Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 3016–3020, May 1989
Biochemistry

Isolation and structural characterization of cDNA clones encoding the mating pheromone Er-1 secreted by the ciliate *Euplotes raikovi*

(protozoa/cell–cell recognition/mating type locus/codon usage/cDNA cloning)

CRISTINA MICELI*, ANTONIETTA LA TERZA*, AND MARIALUISA MELLI†

*Department of Cell Biology, University of Camerino, 62032 Camerino (MC), Italy; and †Sclavo Research Center, 53100 Siena, Italy

Communicated by John R. Preer, Jr., December 16, 1988

ABSTRACT **cDNA clones comprising the entire coding region for the mating pheromone Er-1 of *Euplotes raikovi* have been isolated by oligonucleotide screening of two cDNA libraries in the vectors λgt10 and pUC12. The cDNA sequence contains an open reading frame of 75 amino acids that constitute pre-pro-Er-1. The amino acid sequence of secreted Er-1 starts at aspartic acid-36 of pre-pro-Er-1 and completely matches that known by direct Er-1 protein sequencing. The coding region of Er-1 cDNA ends with codon TAA, which specifies glutamine in other ciliates. The 5′- and 3′-noncoding regions contain, respectively, two and one inverted repeats. The 3′-noncoding-region inverted repeat, which includes the unusual polyadenylylation signal AACAAA, has been related to RNA 3′-terminus formation.**

Since Sonneborn's discovery of mating types in *Paramecium aurelia* (1), numerous species of ciliates have been found capable of controlling cell–cell recognition phenomena and cell transformation for mating by complexes of genetically determined mating types (for reviews, see refs. 2–6). However, the chemical substances determining mating types have been purified to homogeneity and partially characterized in only a few species (for reviews, see refs. 6 and 7; also see refs. 8–11), and nothing is yet known about their genetic regulation at the molecular level.

In *Euplotes raikovi*, four of these substances—referred to as mating pheromones or euplomones *r* and abbreviated as Er (12)—have been shown to be small soluble, acidic proteins, biologically active at concentrations of $\approx 10^{-12}$ M (11, 12). According to data from Mendelian genetic analysis, their genetic control is carried out, as in *Euplotes patella* (13, 14) and *Euplotes octocarinatus* (15), by a series of codominant multiple alleles at the highly polymorphic mating type (*mat*) locus (7, 16). Each mating pheromone segregates in one-to-one correspondence with one *mat* allele and confers one specific molecular phenotype (mating type) to one cell class (16).

Here we describe the molecular cloning and sequencing of cDNAs encoding the mating pheromone Er-1, which is secreted by *E. raikovi* of mating type I, homozygous for the allele *mat-1*.‡ The cDNA sequence specifies an Er-1 precursor of 75 amino acids, including the 40-amino acid sequence of mature Er-1 near the carboxyl terminus (17).

Knowledge of the molecular structure and genetic control of *E. raikovi* mating pheromones will broaden the concept of mating type function in ciliates and at the same time yield opportunities for seeking molecular details of cell–cell recognition phenomena in unicellular eukaryotes.

MATERIALS AND METHODS

Cells and Culture Conditions. *E. raikovi* of clone $1aF_113$ homozygous for the allele *mat-1* and secreting the mating pheromone Er-1 was the source of RNA and DNA. Cells were maintained as described (16). For RNA extraction logarithmically growing cells were harvested, resuspended for 4 hr in starvation medium, and pelleted by centrifugation at 800 × *g* for 3 min.

RNA Extraction. Total RNA was extracted by the LiCl method (18). Usually 2 mg of RNA was obtained from 10^8 cells. Poly(A)$^+$ RNA, selected by chromatography on oligo(dT)-cellulose (Boehringer Mannheim), was usually ≈1.2% of the total RNA.

Purification of Macronuclei and DNA Extraction. Macronuclei were purified as described in *Oxytricha* by Swanton *et al.* (19). They were suspended overnight at 50°C in lysing solution (0.5 M EDTA/1% SDS/10 mM Tris·HCl at pH 9.5) containing proteinase K at 0.2 mg/ml (Boehringer Mannheim). DNA was purified by several phenol and chloroform extractions, treated for 2 hr with RNase A (50 μg/ml; Boehringer Mannheim), and dialyzed before being loaded on agarose gels.

Oligonucleotide Synthesis and Labeling. Oligonucleotides of 29 and 36 bases were synthesized on an Applied Biosystem DNA synthesizer (model 308A) with procedures and reagents recommended by the manufacturers. Oligonucleotides were purified by preparative electrophoresis on 8 M urea/20% polyacrylamide gels. Oligonucleotide bands were visualized by direct UV light, excised, and electroeluted on an ISCO apparatus with dialysis membranes of 3500 Da cut-off. Oligonucleotides were radiolabeled with [α-^{32}P]ATP and T4 polynucleotide kinase (Boehringer Mannheim), according to standard procedures.

cDNA Synthesis and Library Construction. The first strand of cDNA was prepared from poly(A)$^+$ RNA (10 μg) with Moloney murine leukemia virus reverse transcriptase (Bethesda Research Laboratories) as suggested by the supplier. The second strand (2 μg) was synthesized according to Gubler and Hoffmann (20). T4 DNA polymerase (Amersham) was then used to fill in overhanging ends. This blunt-ended double-stranded cDNA was then divided into two samples. One sample (1.5 μg) was methylated with *Eco*RI methylase (Amersham) and ligated to phosphorylated *Eco*RI linkers (Amersham). Linker excess was removed by *Eco*RI digestion, followed by 5% polyacrylamide gel electrophoresis and electroelution of molecules >300 base pairs (bp). cDNA with *Eco*RI-linker ends was ligated to *Eco*RI-digested and phosphatase-treated λgt10 arms. This recombinant DNA was packaged with the Promega Biotec packaging system, using ≈0.5 μg of DNA per packaging extract. The other cDNA sample (500 ng) was ligated into the vector pUC12, previously digested at the *Sma* I site and treated with phosphatase. After ligation, the *Escherichia coli* strain JM83 was transformed with the $CaCl_2$ procedure (21).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Er-1, *Euplotes raikovi* mating pheromone secreted by cells of mating type I; nt, nucleotide(s).

‡The sequence reported in this paper is being deposited in the EMBL/GenBank data base (accession no. J04141).

Screening of Libraries. The libraries in the vectors λgt10 and pUC12 were screened by using labeled oligonucleotides by plaque and colony hybridization, respectively. Plaque or colony filters, prepared according to Maniatis *et al.* (21), were prehybridized 4 hr at 65°C in 5× Denhardt's solution (1× Denhardt's solution is 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.02% bovine serum albumin)/10 mM EDTA/0.5% SDS, pH 7.5. Hybridizations were done in 5× Denhardt's solution/6× NET (20× NET is 3 M NaCl/20 mM EDTA/0.3 M Tris·HCl, pH 8)/0.5% SDS containing yeast tRNA at 100 μg/ml. Filters were then washed twice in 100 ml of 0.5% SDS in 4× NET.

Preliminary experiments to optimize temperatures of filter hybridization and washing were done. Total and poly(A)$^+$ RNA were spotted or electrophoresed and blotted on Hybond N membrane (Amersham). The same blots were hybridized with every oligonucleotide and washed several times under increasingly stringent conditions. Optimal temperatures for 29-base and 36-base oligonucleotides were 42°C and 50°C for hybridization and 50°C and 55°C for washing, respectively.

Analysis of RNA Blot and Southern Blot Hybridizations. For RNA blotting, samples of total and poly(A)$^+$ RNA were loaded on a 1.5% agarose gel electrophoresis containing formaldehyde. For Southern blotting, samples of undigested macronuclear DNA were loaded on 1% agarose gel. RNA and DNA gels were then blotted on Hybond N membrane (Amersham) as specified by the supplier. Prehybridization, hybridization, and washing were done according to Maniatis *et al.* (21), with radiolabeled cDNA inserts as probes; this radiolabeling was according to the method of Feinberg and Vogelstein (22).

Subcloning and DNA Sequencing. Minipreparations of phage and plasmid DNA from clones of λgt10 and pUC12 libraries were prepared as described (21, 23). These DNAs were then digested with the required enzyme (according to the supplier) and run on 1% agarose gel. Digested fragments were excised from the gel, purified by electroelution, and subcloned in M13mp8 or M13mp9; pUC12 vectors were digested as required. Sequencing reactions of M13 and pUC subclones were performed by the methods of Sanger *et al.* (24) and Chen and Seeburg (25), respectively. Sequenase (United States Biochemical) and pUC sequencing (Boehringer Mannheim) kits were used.

RESULTS

Strategy for Designing E*r*1-Specific Oligonucleotide Probes. The approach for isolating cloned DNA sequences specific for E*r*-1 involved the synthesis of a set of oligonucleotides as probes, complementary to some possible coding sequences for portions of E*r*-1. The oligonucleotides were designed for two E*r*-1 stretches, one at the amino terminus and the other at the carboxyl terminus, the amino acid sequences of which were known (ref. 26 and personal communication by S. Raffioni, S. S. Disper, and R. A. Bradshaw). To minimize the problem of the redundancy of the genetic code, we applied the codons most commonly used in the few genes so far sequenced in ciliates.

Table 1 shows that we synthesized a specific 36-mer oligonucleotide (E*r*1-36) for the E*r*-1 amino-terminus peptide and a set of four 29-mer oligonucleotides for the E*r*-1 carboxyl-terminus peptide, using, alternatively, the codons AUU and AUC for isoleucine and UCA and UCG for serine. The four 29-mer oligonucleotides were then separately tested on a poly(A)$^+$ RNA blot to identify which one could produce one or more defined bands at the highest stringency. The 29-mer (E*r*1-29c) containing AUU for isoleucine and UCG for serine appeared to be the most appropriate oligonucleotide, as it produced one distinct band (data not shown). The 29-mer was therefore used as probe, together with E*r*1-36, for screening recombinant clones.

Isolation and Characterization of E*r*-1 cDNA Clones. Total RNA was extracted from early stationary-phase cells and used to recover poly(A)$^+$ RNA by oligo-(dT) chromatography for the preparation of double-stranded cDNA. After *Eco*RI-linker addition, a sample of double-stranded cDNA was inserted in the *Eco*RI site of λgt10 arms and, after phage packaging, used to infect the C-600 hf1 *E. coli* strain. The constructed library contained 1.3×10^6 independent recombinant plaques in a background of 40% nonrecombinant plaques.

Another sample of blunt-ended, double-stranded cDNA was directly cloned in the *Sma* I site of the vector pUC12. It produced 5×10^3 recombinant colonies.

Screening of all plaques and colonies by both oligonucleotides E*r*1-36 and E*r*1-29c revealed five positive recombinant clones, which were shown to contain cross-hybridizing inserts by dot-blot analysis. These inserts, 300–500 bp in length, consisted of overlapping segments as shown by restriction analysis.

Table 1. Synthetic oligonucleotides for the isolation of E*r*-1 cDNA

Amino-terminus peptide sequence			N–Asp–	Ala–	Cys–	Glu–	Gln–	Ala–	Ala–	Ile–	Gln–	Cys–	Val–	Glu–
Possible codons	5'		GAU	GCU	UGU	GAA	CAA	GCU	GCU	AUU	CAA	UGU	GUU	GAA 3'
			C	C	C	G	G	C	C	C	G	C	C	G
			A				A	A	A			A		
			G				G	G	G			G		
Oligonucleotide E*r*1-36	3'		CTG	CGG	ACG	CTT	GTT	CGG	CGG	TAG	GTT	ACG	CAG	CTT 5'
Carboxyl-terminus peptide sequence			–Met–	Tyr–	Ile–	Tyr–	Ser–	Asn–	Cys–	Pro–	Pro–	Tyr–	Val–	C
Possible codons	5'		AUG	UAU	AUU	UAU	UCU	AAU	UGU	CCU	CCU	UAU	GUU	3'
				C	C	C	C	C	C	C	C	C	C	
					A		A			A	A			
							G			G	G			
							AGU							
							C							
Oligonucleotide														
E*r*1-29a	3'		TAC	ATG	TAG	ATG	AGT	TTG	ACG	GGT	GGT	AT	5'	
E*r*1-29b	3'		TAC	ATG	TAA	ATG	AGT	TTG	ACG	GGT	GGT	AT	5'	
E*r*1-29c	3'		TAC	ATG	TAA	ATG	AGC	TTG	ACG	GGT	GGT	AT	5'	
E*r*1-29d	3'		TAC	ATG	TAG	ATG	AGC	TTG	ACG	GGT	GGT	AT	5'	

Triplets changed in the E*r*1-29 set are boxed.

Two of the five positive clones—namely, clone λ4 from the λgt10 library and clone p3/5b from the pUC library, contained the entire coding region of E*r*-1. Their restriction maps and the sequencing strategy are illustrated in Fig. 1. The other three positive clones contained only portions of the inserts of λ4 and p3/5b clones.

The inserts of λ4 and p3/5b clones were sequenced in both strands, using subcloned fragments and E*r*1-29c oligonucleotide as well as the standard sequencing primers. Their sequences allowed reconstruction of the full-length sequence of E*r*-1 cDNA which, as illustrated in Fig. 2*a*, proved to correspond to a band of ≈680 bases by RNA blot analysis. The same RNA blot also revealed a weaker band of ≈1200 bases corresponding to a mRNA of interest (C.M. and A.L.T., unpublished data).

The two bands evidenced by RNA blot were completely destroyed by RNase A treatment but were not affected by DNase I treatment (Fig. 2*b*).

A hybridization pattern consisting of two bands of different intensity was also produced by Southern blot analysis of undigested macronuclear DNA, which was performed using clones λ4 and p3/5b as probes (Fig. 3). To interpret this result we must consider that macronuclear DNA of hypotrich ciliates is organized in linear "gene-sized" molecules that represent independent transcribing units (27). The stronger band at 1100 bp probably corresponds to the macronuclear DNA molecule that transcribes E*r*-1 mRNA. This DNA molecule is ≈450 bases longer than E*r*-1 mRNA—that is, the number of bases expected to include the sequences flanking the transcribing region of every macronuclear DNA molecule (27). The other band at 1300 bp seems to correspond to the DNA molecule that transcribes the mRNA of 1200 bases evidenced by RNA blot analysis.

E*r*-1 cDNA Sequence. The full-length E*r*-1 cDNA sequence is shown in Fig. 4. This sequence includes (*i*) a 225-nucleotide (nt) coding region containing an open reading frame capable of encoding 75 amino acids, (*ii*) a 5′-noncoding region extending for 258 nt, and (*iii*) a 3′-noncoding region consisting of 105 nt ending with the poly(A) tail.

The hydropathy profile deduced from the E*r*-1 cDNA coding region, predicted by the method of Hopp and Woods (28), is shown in Fig. 5. The first 19 amino acids constitute a hydrophobic domain. Starting from amino acid 20 the profile

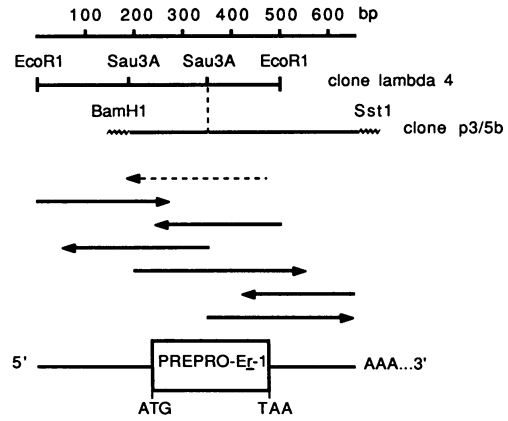

FIG. 1. Restriction maps of λ4 and p3/5b clones and strategy for determining their nucleotide sequences. The two lines below the bp scale represent the restriction maps of λ4 and p3/5b clones. Only the restriction sites used for subcloning and sequencing are reported. Wavy lines in the clone p3/5b map stand for the pUC12 regions flanking the insert. Direction and extent of nucleotide-sequence determination are indicated by arrows. Broken and solid arrows refer to E*r*1-29 oligonucleotide and pUC oligonucleotides as sequencing primers, respectively. The reconstructed full-length E*r*-1 cDNA sequence is diagrammed at the bottom; the coding region is boxed.

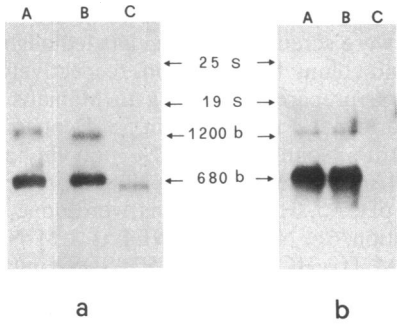

FIG. 2. Blot-hybridization analysis of RNA samples obtained from mating type I *E. raikovi* and fractionated on 1.5% agarose gels containing formaldehyde. (*a*) A sample (2 μg) of poly(A)$^+$ RNA was hybridized with the labeled insert of clone p3/5b (lane A) and of clone λ4 (lane B); a sample (20 μg) of total RNA was hybridized with clone λ4 (lane C). (*B*) Samples (2 μg) of poly(A)$^+$ RNA were treated with DNase I (1 μg), RNase A (1 μg), or untreated and loaded on the gel (lanes B, C, and A, respectively). Hybridization was performed with the labeled insert of clone λ4. An identical hybridization pattern was obtained using the labeled insert of clone p3/5b (data not shown). The position of markers, 25S and 19S rRNA, and the size of hybridizing bands, reported in bases (*b*), are indicated by arrows.

shifts to a hydrophilic domain extending for 20 amino acids. An even more hydrophilic domain involves amino acids 55–62.

The translation initiation site in the E*r*-1 cDNA sequence was assumed to be the ATG at the +1 nt of Fig. 4. The sequence upstream from this ATG codon shows an adenosine residue in position −3. Thus it falls into the group of known eukaryotic mRNA sequences, which have a purine in position −3 but otherwise differ completely from the CCA/GCCATG consensus sequence proposed by Kozak (29) to function in eukaryotes as translation initiation site.

Two other ATG codons, corresponding to nt −72 and −87 are unlikely to function as initiation codons. Both of them, in fact, are followed by only short open reading frames.

The amino acid codon usage of the E*r*-1 cDNA appears to diverge from the standard eukaryotic codon usage. When possible, codons ending with adenosine are preferred and codons ending with guanosine are avoided. Consistently, AGA is the only one of the potential six arginine codons to be used. Both the 5′- and the 3′-noncoding regions are rich in adenosine and thymine nucleotides (67% and 72%, respectively, versus 57% of the coding region) and contain three inverted repeats that may form hairpin structures. Two of

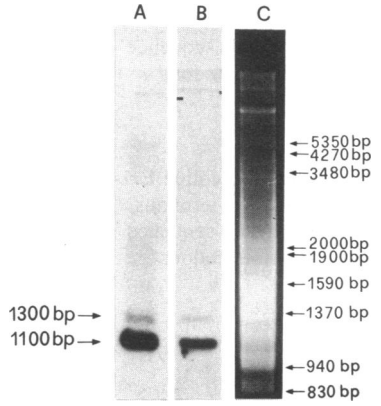

FIG. 3. Southern-blot analysis of undigested macronuclear DNA obtained from mating type I *E. raikovi* and hybridized with the labeled insert of clones λ4 (lane A) and p3/5b (lane B). Lane C was stained with ethidium bromide before blotting. The position of markers and the size of hybridizing bands are indicated at the right and left, respectively.

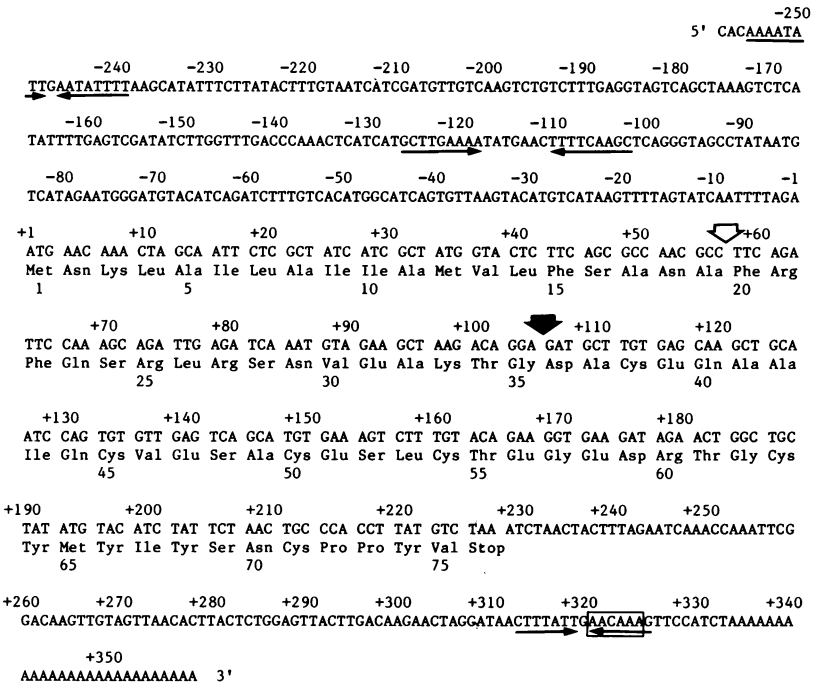

FIG. 4. Er-1 cDNA sequence showing the predicted amino acid sequence of pre-pro-Er-1. Numbering above the nucleotide sequence indicates nucleotide residues and is shown as + or − relative to the translation initiation codon. Numbering below the amino acid sequence indicates amino acid residues. The vertical open and solid arrowheads indicate the presumptive signal peptide cleavage site and the amino-terminus of the secreted Er-1, respectively. Inverted repeats are underlined by arrows. The putative polyadenylylation signal is boxed.

these inverted repeats are located in the 5'-noncoding region and involve the sequence from nt −255 to −239 and from −126 to −103. The other inverted repeat is located close to the 3' end, from nt +312 to +330, just upstream from the poly(A) tail.

The typical polyadenylylation signal AATAAA is not present in the Er-1 cDNA sequence; the corresponding position is occupied by the sequence AACAAA.

DISCUSSION

We have determined the sequence of the cDNA for the mating pheromone Er-1 of *E. raikovi*. The sequence fits in size with the poly(A)$^+$ Er-1 mRNA of 680 bases, which appears to be transcribed by a macronuclear DNA molecule of 1100 bp.

The Er-1 cDNA sequence contains an open reading frame encoding a 75-amino acid polypeptide chain. The portion from Asp-36 to Val-75 constitutes the secreted Er-1. This sequence completely matches the sequence determined before this reported cDNA sequence by chemical analysis on Er-1 samples purified from cell supernatant (17).

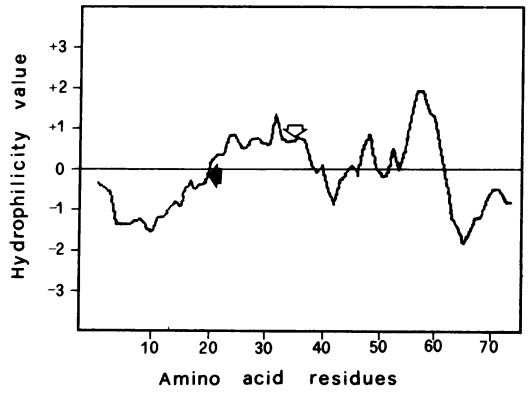

FIG. 5. Hydropathy profile of pre-pro-Er-1 predicted by the method of Hopp and Woods (28), using six residues as window size. The cleavage site of the presumptive signal peptide of pre-pro-Er-1 and the amino terminus of mature Er-1 are marked by solid and open arrowheads, respectively.

The sequence upstream from Asp-36 most likely includes a signal peptide, which might serve to initiate Er-1 export across the membrane of the rough endoplasmic reticulum. Support for this assumption derives from the analysis of the hydropathy profile of the predicted amino acid sequence that shows a hydrophobic domain involving amino acids 1–19. In this domain, three distinct regions typical of signal peptides of secretory proteins are identifiable according to criteria established by von Heijne (30): (*i*) one short and basic region containing lysine (amino acids 1–3); (*ii*) one central and strongly hydrophobic region (amino acids 4–15); and (*iii*) one short and polar region containing asparagine and serine (amino acids 16–19).

The cleavage of the signal peptide by a specific peptidase appears to be a general prerequisite for protein transfer into the endoplasmic reticulum. On the basis of the (−3, −1) rule of von Heijne (30), the signal peptide cleavage site in the Er-1 nascent polypeptide chain is predicted to reside in the Ala-Asn-Ala motif (amino acids 17–19). No other region appears a good candidate for a cleavage site. The motif Lys-Thr-Gly, which immediately precedes the amino terminus of the secreted Er-1, is unlikely to be a cleavage site due to the presence of a lysine residue.

In conclusion, we assume that Er-1 mRNA is translated as pre-pro-Er-1 of 75 amino acids. After cleavage of the secretory signal peptide of 19 amino acids, pre-pro-Er-1 is exported into the rough endoplasmic reticulum as pro-Er-1 of 56 amino acids.

Pro-Er-1 has not been detected, at least in relevant amounts, in the cell supernatant from which only the mature form of 40 amino acids was isolated (9, 17). Therefore, as also occurs in secretion of polypeptide hormones and growth factors of higher eukaryotes (31), pro-Er-1 appears to be processed proteolytically before being secreted.

A noteworthy feature of the Er-1 cDNA sequence is the use of TAA as stop codon. In all ciliate genes sequenced so far, only one termination codon has been identified—TGA. TAA, as well as TAG, have been found to specify glutamine in genes of *Paramecium* (32, 33), *Tetrahymena* (34), and *Stylonychia* (35, 36). In addition, glutamine tRNAs recognizing TAA and TAG were purified in *Tetrahymena* (37).

On the other hand, it is certain that TAA functions as a termination codon in the Er-1 cDNA: one TAA codon is

located adjacent to the carboxyl terminus of the amino acid sequence of the secreted Er-1, and downstream from it there are two other TAA codons and no TAG or TGA before the poly(A) tail. This in-frame repetition of stop codons is not unique to *Euplotes*. It also occurs in immobilizing antigen genes of *Paramecium* (38) and in tubulin genes of *Stylonychia* (36). It might therefore constitute a distinctive trait of certain ciliate genes.

The deviation of ciliates from the universal genetic code deserves discussion about the position of this group in the evolution of eukaryotes (for review, see refs. 34 and 39). This deviation has been supposed to be a consequence of the antiquity of ciliates, which have been thought an evolutionary line branching off from the eukaryotic lineage before the genetic code was fixed (34).

Our finding that *Euplotes* uses TAA as a stop codon weakens this hypothesis. Instead, our results suggest that the eccentric genetic code of some ciliates is a derived and not a primitive character. In accordance with information most recently obtained by Baroin *et al.* (40) from comparisons among 28S rRNA sequences of protists, the eccentricity of the genetic code of some ciliates might represent a secondary event. As first suggested by Hanyu *et al.* (37), this event might be from the appearance of mutated glutamine tRNAs recognizing TAA and TAG codons that have been preserved for protection against point mutations, giving rise to TAA and TAG triplets within coding regions. These triplets, being recognized by glutamine tRNAs, should not have the deleterious effect of ending translation in the middle of a coding region.

Both noncoding regions of the Er-1 cDNA sequence show interesting structural features. The 5′-noncoding region contains two inverted repeats, very rich in adenosine and thymidine nucleotides. They might represent DNA replication origins. Consistently, nonconserved inverted repeats of *Stylonychia* macronuclear DNA molecules, similarly rich in adenosine and thymidine nucleotides and located immediately upstream of the translation initiation site, have been found capable of initiating autonomous DNA replication in a heterologous mouse L-cell system (E. Helftenbein, personal communication).

In the 3′-noncoding region, the sequence AACAAA appears inserted in one inverted repeat 10 bases upstream of the poly(A) tail. This complex structure might be involved in RNA 3′-terminus formation. Strong support for this assumption derives from the following considerations. (*i*) As occurs in mRNAs of human apolipoprotein B48 (41) and of platelet-derived growth factor A chain (42), the sequence AACAAA may replace the canonical polyadenylylation signal AATAAA in determining the addition of the poly(A) tail to the Er-1 mRNA. (*ii*) Nonconserved inverted repeats are always present at the RNA 3′ end of histone genes of high eukaryotes (43), plastid genes (44), and *Stylonychia* tubulin genes (35); however, in none of these genes has polyadenylylation been observed. To verify our assumption, we need to compare the Er-1 cDNA sequence with sequences of other *Euplotes* genes.

This study was begun by C.M. during sabbatical leave at the University of California, Santa Barbara, in the laboratory of Dr. E. Orias, whom we wish to thank for advice in approaching this research and for providing facilities for synthesizing the 36-mer oligonucleotide; we also thank Dr. S. Zhang for his generous help during that period. We express our gratitude to Dr. G. Ratti and Mr. S. Ricci for synthesizing the 29-mer oligonucleotides and to Dr. E. Palla for helpful suggestions about cDNA synthesis. We also thank Dr. E. Helftenbein for communicating results before publication, and we are also particularly indebted to Drs. S. Raffioni, S. S. Disper, and R. A. Bradshaw for communicating to us the amino acid sequence of the mature form of Er-1 before publication. Special gratitude is addressed to Dr. P. Luporini for his stimulating discussions during the work and his constructive criticisms of the manuscript. Finally, we thank Drs. G. H. Beale and D. L. Nanney for helpful comments on the manuscript. Financial support was mainly provided by a grant of the Italian "Ministero della Pubblica Istruzione" to P. Luporini.

1. Sonneborn, T. M. (1937) *Proc. Natl. Acad. Sci. USA* **23**, 378–395.
2. Preer, J. R., Jr. (1969) in *Research in Protozoology*, ed. Chen, T. (Pergamon, Oxford), Vol. 3, pp. 130–278.
3. Nanney, D. L. (1980) *Experimental Ciliatology* (Wiley, New York).
4. Vivier, E. (1984) in *Traité de Zoologie*, ed. Grassé, P. P. (Masson, Paris), Vol. 2, pp. 621–760.
5. Nobili, R., Luporini, P. & Esposito, F. (1986) in *Invertebrate Models: Cell Receptors and Cell Communication*, ed. Greenberg, A. H. (Karger, Basel), pp. 6–28.
6. Miyake, A. (1981) in *Biochemistry and Physiology of Protozoa*, eds. Levandowsky, M. & Hutner, S. H. (Academic, New York), Vol. 4, pp. 125–198.
7. Luporini, P. & Miceli, C. (1986) in *The Molecular Biology of Ciliated Protozoa*, ed. Gall, J. G. (Academic, New York), pp. 263–299.
8. Akada, R. (1986) *J. Exp. Zool.* **237**, 287–290.
9. Concetti, A., Raffioni, S., Miceli, C., Barra, D. & Luporini, P. (1986) *J. Biol. Chem.* **261**, 10582–10586.
10. Schulze-Dieckhoff, H., Freiburg, M. & Heckmann, K. (1987) *Eur. J. Biochem.* **168**, 89–94.
11. Raffioni, S., Miceli, C., Concetti, A., Barra, D. & Luporini, P. (1987) *Exp. Cell Res.* **172**, 417–424.
12. Miceli, C., Concetti, A. & Luporini, P. (1983) *Exp. Cell Res.* **149**, 593–598.
13. Kimball, R. F. (1942) *Genetics* **27**, 269–285.
14. Akada, R. (1985) *Genet. Res.* **46**, 125–132.
15. Heckmann, K. & Kuhlmann, H. W. (1986) *J. Exp. Zool.* **237**, 87–96.
16. Luporini, P., Raffioni, S., Concetti, A. & Miceli, C. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2889–2893.
17. Raffioni, S., Luporini, P., Chait, B. T., Disper, S. S. & Bradshaw, R. A. (1988) *J. Biol. Chem.* **263**, 18152–18159.
18. Cathala, G., Savouret, J. F., Mendez, B., West, B. L., Karin, M., Martial, J. A. & Baxter, J. D. (1983) *DNA* **2**, 329–335.
19. Swanton, M. T., Greslin, A. F. & Prescott, D. M. (1980) *Chromosoma* **77**, 203–215.
20. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25**, 263–269.
21. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
22. Feinberg, A. P. & Vogelstein, B. (1984) *Anal. Biochem.* **137**, 266–270.
23. Perbal, B. (1984) *A Practical Guide to Molecular Cloning* (Wiley, New York).
24. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
25. Chen, E. L. & Seeburg, P. H. (1985) *DNA* **4**, 165–170.
26. Disper, S. S., Bradshaw, R. A., Raffioni, S., Miceli, C. & Luporini, P. (1986) *Eur. J. Cell Biol.* **42**, Suppl. 16, 40.
27. Klobutcher, L. A. & Prescott, D. M. (1986) in *The Molecular Biology of Ciliated Protozoa*, ed. Gall, J. G. (Academic, New York), pp. 111–149.
28. Hopp, T. P. & Woods, K. R. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3824–3828.
29. Kozak, M. (1986) *Cell* **44**, 283–292.
30. von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690.
31. Lynn Burgess, T. & Kelly, R. B. (1987) *Annu. Rev. Cell Biol.* **3**, 243–293.
32. Caron, F. & Meyer, E. (1985) *Nature (London)* **314**, 185–188.
33. Preer, J. R., Jr., Preer, L. B., Rudman, B. M. & Barnett, A. (1985) *Nature (London)* **314**, 188–190.
34. Horowitz, S. & Gorovsky, M. A. (1985) *Proc. Natl. Acad. Sci. USA* **83**, 2452–2455.
35. Helftenbein, E. (1985) *Nucleic Acids Res.* **13**, 415–432.
36. Conzelmann, K. K. & Helftenbein, E. (1987) *J. Mol. Biol.* **198**, 643–653.
37. Hanyu, N., Kuchino, Y., Nichimura, S. & Beier, H. (1986) *EMBO J.* **5**, 1307–1311.
38. Preer, J. R., Jr., Preer, L. B., Rudman, B. M. & Barnett, A. (1987) *J. Protozool.* **34**, 418–423.
39. Caron, F. (1986) *Microbiol. Sci.* **3**, 36–40.
40. Baroin, A., Perasso, R., Qu, L. H., Brugerolle, G., Bachellerie, J. & Adoutte, A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 3474–3478.
41. Powell, L. M., Wallis, S. C., Pease, J. R., Edwards, Y. H., Knott, T. J. & Scott, J. (1987) *Cell* **50**, 831–840.
42. Rorsman, F., Bywater, M., Knott, T. J., Scott, J. & Betsholtz, C. (1988) *Mol. Cell Biol.* **8**, 571–577.
43. Schaufele, F., Gilmartin, G. M., Bannwarth, W. & Birnstiel, M. L. (1986) *Nature (London)* **323**, 777–781.
44. Stern, D. B. & Gruissem, W. (1987) *Cell* **51**, 1145–1157.