Volume 245, number 1,2, 4–8 FEB 06883 March 1989

# Amino acid sequence of locust neuroparsins

Josiane Girardie, Adrien Girardie, Jean-Claude Huet* and Jean-Claude Pernollet*

*Laboratoire de Neuroendocrinologie, UA CNRS 1138, Université de Bordeaux I, 33405 Talence Cedex* and **Laboratoire d'Etude des Protéines, Centre INRA, route de St Cyr, 78026 Versailles Cedex, France*

Received 13 December 1988

Neuroparsins A and B were isolated from the nervous part of the corpus cardiaca of *Locusta migratoria* via a two-step purification procedure. Both consist of two polypeptide chains linked by disulfide bridges. The N-terminal sequence of both native neuroparsins was determined: the N-terminal end of neuroparsin B was unique while that of neuroparsin A showed three different sequences. These sequences were that of neuroparsin B and two others having five and two extra N-terminal residues. Neuroparsin B was found as a homodimer and the complete sequence of the monomer, determined from peptide fragments generated by treatment with cyanogen bromide and endoprotease Glu-C, comprises 78 residues.

Amino acid sequence; Neurohormone; Insect; (*Locusta migratoria*)

## 1. INTRODUCTION

Neuroparsins A (NPA) and B (NPB) are two proteins ($M_r \sim 14000$) formed by two polypeptide chains ($M_r \sim 7000$) linked by disulfide bonds. They have recently been isolated [1] from nervous corpora cardiaca of *Locusta migratoria* using electroelution of excised electrophoretic bands. These two proteins are produced by the A1 type of protocerebral median neurosecretory cells [1]. They are multifunctional neurohormones: they inhibit the effects of juvenile hormone [1], stimulate fluid reabsorption of isolated recta [2] and induce an increase in hemolymph lipid and trehalose levels [3].

Here, we report the isolation of neuroparsins using sequentially anion-exchange and reversed-phase liquid chromatography, and their N-terminal sequencing. Neuroparsin B was shown to be formed by the association of two identical chains which were completely sequenced.

*Correspondence address:* J. Girardie, Laboratoire de Neuroendocrinologie, UA CNRS 1138, Université de Bordeaux I, 33405 Talence Cedex, France

*Abbreviations:* NBRF, National Biomedical Research Foundation; NPA(B), neuroparsin A(B); RPLC, reversed-phase liquid chromatography; TFA, trifluoroacetic acid

## 2. MATERIALS AND METHODS

### 2.1. *Materials*

Adults of both sexes of *Locusta migratoria migratorioides* were reared as in [1]. Nervous corpora cardiaca were carefully separated and homogenized in 70% methanol at 4°C using an ultrasonic probe. The solution was centrifuged at 33000 × *g* for 30 min. The supernatant was freeze-dried following evaporation of methanol under $N_2$ and stored at −80°C.

### 2.2. *Purification procedures*

Freeze-dried extracts were resuspended in 25 mM Tris-HCl (pH 8) and chromatographed on an anion-exchange column (Mono Q HR5/5, Pharmacia). A linear gradient starting at 25 mM Tris-HCl (pH 8) and ending at 0.5 M NaCl in 33 min was employed at a flow rate of 1 ml/min. Fractions were manually collected following the UV absorbance at 225 nm. The neuroparsin fractions were checked at this step via their electrophoretic mobility in 7.5% PAGE following the procedure in [4]. Each neuroparsin fraction was desalted and purified on a reversed-phae C8 30 nm column (Pro RPC HR5/10, Pharmacia). A linear gradient from 28 to 64% of $CH_3CN$ in 40 min at a flow rate of 0.5 ml/min. The purity of each neuroparsin peak was monitored by 10–20% SDS-PAGE [4].

### 2.3. *Reduction and alkylation*

Neuroparsin B was reduced and alkylated according to [5],



*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/89/$3.50 

using 4-vinylpyridine as alkylating agent. The sample was desalted directly on the same RP column under the conditions described above.

### 2.4. *Cyanogen bromide cleavage*

Reduced and alkylated NPB was treated according to [6], using 70% TFA instead of formic acid. The reaction mixture was diluted with water (3.5 times) and half evaporated under $N_2$. The peptides were recovered using the RPLC conditions previously described.

### 2.5. *Enzymatic digestion*

Reduced and alkylated NPB was digested with endoprotease Glu-C as in [6] for 24 h. The reaction mixture was applied to an Aquapore RP 300 column (4.6 × 30 mm), and fragmented peptides were eluted with a 30 min gradient of 0–40% $CH_3CN$ in 0.1% TFA at a flow rate of 0.5 ml/min.

### 2.6. *Sequence analysis*

Purified neuroparsins and all peptide fragments were sequenced using an Applied Biosystems model 470A gas-phase protein sequencer [7]. Released phenylthiohydantoin amino acids were analyzed using an on-line analyzer (Applied Biosystems, model 120A).

### 2.7. *Predictions of structure*

The hydropathy profile was determined using the algorithm

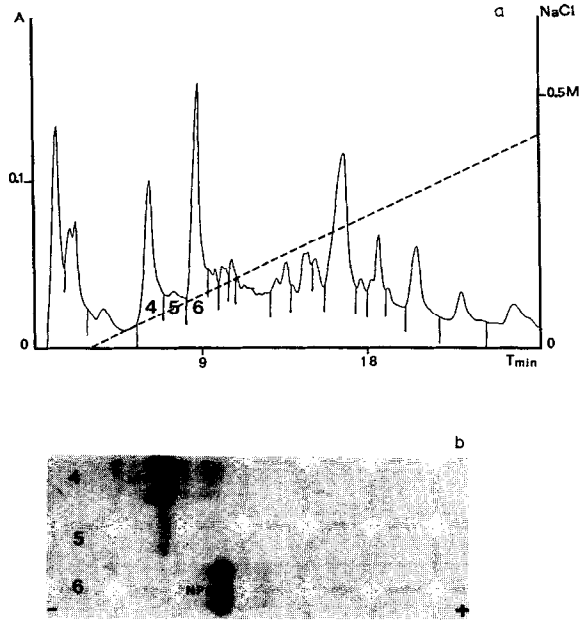


Fig.1. Purification of neuroparsins. (a) Anion-exchange LC of nervous corpus cardiaca extracts. (——) Absorbance at 225 nm; (– – –) NaCl gradient. Numbers indicate fractions containing neuroparsins. (b) 7.5% PAGE of fraction 4–6. Fractions 4,5 contain a major protein with the electrophoretic mobility of NPA. Fraction 6 contains a major protein with the electrophoretic mobility of NPB.

Table 1

Edman amino acid sequencing of the monomer of neuroparsin B

| Cycle | N-term | B1 | B2 | S1 | S2 | S3 |
|---|---|---|---|---|---|---|
| Elution position ($CH_3CN$ %) | 33 | 29 | 30 | 15.5 | 22.5 | 32.5 |
| 1 | S(329) | S(538) | D(415) | G(145) | S(145) | Y(131) |
| 2 | – | C(382) | C(437) | G(−47) | C(58) | G(113) |
| 3 | E(379) | E(414) | R(159) | L(96) | E(82) | D(101) |
| 4 | G(144) | G(307) | C(377) | P(96) | G(72) | V(132) |
| 5 | A(406) | A(391) | G(356) | S(75) | A(66) | T(143) |
| 6 | N(341) | N(363) | L(436) | S(−13) | N(52) | D(79) |
| 7 | – | C(377) | C(415) | C(13) | C(41) | F(117) |
| 8 | V(283) | V(319) | S(279) | | V(47) | F(10) |
| 9 | V(95) | V(56) | G(248) | | V(3) | G(99) |
| 10 | D(225) | D(365) | C(237) | | D(36) | R(58) |
| 11 | L(164) | L(290) | S(158) | | L(36) | K(69) |
| 12 | T(63) | T(291) | L(138) | | T(36) | V(81) |
| 13 | R(82) | R(293) | H(56) | | R(31) | C(80) |
| 14 | – | C(241) | N(153) | | | A(96) |
| 15 | E(115) | E(226) | L(121) | | | K(59) |
| 16 | Y(109) | Y(201) | Q(108) | | | G(63) |
| 17 | G(83) | G(159) | C(96) | | | P(80) |
| 18 | D(83) | D(189) | F(76) | | | G(37) |
| 19 | V(106) | V(143) | F(11) | | | D(44) |
| 20 | T(37) | T(128) | F(−9) | | | K(36) |
| 21 | D(36) | D(118) | E(52) | | | C(43) |
| 22 | F(67) | F(108) | G(34) | | | G(33) |
| 23 | F(48) | F(43) | G(11) | | | G(18) |
| 24 | G(27) | G(79) | L(35) | | | P(36) |
| 25 | R(47) | R(115) | P(37) | | | Y(24) |
| 26 | K(35) | K(99) | S(22) | | | E(25) |
| 27 | V(51) | V(73) | S(2) | | | L(25) |
| 28 | – | C(76) | | | | H(12) |
| 29 | A(56) | A(69) | | | | G(17) |
| 30 | K(24) | K(76) | | | | K(11) |
| 31 | G(28) | G(41) | | | | C(13) |
| 32 | P(41) | P(68) | | | | G(9) |
| 33 | G(14) | G(26) | | | | V(9) |
| 34 | D(26) | D(55) | | | | G(5) |
| 35 | K(13) | K(55) | | | | M(10) |
| 36 | – | C(41) | | | | D(6) |
| 37 | G(16) | G(34) | | | | C(5) |
| 38 | G(14) | G(23) | | | | R(5) |
| 39 | P(17) | P(32) | | | | C(4) |
| 40 | Y(18) | Y(32) | | | | G(4) |
| 41 | E(18) | E(34) | | | | L(4) |
| 42 | L(15) | L(24) | | | | C(2) |
| 43 | H(14) | H(26) | | | | |
| 44 | G(15) | G(18) | | | | |
| 45 | | K(15) | | | | |
| 46 | | C(14) | | | | |
| 47 | | G(3) | | | | |
| 48 | | V(4) | | | | |

N-term, sequencing of the N-terminal end of the native (dimeric) NPB. Values in parentheses indicate the differential amount (pmol) of phenylthiohydantoin amino acids determined at each step vs the former. Amino acids are indicated using the single-letter code

of Kyte and Doolittle [8], with a window width of 7. Secondary structure prediction was carried out with the GOR algorithm, version III [9].

## 3. RESULTS

Crude extracts of nervous corpus cardiaca were separated on Mono Q column into 20 fractions (fig.1a). Electrophoresis revealed that NPA was mainly found in fraction 4 and somehow in fraction 5, whereas NPB was eluted in fraction 6 (fig.1b). Fractions 4 and 6 gave two peaks after reversed-phase chromatography, with the same retention time whatever the fraction used. Only the major peak eluting at 35% $CH_3CN$ corresponded to the 14 kDa NPA or NPB, as determined through 10–20% PAGE. Sequencing of native NPB yielded a single sequence up to 44 residues (table 1), suggesting that both chains were identical. In contrast to NPB, NPA, sequenced up to 29 residues, appeared to be composed of three polypeptide chains. One of the chains, accounting for 30%, was identical to NPB, while the other two, representing 50 and 20%, had respectively five (Asn-Pro-Ile-Ser-Arg) and two (Ile-Arg) extra N-terminal residues.

The major peak of fraction 6 was rechromatographed on a reversed-phase system, after reduction and alkylation to purify NPB. Only one peak ($M_r \sim 7000$) was obtained, consisting of the NPB monomer. As shown in table 1, it was cleaved by CNBr and separated into two peptides (B1,B2) which were sequenced. Using endoprotease Glu-C, three peptides were obtained (S1–S3) overlapping with the native N-terminus and B1 and B2 (table 1 and fig.2). No microheterogeneity was found at any step during the sequencing of these peptides and the C-terminus was free. The NPB monomer is characterized by the absence of Trp and Ile and by the very high frequency of Cys (12 out of 78 residues) regularly distributed along the sequence. These residues were usually preceded by a positively charged amino acid and followed by Gly or Glu. The calculated $pH_i$ was found to be 5.98, in agreement with previous experimental determination [10].

As regards the secondary structure of NPB, no periodic structure (either $\alpha$-helix or $\beta$-sheet) was ever predicted all along the sequence. The hydropathy profile confirmed the general hydrophobicity of the molecule (fig.3) which is extractable in 70% methanol. Only the central region of the sequence, from residue 30 to 45, appears to be hydrophilic and can be expected to be an antigenic determinant. Another could involve the region around Arg 13. The NPB sequence, compared with those of the NBRF data bank (release

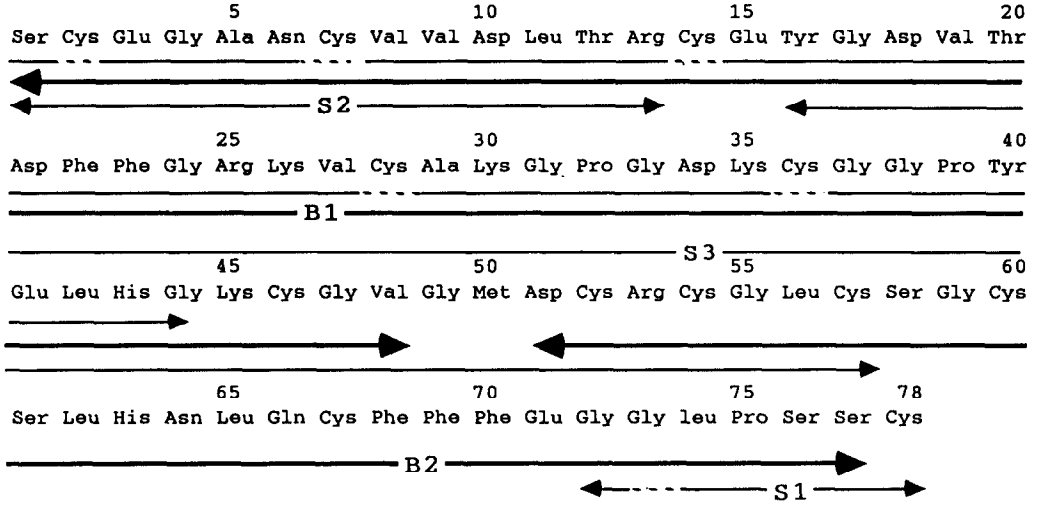


Fig.2. Complete amino acid sequence of the monomer of neuroparsin B of *Locusta migratoria*. (——→) N-terminus of the native protein. (←——→) Cleaved peptides (B, cyanogen bromide peptides; S, endoprotease Glu-C peptides). Dashed lines indicate undetermined residues.

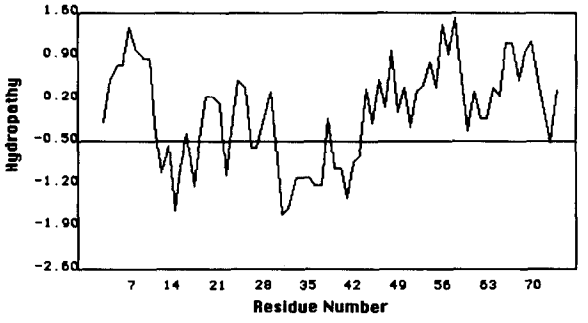


Fig.3. Hydropathy profile of neuroparsin B monomer.

16) and with the translations of the open reading frames of Genbank (release 54), appeared as an original protein. Nevertheless, some partial homologies were found between some NPB regions and several vertebrate hormones, mainly the hypothalamic pituitary hormones and insulin (to be published).

## 4. DISCUSSION

Neuroparsins were obtained in a purity grade permitting their sequencing after two steps of liquid chromatography. Purified using the same procedure, they were previously demonstrated to have antidiuretic [2], hyperglycemic and hyperlipemic [3] and antigonadotropic [11] actions. We demonstrate that all these activities are triggered by a single molecule; the neuroparsins are obviously polytropic locust neurohormones as it is of common occurrence with most neuropeptides of both invertebrates and vertebrates.

The absence of microheterogeneity all along the NPB sequence establishes that NPB is composed of two identical chains originating in a single gene. Nevertheless, the NPB sequence is found not only as the major part of the two NPA typical sequences, but also as a component of native NPA, so NPB appears to be a post-translational product of the maturation of the other two chains. The processing of these polypeptides involves several steps to reach the final NPB homodimeric molecule, as suggested by the N-terminal heterogeneity of NPA. This sequence homology between NPB and NPA N-termini corroborates that NPB originates from NPA by enzymatic cleavage as has been previously demonstrated [12]. In both cases, it occurs after an arginyl residue. The occurrence of extra arginyl residues at the N-terminus of NPA is in agreement with the clear separation of both neuroparsins under ion-exchange chromatography and native electrophoresis.

The calculated $M_r$ of the NPB monomer is 8188, in agreement with the two previous estimations of $M_r$ 9000 [10] and $M_r$ 7000 [4]. It is worth noting that with 12 Cys per monomer, NPB is currently the insect neurohormone that is highest in sulfur content. This high cysteine content is in agreement with the strong in vitro labelling of neuroparsins following injection of [$^{35}$S]cysteine [13]. It is likely that numerous and frequent intrachain putative disulfide bridges only allow the occurrence of loops and turns beside aperiodic arrangements and prevent the formation of a periodic structure. As deduced from the hydropathy profile of NPB, both termini are hydrophobic. They could well be involved in the oligomeric association of the native protein.

A comparison of the NPB sequence with the incomplete N-terminal sequence of the locust corpus cardiacum peptide [14] reveals a clear homology, suggesting the identity of the two molecules. Moreover, the partial homology of neuroparsins with vertebrate hormones is in favor of a common ancestral origin of neuroparsins. Whether the neuroparsins originate from a single precursor such as insulin remains a questionable point.

*Acknowledgement:* This study was supported by an INRA Action Incitative Programmée no.8/4625.

## REFERENCES

[1] Girardie, J., Bourême, D., Couillaud, F., Tamarelle, M. and Girardie, A. (1987) Insect. Biochem. 17, 977–983.
[2] Fournier, B. and Girardie, J. (1988) J. Insect. Physiol. 34, 309–313.
[3] Moreau, R., Gourdoux, L. and Girardie, J. (1988) Arch. Insect Biochem. Physiol. 8, 135–145.
[4] Bourême, D. and Girardie, J. (1986) CR Acad. Sci. Paris 302, 655–660.
[5] Henschen, A. (1986) in: Advanced Methods in Protein Microsequence Analysis (Wittmann-Liebold, B. et al. eds) pp.244–255, Springer, Berlin.
[6] Kamp, R.M. (1986) in: Advanced Methods in Protein Microsequence Analysis (Wittmann-Liebold, B. et al. eds) pp.8–20, Springer, Berlin.

[7] Hewick, R.N., Hunkapiller, M.W., Hood, L.E. and Dreyer, W.J. (1981) J. Biol. Chem. 256, 7990–7997.
[8] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[9] Gibrat, J.F., Garnier, J. and Robson, B. (1987) J. Mol. Biol. 198, 425–443.
[10] Girardie, J., Faddoul, A. and Girardie, A. (1985) Insect Biochem. 15, 85–92.
[11] Couillaud, F., Girardie, A. and Girardie, J. (1988) Int. J. Invert. Reprod. Dev., in press.
[12] Bourême, D. and Girardie, J. (1988) Ann. Soc. Ent. Fr., in press.
[13] Girardie, J., Bourême, D. and Girardie, A. (1987) Insect Biochem. 17, 29–36.
[14] Mordue, W., Morgan, P.J. and Siegert, K.J. (1985) Peptides 6, 407–410.