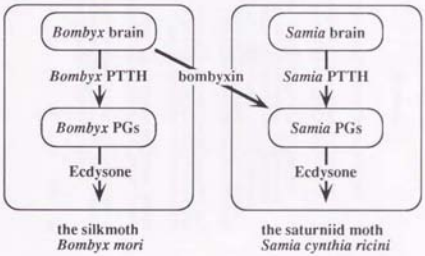


Figure 1-3. Primary structure (**A**) and biological function (**B**) of the *Bombyx* PTTH, the *Samia* PTTH and bombyxin-II. PGs, prothoracic glands. The disulfide-bond location in the *Samia*PTTH is not determined yet, but is probably the same as that in the *Bombyx* PTTH.

Figure 1-4. Amino acid sequences of insulin-superfamily peptides. Invertebrate insulin-related peptides have been identified in a few phyla such as Porifera [sponge insulin from *Geodia cydonium* (Robitzki *et al.*, 1989)], Mollusca [molluscan insulin-related peptides (MIPs) from *Lymnaea stagnalis* (Smit *et al.*, 1988; Smit *et al.*, 1991)] and Arthropoda [bombyxin from the silkmoth *Bombyx mori* (Nagasawa *et al.*, 1984a; Nagasawa *et al.*, 1986; Maruyama *et al.*, 1988; Jhoti *et al.*, 1987); Samia bombyxin-related peptides (SBRPs) from the satumiid moth *Samia cynthia ricini* (Kimura-Kawakami *et al.*, 1992); Agrius bombyxin-related peptide (ABRP) from the potato hornmoth *Agrius convolvuli* (M.Iwami, in preparation); locust insulin-related peptide (LIRP) from the grasshopper *Locusta migratoria* (Lagueux *et al.*, 1990)]. Five molecular species of bombyxin have so far been isolated from the heads of the silkmoth *Bombyx mori* using the *Samia* pupal assay (Ishizaki and Ichikawa, 1967); the primary structure is determined completely for bombyxin-II, -IV and partially for bombyxin-I, -III, -V (Nagasawa *et al.*, 1986 ; Jhoti *et al.*, 1987; Maruyama *et al.*, 1988). Numbering of residues is based on the insulin sequence. Cys residues are boxed. Dashes, gaps in the sequence inserted for best alignment.

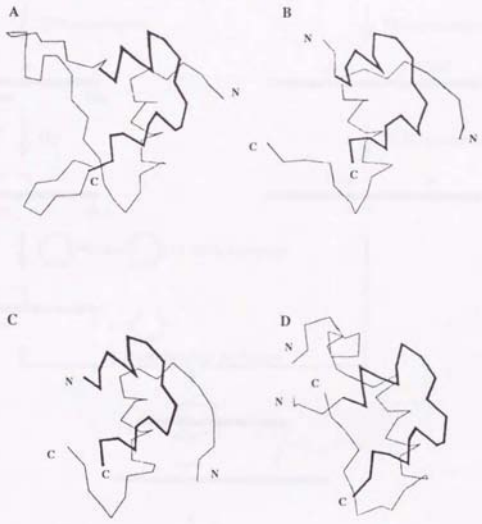


Figure 1-5. Computer models of members of the insulin superfamily. The models were built on computer graphic systems using the coordinates of the insulin crystal structures as a starting point. The regions corresponding to the insulin A-chains are shown as heavy lines. (A) IGF-I, (B) relaxin, (C) bombyxin-II, (D) molluscan insulin-like peptide-1 (From Murray-Rust *et al.*, 1992).

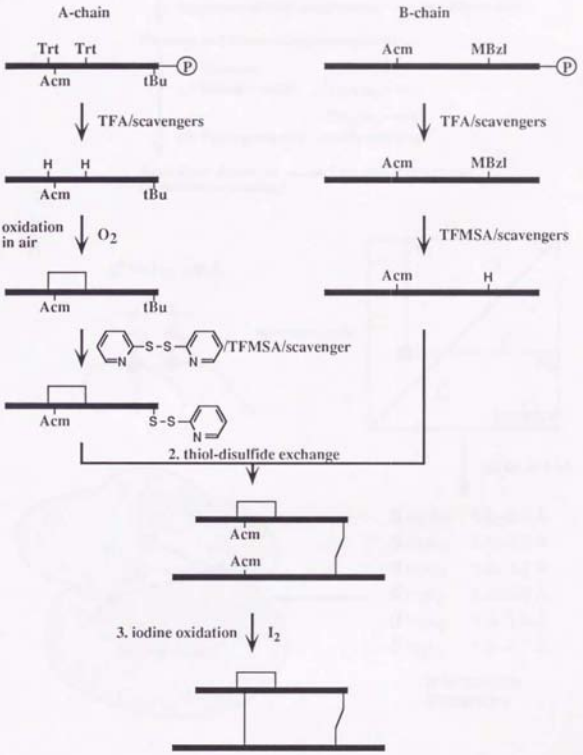


Figure 1-6. Stratery for the synthesis of bombyxin-II and its analogs using a regioselective disulfide bond formation (Maruyama *et al.*, 1992; Nagata *et al.*, 1992b).

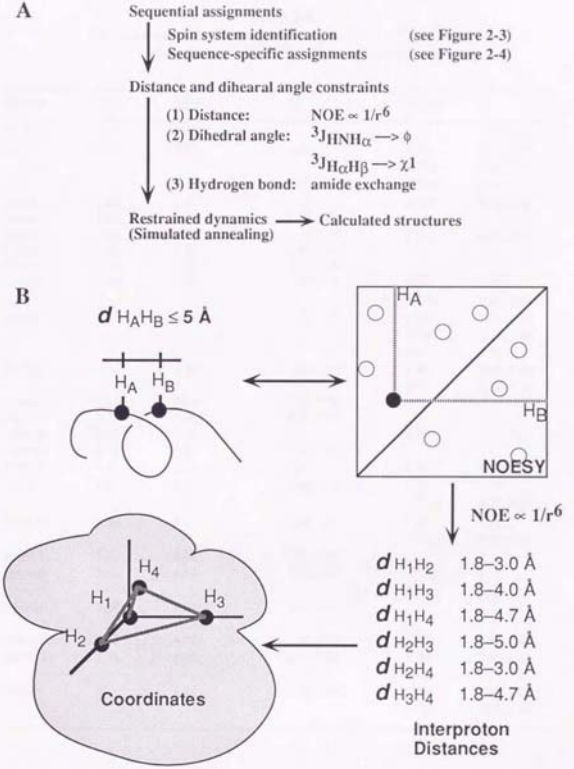


Figure 1-7. (**A**) Steps involved in determination of the three-dimensional structure of a protein by two-dimensional ^{1}H NMR and simulated annealing calculations (Adapted from Wright, 1990). (**B**) Three-dimensional structure determination of a peptide using NOE-derived distance constraints. The cross peaks in the NOESY spectrum arise from through-space correlations and provide information about the spatial proximity of pairs of protons. NOE gives information about inter-proton distances up to a maximum of about 5 Å. A family of structures consistent with the NOE distance constraints are generated.

Table 2-1

^1H chemical shifts in ppm for bombyxin-II at pH 2.0, 28°C, dissolved in 70%/30% water/acetic acid

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	Others	
GlyA1		3.87, 3.75			
IleA2	8.51	3.79	1.20	$C^{\gamma}H$	1.14, 0.94
				$C^{\gamma}H_3$	0.78
				$C^{\delta}H_3$	0.68
ValA3	8.09	3.67	1.98	$C^{\gamma}H3$	0.95, 0.88
AspA4	8.01	4.49	2.88, 2.88		
GluA5	7.94	4.21	2.16, 2.12	$C^{\gamma}H$	2.62, 2.56
CysA6	8.13	5.37	3.39, 2.89		
CysA7	7.93	4.85	3.53, 3.32		
LeuA8	8.10	4.32	1.87, 1.73	$C^{\gamma}H$	1.67
				$C^{\delta}H_3$	0.94, 0.88
ArgA9	7.64	4.61	1.87, 1.72	$C^{\gamma}H$	1.64, 1.64
				$C^{\delta}H$	3.24, 3.24
				$N^{\varepsilon}H$	7.18
ProA10		4.54	2.13, 1.78	$C^{\gamma}H$	2.02, 1.88
				$C^{\delta}H$	3.56, 3.56
CysA11	8.37	5.09	3.61, 3.30		
SerA12	8.73	4.63	4.36, 4.01		
ValA13	8.73	3.76	2.13	$C^{\gamma}H_3$	1.13, 1.02
AspA14	8.18	4.42	2.86, 2.79		
ValA15	7.51	3.75	2.31	$C^{\gamma}H_3$	1.10, 1.10
LeuA16	8.01	4.11	1.94, 1.94	$C^{\gamma}H$	1.74
				$C^{\delta}H_3$	0.77, 0.74
LeuA17	8.33	4.38	1.84, 1.50	$C^{\gamma}H$	1.93
				$C^{\delta}H_3$	0.86, 0.84
SerA18	7.71	4.26	3.89, 3.89		
TyrA19	7.95	4.54	3.34, 3.07	$C^{\delta}H_2$	7.33
				$C^{\varepsilon}H_2$	6.69
CysA20	7.58	4.60	3.23, 2.93		
pGluB(-2)	7.82	4.40	2.55, 2.08	$C^{\gamma}H$	2.47, 2.47
GlnB(-1)	8.38	4.69	2.17, 2.01	$C^{\gamma}H$	2.45, 2.45
				$N^{\varepsilon}H$	7.52, 6.86
ProB0		4.44	2.34, 1.94	$C^{\gamma}H$	2.11, 2.05
				$C^{\delta}H$	3.87, 3.74

Table 2-1
(Continued)

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	Others	
GlnB1	8.46	4.34	2.15, 2.03	$C^{\gamma}H$	2.45, 2.45
				$N^{\epsilon}H$	7.54, 6.87
AlaB2	8.20	4.35	1.42		
ValB3	7.82	4.13	1.97	$C^{\gamma}H_3$	0.88, 0.88
HisB4	8.42	4.79	3.22, 3.13	$C^{\delta}H$	7.27
	8.63			$C^{\epsilon}H$	
ThrB5	7.99	4.57	4.31	$C^{\gamma}H_3$	1.17
TyrB6	8.61	4.47	2.95, 2.95	$C^{\delta}H_2$	7.03
				$C^{\epsilon}H_2$	6.77
CysB7	8.60	4.72	3.16, 3.28		
GlyB8	8.38	4.02, 3.83			
ArgB9	8.48	4.17	1.85, 1.78	$C^{\gamma}H$	1.66, 1.66
				$C^{\delta}H$	3.22, 3.22
				$N^{\epsilon}H$	7.26
HisB10	8.43	4.61	3.42, 3.28	$C^{\delta}H$	7.36
				$C^{\epsilon}H$	8.64
LeuB11	8.17	4.06	1.75, 1.75	$C^{\gamma}H$	1.58
				$C^{\delta}H_3$	0.84, 0.84
AlaB12	8.06	4.04	1.50		
ArgB13	7.97	4.12	1.95, 1.95	$C^{\gamma}H$	1.75, 1.75
				$C^{\delta}H$	3.23, 3.23
				$N^{\epsilon}H$	7.25
ThrB14	7.97	3.99	4.23	$C^{\gamma}H_3$	1.17
LeuB15	8.59	4.06	1.71, 1.61	$C^{\gamma}H$	1.70
				$C^{\delta}H_3$	0.84, 0.84
AlaB16	7.87	4.07	1.50		
AspB17	7.93	4.67	3.12, 3.01		
LeuB18	8.45	4.27	1.84, 1.67	$C^{\gamma}H$	1.84
				$C^{\delta}H_3$	0.90, 0.90
CysB19	8.84	4.43	3.11, 2.96		
TrpB20	7.94	4.58	3.49, 3.49	$N^{1}H$	9.93
				$C^{2}H$	7.31
				$C^{4}H$	7.62
				$C^{5}H$	7.11
				$C^{6}H$	7.19
				$C^{7}H$	7.45
GluB21	8.24	3.95	2.15, 2.15	$C^{\gamma}H$	2.54, 2.43
AlaB22	8.07	4.29	1.50		
GlyB23	7.84	3.97, 3.89			
ValB24	7.69	4.15	2.08	$C^{\gamma}H_3$	0.92, 0.88
AspB25	8.09	4.70	2.90, 2.90		

Table 2-2. *Structural statistics*

	<SA>	$(SA)_r$
Root-mean-square deviations from experimental distance constraints (Å) (535)	0.067 ± 0.001	0.066
Number of distance constraint violations > 0.3 Å	3 - 9 (Maximum 0.51 Å)	5 (Maximum 0.48 Å)
Root-mean-square deviations from experimental dihedral constraints (deg) (24)	3.49 ± 0.99	0.80
Number of dihedral constraint violations > 5°	1 - 4 (Maximum 18.6°)	1 (Maximum 17.5°)
FNOE (kcal/mol)*	118.2 ± 4.3	116.4
Frepel (kcal/mol)*	59.3 ± 5.9	48.6
EL-J (kcal/mol)†	-154.8 ± 11.1	-162.2
RMSDs from idealized geometry		
Bonds (Å) (724)	0.008 ± 0.0000	0.008
Angles (degrees) (1305)	2.27 ± 0.02	2.23
Impropers (degrees) (298)‡	1.20 ± 0.05	1.14

<SA> are the 10 refined simulated annealing structures; $(SA)_r$ is the restrained minimized mean structure, where the mean structure was obtained by averaging the coordinates of the individual <SA> structures best-fitted to each other.

* The value of the square-well NOE potential, F_{NOE}, is calculated with a force constant of 50 kcal/mol per $Å^{-2}$. The value of F_{repel} is calculated with a force constant of 4 kcal/mol per $Å^{-4}$ with the van der Waals radii scaled by a factor of 0.8 of the standard values used in the CHARMm empirical function.

† E_{L-J} is the Lennard-Jones van der Waals energy calculated with the CHARMm empirical energy function, which was not included in the simulated annealing calculations.

‡ The improper torsion term is used to maintain the planar geometry and chirality.

Figure 2-1. Amino acid sequences of insulin-superfamily peptides. Invertebrate insulin-related peptides have been identified in a few phyla such as Porifera [sponge insulin from *Geodia cydonium* (Robitzki *et al.*, 1989)], Mollusca [molluscan insulin-related peptides (MIPs) from *Lymnaea stagnalis* (Smit *et al.*, 1988; Smit *et al.*, 1991)] and Arthropoda [bombyxin from the silkmoth *Bombyx mori* (Nagasawa *et al.*, 1984a; Nagasawa *et al.*, 1986; Maruyama *et al.*, 1988; Jhoti *et al.*, 1987); Samia bombyxin-related peptides (SBRPs) from the saturniid moth *Samia cynthia ricini* (Kimura-Kawakami *et al.*, 1992); Agrius bombyxin-related peptide (ABRP) from the potato hornmoth *Agrius convolvuli* (M.Iwami, in preparation); locust insulin-related peptide (LIRP) from the grasshopper *Locusta migratoria* (Lagueux *et al.*, 1990)]. Five molecular species of bombyxin have so far been isolated from the heads of the silkmoth *Bombyx mori* using the *Samia* pupal assay (Ishizaki and Ichikawa, 1967); the primary structure is determined completely for bombyxin-II, -IV and partially for bombyxin-I, -III, -V (Nagasawa *et al.*, 1986 ; Jhoti *et al.*, 1987; Maruyama *et al.*, 1988). Numbering of residues is based on the insulin sequence. Cys residues are boxed. Dashes, gaps in the sequence inserted for best alignment.

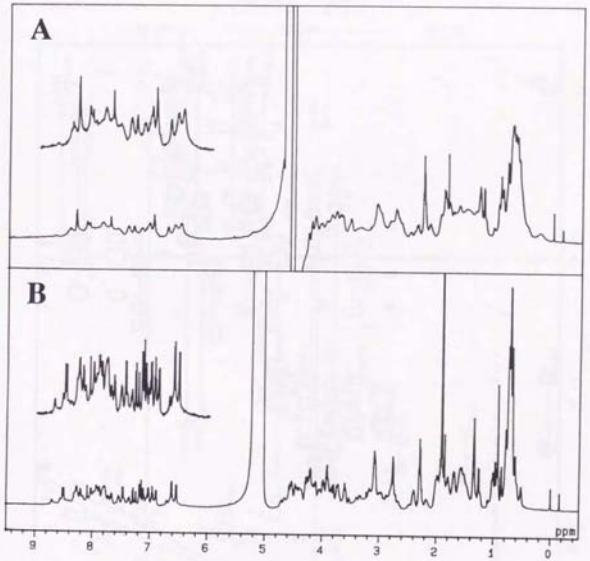


Figure 2-2. ^{1}H NMR spectra of bombyxin-II (2 mM) at 28°C in 90%/10% (v/v) $H_2O/^2H_2O$ (pH*2.0) (**A**) and in 70%/30% (v/v) $H_2O/C^2H_3CO_2{}^2H$ (pH*2.0) (**B**). pH* indicates direct pH-meter reading measured at 25°C.

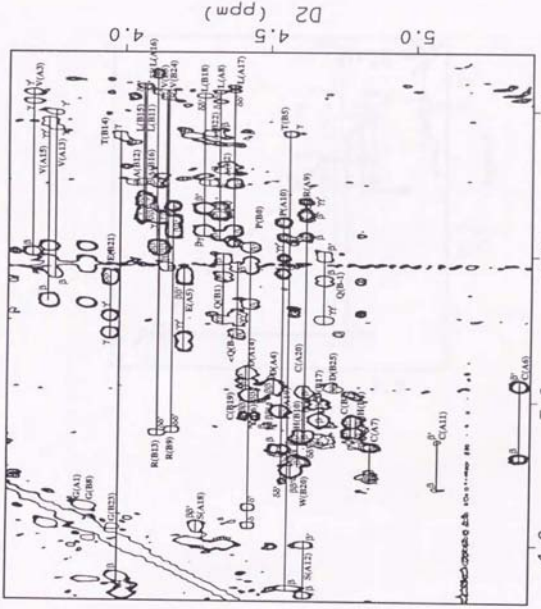


Figure 2-3. Identification of the amino acid ^1H spin systems in a TOCSY spectrum of bombyxin-II (3 mM, 70%/30% (v/v) $^2H_2O/C^2H_3CO_2{}^2H$, pH*2.0, 28°C; 600 MHz; τ_m = 45 ms; absorption mode; digital resolution 3.2 Hz/point). (A) Aliphatic side chains.

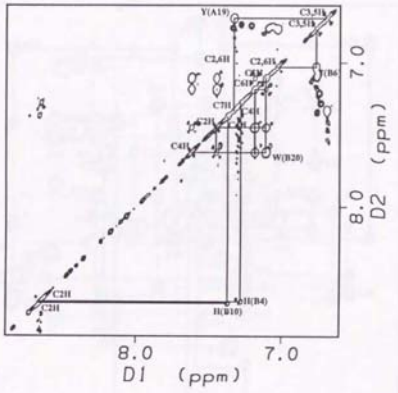


Figure 2-3. (Continued) Identification of the amino acid ^1H spin systems in a TOCSY spectrum of bombyxin-II (3 mM, 70%/30% (v/v) $^2H_2O/C^2H_3CO_2{}^2H$, pH*2.0, 28°C; 600 MHz; τ_m = 45 ms; absorption mode; digital resolution 3.2 Hz/point). (B) Aromatic side chains.

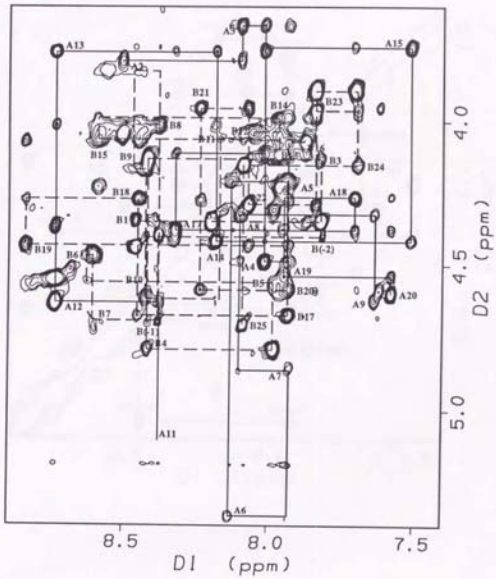


Figure 2-4. Sequential assignments via ^1H-^1H nuclear Overhauser effects in a NOESY spectrum of bombyxin-II (4 mM, 70%/30% (v/v) $H_2O/C^2H_3CO_2{}^2H$, pH*2.0, 28°C; 600 MHz; τ_m = 75 ms; absorption mode; digital resolution 3.2 Hz/point). (A) Pathway of sequential assignments via $d_{\alpha N}$.

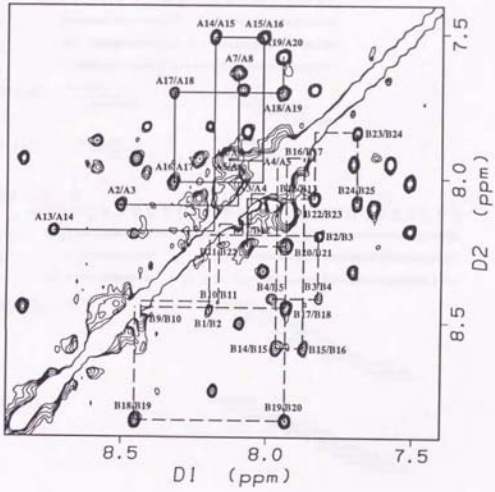


Figure 2-4. (Continued) Sequential assignments via ^1H-^1H nuclear Overhauser effects in a NOESY spectrum of bombyxin-II (4 mM, 70%/30% (v/v) $H_2O/C^2H_3CO_2{}^2H$, pH*2.0, 28°C; 600 MHz; τ_m = 75 ms; absorption mode; digital resolution 3.2 Hz/point). (B) Pathway of sequential assignments via d_{NN}.

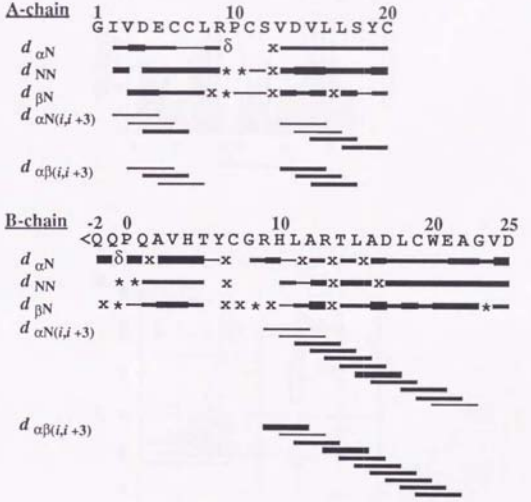


Figure 2-4. (Continued) (C) Sequential NOE connectivities. The height of the bars indicates the approximate intensity of the NOESY cross-peaks recorded with a mixing time of 75 ms. * indicates an undefined connectivity; x indicates an NOE connecitivity which is not clearly observed due to overlapping with other NOE peaks.

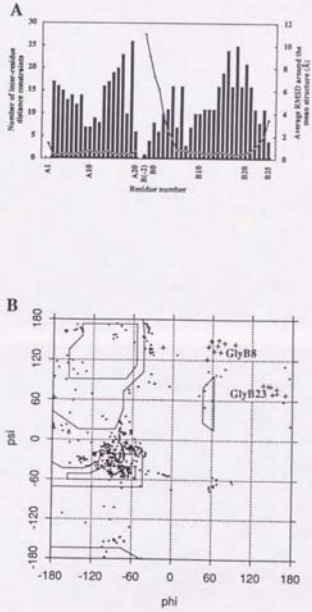


Figure 2-5. (A) Number of NOE constraints and RMSDs for each residue. The number of interresidue distance constraints (columns) and the average values of the main-chain (N, C^{α}, C') RMSDs (circles) were plotted as a function of residue number. (B) Ramachandran plot for the final 10 structures. Gly residues are plotted with +.

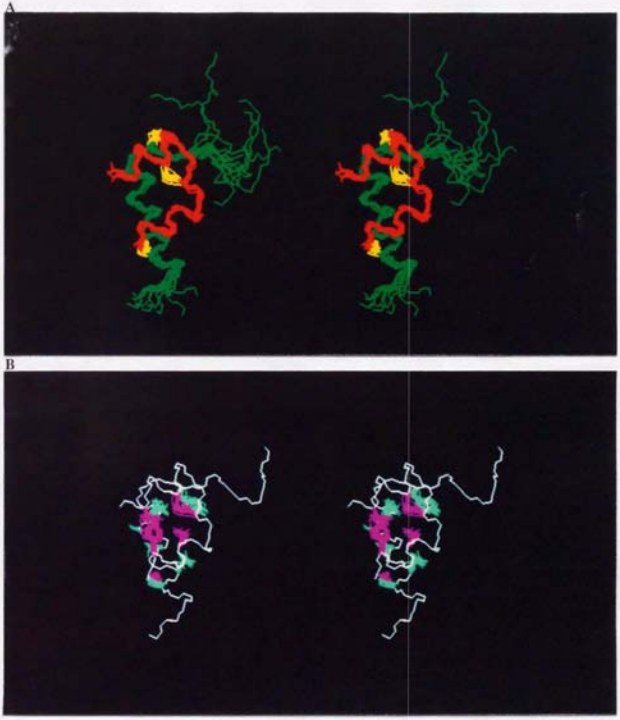


Figure 2-6. Solution structure of bombyxin-II (stereo view). (**A**) Main-chain atoms (N, C^{α}, C') and disulfide bonds of the 10 converged structures are superimposed. A-chain is shown in red, B-chain in green and disulfide bonds in yellow. (**B**) Hydrophobic core. Averaged main chains and overlaid side chains involved in the hydrophobic core are shown. Main chain is colored in white, hydrophobic side chains in the A- and B-chains are in purple and light green, respectively.

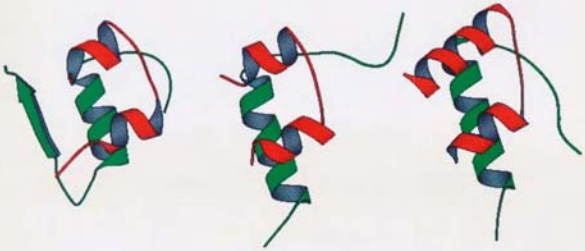


Figure 2-7. Comparison of the solution structure of bombyxin-II with the solution structure of human insulin (Brookhaven Protein Data Bank entry 1HIU; Hua *et al.*, 1991) and crystal structure ofhuman relaxin 2 (6RLX; Eigenbrot *et al.*, 1991). Schematic representation of the main-chainfold (figure created with Molscript) (Kraulis, 1991). A-chain is shown in red and B-chain in green.

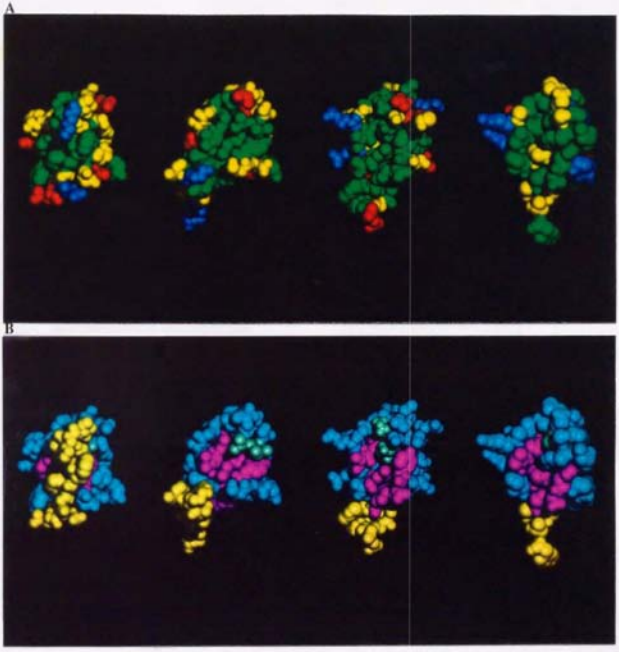
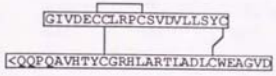
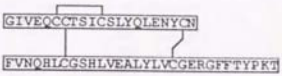


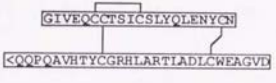
Figure 2-8. Comparison of the proposed receptor-recognition surface (viewed from the left side of Figures 2-6 and 2-7). Space-filling model. All non-hydrogen atoms are shown. (**A**) Distribution of side-chain functional groups. Acidic, basic, polar and hydrophobic residues are colored in red, blue, yellow and green, respectively. The Arg residues at B9 and B13 in bombyxin-II and human relaxin 2 are exposed on the B-chain helix. (**B**) Locked and unlocked states. The solution structure of human insulin (Brookhaven Protein Data Bank entry 1HIU; Hua et al., 1991) corresponds to the locked, inactive state, while the solution structure of [GlyB24]human insulin (1HIT; Hua et al., 1991) represents a model of the unlocked, active state (Hua et al., 1991). The solution structure of bombyxin-II and the crystal structure of human relaxin 2 (6RLX; Eigenbrot et al., 1991) corresponds to the unlocked state. The hydrophobic surface is exposed in the unlocked state, while it is covered by the B-chain C-terminal section in the locked state. B-chain C-terminal segment (from the residue at B20 to the C-terminus) is colored in yellow; hydrophobic residues at A2 and A3 in light green; other hydrophobic core residues in purple; other residues in light blue.



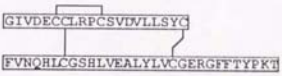
bombyxin-II



human insulin



imbyxin



bonsulin

Peptide	Insulin Activity			Bombyxin Activity
	Binding to Human Insulin Receptor (Kd, *M*)	2-Deoxyglucose Uptake (ED50, *M*)	Thymidine Incorporation (ED50, *M*)	Prothoracicotropic Activity to *Samia* (ng/*Samia* unit*)
Human Insulin	4×10^{-10}	6×10^{-12}	2×10^{-11}	> 5,000
Bonsulin	2×10^{-7}	3×10^{-10}	3×10^{-9}	> 5,000
Imbyxin	$> 1 \times 10^{-6}$	$> 1 \times 10^{-9}$	$> 1 \times 10^{-9}$	250
Bombyxin-II	$> 1 \times 10^{-6}$	$> 1 \times 10^{-9}$	$> 1 \times 10^{-9}$	0.5

* One *Samia* unit is defined as the minimal dose necessary to cause the adult formation in an assay pupa of *Samia cynthia ricini*.

Table

Figure 3-1. Primary structure and biological activity of bombyxin-II, human insulin and their hybrid molecules, bonsulin and imbyxin.

Table 3-2

^1H chemical shifts in ppm for bonsulin at pH 2.0, 28°C, dissolved in 70%/30% water/acetic acid

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	Others	
GlyA1		4.07, 4.04			
IleA2	8.53	3.75	1.20	$C^{\gamma}H$	1.14, 0.94
				$C^{\gamma}H_3$	0.78
				$C^{\delta}H_3$	0.68
ValA3	8.09	3.68	2.03	$C^{\gamma}H_3$	0.99, 0.92
AspA4	8.12	4.49	2.93, 2.93		
GluA5	7.82	4.19	2.20, 2.09	$C^{\gamma}H$	2.67, 2.56
CysA6	8.12	5.29	3.26, 2.99		
CysA7	7.98	4.85	3.72, 3.38		
LeuA8	8.04	4.25	1.93, 1.76	$C^{\gamma}H$	1.66
				$C^{\delta}H_3$	0.96, 0.89
ArgA9	7.50	4.63	1.88, 1.72	$C^{\gamma}H$	1.72, 1.66
				$C^{\delta}H$	3.26, 3.26
				$N^{\varepsilon}H$	7.19
ProA10		4.50	1.95, 1.95	$C^{\gamma}H$	2.04, 1.75
				$C^{\delta}H$	3.58, 3.52
CysA11	8.89	5.14	3.57, 3.47		
SerA12	8.07	4.64	4.32, 4.02		
ValA13	8.71	3.76	2.09	$C^{\gamma}H_3$	1.07, 0.98
AspA14	8.11	4.43	2.85, 2.80		
ValA15	7.54	3.76	2.31	$C^{\gamma}H_3$	1.12, 1.07
LeuA16	7.87	4.10	2.02, 1.59	$C^{\gamma}H$	1.83
				$C^{\delta}H_3$	0.89, 0.86
LeuA17	8.26	4.35	1.81, 1.49	$C^{\gamma}H$	1.79
				$C^{\delta}H_3$	0.87, 0.87
SerA18	7.65	4.25	3.89, 3.86		
TyrA19	7.97	4.57	3.38, 2.99	$C^{2,6}H$	7.32
				$C^{3,5}H$	6.69
CysA20	7.51	4.65	3.14, 2.78		
PheB1		4.40	3.28, 3.18	$C^{2,6}H$	7.29
				$C^{3,5}H$	7.40
				$C^{4}H$	7.35
ValB2	8.35	4.18	2.02	$C^{\gamma}H_3$	0.92, 0.92
AsnB3	8.42	4.72	2.84, 2.79	$N^{\delta}H$	7.62, 6.98
GlnB4	8.42	4.39	2.32, 1.93	$C^{\gamma}H$	2.11, 2.11
				$N^{\varepsilon}H$	7.54, 6.85
HisB5	8.60	4.63	3.39, 3.29	$C^{4}H$	7.35
				$C^{2}H$	8.66

Table 3-2
(Continued)

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	Others	
LeuB6	8.69	4.46	1.72, 1.72	$C^{\gamma}H$	1.51
				$C^{\delta}H_3$	0.89, 0.82
CysB7	8.22	4.85	3.22, 3.07		
GlyB8	9.02	4.03, 3.87			
SerB9	8.78	4.17	3.93, 3.93		
HisB10	8.08	4.57	3.57, 3.40	$C^{4}H$	7.48
				$C^{2}H$	8.72
LeuB11	7.44	4.08	1.87, 1.45	$C^{\gamma}H$	1.50
				$C^{\delta}H_3$	0.85, 0.85
ValB12	7.49	3.52	2.08	$C^{\gamma}H_3$	1.01, 0.92
GluB13	7.93	4.13	2.23, 2.14	$C^{\gamma}H$	2.59, 2.56
AlaB14	7.96	4.14	1.53		
LeuB15	8.24	4.07	1.70, 1.58	$C^{\gamma}H$	1.67
				$C^{\delta}H_3$	0.82, 0.82
TyrB16	8.17	4.21	3.16, 3.16	$C^{2,6}H$	7.10
				$C^{3,5}H$	6.77
LeuB17	8.02	4.15	1.95, 1.71	$C^{\gamma}H$	1.86
				$C^{\delta}H_3$	0.95, 0.95
ValB18	8.60	3.97	2.19	$C^{\gamma}H_3$	1.12, 0.99
CysB19	8.71	4.65	3.11, 3.00		
GlyB20	7.87	3.97, 3.95			
GluB21	7.99	4.31	2.18, 2.16	$C^{\gamma}H$	2.60, 2.52
ArgB22	7.80	4.35	2.00, 1.87	$C^{\gamma}H$	1.75, 1.75
				$C^{\delta}H$	3.25, 3.25
				$N^{\varepsilon}H$	7.31
GlyB23	8.05	3.97, 3.94			
PheB24	7.87	4.58	2.98, 2.89	$C^{2,6}H$	7.12
				$C^{3,5}H$	7.25
				$C^{4}H$	7.23
PheB25	7.95	4.63	3.10, 2.92	$C^{2,6}H$	7.18
				$C^{3,5}H$	7.30
				$C^{4}H$	7.24
TyrB26	7.95	4.62	2.96, 2.96	$C^{2,6}H$	7.06
				$C^{3,5}H$	6.80
ThrB27	7.85	4.56	4.10	$C^{\gamma}H_3$	1.22
ProB28		4.46	2.32, 1.97	$C^{\gamma}H$	2.06, 2.06
				$C^{\delta}H$	3.70, 3.70
LysB29	8.24	4.46	1.91, 1.81	$C^{\gamma}H$	1.53, 1.53
				$C^{\delta}H$	1.70, 1.70
				$C^{\varepsilon}H$	3.01, 3.01
				$N^{\varepsilon}H$	7.98, 7.98
ThrB30	8.02	4.50	4.45	$C^{\gamma}H_3$	1.21

Table 3-3

^1H chemical shifts in ppm for imbyxin at pH 2.0, 28°C, dissolved in 70%/30% water/acetic acid

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	Others	
GlyA1		3.95, 3.95			
IleA2	8.42	4.32	1.86	$C^{\gamma}H$	1.49, 1.17
				$C^{\gamma}H_3$	0.92
				$C^{\delta}H_3$	0.87
ValA3	8.19	4.29	2.05	$C^{\gamma}H_3$	0.95, 0.93
GluA4	8.29	4.49	2.11, 1.97	$C^{\gamma}H$	2.45, 2.45
GlnA5	8.37	4.35	2.08, 1.95	$C^{\gamma}H$	2.30, 2.30
				$N^{\varepsilon}H$	7.41, 6.86
CysA6	8.60	4.67	3.17, 3.11		
CysA7	8.89	5.08	3.44, 2.95		
ThrA8	7.54	4.63	4.71	$C^{\gamma}H_3$	1.32
SerA9	8.71	4.33	4.00, 4.00		
IleA10	7.72	3.94	1.80	$C^{\gamma}H$	1.46, 1.20
				$C^{\gamma}H_3$	0.83
				$C^{\delta}H_3$	0.83
CysA11	7.93	4.36	3.45, 2.99		
SerA12	8.50	4.36	4.05, 3.95		
LeuA13	7.97	4.18	1.69, 1.69	$C^{\gamma}H$	1.64
				$C^{\delta}H_3$	0.92, 0.85
TyrA14	8.03	4.40	3.13, 3.04	$C^{2,6}H$	7.05
				$C^{3,5}H$	6.79
GlnA15	8.06	4.19	2.19, 2.19	$C^{\gamma}H$	2.47, 2.47
				$N^{\varepsilon}H$	7.37, 6.87
LeuA16	8.15	4.24	1.80, 1.76	$C^{\gamma}H$	1.76
				$C^{\delta}H_3$	0.91, 0.87
GluA17	8.22	4.14	2.09, 2.09	$C^{\gamma}H$	2.58, 2.47
AsnA18	7.91	4.62	2.68, 2.64	$N^{\delta}H$	7.28, 6.73
TyrA19	8.03	4.57	3.17, 3.06	$C^{2,6}H$	7.12
				$C^{3,5}H$	6.81
CysA20	8.19	4.69	3.23, 2.93		
AsnA21	7.99	4.80	2.92, 2.87	$N^{\delta}H$	7.55, 6.88
pGluB(-2)	7.81	4.40	2.55, 2.08	$C^{\gamma}H$	2.43, 2.40
GlnB(-1)	8.36	4.69	2.15, 1.99	$C^{\gamma}H$	2.44, 2.44
				$N^{\varepsilon}H$	7.53, 6.86
ProB0		4.44	2.32, 1.93	$C^{\gamma}H$	2.11, 2.05
				$C^{\delta}H$	3.86, 3.73

Table 3-3
(Continued)

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	Others
GlnB1	8.40	4.35	2.13, 2.00	$C^{\gamma}H$ 2.42, 2.42
				$N^{\varepsilon}H$ 7.51, 6.87
AlaB2	8.19	4.35	1.37	
ValB3	7.86	4.11	2.02	$C^{\gamma}H_3$ 0.90, 0.87
HisB4	8.50	4.88	3.27, 3.19	$C^{4}H$ 7.29
				$C^{2}H$ 8.66
ThrB5	8.10	4.77	4.18	$C^{\gamma}H_3$ 1.15
TyrB6	8.30	4.69	3.02, 2.98	$C^{2,6}H$ 7.03
				$C^{3,5}H$ 6.75
CysB7	8.54	4.58	3.18, 3.12	
GlyB8	8.32	3.99, 3.96		
ArgB9	7.97	4.37	1.84, 1.76	$C^{\gamma}H$ 1.66, 1.66
				$C^{\delta}H$ 3.18, 3.18
				$N^{\varepsilon}H$ 7.20
HisB10	8.47	4.75	3.30, 3.18	$C^{4}H$ 7.32
				$C^{2}H$ 8.65
LeuB11	8.25	4.37	1.64, 1.64	$C^{\gamma}H$ 1.60
				$C^{\delta}H_3$ 0.89, 0.85
AlaB12	8.27	4.30	1.44	
ArgB13	8.18	4.35	1.92, 1.83	$C^{\gamma}H$ 1.70, 1.70
				$C^{\delta}H$ 3.22, 3.22
				$N^{\varepsilon}H$ 7.20
ThrB14	7.86	4.35	4.35	$C^{\gamma}H_3$ 1.23
LeuB15	8.15	4.20	1.64, 1.57	$C^{\gamma}H$ 1.64
				$C^{\delta}H_3$ 0.85, 0.79
AlaB16	8.06	4.19	1.42	
AspB17	7.98	4.66	3.18, 2.97	
LeuB18	7.98	4.28	1.79, 1.65	$C^{\gamma}H$ 1.71
				$C^{\delta}H_3$ 0.88, 0.85
CysB19	8.22	4.59	3.18, 2.97	
TrpB20	8.00	4.64	3.36, 3.36	$N^{1}H$ 10.00
				$C^{2}H$ 7.28
				$C^{4}H$ 7.55
				$C^{5}H$ 7.10
				$C^{6}H$ 7.18
				$C^{7}H$ 7.43
GluB21	8.07	4.14	1.96, 1.96	$C^{\gamma}H$ 2.32, 2.04
AlaB22	7.95	4.28	1.44	
GlyB23	8.00	3.97, 3.96		
ValB24	7.70	4.18	2.09	$C^{\gamma}H_3$ 0.93, 0.89
AspB25	8.20	4.74	2.92, 2.92	

Table 3-4. *Structural statistics for bonsulin*

	<SA>	$(SA)_r$
Root-mean-square deviations from experimental distance constraints (Å) (623)	0.048 ± 0.003	0.048
Number of distance constraint violations > 0.3 Å	1 - 5 (Maximum 0.41 Å)	2 (Maximum 0.40 Å)
Root-mean-square deviations from experimental dihedral constraints (deg) (34)	2.02 ± 0.28	0.71
Number of dihedral constraint violations > 5°	1 - 3 (Maximum 11.9°)	1 (Maximum 9.0°)
F_{NOE} (kcal/mol)*	71.6 ± 7.4	71.1
F_{repel} (kcal/mol)*	43.6 ± 5.6	37.7
E_{L-J} (kcal/mol)†	-204.2 ± 10.7	-194.3
RMSDs from idealized geometry		
Bonds (Å) (779)	0.006 ± 0.0000	0.006
Angles (degrees) (1405)	2.15 ± 0.008	2.12
Impropers (degrees) (344)‡	1.02 ± 0.01	0.99

<SA> are the 10 refined simulated annealing structures; $(SA)_r$ is the restrained minimized mean structure, where the mean structure was obtained by averaging the coordinates of the individual <SA> structures best-fitted to each other.

* The value of the square-well NOE potential, F_{NOE}, is calculated with a force constant of 50 kcal/mol per $Å^{-2}$. The value of F_{repel} is calculated with a force constant of 4 kcal/mol per $Å^{-4}$ with the van der Waals radii scaled by a factor of 0.8 of the standard values used in the CHARMm empirical function.

† E_{L-J} is the Lennard-Jones van der Waals energy calculated with the CHARMm empirical energy function, which was not included in the simulated annealing calculations.

‡ The improper torsion term is used to maintain the planar geometry and chirality.

Table 3-5. *Structure statistics for imbyxin*

	<SA>	$(SA)_r$
Root-mean-square deviations from		
experimental distance constraints (Å) (509)	0.068 ± 0.002	0.057
Number of distance constraint	2 - 7	2
violations > 0.3 Å	(Maximum 0.51 Å)	(Maximum 0.47 Å)
Root-mean-square deviations from		
experimental dihedral constraints (deg) (21)	2.66 ± 1.01	0.97
Number of dihedral constraint	0 - 3	2
violations > 5°	(Maximum 12.9°)	(Maximum 8.3°)
F_{NOE} (kcal/mol)*	117.7 ± 7.4	83.6
F_{repel} (kcal/mol)*	47.3 ± 4.8	40.0
E_{L-J} (kcal/mol)†	-143.9 ± 11.5	-122.5
RMSDs from idealized geometry		
Bonds (Å) (742)	0.007 ± 0.0005	0.007
Angles (degrees) (1328)	2.27 ± 0.02	2.22
Impropers (degrees) (327)‡	1.13 ± 0.04	1.07

<SA> are the 10 refined simulated annealing structures; $(SA)_r$ is the restrained minimized mean structure, where the mean structure was obtained by averaging the coordinates of the individual <SA> structures best-fitted to each other.

* The value of the square-well NOE potential, F_{NOE}, is calculated with a force constant of 50 kcal/mol per $Å^{-2}$. The value of F_{repel} is calculated with a force constant of 4 kcal/mol per $Å^{-4}$ with the van der Waals radii scaled by a factor of 0.8 of the standard values used in the CHARMm empirical function.

† E_{L-J} is the Lennard-Jones van der Waals energy calculated with the CHARMm empirical energy function, which was not included in the simulated annealing calculations.

‡ The improper torsion term is used to maintain the planar geometry and chirality.

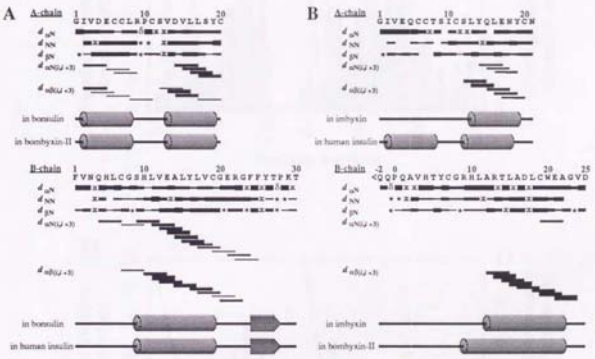


Figure 3-1. Sequential NOE connectivities and secondary structures in bonsulin (**A**) and imbyxin (**B**). The height of the bars indicates the approximate intensity of the NOESY cross-peaks recorded with a mixing time of 75 ms. * indicates an undefined connectivity; x indicates an NOE connecitivity which is not clearly observed due to overlapping with other NOE peaks. The secondary structures of bonsulin and imbyxin are compared with those of the corresponding peptide chains in bombyxin-II and human insulin. A cylinder represents an a-helix, an arrow represents a β-strand.

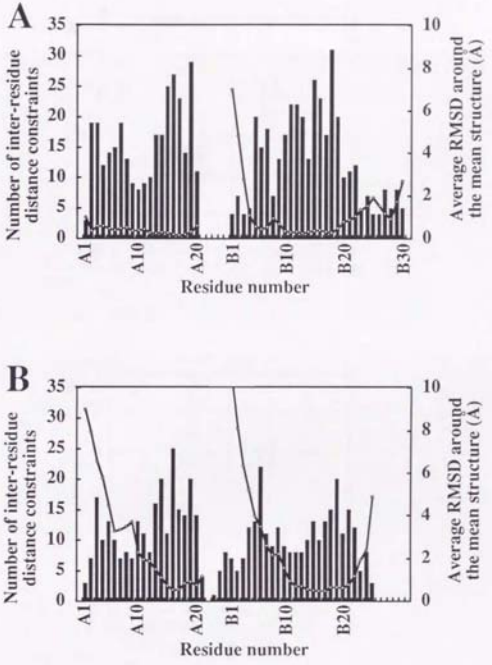


Figure 3-2. Number of NOE constraints and RMSDs for each residue in bonsulin (**A**) and imbyxin (**B**). The number of interresidue distance constraints (columns) and the average values of the main-chain (N, C^{α}, C') RMSDs (circles) are plotted as a function of residue number.

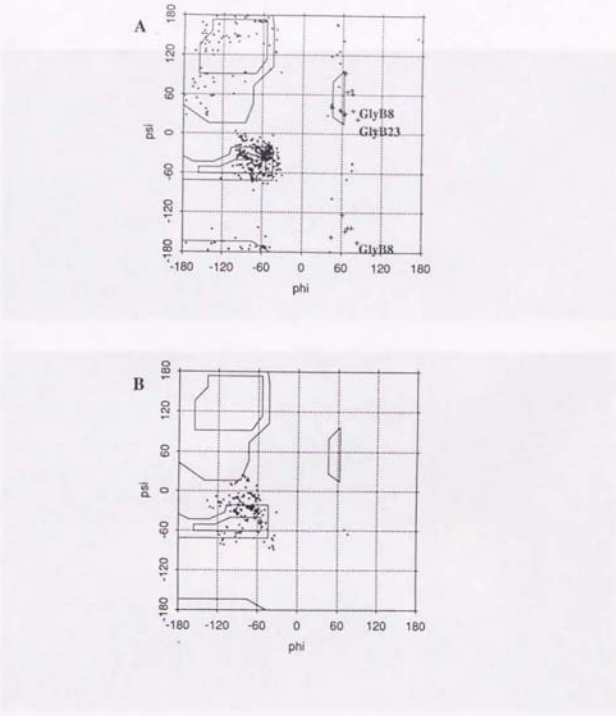


Figure 3-3. Ramachandran plot for the well-defined regions in the final 10 structures of bonsulin (IleA2 to TyrA19 and GlnB4 to ProB28) (**A**) and imbyxin (LeuA13 to TyrA19 and AlaB12 to AlaB22) (**B**). Gly residues are plotted with +.

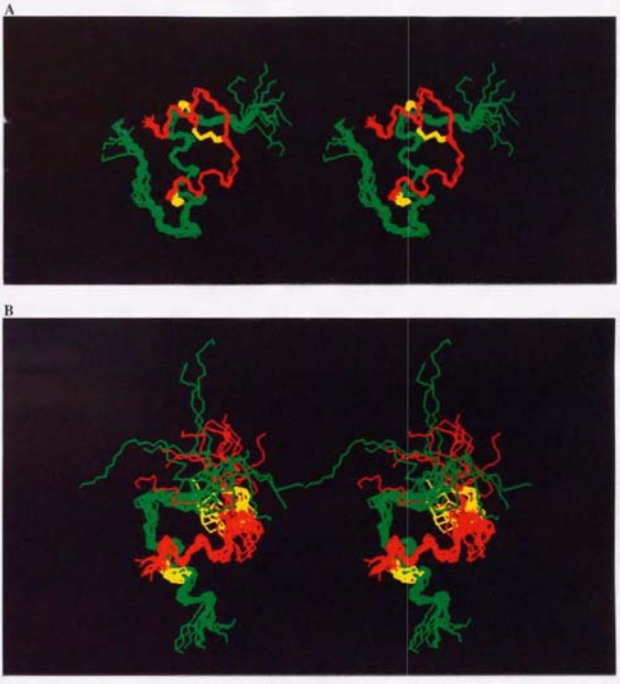


Figure 3-4. Solution structure of bonsulin (**A**) and imbyxin (**B**) (stereo view). Main-chain atoms (N, C^{α}, C') and disulfide bonds of the 10 converged structures are superimposed. A-chain is shown in red, B-chain in green and disulfide bonds in yellow.

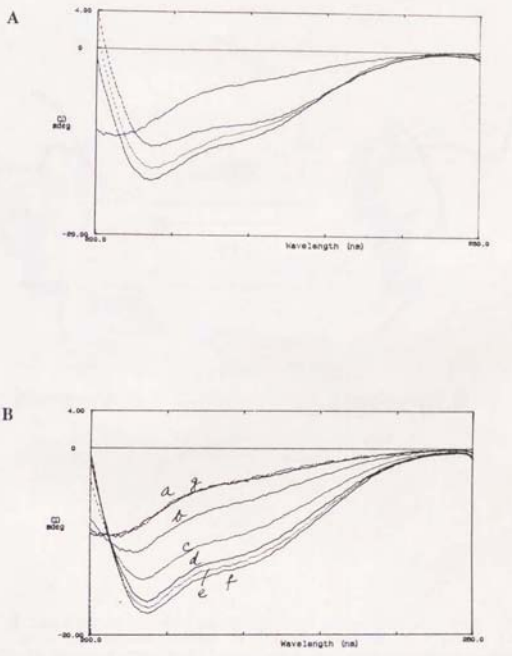


Figure 3-5. (A) CD spectrum of bombyxin-II (BBX), human insulin (ISL), bonsulin (BSL) and imbyxin (IBX) dissolved in 50 mM sodium phosphate buffer (pH 6.8). (B) TFE-dependent change of the CD spectrum of imbyxin. TFE concentration: a 0%; b 10%; c 20%; d 30%; e 40%; f 50%; g TFE removed (50% —> 0%).

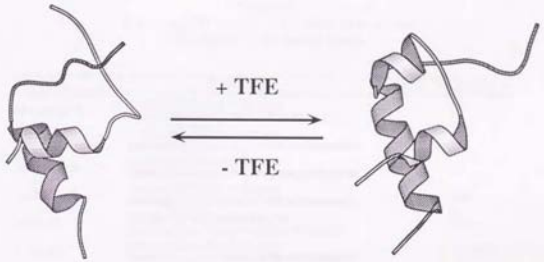
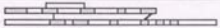
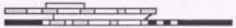
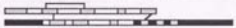
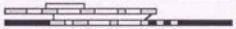
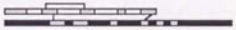


Figure 3-6. Proposed solvent-dependent equilibrium in the conformation of imbyxin. The conformer I is a representative structure of imbyxin dissolved in 70%/30% (v/v) $H_2O/C^2H_3CO_2{}^2H$ elucidated by the NMR method, and is considered to represent the major conformation of imbyxin in aqueous solution without TFE as indicated by CD. The conformer II is a modeled structure of imbyxin calculated using energy minimization techniques, assuming conformational homology with the A-chain of human insulin and the B-chain of bombyxin-II, the structures of which were determined by the NMR method (Hua *et al.*, 1991; Nagata *et al.*, submitted), and is proposed to be the major conformation of imbyxin in aqueous solution containing 30% or more of TFE based on the CD spectra of imbyxin shown in Figure 3-5.

Table 4-1
Bombyxin-like activity of the chimera molecules of bombyxin-II and human insulin

Peptide	Structure†	Relative potency‡
bombyxin-II		1
bonsylin-(6-18)		1
bonsylin-(6-17)		0.3
bonsylin-(7-18)		0.05
imbyxin		0.002
bonsulin		< 0.0001
human insulin		< 0.0001

† The white and black bars represent amino acid residues unique to bombyxin-II and human insulin, respectively. Gray bar represents amino acid residues common to bombyxin-II and human insulin.

‡ Relative potency of a peptide was derived as follows:
(ED_{50} of bombyxin-II [mol/pupa])/(ED_{50} of the peptide [mol/pupa]).

Table 4-2

^1H chemical shifts in ppm for bonsylin-(6-18) at pH 2.0, 28°C, dissolved in 70%/30% water/acetic acid

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	Others	
GlyA1		4.04, ?			
IleA2	8.51	3.90	1.34	$C^{\gamma}H$	1.18, 0.98
				$C^{\gamma}H_3$	0.80
				$C^{\delta}H_3$	0.69
ValA3	8.07	3.72	1.98	$C^{\gamma}H_3$	0.94, 0.88
AspA4	8.07	4.54	2.92, 2.89		
GluA5	8.05	4.19	2.15, 2.15	$C^{\gamma}H$	2.59, 2.59
CysA6	8.15	5.32	3.40, 2.86		
CysA7	7.94	4.82	3.47, 3.37		
LeuA8	8.10	4.35	1.86, 1.73	$C^{\gamma}H$	1.66
				$C^{\delta}H_3$	0.94, 0.88
ArgA9	7.63	4.70	1.87, 1.75	$C^{\gamma}H$	1.75, 1.67
				$C^{\delta}H$	3.26, 3.26
				$N^{\varepsilon}H$	7.19
ProA10		4.50	2.01, 1.77	$C^{\gamma}H$	2.06, 1.90
				$C^{\delta}H$	3.60, 3.60
CysA11	8.65	4.97	3.52, 3.20		
SerA12	8.35	4.57	4.33, 4.03		
ValA13	8.69	3.72	2.10	$C^{\gamma}H_3$	1.09, 0.98
AspA14	8.16	4.41	2.88, 2.81		
ValA15	7.51	3.74	2.29	$C^{\gamma}H_3$	1.09, 1.09
LeuA16	8.02	4.10	1.91, 1.43	$C^{\gamma}H$	1.71
				$C^{\delta}H_3$	0.76, 0.72
LeuA17	8.32	4.34	1.84, 1.50	$C^{\gamma}H$	1.92
				$C^{\delta}H_3$	1.07, 0.86
SerA18	7.70	4.26	3.90, 3.90		
TyrA19	7.95	4.54	3.33, 3.07	$C^{2,6}H$	7.32
				$C^{3,5}H$	6.70
CysA20	7.56	4.66	3.19, 2.85		
PheB1	4.40		3.28, 3.19	$C^{2,6}H$	7.30
				$C^{3,5}H$	7.41
				$C^{4}H$	7.37
ValB2	8.36	4.17	2.02	$C^{\gamma}H_3$	0.92, 0.92
AsnB3	8.42	4.73	2.84, 2.84	$N^{\delta}H$	7.66, 7.02
GlnB4	8.54	4.27	2.04, 1.89	$C^{\gamma}H$	2.32, 2.32
				$N^{\varepsilon}H$	7.41, 6.86
HisB5	8.47	4.72	3.26, 3.26	$C^{4}H$	7.30
				$C^{2}H$	8.66

Table 4-2
(Continued)

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	Others
TyrB6	8.48	4.44	3.01, 2.94	$C^{2,6}H$ 7.03
				$C^{3,5}H$ 6.80
CysB7	8.48	4.58	3.28, 3.18	
GlyB8	8.69	4.03, 3.82		
ArgB9	8.30	4.16	1.87, 1.87	$C^{\gamma}H$ 1.77, 1.69
				$C^{\delta}H$ 3.22, 3.22
				$N^{\varepsilon}H$ 7.27
HisB10	8.38	4.54	3.42, 3.26	$C^{4}H$ 7.37
				$C^{2}H$ 8.67
LeuB11	8.31	4.04	1.73, 1.66	$C^{\gamma}H$ 1.61
				$C^{\delta}H_3$ 0.81, 0.81
AlaB12	8.01	4.03	1.51	
ArgB13	7.95	4.09	1.96, 1.96	$C^{\gamma}H$ 1.76, 1.76
				$C^{\delta}H$ 3.24, 3.24
				$N^{\varepsilon}H$ 7.23
ThrB14	7.95	4.24	3.95	$C^{\gamma}H_3$ 1.18
LeuB15	8.58	4.06	1.69, 1.69	$C^{\gamma}H$ 1.55
				$C^{\delta}H_3$ 0.81, 0.81
AlaB16	7.95	4.06	1.51	
AspB17	7.88	4.57	3.10, 2.99	
LeuB18	8.45	4.25	1.78, 1.63	$C^{\gamma}H$ 1.82
				$C^{\delta}H_3$ 0.87, 0.87
CysB19	8.84	4.57	3.18, 2.99	
GlyB20	7.93	4.10, 4.02		
GluB21	8.01	4.40	2.24, 2.17	$C^{\gamma}H$ 2.61, 2.56
ArgB22	7.92	4.35	2.00, 1.88	$C^{\gamma}H$ 1.75, 1.75
				$C^{\delta}H$ 3.26, 3.26
				$N^{\varepsilon}H$ 7.30
GlyB23	8.10	3.95, 3.87		
PheB24	7.88	4.58	2.99, 2.90	$C^{2,6}H$ 7.14
				$C^{3,5}H$ 7.29
				$C^{4}H$ 7.21
PheB25	7.96	4.62	3.10, 2.93	$C^{2,6}H$ 7.20
				$C^{3,5}H$ 7.33
				$C^{4}H$ 7.27
TyrB26	7.95	4.63	2.97, 2.97	$C^{2,6}H$ 7.09
				$C^{3,5}H$ 6.81
ThrB27	7.88	4.58	4.10	$C^{\gamma}H_3$ 1.23
ProB28		4.41	2.32, 1.98	$C^{\gamma}H$ 2.06, 2.06
				$C^{\delta}H$ 3.70, 3.70
LysB29	8.26	4.45	1.92, 1.81	$C^{\gamma}H$ 1.52, 1.52
				$C^{\delta}H$ 1.73, 1.73
				$C^{\varepsilon}H$ 3.03, 3.03
				$N^{\zeta}H$?, ?
ThrB30	8.03	4.51	4.43	$C^{\gamma}H_3$ 1.21

Table 4-3. *Structural statistics for bonsylin-(6-18)*

	<SA>	$(SA)_r$
Root-mean-square deviations from experimental distance constraints (Å) (523)	0.055 ± 0.002	0.047
Number of distance constraint violations > 0.3 Å	2 - 5 (Maximum 0.42 Å)	1 (Maximum 0.32 Å)
Root-mean-square deviations from experimental dihedral constraints (deg) (32)	2.45 ± 0.22	1.18
Number of dihedral constraint violations > 5°	1 - 4 (Maximum 11.0°)	4 (Maximum 8.8°)
F_{NOE} (kcal/mol)*	78.4 ± 5.2	57.5
F_{repel} (kcal/mol)*	41.2 ± 3.1	30.9
E_{L-J} (kcal/mol)†	-188.0 ± 11.6	-195.0
RMSDs from idealized geometry		
Bonds (Å) (786)	0.006 ± 0.0000	0.005
Angles (degrees) (1411)	2.12 ± 0.009	2.08
Impropers (degrees) (357)‡	1.04 ± 0.02	0.99

<SA> are the 10 refined simulated annealing structures; $(SA)_r$ is the restrained minimized mean structure, where the mean structure was obtained by averaging the coordinates of the individual <SA> structures best-fitted to each other.

* The value of the square-well NOE potential, F_{NOE}, is calculated with a force constant of 50 kcal/mol per $Å^{-2}$. The value of F_{repel} is calculated with a force constant of 4 kcal/mol per $Å^{-4}$ with the van der Waals radii scaled by a factor of 0.8 of the standard values used in the CHARMm empirical function.

† E_{L-J} is the Lennard-Jones van der Waals energy calculated with the CHARMm empirical energy function, which was not included in the simulated annealing calculations.

‡ The improper torsion term is used to maintain the planar geometry and chirality.

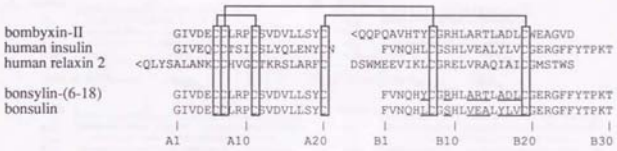


Figure 4-1. Amino acid sequences of bombyxin-II, human insulin and human relaxin 2 (upper) and those of the chimera molecules of bombyxin-II and human insulin (lower). Bonsulin is a chimera molecules consisting of the bombyxin-II A-chain and the human insulin B-chain. Bonsylin-(6-18), another chimera molecule, is different in amino-acid sequence from bonsulin in the B-chain middle part (B6 to B18) exclusively, where bonsylin-(6-18) possesses the bombyxin-II-type residues instead of the human insulin-type residues. The six residues underlined are the different residues between bonsylin-(6-18) and bonsulin. Numbering of residues is based on the insulin sequence. Cys residues are boxed.

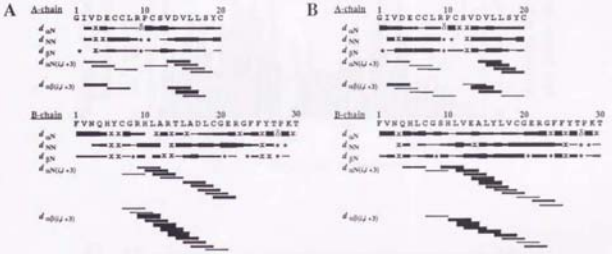


Figure 4-2. Sequential NOE connectivities in bonsylin-(6-18) (A) and bonsulin (B). The height of the bars indicates the approximate intensity of the NOESY cross-peaks recorded with a mixing time of 75 ms. * indicates an undefined connectivity; x indicates an NOE connecitivity which is not clearly observed due to overlapping with other NOE peaks.

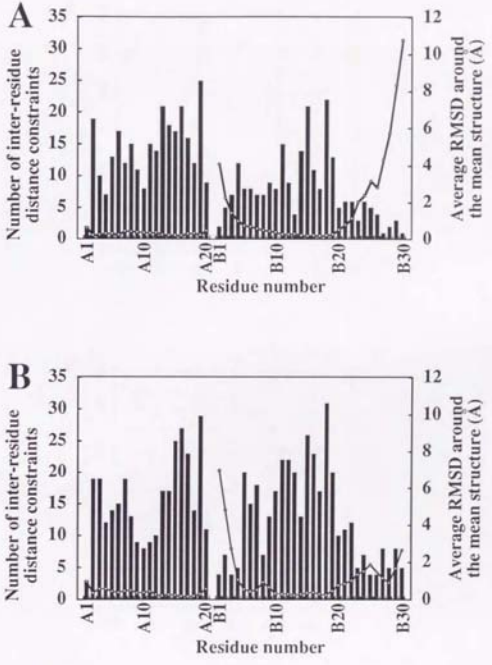


Figure 4-3. Number of inter-residue NOE constraints and RMSDs for each residue of bonsylin-(6-18) (**A**) and bonsulin (**B**). The number of interresidue distance constraints (columns) and the average values of the main-chain (N, C^{α}, C') RMSDs (circles) were plotted as a function of residue number.

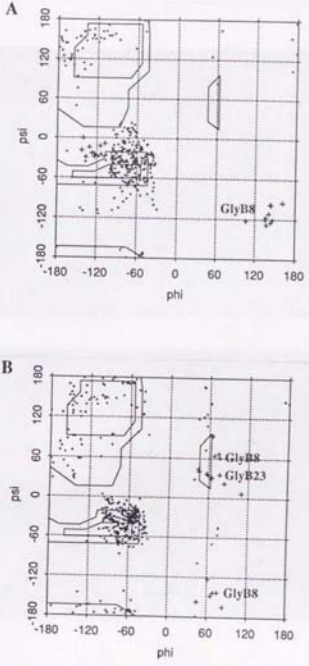


Figure 4-4. Ramachandran plot for the well-defined regions in the final 10 structures of bonsylin-(6-18) (IleA2 to TyrA19 and GlnB4 to ArgB22) (**A**) and bonsulin (IleA2 to TyrA19 and GlnB4 to ProB28) (**B**). Gly residues are plotted with +.

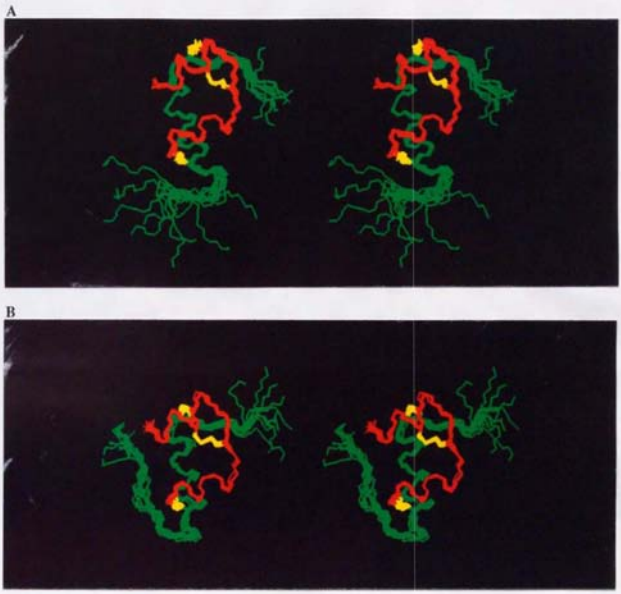


Figure 4-5. Solution structure of bonsylin-(6-18) (**A**) and bonsulin (**B**) (stereo view). Main-chain atoms (N, C^{α}, C') and disulfide bonds of the 10 converged structures are superimposed. A-chain is shown in red, B-chain in green and disulfide bonds in yellow.

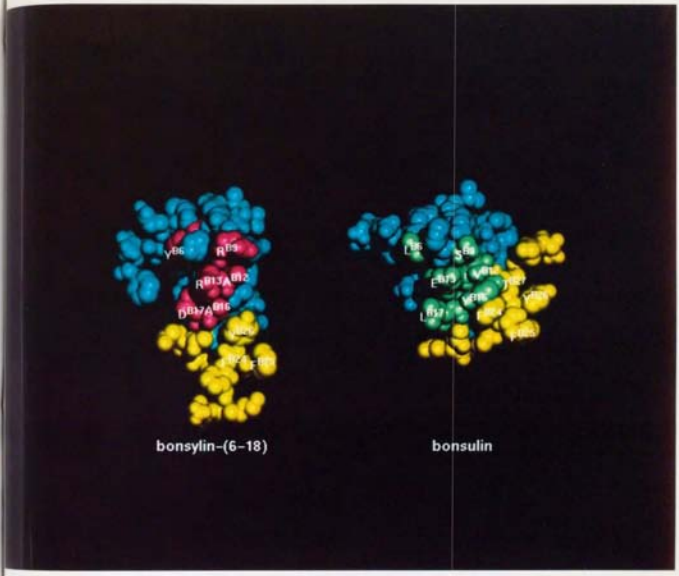


Figure 4-6. Comparison of the patches in the B-chain central parts of bonsylin-(6-18) and bonsulin (**A**) and, in addition, bombyxin-II (in solution; Nagata et al., submitted; Brookhaven PDB ident code 1BOM) and human insulin (in solution; Hua et al., 1991; 1HIU) (**B**). Space filling models of the mean structures are viewed from the left-back side of Figure 5. (**A**) The patches unique to bonsylin-(6-18) and bonsulin are colored in red and green, respectively. The conformationally different B-chain C-terminal parts are colored yellow, and the other parts blue. The Y/RRD/AA patch (red) is of critical importance in the recognition of the Samia bombyxin receptor. (**B**) Amino acid residues are colored by their characters: acidic, red; basic, blue; neutral, yellow; and hydrophobic, green. The Y/RRD/AA patch is found in bombyxin-IIand bonsylin-(6-18), which possess full potency in the Samia bombyxin-receptor recognition, whereasthe L/SEL/VY patch is found in human insulin and bonsulin, which possess no detectablepotency in the Samia bombyxin-receptor recognition.

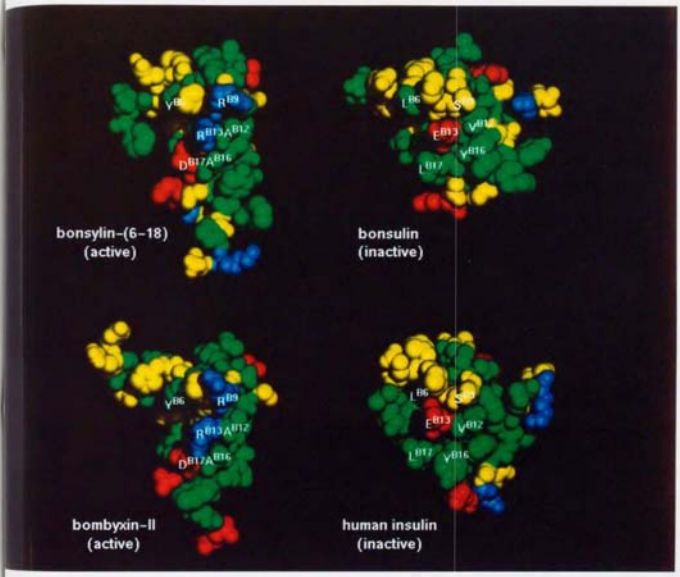


Figure 4-6. Comparison of the patches in the B-chain central parts of bonsylin-(6-18) and bonsulin (**A**) and, in addition, bombyxin-II (in solution; Nagata et al., submitted; Brookhaven PDB ident code 1BOM) and human insulin (in solution; Hua et al., 1991; 1HIU) (**B**). Space filling models of the mean structures are viewed from the left-back side of Figure 5. (**A**) The patches unique to bonsylin-(6-18) and bonsulin are colored in red and green, respectively. The conformationally different B-chain C-terminal parts are colored yellow, and the other parts blue. The Y/RRD/AA patch (red) is of critical importance in the recognition of the Samia bombyxin receptor. (**B**) Amino acid residues are colored by their characters: acidic, red; basic, blue; neutral, yellow; and hydrophobic, green. The Y/RRD/AA patch is found in bombyxin-IIand bonsylin-(6-18), which possess full potency in the Samia bombyxin-receptor recognition, whereasthe L/SEL/VY patch is found in human insulin and bonsulin, which possess no detectable potency in the Samia bombyxin-receptor recognition.

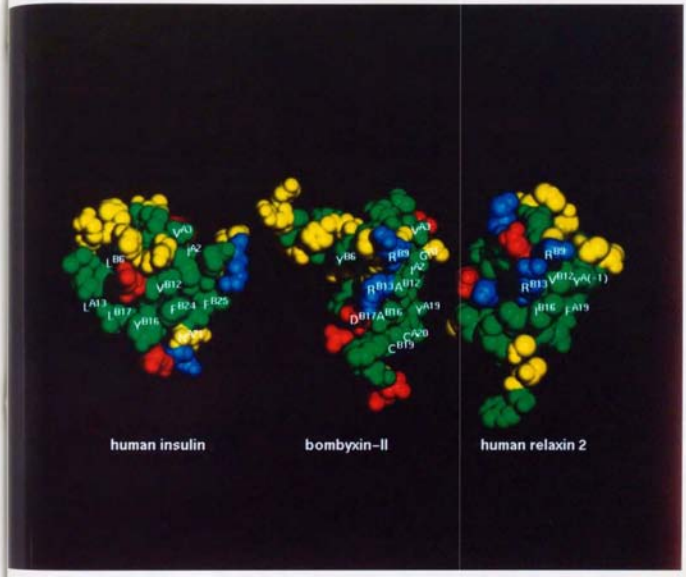


Figure 4-7. Comparison of the proposed receptor-recognition surfaces of bombyxin-II (in solution; Nagata et al., submitted; 1BOM), human insulin (in solution; Hua et al., 1991; 1HIU) and human relaxin 2 (in the crystalline state; Eigenbrot et al., 1991; 6RLX). Space-filling models of the mean structures are viewed from the same side as Figure 6. Amino acid residues are colored by their characters: acidic, red; basic, blue; neutral, yellow; and hydrophobic, green. Theresidueshown or proposed to be important for recognition of respective receptors arelabeled. Human insulin is proposed to have two receptor-recognition sites in this surface (Schäffer, 1994): the binding-site 1 involving GlyA1, AsnA21, ValB12, PheB24, PheB25 and the binding-site 2 involving LeuA13 and LeuB17. The binding-site 1 is partly covered by the B-chain C-terminal part.

Table 5-1

Properties and yields of the [Cys(Acm)^{B7},Cys^{B19}]bombyxin-II B-chain analogs

Peptide	RT on RP-HPLC[a] (min)	Average molecular mass[b] observed (u)/calculated (u)	Yield[c] (mg)
native	30.0	3194/3194	
ThrB 5Ala	30.3	3164/3164	8.7 (9.2%)
TyrB 6Ala	30.1	3101/3101	12.2 (13.1%)
ArgB 9Ala	32.7	3108/3108	13.3 (14.3%)
HisB10Ala	34.4	3128/3128	7.9 (8.4%)
LeuB11Ala	26.8	3151/3151	12.8 (13.5%)
AlaB12Val	31.2	3222/3222	7.9 (8.2%)
ArgB13Ala	32.1	3108/3108	14.8 (15.9%)
ThrB14Ala	30.8	3164/3164	26.4 (27.8%)
LeuB15Ala	24.7	3151/3151	11.0 (11.6%)
AlaB16Tyr	30.6	3286/3286	11.1 (11.3%)
AspB17Ala	31.2	3150/3150	16.9 (17.9%)
LeuB18Ala	25.9	3151/3151	9.4 (9.4%)

[a] RP-HPLC was run on a JASCO LC-900 HPLC system under the following conditions: column, SenshuPak Pegasil ODS (4.6 x 150 mm); eluent, 10–50% CH_3CN (40 min, linear gradient) in 0.1% TFA; flow rate, 1.0 ml/min; column temperature, 40°C; detection, absorbance at 280 nm.

[b] FAB-MS was measured on a JEOL JMS-SX102 mass spectrometer. Both the observed and calculated values of average molecular mass are those of the molecular ion, $(m + H)^+$, where m is the average molecular mass of the peptide.

[c] Both the total coupling yield and the recovery during purification were taken into account.

Table 5-2

Properties and yields of the [Cys(Acm)^{A7,B7},A6-A11,A20-B19-cystine]bombyxin-II analogs

Peptide	RT on RP-HPLC[a] (min)	Average molecular mass[b] observed (u)/calculated (u)	Yield[c] (μg)	
native				
ThrB 5Ala	31.1	5418/5418	85	(47%)
TyrB 6Ala	31.0	5356/5356	83	(46%)
ArgB 9Ala	32.3	5363/5363	71	(39%)
HisB10Ala	33.6	5382/5382	57	(32%)
LeuB11Ala	29.7	5406/5406	78	(43%)
AlaB12Val	31.4	5476/5476	81	(45%)
ArgB13Ala	32.1	5364/5363	79	(44%)
ThrB14Ala	31.7	5419/5418	78	(43%)
LeuB15Ala	28.9	5406/5406	77	(43%)
AlaB16Tyr	31.1	5540/5540	82	(46%)
AspB17Ala	31.6	5405/5404	72	(40%)
LeuB18Ala	29.2	5406/5406	77	(43%)

[a] RP-HPLC was run on a JASCO LC-900 HPLC system under the following conditions: column, SenshuPak Pegasil ODS (4.6 x 150 mm); eluent, 10–50% CH_3CN (40 min, linear gradient) in 0.1% TFA; flow rate, 1.0 ml/min; column temperature, 40°C; detection, absorbance at 280 nm.

[b] FAB-MS was measured on a JEOL JMS-SX102 mass spectrometer. Both the observed and calculated values of average molecular mass are those of the molecular ion, $(m + H)^+$, where m is the average molecular mass of the peptide.

[c] Both the yield of the chain combination reaction and the recovery during purification were taken into account.

Table 5-3
Properties and yields of the bombyxin-II analogs

Peptide	RT on RP-HPLC[a] (min)	Average molecular mass[b] observed (u)/calculated (u)	Yield[c] (μg)	
native				
ThrB 5Ala	29.0	5275/5274	43	(58%)
TyrB 6Ala	28.5	5212/5212	46	(63%)
ArgB 9Ala	29.6	5220/5219	34	(55%)
HisB10Ala	30.2	5239/5238	28	(56%)
LeuB11Ala	28.5	5262/5262	45	(66%)
AlaB12Val	29.7	5332/5332	41	(58%)
ArgB13Ala	30.0	5220/5219	38	(55%)
ThrB14Ala	29.3	5274/5274	43	(63%)
LeuB15Ala	28.4	5262/5262	41	(61%)
AlaB16Tyr	29.0	5396/5396	39	(54%)
AspB17Ala	28.9	5260/5260	35	(55%)
LeuB18Ala	28.5	5262/5262	38	(56%)

[a] RP-HPLC was run on a JASCO LC-900 HPLC system under the following conditions: column, SenshuPak Pegasil ODS (4.6 x 150 mm); eluent, 10–50% CH_3CN (40 min, linear gradient) in 0.1% TFA; flow rate, 1.0 ml/min; column temperature, 40°C; detection, absorbance at 280 nm.

[b] FAB-MS was measured on a JEOL JMS-SX102 mass spectrometer. Both the observed and calculated values of average molecular mass are those of the molecular ion, $(m + H)^+$, where m is the average molecular mass of the peptide.

[c] Both the yield of the iodine oxidation reaction and the recovery during purification were taken into account.

Table 5-4

Bombyxin-like prothoracicotropic activity to Samia cynthia ricini of bombyxin-II and bombyxin-II analogs

Peptide	Amino acid sequence of the B-chain	*Samia* unit† (ng/pupa)	Relative potency‡
native	<QQPQAVHTYCGRHLARTLADLCWEAGVD	0.125	1
ThrB 5Ala	<QQPQAVHAYCGRHLARTLADLCWEAGVD	0.125	1
TyrB 6Ala	<QQPQAVHTACGRHLARTLADLCWEAGVD	5	0.025
ArgB 9Ala	<QQPQAVHTYCGAHLARTLADLCWEAGVD	0.125	1
HisB10Ala	<QQPQAVHTYCGRALARTLADLCWEAGVD	0.125	1
LeuB11Ala	<QQPQAVHTYCGRHARTLADLCWEAGVD	0.5	0.25
AlaB12Val	<QQPQAVHTYCGRHLVRTLADLCWEAGVD	2.5	0.05
ArgB13Ala	<QQPQAVHTYCGRHLAATLADLCWEAGVD	0.125	1
ThrB14Ala	<QQPQAVHTYCGRHLARALADLCWEAGVD	0.25	0.5
LeuB15Ala	<QQPQAVHTYCGRHLARTAADLCWEAGVD	1	0.125
AlaB16Tyr	<QQPQAVHTYCGRHLARTLYDLCWEAGVD	0.5	0.25
AspB17Ala	<QQPQAVHTYCGRHLARTLAALCWEAGVD	0.5	0.25
LeuB18Ala	<QQPQAVHTYCGRHLARTLADACWEAGVD	0.25	0.5

† One *Samia* unit is defined as the minimal effective dose to induce adult development in half or more numbers of assay pupae of *Samia cynthia ricini*. $n \geq 8$.

‡ Relative potency of a peptide is defined as follows:
(ED_{50} of bombyxin-II [mol/pupa])/(ED_{50} of the peptide [mol/pupa]).

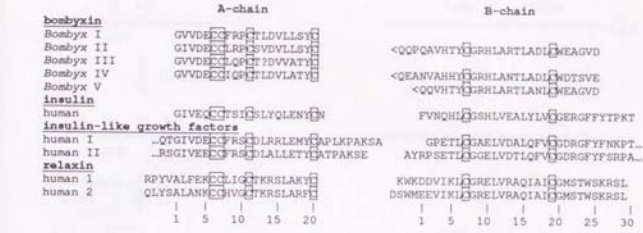


Figure 5-1. Amino acid sequences of bombyxin, insulin, insulin-like growth factors and relaxin. Numbering of residues is based on the insulin sequence. Cys residues are boxed. Dashes, gaps in the sequence inserted for best alignment.

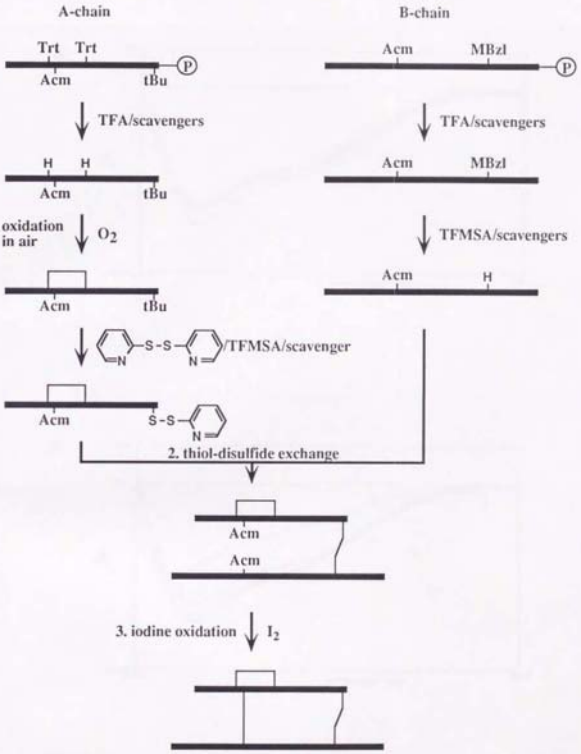


Figure 5-2. Stratery for the synthesis of bombyxin-II and its analogs using a regioselective disulfide bond formation (Maruyama *et al.*, 1992; Nagata *et al.*, 1992b).

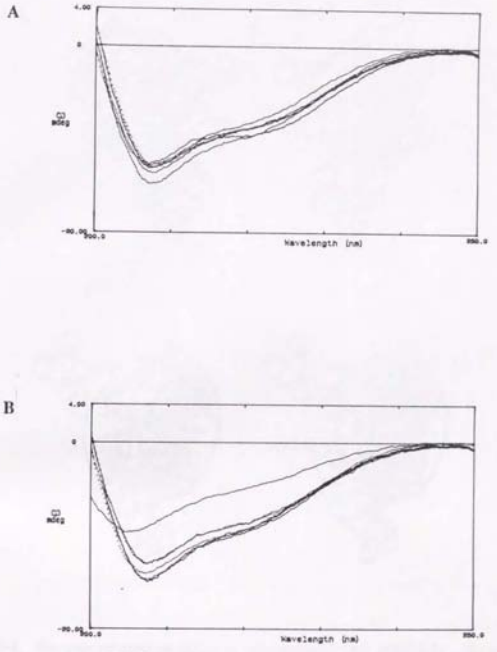


Figure 5-3. CD spectra of the Ala-scanning analogs and bombyxin-II. (A) Superimposition of the CD spectra of TyrB8Ala, ArgB9Ala, HisB10Ala, LeuB11ALa, AlaB12Val and ArgB13Ala analogs. (B) Superimposition of the CD spectra of ThrB14Ala, LeuB15Ala, AlaB16Tyr, AspB17Ala, LeuB18Ala analogs and bombyxin-II. The CD spectrum of the LeuB15Ala analog showed a peculiar curve in the CD, indicating a marked difference in main-chain conformation.

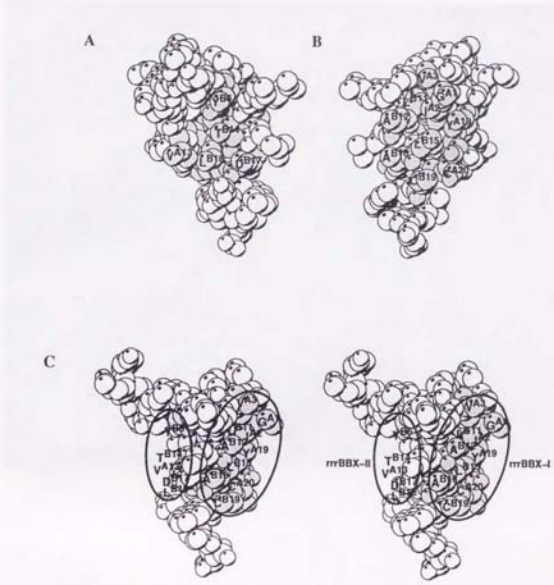


Figure 5-4. Receptor-recognition region of bombyxin-II. (A) rrrBBX-II. (B) rrrBBX-I. The rrrBBX-I contains important residues, GlyA1, ValA3, CysA20-CysB19, LeuB11, AlaB12, LeuB15 and AlaB16, and highly conserved hydrophobic residues, IleA2 and TyrA19. The rrrBBX-II contains important residues, TyrB6, ThrB14, AspB17 and LeuB18, and a highly conserved hydrophobic residue, ValA13. (C) The rrrBBX-I and the rrrBBX-II are separated by the two Arg residues at B9 and B13, which are not important for the recognition of the *Samia* bombyxin receptor (stereo pair). The amino-acid residues involved in the rrrBBX-I or the rrrBBX-II are shown in gray, other residues are shown in white.

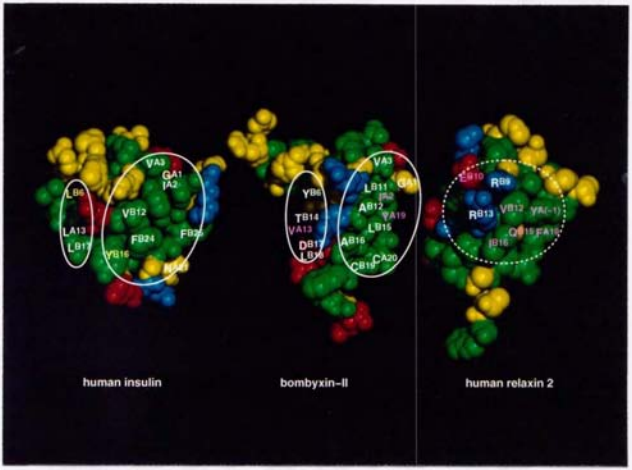


Figure 5-5. Comparison of the receptor-recognition regions of bombyxin-II, human insulin and human relaxin 2. Important residues for activity were labeled in white. In the case of insulin, important residues are labeled according to Schäffer (1994). Other important residues are labeled in yellow. In the cases of bombyxin-II and human relaxin 2, residues proposed to be important are labeled in pink. The receptor-recognition sites of bombyxin-II and human insulin consist of two regions, which are separated by the amino acid residues at B9 and B13. Whereas in relaxin, the two Arg residues at B9 and B13 are essential to its activity. Thus, the receptor-recognition region of relaxin may be different from those of bombyxin and insulin.

Table 6-1

Relative potencies in bombyxin-like prothoracicotropic activity to Samia cynthia ricini of bombyxin-II and its analogs

Peptide	Structure	Relative potency
bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGRHLARTLADLCWEAGVD	1
[LeuA3]-bombyxin-II	GILDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGRHLARTLADLCWEAGVD	0.1
[AlaB5]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHAYCGRHLARTLADLCWEAGVD	1
[AlaB6]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTACGRHLARTLADLCWEAGVD	0.03
[AlaB9]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGAHLARTLADLCWEAGVD	1
[AlaB10]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGRALARTLADLCWEAGVD	1
[AlaB11]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGRHAARTLADLCWEAGVD	0.3
[ValB12]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGRHLVRTLADLCWEAGVD	0.05
[AlaB13]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGRHLAATLADLCWEAGVD	1
[AlaB14]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGRHLARALADLCWEAGVD	0.5
[AlaB15]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGRHLARTAADLCWEAGVD	0.1
[TyrB16]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGRHLARTLYDLCWEAGVD	0.3
[AlaB17]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGRHLARTIAALCWEAGVD	0.3
[AlaB18]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGRHLARTIADACWEAGVD	0.5
[CitB9,B13]-bombyxin-II	GIVDEC<u>CLRPC</u>SVDVLLSYC <QQPQAVHTYCGBHLABTIADACWEAGVD	1

Table 6-1
(Continued)
Relative potencies in bombyxin-like prothoracicotropic activity to Samia cynthia ricini of bombyxin-II, human insulin and their chimera molecules

Peptide	Structure	Relative potency
bombyxin-II	GIVDECCLRPCSVDVLLSYC <QQPQAVHTYCGRHLARTLADLCWEAGVD	1
human insulin	GIVEQCCTSICSLYQLENYCN FVNQHLCGSHLVEALYLVCGERGFFYTPKT	< 0.0001
imbyxin	GIVEQCCTSICSLYQLENYCN <QQPQAVHTACGRHLARTLADLCWEAGVD	0.002
bonsulin	GIVDECCLRPCSVDVLLSYC FVNQHLCGSHLVEALYLVCGERGFFYTPKT	< 0.0001
imbylin	GIVEQCCTSICSLYQLENYCN <QQPQAVHTYCGRHLARTLADLCGERGFFYTPKT	0.002
bombylin	GIVDECCLRPCSVDVLLSYC <QQPQAVHTYCGRHLARTLADLCGERGFFYTPKT	1
bonsylin-(6-18)	GIVDECCLRPCSVDVLLSYC FVNQHYCGRHLARTLADLCGERGFFYTPKT	1
bonsylin-(7-18)	GIVDECCLRPCSVDVLLSYC FVNQHLCGRHLARTLADLCGERGFFYTPKT	0.05
bonsylin-(6-17)	GIVDECCLRPCSVDVLLSYC FVNQHYCGRHLARTLADVCGERGFFYTPKT	0.3
bonsylin-(6-16)	GIVDECCLRPCSVDVLLSYC FVNQHYCGRHLARTLALVCGERGFFYTPKT	0.01
bonsylin-(6-15)	GIVDECCLRPCSVDVLLSYC FVNQHYCGRHLARTLYLVCGERGFFYTPKT	0.005

Table 6-1
(Continued)
Relative potencies in bombyxin-like prothoracicotropic activity to Samia cynthia ricini of the truncated analogs and disulfide bond isomers of bombyxin-II

Peptide	Structure	Relative potency
bombyxin-II	GIVDECCLRPCSVDVLLSYC <QQPQAVHTYCGRHLARTLADLCWEAGVD	1
des-A1-bombyxin-II	IVDECCLRPCSVDVLLSYC <QQPQAVHTYCGRHLARTLADLCWEAGVD	0.1
des-(B(-2)-B1)-bombyxin-II†	GIVDECCLRPCSVDVLLSYC AVHTYCGRHLARTLADLCWEAGVD	1
des-(B22-B25)-bombyxin-II‡	GIVDECCLRPCSVDVLLSYC <QQPQAVHTYCGAHLARTLADLCWE	1
[A7-A11,A6-B7-cystine]-bombyxin-II	GIVDECCLRPCSVDVLLSYC <QQPQAVHTYCGRALARTLADLCWEAGVD	0.5
[A6-A7,A11-B7-cystine]-bombyxin-II	GIVDECCLRPCSVDVLLSYC <QQPQAVHTYCGRHAARTLADLCWEAGVD	0.5
[A11-A20,A6-B19-cystine]-bombyxin-II	GIVDECCLRPCSVDVLLSYC <QQPQAVHTYCGRHLAATLADLCWEAGVD	0.02
[A6-A20,A11-B19-cystine]-bombyxin-II	GIVDECCLRPCSVDVLLSYC <QQPQAVHTYCGRHLARALADLCWEAGVD	0.05

† Minoru Tanaka, personal communication.
‡ Maruyama, 1991.

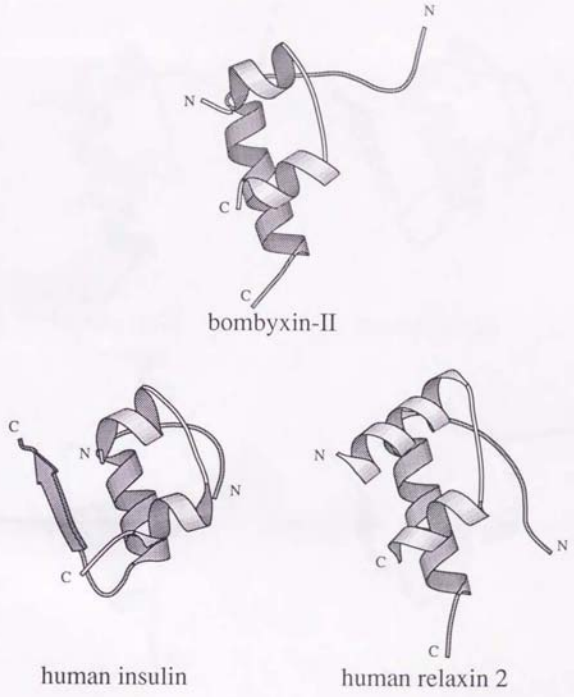


Figure 6-1. Three-dimensional structures of bombyxin-II, human insulin determined by the NMR method and human relaxin 2 determined by the X-ray crystallography. The structures of human insulin and human relaxin 2 shown were determined by Hua *et al.* (1991) and Eigenbrot *et al.* (1991), respectively. The A-chains are shown in light gray, and the B-chains in dark gray.

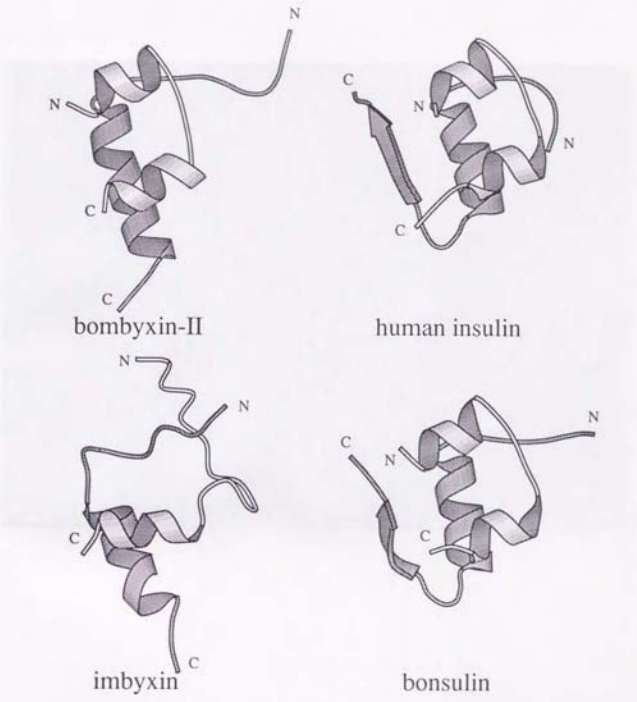


Figure 6-2. Three-dimensional structures of bombyxin-II, human insulin and their hybrid molecules, bonsulin (bombyxin-II A-chain + human insulin B-chain) and imbyxin (human insulin A-chain + bombyxin-II B-chain) determined by the NMR method. The structure of human insulin shown was determined by Hua *et al.* (1991). The A-chains are shown in light gray, and the B-chains in dark gray.

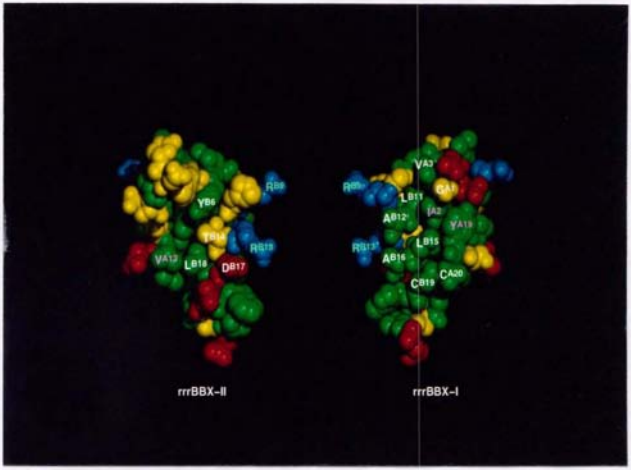


Figure 6-3. Receptor-recognition regions of bombyxin-II. Residues shown to be important for activity were labeled in white. Residues proposed to be important are labeled in pink. The rrrBBX-I and the rrrBBX-II are separated by two Arg residues at B9 and B13, which are not important for bombyxin actvity.

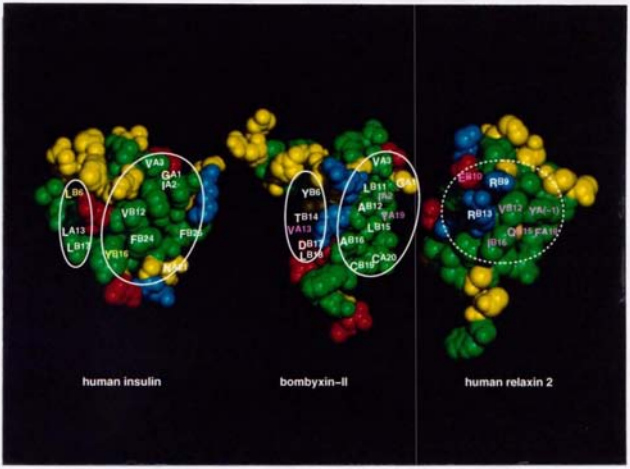


Figure 6-4. Comparison of the receptor-recognition regions of bombyxin-II, human insulin and human relaxin 2. Important residues for activity were labeled in white. In the case of insulin, important residues are labeled according to Schäffer (1994). Other important residues are labeled in yellow. In the cases of bombyxin-II and human relaxin 2, residues proposed to be important are labeled in pink. The receptor-recognition sites of bombyxin-II and human insulin consist of two regions, which are separated by the amino acid residues at B9 and B13. Whereas in relaxin, the two Arg residues at B9 and B13 are essential to its activity. Thus, the receptor-recognition region of relaxin may be different from those of bombyxin and insulin.

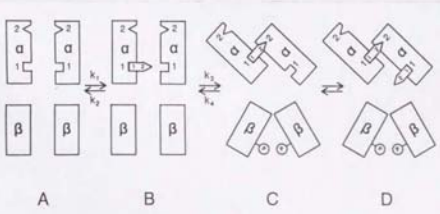
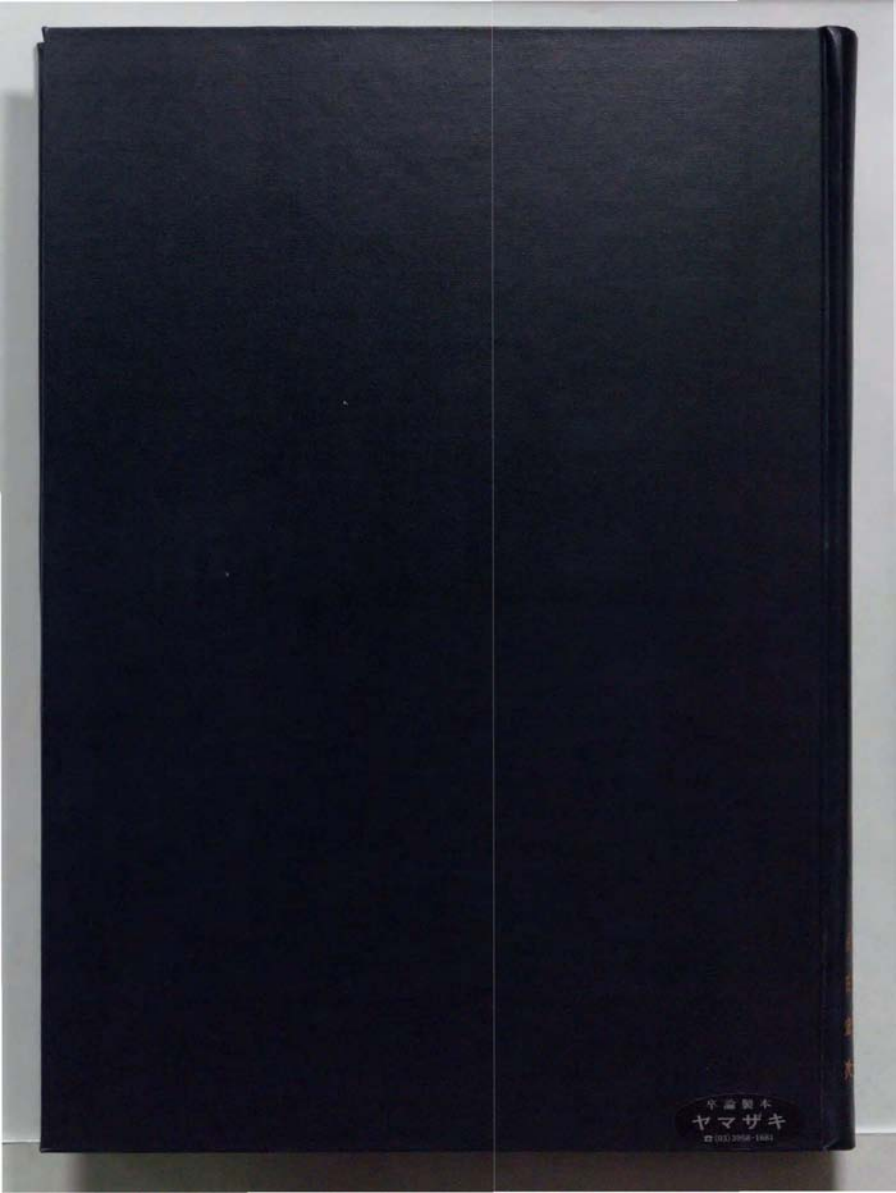
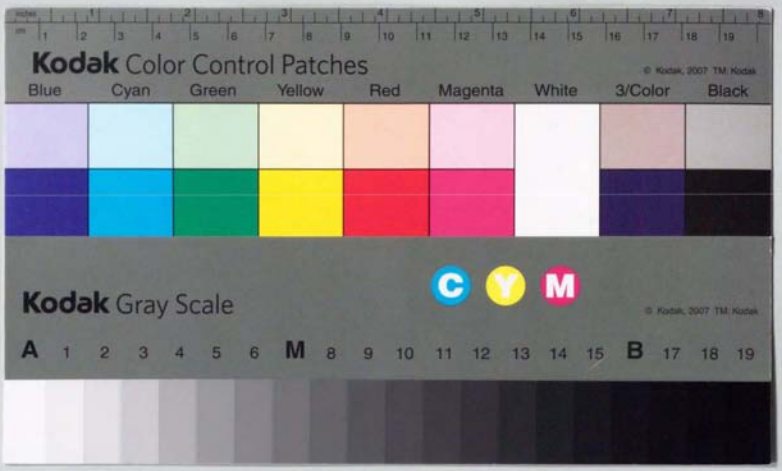


Figure 6-5. Model for insulin binding to its receptor proposed by Schäffer (1994). **(A)** Empty receptor, **(B)** initial binding, **(C)** high-affinity binding — active complex, **(D)** low-affinity binding of second insulin molecule. The phosphorylation of the activated receptor is indicated in **(C)** and **(D)**. The initially formed complex is able to proceed to form a high affinity complex where the insulin molecule bridges the two a subunits (stage **C**). In this state, the binding energy arises from both binding site 1 and binding site 2 interactions, and the contribution from each of the two sites may be different from the binding energy pertaining to one-site binding. The bridging step may be the biologically important one, giving rise to a conformational change which is transmitted through the cell membrane, activating the tyrosine kinase. According to the model, the binding of insulin to the high-affinity site leaves a free binding site 1 on one a subunit and a free binding site 2 on the other (Schäffer, 1994).


製本
ヤマザキ


Kodak Color Control Patches
© Kodak, 2007 TM: Kodak
Blue
Cyan
Green
Yellow
Red
Magenta
White
3/Color
Black
Kodak Gray Scale
C
Y
M
© Kodak, 2007 TM: Kodak
A 1 2 3 4 5 6 M 8 9 10 11 12 13 14 15 B 17 18 19