Vol. 140, No. 1, 1986 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 15, 1986 Pages 357-364

LEUCOSULFAKININ-II, A BLOCKED SULFATED INSECT NEUROPEPTIDE WITH HOMOLOGY TO CHOLECYSTOKININ AND GASTRIN

Ronald J. Nachman*†§, G. Mark Holman¶, Benjamin J. Cook¶, William F. Haddon§ and Nicholas Ling≠

§Western Regional Research Center, Agricultural Research Service, U. S. Dept. of Agriculture, 800 Buchanan St., Berkeley, CA 94710

¶ Veterinary Toxicology and Entomology Research Laboratory, Agricultural Research Service, U. S. Dept. of Agriculture, F and B Road, College Station, TX 77841

≠ Laboratories for Neuroendocrinology, The Salk Institute, La Jolla, CA 92037

Received August 29, 1986

A sulfated neuropeptide [pGlu-Ser-Asp-Asp-Tyr(SO_3H)-Gly-His-Met-Arg-Phe-NH_2], with a blocked N-terminus and related to the undecapeptide leucosulfakinin, has been isolated from head extracts of the cockroach, *Leucophaea maderae*. It exhibits sequence homology with the hormonally-active portion of vertebrate hormones cholecystokinin, human gastrin II and caerulin. This peptide, termed leucosulfakinin-II, shares a common C-terminal heptapeptide fragment with leucosulfakinin and a comparison of the two sequences provides an assessment of the importance of the constituent amino acids to biological activity. Leucosulfakinin-II shows a greater resemblance to cholecystokinin than does leucosulfakinin. Leucosulfakinin-II and leucosulfakinin are the only two reported invertebrate sulfated neuropeptides. As with leucosulfakinin, the intestinal myotropic activity of leucosulfakinin-II is analogous to that of gastrin and cholecystokinin. The sequence homology between the leucosulfakinins and the vertebrate hormones, as well as their analogous myotropic activity, suggest that gastrin/cholecystokinin-like neuropeptides are not confined to vertebrates, but also occur in invertebrates.

Insects utilize vertebrate-like neuropeptides to carry out a variety of regulatory functions. Sequence homology has been noted between partial sequences of silkworm neuropeptides with prothoracicotropic hormone activity and the insulin A chain (1), between partial sequences of the melanization and reddish coloration hormone and insulin-like growth factor II (2), and between the cockroach cardioaccelerator-hypertrehalosemic neuropeptide CC-2 and the vertebrate peptide glucagon (3). The presence of C-terminal gastrin or cholecystokinin(CCK)-like immunoreactivity has been demonstrated in tissues of the tobacco hornworm moth *Manduca*, silkworm *Bombyx*, drone fly *Eristalis*, cockroach *Periplaneta* and in neurons of the central nervous and

* To whom correspondence should be addressed.

† Work completed in part during a residence by R.J.N. as Visiting Scientist in the Laboratories for Neuroendocrinology, The Salk Institute, La Jolla, CA 92037.

0006-291X/86 $1.50

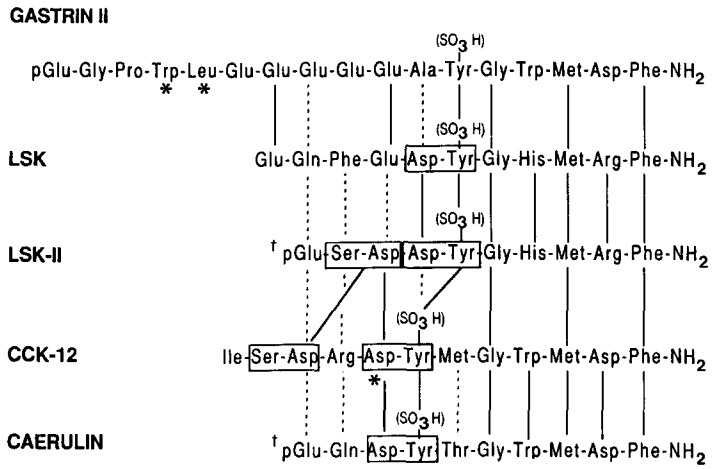

Figure 1. Sequence homology between leucosulfakinin (LSK), LSK-II and human gastrin II, cholecystokinin-12 (CCK-12) and caerulin; * denotes the N-termini of active naturally-occurring, truncated forms of gastrin and CCK. Solid lines indicate identical amino acids and dashed lines correspond to residues in which the nucleotide codons could differ by a single nucleotide. The † indicates that $pGlu^1$ of LSK-II matches the $pGlu^1$ of caerulin.

neuroendocrine systems of the blowfly *Caliphora* (4-7). Other invertebrates that contain gastrin-CCK immunoreactive substances include the crab *Cancer*, sea hare *Aplysia*, and protochordate *Ciona* (8,9). Moreover, we have recently reported that the structure of a sulfated myotropic neuropeptide, leucosulfakinin (LSK), from the Madeira cockroach (*Leucophaea maderae*) has sequence homology with the active C-terminal portion of vertebrate hormones human gastrin II and cholecystokinin (Fig. 1). The gastrointestinal myotropic activity of LSK is analogous to that of gastrin and CCK (10,11).

We now report here the isolation, characterization, and synthesis of a new sulfated myotropic neuropeptide, Leucosulfakinin-II (LSK-II), that has a blocked N-terminus but is related to LSK, from head extracts of the cockroach *Leucophaea maderae*. In addition, LSK-II also exhibits homology with vertebrate hormones cholecystokinin, human gastrin II, and caerulin. The N-terminus of this neuropeptide is blocked by a pyroglutamyl residue, a structural feature found in all insect adipokinetic hormones and prevalent among vertebrate hormones. It may serve the function of protecting the neuropeptide from degradation by aminopeptidases during transport *in vivo* (12). LSK-II shows a greater resemblance to CCK than does LSK. The sequence homology and analogous intestinal myotropic activity suggest that the distribution of gastrin/CCK-like neuropeptides in animals is greater than previously demonstrated. Distinct from other known neuropeptides from invertebrates, the two LSKs are sulfated. Like LSK, LSK-II also has been

identified in extracts of *L. maderae* corpora cardiaca, the major neurohumoral organs of insects, which are analogous to the vertebrate hypothalamus-hypophyseal system.

MATERIALS AND METHODS

The colonization of *Leucophaea maderae*, preparation of isolated hindguts for bioassay, threshold concentration and head peptide content determination (13,14), acidic (13) and basic (15) barium hydroxide [$Ba(OH)_2$] amino acid analysis, enzymatic deblocking (16), fast atom bombardment mass spectrometry (10), extraction and HPLC fractionation (10,13,18), and microsequencing (13,16) procedures and conditions have been described previously.

Peptide synthesis. Solid phase synthesis of the peptide-amide and peptide-acid utilized p-methylbenzhydrylamine and Merrifield resins, respectively, on a Beckman (Beckman Inc., Palo Alto, CA) Model 990C peptide synthesizer (13,16). Derivatized amino acids were purchased from Peninsula Laboratories (Belmont, CA). Amino acid side chain groups were protected as follows: His and Arg, tosyl; Tyr, 2,6-dichlorobenzyl; Asp, cyclohexyl; pGlu, carbobenzoxyl; Ser, benzyl or acetyl (17). The protected peptide-resins were cleaved as previously described (10). Incorporation of the sulfate ester at the tyrosine residue was effected by treatment of the peptide-amide acetate (Ac), pGlu-Ser(OAc)-Asp-Asp-Tyr-Gly-His-Met-Arg-Phe-NH_2, with con. sulfuric acid at -5°C for 30 min. The crude mixture was treated as previously described (17). The crude product was dissolved in distilled water and brought to pH 11 with con. ammonium hydroxide. After 3 hrs at ambient temperature, the solution was frozen, lyophilized, and the resulting powder purified on a Vydac C18 column with a Beckman 332 HPLC system followed by a μ-Bondapak phenyl column on a Waters ALC-100 HPLC system under previously described conditions (10). The synthetic peptide sulfate gave the following analyses: acid hydrolysis, B(2.1), F(1.0), G(1.0), H(1.1), M(1.0), R(1.1), S(1.0), Y(0.9), Z(1.1); $Ba(OH)_2$ hydrolysis, F(1.0), Y[SO_3H](0.8).

RESULTS AND DISCUSSION

We isolated LSK-II from a methanol/water/acetic acid (90:9:1) extract of 3000 *L. maderae* heads by a four-step HPLC fractionation procedure using Waters μ-Bondapak phenyl, Rainin Microsorb C1, Techsphere 3 C18 and Waters I-125 Protein-Pac columns. Biological activity was monitored by observing the myotropic effect of column fractions, that is, changes in the frequency and/or amplitude of spontaneous contractions of the cockroach proctodeum (hindgut) (13). Initial separation of the extracts on a μ-Bondapak phenyl column yielded five fractions of active material. Four of these fractions were also found in brain-corpora cardiaca extracts of *L. maderae* (18). One of those four, eluting at 54 to 56 min, was purified on a Microsorb C1 column and an active fraction eluting at 32 to 35 min was recovered. The active fraction was further purified on a Techsphere 3 C18 column and an active peak eluting at 32.6 min was isolated (Fig. 2). Final purification was effected on a Waters I-125 Protein-Pac column and 4.5 μg of pure peptide (eluting at 54.3 min) was obtained (Fig. 3). Amino acid analysis of the peptide revealed the molar ratio composition as Arg(1), Asx(2), Gly(1), Glx(1), His(1), Met(1), Phe(1), Ser(1) and Tyr(1). The inability of aminopeptidase M to destroy the biological activity of the peptide and the presence of glutamic acid in the amino acid composition suggested that the N-terminal might be blocked by pyroglutamic acid (pGlu). Incubation of LSK-II with pyroglutamyl aminopeptidase followed by HPLC fractionation resulted in isolation of the undigested parent compound and a fragment eluting

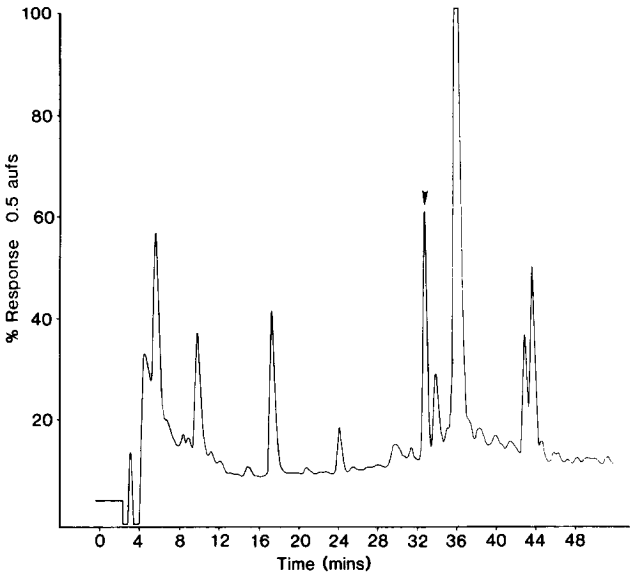

Figure 2. Fractionation on Techsphere 3 C18 of area eluting 32 to 35 min from Microsorb C1 (10). Peak eluting at 32.6 min contained myotropic activity.

at 37.5 min on Techsphere 3 C18. Microsequence analysis of the peptide fragment yielded the sequence Ser-Asp-Asp-Tyr-Gly-His-Met-Arg-Phe. Evaluation of the enzyme digestion data led to the following primary structure pGlu-Ser-Asp-Asp-Tyr-Gly-His-Met-Arg-Phe, which accounted for all the amino acids. However, two synthetic replicates of the established sequence, one with a C-terminal amide and the other a carboxylic acid, behaved differently from the native substance on HPLC analysis and were also inactive in the hindgut bioassay up to 2×10^{-6} M.

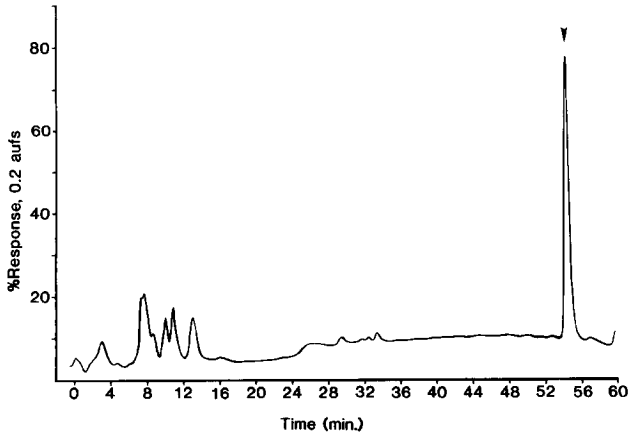

Figure 3. Fractionation on Waters I-125 Protein-Pac column (10) of active peak collected from Tecsphere 3 C18 (Fig. 2). Myotropic activity eluted in a peak at 54.3 min.

Our realization that LSK exhibited sequence homology with human gastrin and CCK (Fig. 1) provided an important clue that it contained a sulfate ester at the tyrosine residue (10).

Furthermore, the sequence homology between LSK-II and LSK strongly suggested that the former also contained a tyrosine sulfate ester. Both share a common C-terminal heptapeptide Asp-Tyr-Gly-His-Met-Arg-Phe and comparison of the remainder of the N-terminal sequences of the LSKs reveal differences that could arise from single-base substitution in the respective codons (Fig. 1). In addition, the differences in the N-terminal sequence between LSK and LSK-II further suggest that the active core of these neuropeptides required to stimulate hindgut contractile activity can be no longer than either [5-11]LSK or [4-11]LSK; the latter because of the similarity of the acidic Glu and Asp residues preceding the Asp-Tyr pair in the two LSK sequences (Fig. 1).

Incorporation of the sulfate group on the tyrosine of the synthetic LSK-II peptide-amide proved to be a greater challenge than that of LSK, due to the presence of serine, a residue highly susceptible to sulfation. The problem was circumvented by utilizing a synthetic sequence containing an acetyl group on the serine hydroxyl. Treatment of this peptide-amide with con. sulfuric acid followed by hydrolysis of the acetate in aqueous ammonium hydroxide (17) led to a crude sulfated product. Purification by reverse phase HPLC on a Vydac C18 column (product eluted at 76 to 78 min) followed by a μ-Bondapak phenyl column yielded a peak at 54.7 min, corresponding to the retention time of natural LSK-II. The synthetic peptide caused an increase in the spontaneous contractile activity of the cockroach hindgut at a threshold concentration of $4.5 \pm 0.4 \times 10^{-11}$ M ($\bar{X} \pm$ SD, n=5), virtually identical to the value recorded for the natural product ($4.2 \pm 1.4 \times 10^{-11}$ M) (Fig. 4) and ~ 5-fold more potent than that of LSK (10). The sulfated peptide had a retention time of 32.4 min on a Techsphere 3 C18 column, again similar to the value obtained for the natural product (32.6 min). Amino acid analysis of the synthetic confirmed that the residues and molar ratios were identical to those of LSK-II. Fast atom bombardment mass spectra of natural and synthetic LSK-II showed molecular ions at 1317.2 (MH^+, 87%) and 1237.3 ($MH^+ - SO_3H$, 100%) consistent with the presence of a single sulfate group (19) and a pyroglutamyl residue in the structure. The identification of tyrosine sulfate in the $Ba(OH)_2$ hydrolysate of natural LSK-II pinpointed the location of the sulfate ester (15, 17). The complete structure of LSK-II was therefore established as pGlu-Ser-Asp-Asp-Tyr(SO_3H)-Gly-His-Met-Arg-Phe-NH_2. The sulfate ester could have been readily removed under the acidic

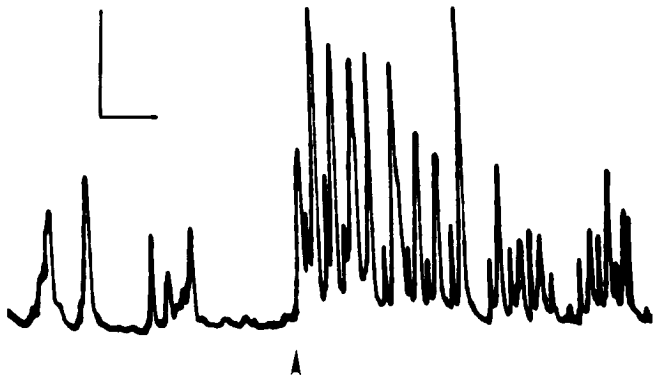

Figure 4. Response of the hindgut preparation of *Leucophaea maderae* (14) to leucosulfakinin-II (LSK-II). Concentration, 1.25 x 10^{-10} M (3 x threshold); time base, 1 min; vertical calibration, 2-mm tissue movement. The point of LSK-II application is indicated by the arrow.

conditions of amino acid analysis and sequencing. The content of LSK-II was calculated to be ~1.6 ng/head (0.0012 nm/head) in the insect, ~ 3-fold greater than that of LSK (10).

Sequence comparison reveals that residues Tyr(SO_3H)^5, Gly^6, Met^8 and Phe^{10} of LSK-II are identical to those at positions 12, 13, 15 and 17 of human gastrin II (Fig. 1). The blocked residue pGlu^1 of LSK-II is structurally related to Glu^8 in the analogous position of gastrin II, the former being a cyclic analog of the latter. Sequence homology is potentially greater, as the differences between Asp^3, Asp^4 of LSK-II and Glu^{10}, Ala^{11} of gastrin II could arise from single-base substitution in the respective codons (Fig.1) .

While LSK-II shows less resemblance to gastrin II than does LSK, it is more closely related to cholecystokinin. Residues Asp^3, Gly^6, Met^8, and Phe^{10} are identical to those in the analogous 5, 8, 10 and 12 positions of active CCK-12 (17). Lateral homologies exist between Ser^2-Asp^3 and Asp^4-Tyr(SO_3H)^5 of LSK-II and the corresponding amino acid pairs in positions 2-3 and 5-6 of CCK-12 (Fig. 1). Intervening residues Arg^4 and Met^7 of CCK prevent a direct alignment with the amino acid block Ser-Asp-Asp-Tyr(SO_3H) of LSK-II. In addition, sequence differences pGlu^1,Ser^2,Asp^4 of LSK-II and Asp^3,Arg^4,Tyr^6 of CCK-12 could arise from single-base substitution in the respective codons. Furthermore, the pGlu^{1}of LSK-II is identical to the residue in the same position of caerulin, a frog peptide related to CCK. Thus, 5 of 10 (50%) of the residues in LSK-II are identical to those in the decapeptide caerulin (Fig. 1). The presence of a sulfate ester on the tyrosine residues of LSK-II and LSK, a modification present in only a small number of isolated peptides (15,20), introduces significant resemblance beyond amino acid residue matches. Both LSKs contain an Asp preceding the Tyr(SO_3H) moiety, a structural feature proposed to serve

as a component of the recognition site for tyrosylprotein sulfotransferase (20) in CCK, caerulin and other sulfated peptides. The sequence homology and analogous biological activity exhibited by the LSKs and the gastrin-CCK neurohormones suggest that the probability of coincidental structural resemblance is remote. Furthermore, the presence of both LSK-II and LSK in the neurosecretory corpora cardiaca and the brain, their myotropic activity at the distant hindgut, and homology with vertebrate hormones suggests a hormonal function for the LSKs in *L. madera*.

Synthetic CCK-8, caerulin (Peninsula Laboratories), and gastrin II are inactive on the isolated cockroach hindgut up to 10^{-6} M (10). This was expected because of the disparity between the basic His and Arg residues at the C-terminal region of the LSKs and the analogous nonpolar Trp and acidic Asp residues in the same relative positions of gastrin and CCK (Fig. 1). This disparity is consistent with the observation that gastrin-CCK immunoreactive substances from the crab *Cancer magister* and sea hare *Aplysia californica* are inactive in mammalian bioassays. Presumably, the presence of sequence differences in the C-terminal region of invertebrate CCK-gastrins affect biological activity more than immunoreactivity (8).

The biological actions of CCK in mammalian systems include the stimulation of pancreatic enzyme secretion, gall bladder contraction, intestinal motility, and involvement in the mechanism of satiety (17,21). Nonsulfated CCKs exhibit a 250 to 350 fold decrease in amylase secretion and gall bladder contraction as compared with the sulfated peptide (17,22), demonstrating the importance of the sulfate ester. Caerulin exhibits activity similar to CCK in mammalian bioassays.

Although the primary gastrointestinal function of gastrin in mammals is to induce gastric acid secretion, it also stimulates smooth muscle contraction, increases blood circulation and water secretion in the stomach and intestine, and stimulates pancreatic secretion. The sulfate ester is not required in gastrin to elicit gastric acid secretion (23,24). Contraction of visceral muscle promoted by the LSKs in the cockroach hindgut and its attendant stimulation of haemolymph (blood) circulation (25), are analogous to gastrin-induced motility and increased blood circulation in mammalian intestines. It is unknown whether the LSKs are also involved in the secretion of digestive enzymes and regulation of water balance or satiety in the cockroach, other insects, or other invertebrates. The LSK structures may be of significance to the field of peptide hormone origins.

ACKNOWLEDGMENTS

We thank T. Peterson, R. Wagner and M. Wright for technical support; and J. Menn of the National Program Staff of the U.S. Dept. of Agriculture for support of this program.

REFERENCES

1. Nagasawa, H., Kataoka, H., Isogai, A., Tamara, S., Suzuki, A., Ishizaki, H., Mizoguchi, A Fujiwara, Y. and Suzuki, A. (1984) *Science* **226**, 1344-1345.
2. Matsumoto, S., Isogai, A., and Suzuki, A. (1985) *FEBS Letters* **189**, 115-118.
3. Scarborough, R.M., Jamieson, G.C., Kalish, F., Kramer, S.J., McEnroe, G.A., Miller, C.A., and Schooley, D.A. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5575-5579.
4. Witten, J.L., Schaffer, M.H., O'Shea, M., Cook, J.C., Hemling, M.E. and Rinehart, K.L. (1984) *Biochem. Biophys. Res. Commun.* **124**, 350-358.
5. Truman, J.W. and Taghert, P.H. (1983) in *Brain Peptides* (Krieger, D.T., Brownstein, M.J., eds) pp. 165-181, John Wiley, New York.
6. Duve, H. and Thorpe, A. (1984) in *Insect Neurochemistry and Neurophysiology* (Borkovec, A.B. and Kelley, T.J., eds) pp. 171-195, Plenum, New York.
7. Witten, J., Worden, M.K., Schaffer, M.H., O'Shea, M. (1984) *Soc. Neurosci. Abstr.* **10**, 46.
8. Vigna, S.R. (1985) *Peptides* **6** (suppl. 3), 283-287.
9. Thorndyke, M. and Dockray, G.J. (1986) *Abstr. Sixth Interntl. Symp. Gastrointest. Hormones, Can. J. Physiol. Pharm.*, 24-25.
10. Nachman, R.J., Holman, G.M., Haddon, W.F. and Ling, N. (1986) *Science*, in press.
11. Takahashi, T., Kantoh, M., Kusunoki, M., Yamamura, T., and Utsunomiya, J. (1986) *Abstr. Sixth Internatl. Symp. Gastrointest. Hormones, Can. J. Physiol. Pharm.*, 46-47.
12. Gregory, H., Hardy, P.M., Jones, D.S., Kenner, G.W. and Sheppard, R.C. (1964) *Nature* **204**, 931-933.
13. Holman, G.M., Cook, B.J. and Nachman, R.J. (1986) *Comp. Biochem. Physiol., 84C*, 205-211.
14. Cook, B.J. and Holman, G.M. (1978) *Comp. Biochem. Physiol.* **61C**, 291-295.
15. Doolittle, R.F. (1965) *Biochem. J.* **94**, 742-750.
16. Holman, G.M., Cook, B.J. and Nachman, R.J.(1986) *Comp. Biochem. Physiol.*, in press.
17. Ondetti, M.A., Pluscec, J., Sabo, E.F., Sheehan, J.T. and Williams, N. (1970) *J. Am. Chem. Soc.* **92**, 195-199.
18. Holman, G.M., Cook, B.J. and Wagner, R.M. (1984) *Comp. Biochem. Physiol.* **77C**, 1-5.
19. Gibson, B.W., Poulter, L. and Williams, D.H. (1986) *J. Nat.Products (Lloydia)* **49**, 26-34.
20. Lee, R.W. and Huttner, W.B. (1983) *J. Biol. Chem.* **258**, 11326-11334.
21. Deutsh, J.A. (1983) *Prog. Neurobiol.* **20**, 313-320.
22. Yajima, H., Mori, Y., Koyama, K., Tobe, T., Setoyama, M., Adachi, H., Kanno, T. and Saito, A. (1976) *Chem. Pharm. Bull.* **24**, 2794-2802.
23. Grossman, M.I. (1970) *Nature* **228**, 1147-1150.
24. Fasth, S., Filippsson, S., Hulten, L. and Martinson, *Experientia* (1973) **29**, 982-984.
25. Chapman, R.F. (1971) *The Insects: Structure and Function*, p. 670, Elsevier, New York.