Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochimica et Biophysica Acta 1621 (2003) 17–21
BIOCHIMICA ET BIOPHYSICA ACTA
BBA
www.bba-direct.com

# Structural characterization of E*n*-1, a cold-adapted protein pheromone isolated from the Antarctic ciliate *Euplotes nobilii*

Claudio Alimenti[a], Claudio Ortenzi[a], Vito Carratore[b], Pierangelo Luporini[a,*]

[a] *Dipartimento di Biologia Molecolare Cellulare e Animale, University of Camerino, Via Camerini 5, 62032 Camerino (MC), Italy*
[b] *Istituto di Biochimica delle Proteine ed Enzimologia, CNR/IBPE, Via G. Marconi, Naples 80125, Italy*

Received 19 November 2002; received in revised form 7 January 2003; accepted 14 January 2003

**Abstract**

The second of two diffusible cell signal proteins (pheromones) purified from a wild-type strain of the Antarctic ciliate, *Euplotes nobilii*, has been determined by automated Edman degradation of the whole molecule and peptides generated by its chymotryptic digestion. The proposed sequence of 52 amino acids of this new pheromone, designated E*n*-1, is: NPEDWFTPDT$_{10}$CAYGDSNTAW$_{20}$TTCTTPGQT-C$_{30}$YTCCSSCFDV$_{40}$VGEQACQMSA$_{50}$QC. In common with the previously determined 60-amino-acid sequence of the other pheromone, E*n*-2, it bears eight cysteines in conserved positions (presumably linked into four conserved intrachain disulfide bonds), and physicochemical features of potential significance for cold adaptation, such as a reduced hydrophobicity, an increased solvent accessibility, and an improved local backbone flexibility. However, E*n*-1 diverges from E*n*-2 for having evolved a threonine cluster in the place of a glycine cluster to apparently make more flexible a region that is likely functionally important.
© 2003 Elsevier Science B.V. All rights reserved.

*Keywords:* Cold-adapted protein; Threonine repetition; Water-borne protein hormone; Ciliate biology; Antarctic biology

## 1. Introduction

Protein signals (pheromones), acting in autocrine fashion to promote the vegetative (mitotic) cell growth and in paracrine fashion to induce mating pair formation [1,2], are constitutively synthesized and diffused into the extracellular environment by free-living species of protozoan ciliates (Refs. [3,4], as reviews). The study of their structure and function has for long been limited to temperate water species of *Blepharisma* [5,6], *Dileptus* [7], and *Euplotes* [8–12], that can be easier domesticated to laboratory conditions and perpetuated through mass cultures.

Two pheromones (E*n*-1 and E*n*-2) have recently been isolated from a species, *Euplotes nobilii*, collected from the freezing coastal waters of Antarctica and currently represented in the laboratory by a number of strains growing well in a cold room at 4 °C [13]. Both pheromones are released from cells of the strain AC-1 that is heterozygous at the *mat* (mating-type) locus for two of a multiple series of pheromone genes regulated by relationships of co-dominance [13], as in other pheromone-secreting species of *Euplotes* [4,14]. The sequence of E*n*-2 of 60 amino acids was first determined, and its most distinctive feature was shown to be an adaptive insertion of a Gly-rich motif potentially capable to make more flexible a functionally critical domain of the molecule [15]. The determination of the complete E*n*-1 sequence of 52 amino acids now provides a wider basis to discern the structural solutions that akin proteins from the same organism have adopted to become cold adapted.

## 2. Materials and methods

### 2.1. Cell cultures and pheromone purification

Cultures of strain AC-1, source of the pheromone E*n*-1, and strain AC-4, used in assays of E*n*-1 activity, were fed on the green alga *Dunaliella tertiolecta* grown in natural seawater, previously sterilized, and enriched with Walne medium. To be used for the pheromone purification, they were first allowed to multiply for 1–2 weeks in the presence of food to reach a cell density of about $10^4$/ml, and then concentrated and resuspended, at the same density and for 3–4 days, in fresh seawater without food. Afterwards, the

* Corresponding author. Fax: +39-737-636216.
*E-mail address:* piero.luporini@unicam.it (P. Luporini).

0304-4165/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0304-4165(03)00011-4

supernatant of these cultures was collected and passed, first, through paper filters to remove gross debris in suspension and, then, through filters of 0.2 μm pore size to remove fine particulate material. Pheromone was purified from these supernatant preparations essentially according to a standard protocol originally devised to purify pheromones of *Euplotes raikovi* [14]. It is based on three steps, the first of which is pheromone adsorption from the supernatant onto Sep-Pak $C_{18}$ cartridges (Water, Milford, MA), and the other two are chromatographic fractionations of the cartridge eluate on Superose-12 and Mono-Q columns (Amersham Pharmacia Biotech, Piscataway, NJ). Usually, 150 μg of homogeneous protein was prepared from a 10-l volume of cell supernatant. (However, recent optimization trials indicate that up to 10-fold, this amount can be obtained from supernatant that is directly recovered from cultures grown uninterruptedly for 2–3 weeks, at 8–10 °C.) The purified material could be stored at −20 °C after lyophilisation without losing activity. This was measured to be of the order of $10^{-8}$ M in assays carried out on cells of strain AC-4 and based on the conventional and practical (yet questionable) criterion of mating induction [16,17].

### 2.2. Protein modification and cleavage

To carry out E*n*-1 chymotrypsin digestion, a purified preparation was (1) resuspended at a final concentration of 0.5 μg/μl in 0.1 M Tris, pH 8.1, containing 8 M urea (Mallinckrodt Baker, Phillipsburg, NJ), (2) reduced with 100 mM dithiothreitol (DTT) (Sigma-Aldrich, St. Louis, MO) for 2 h at 50 °C, (3) *S*-pyridyl-ethylated with 1% (v/v) 4-vinylpyridine (Sigma-Aldrich) for 2 h at room temperature, (4) immediately desalted by gel-filtration chromatography on a Superdex Peptide HR 10/30 column (Amersham Pharmacia Biotech) equilibrated with 30% acetonitrile (Mallinckrodt Baker), (5) lyophilised, and (6) incubated at a final concentration of 0.2 μg/μl, for 2 h at room temperature, in 0.1 M Tris, pH 8.1, with 2 M urea, and chymotrypsin (Sigma-Aldrich) (enzyme/substrate ratio, 1:100 w/w).

### 2.3. Peptide separation

The E*n*-1 enzymatic digest was fractionated by reverse-phase-high-performance liquid chromatography (RP-HPLC) on a 2.1-mm × 250-mm $C_{18}$ column (Supelco, Bellefonte, PA), eluted with acetonitrile in the presence of 0.06% (v/v) trifluoroacetic acid and at a flow rate of 200 μl/min. The separated peptides were lyophilised and stored at −20 °C before use.

### 2.4. Mass spectrometry and amino acid sequence analysis

Molecular mass determinations were carried out on a matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) Voyager DE-PRO apparatus (Perseptive Biosystems, Framingham, MA) equipped with a nitrogen laser (operated at 337 nm). Automated Edman degradation was carried out in a Procise 492 protein sequencer (Applied Biosystem, Perkin Elmer Division, Foster City, CA).

## 3. Results and discussion

### 3.1. Sequence determination

Preliminary to the analysis of the E*n*-1 sequence, the number of Cys residues present in this sequence, earlier deduced to be eight from E*n*-1 amino acid composition data (not shown), was verified by measuring the variation in the molecular mass that occurs between samples of native E*n*-1 and E*n*-1 reduced with DTT and alkylated with 4-vinylpyridine. The difference observed was 845.2 Da (native E*n*-1 and alkylated E*n*-1 having masses of 5617.6 and 6462.8 Da, respectively), that is, a value consistent with the conversion of eight cysteines into eight 4-pyridylethyl cysteine residues (each modification causing a mass increase of 105 Da).

Initial knowledge of the amino acid sequence of E*n*-1 was provided by automated Edman degradation of an alkylated preparation of intact E*n*-1. By this analysis, the following 31 residues at the N terminus of the molecule were unequivocally identified: NPEDWFTPDT$_{10}$CAYGDSNTAW$_{20}$TTCTTPGQTC$_{30}$Y.

As shown in Fig. 1, to complete the sequence, it was sufficient to analyse by automated Edman degradation two peptides generated from a mild chymotryptic digestion of alkylated E*n*-1, which appeared as the most pronounced, along with a third peak corresponding to uncut E*n*-1, in the chromatographic profile of the chymotryptic digest. One of the two peaks (denoted *a*) showed a mass of 2723.0 Da; the other (denoted *b*), a mass of 3752.2 Da. Because the sum of these masses was practically equivalent to the mass of 6462.8

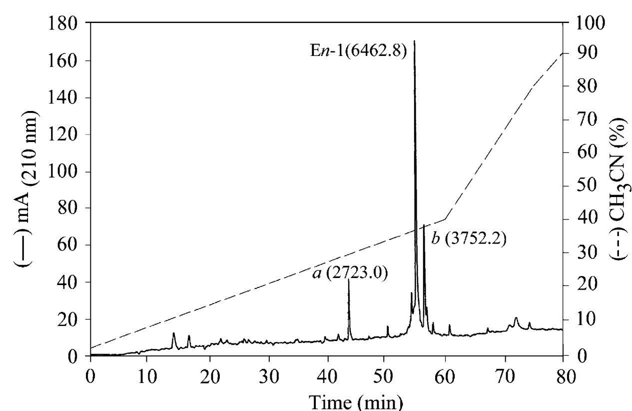


Fig. 1. Chymotryptic map of E*n*-1. A purified E*n*-1 preparation, reduced with DTT and alkylated with 4-vinylpyridine, was exposed to digestion with chymotrypsin and the digest was fractionated by RP-HPLC. The fragments were detected by their absorbance at 210 nm, and those relative to the peaks designated *a* and *b* were analysed by mass spectrometry to determine their molecular masses (reported in parentheses), and by automated Edman degradation to determine their sequences.

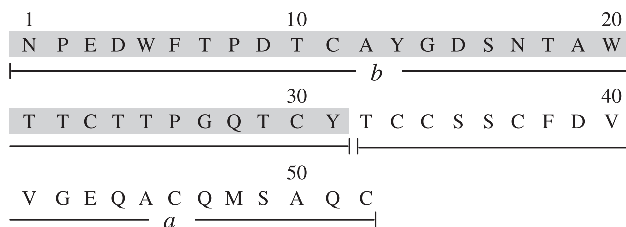


Fig. 2. Proposed amino acid sequence of E*n*-1. Shading indicates residues determined by analysis of intact E*n*-1. The underlined segments *a* and *b* indicate the E*n*-1 chymotryptic peptides corresponding to the equally labeled peaks in the chromatographic profile of Fig. 1. The sequence is available from the Swiss-Prot database under the accession number P83441.

Da of an intact and alkylated E*n*-1 molecule, it was evident that the entire E*n*-1 sequence could be derived from only the two sequences of peptides *a* and *b*. As shown in Fig. 2, the sequence of peptide *b* ending with a Tyr residue (i.e., a potential site of chymotrypsin activity) was found to be the same as the 31-amino-acid sequence previously determined by chemical analysis of undigested E*n*-1, while the sequence of peptide *a* was determined to be composed of the following 21 residues (numbered in relation with those of the peptide *b*): $T_{32}CCSSCFDVVG_{42}EQACQMSAQC_{52}$. The fact that five of these residues were cysteines provided further evidence of the completeness of the E*n*-1 sequence. Added to the three cysteines found in peptide *b*, the total number of Cys residues in the E*n*-1 sequence turned out to be eight as initially calculated.

### 3.2. Sequence relationships and properties

The alignment of the proposed 52-amino-acid sequence of E*n*-1 with the 60-residue sequence of the other *E. nobilii* pheromone, E*n*-2, is illustrated in Fig. 3. An insertion of five gaps was sufficient to put in register all the eight Cys residues. These likely represent the sequence feature that is most retained across the whole *E. nobilii* pheromone family, as is the case of the 6 cysteines paired into three conserved intrachain disulfide bonds in most of the α-helical *E. raikovi* pheromones [22,23], and of the 10 cysteines of most *Euplotes octocarinatus* pheromones [11,12]. The structural homology between E*n*-1 and E*n*-2 is further denoted by 14 other sequence identities. These are largely localized in the N-terminal half of the two molecules, which may thus be regarded as the region that is more conserved and probably correlated with common functions such as oligomer forma-

A. Amino acid sequence alignment of E*n*-1 and E*n*-2

B. Predicted features of E*n*-1 and E*n*-2 structures

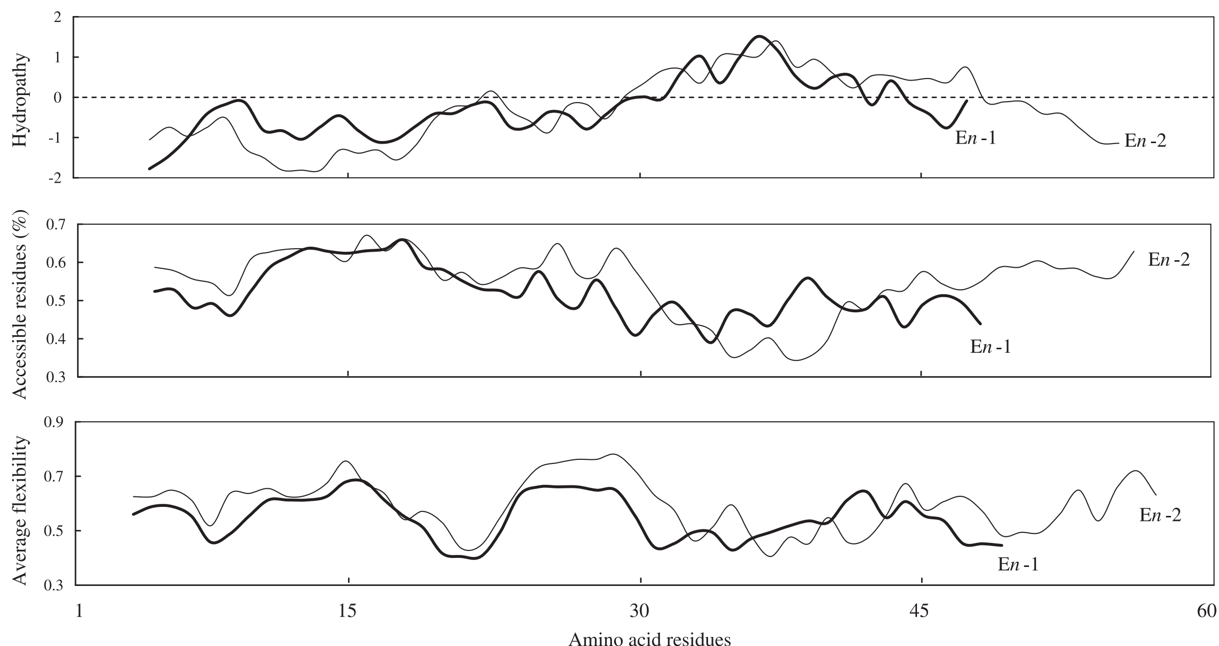


Fig. 3. Alignment and predicted properties of the E*n*-1 and E*n*-2 amino acid sequences. (A) The alignment was based on the CLUSTAL-X algorithm [18]; sequence identities are marked by asterisks, and Cys residues are printed in bold. (B) The profiles of hydropathy, solvent accessibility, and average flexibility were produced according to Kyte and Doolittle [19], Janin [20], and Bhaskaran and Ponnuswanny [21], respectively, using a span setting of seven residues; in the hydropathy profile, values above and below the zero line indicate hydrophobic and hydrophilic regions, respectively.

tion and receptor docking. On the other hand, four of the five gaps added to optimize the alignment involve the C-terminal half, which differs also in length between the two molecules. This region thus appears to be more variable and pheromone specific, and is probably related to unique properties such as receptor binding and activation.

Consistent with this close relatedness of structure, E*n*-1 and E*n*-2 also mimic each other in physicochemical properties distinctive of cold-adapted proteins isolated from a variety of psychrophilic organisms, including Antarctic bacteria and archea [24–26] and notothenioid fishes [27]. One of these properties is related to the content and distribution of the hydrophobic amino acids. In both E*n*-1 and E*n*-2, they represent a component which is markedly reduced with respect to the complex of *E. raikovi* pheromones of temperate sea waters (i.e., 33% versus 42.8%), and are largely distributed within the C-terminal half of the sequence. In *E. raikovi* pheromones, their distribution is rather uniform throughout the whole sequence [8,9]. A second property is solvent accessibility, which is predicted to be largely favoured in both E*n*-1 and E*n*-2. The major accessible surface appears to be provided by the N-terminal region. Indeed, half of the 12 hydrophilic residues of the E*n*-1 sequence lie concentrated in the initial segment spanning from $Asn_1$ to $Asn_{17}$; analogously in E*n*-2, the $Asp_1$–$Asp_{22}$ segment has 9 of the total of 16 hydrophilic residues. Two other less extended areas of solvent accessibility would involve an internal sequence segment containing a repetition of Thr residues in E*n*-1 ($T_{21}TCTT_{25}$) and of Gly residues in E*n*-2 ($G_{26}GTGN_{30}$), as well as the C-terminal tail having in both molecules a complex of four hydrophilic residues.

In cold-active proteins, a reduced hydrophobicity and an enhanced solvent accessibility, along with a weakening of intramolecular stabilizing forces, have commonly been correlated with increased backbone flexibility necessary to counterbalance the lower thermal energy of the environment [28–32]. This correlation apparently holds true also in E*n*-1 and E*n*-2. Of two principal sites of backbone flexibility predicted in these molecules [21], one would in fact involve the N-terminal region, where minimal levels of hydrophobicity overlap with maximal levels of solvent accessibility. The other site would involve part of the Thr-rich motif in E*n*-1, and the Gly cluster in E*n*-2. While there is experimental support for the notion that psychrophilic proteins may utilize Gly clustering to enhance their backbone flexibility in proximity to, or in coincidence with functional domains [33,34], evidence for a possible association of Thr repetitions with local flexibility has only been obtained from mutational analysis and NMR spectrometry of the LIM domain of the quail cysteine and glycine-rich protein CRP2 [35]. However, the fact that in E*n*-1 the Thr-rich segment appears to exactly replace the Gly-rich segment of E*n*-2 suggests that these structures may be functionally convergent to generate local flexibility.

Further insight on the conformational flexibility of E*n*-1 and E*n*-2 is suggested by a comparison of these molecules with the *E. raikovi* pheromones, particularly relative to the amino acid composition of the C-terminal sequence. In E*n*-1, this sequence contains no Pro residues and two Ala, while in E*n*-2, it contains one Pro and three Ala. The opposite situation distinguishes the *E. raikovi* pheromones (except E*r*-23), whose terminal loop conservatively contains two or three Pro residues, and no Ala residue [8,9,36]. This allotypic substitution between Pro and Ala residues has a close counterpart in the lactate dehydrogenase $A_4$ of Antarctic notothenioid fishes [27], and may represent another case of psychrophilic proteins overcoming steric limitations imposed by rigid Pro residues on functional domains by utilizing Ala residues that do not impose rotational constraints.

## Acknowledgements

The Italian National Program of Antarctic Research (PNRA) provided financial support to this work.

## References

[1] A. Vallesi, G. Giuli, R.A. Bradshaw, P. Luporini, Nature 376 (1995) 522–524.
[2] C. Ortenzi, C. Alimenti, A. Vallesi, B. Di Pretoro, A. La Terza, P. Luporini, Mol. Biol. Cell 11 (2000) 1445–1455.
[3] P. Luporini, A. Vallesi, C. Miceli, R.A. Bradshaw, J. Eukaryot. Microbiol. 42 (1995) 208–212.
[4] A. Miyake, in: K. Hausmann, P.C. Bradbury (Eds.), Ciliates—Cells as Organisms, Gustav Fischer, Stuggart, 1996, pp. 243–290.
[5] A. Miyake, in: D.H. O'Day, P.A. Horgen (Eds.), Sexual Interactions in Eukaryotic Microbes, Academic Press, New York, 1981, pp. 95–129.
[6] S. Mayumi, T. Harumoto, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 14446–14451.
[7] S. Afon'kin, A.L. Yudin, Acta Protozool. 26 (1987) 91–100.
[8] S. Raffioni, C. Miceli, A. Vallesi, S.K. Chowdhury, B.T. Chait, P. Luporini, R.A. Bradshaw, Proc. Natl. Acad. Sci. U. S. A. 89 (1992) 2071–2075.
[9] P. Luginbühl, M. Ottiger, S. Mronga, K. Wüthrich, Protein Sci. 3 (1994) 1537–1546.
[10] R. Zahn, F. Damberger, C. Ortenzi, P. Luporini, K. Wüthrich, J. Mol. Biol. 313 (2001) 923–931.
[11] C. Brünen-Nieweler, J.C. Weiligmann, B. Hansen, H.W. Kuhlmann, M. Möllenbeck, K. Heckmann, Eur. J. Protistol. 34 (1998) 124–132.
[12] M. Möllenbeck, K. Heckmann, Eur. J. Protistol. 35 (1999) 225–230.
[13] A. Felici, C. Alimenti, C. Ortenzi, C.P. Luporini, Ital. J. Zool. 66 (1999) 355–360.
[14] P. Luporini, S. Raffioni, A. Concetti, C. Miceli, Proc. Natl. Acad. Sci. U. S. A. 83 (1986) 2889–2893.
[15] C. Alimenti, C. Ortenzi, V. Carratore, P. Luporini, FEBS Lett. 514 (2002) 329–332.
[16] C. Ortenzi, P. Luporini, J. Eukaryot. Microbiol. 42 (1995) 242–248.
[17] A. Miyake, J. Beyer, Exp. Cell Res. 76 (1973) 15–24.
[18] F. Jeammougin, J.D. Thompson, M. Gouy, D.G. Higgins, T.J. Gibson, Trends Biochem. Sci. 23 (1998) 403–405.
[19] J. Kyte, R.F. Doolittle, J. Mol. Biol. 157 (1982) 105–132.
[20] J. Janin, Nature 277 (1979) 491–492.
[21] R. Bhaskaran, P.K. Ponnuswanny, Int. J. Pept. Protein Res. 32 (1988) 242–255.

[22] A.E. Stewart, S. Raffioni, T. Chaudhary, B.T. Chait, P. Luporini, R.A. Bradshaw, Protein Sci. 1 (1992) 777–785.
[23] L.R. Brown, S. Mronga, R.A. Bradshaw, C. Ortenzi, P. Luporini, K. Wüthrich, J. Mol. Biol. 231 (1993) 800–816.
[24] T. Thomas, R. Cavicchioli, FEBS Lett. 439 (1998) 281–286.
[25] A. Galkin, L. Kulakova, H. Ashida, Y. Sawa, N. Esaki, Appl. Environ. Microbiol. 65 (1999) 4014–4020.
[26] D. Georlette, Z.O. Jónsson, F. Van Petegem, J.P. Chessa, J. Van Beeumen, U. Hübscher, C. Gerday, Eur. J. Biochem. 267 (2000) 3502–3512.
[27] P.A. Fields, G.N. Somero, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 11476–11481.
[28] S. Davail, G. Feller, E. Narinx, C. Gerday, J. Biol. Chem. 269 (1994) 17448–17453.
[29] G. Feller, C. Gerday, Cell. Mol. Life Sci. 53 (1997) 830–841.
[30] C. Gerday, M. Aittaleb, J.L. Arpigny, E. Baise, J.P. Chessa, G. Garsoux, I. Petrescu, G. Feller, Biochim. Biophys. Acta 1342 (1997) 119–131.
[31] C.J. Marshall, Trends Biotechnol. 15 (1997) 359–364.
[32] S. D'Amico, C. Gerday, G. Feller, J. Biol. Chem. 276 (2001) 25791–25796.
[33] G. Feller, J.L. Arpigny, E. Narinx, C. Gerday, Comp. Biochem. Physiol., A 118 (1997) 495–499.
[34] H. Kumeta, T. Hoshino, T. Goda, T. Okayama, T. Shimada, S. Ohgiya, H. Matsuyama, K. Ishizaki, Biosci. Biotechnol. Biochem. 63 (1999) 1165–1170.
[35] K. Kloiber, R. Weiskirchen, B. Kräutler, K. Bister, R. Konrat, J. Mol. Biol. 293 (1999) 893–908.
[36] G. Di Giuseppe, C. Miceli, R. Zahn, F. Damberger, K. Wüthrich, P. Luporini, J. Eukaryot. Microbiol. 49 (2002) 86–92.