THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 263, No. 34, Issue of December 5, pp. 18152-18159, 1988
Printed in U.S.A.

Primary Structure of the Mating Pheromone Er-1 of the Ciliate *Euplotes raikovi**

(Received for publication, June 30, 1988)

Simona Raffioni‡, Pierangelo Luporini§, Brian T. Chait¶, Steven S. Disper‡, and Ralph A. Bradshaw‡‖

From the ‡Department of Biological Chemistry, California College of Medicine, University of California, Irvine, California 92717, the §Department of Cell Biology, University of Camerino, 62032 Camerino (MC), Italy, and the ¶Rockefeller University, New York, New York 10021

The complete amino acid sequence of the mating pheromone Er-1 purified from *Euplotes raikovi* homozygous for *mat-1* was determined by automated Edman degradation of the whole protein and peptides generated by cyanogen bromide, trypsin, *Staphylococcus aureus* V8 protease, and chymotrypsin. The proposed sequence is:

```
1                                   10
Asp-Ala-Cys-Glu-Gln-Ala-Ala-Ile-Gln-Cys-
                                    20
Val-Glu-Ser-Ala-Cys-Glu-Ser-Leu-Cys-Thr-
                                    30
Glu-Gly-Glu-Asp-Arg-Thr-Gly-Cys-Tyr-Met-
                                    40
Tyr-Ile-Tyr-Ser-Asn-Cys-Pro-Pro-Tyr-Val
```

The calculated molecular weight is 4411.0, which is in agreement with the averaged mass of 4410.2 obtained by fission fragment ionization mass spectrometry. Previously reported values of the native molecular weight, determined by gel filtration, have ranged from 9,000 to 12,000. Thus, the native structure is likely a dimer (or larger aggregate) of identical subunits with the three disulfide bonds present occurring as intrachain links. Secondary structure predictions suggest a helical structure at the amino terminus. A comparison of the Er-1 amino acid sequence with known protein sequences did not reveal any significant similarities.

Mating pheromones of ciliates are diffusible extracellular communication signals that distinguish different intraspecific classes of cells, commonly referred to as "mating types." Their most obvious function, although not necessarily their primary one, may be readily observed when cells of a given mating type are mixed with the mating pheromone produced by cells of another mating type. Upon suspension, the cells undergo physiological changes and acquire a temporary competence to unite in mating pairs.

Mating pheromones have been isolated and partially characterized from the heterotrich *Blepharisma japonicum* (1-3), and the hypotrichs *Euplotes raikovi* (4-7) and *Euplotes octocarinatus* (8, 9). In these latter two species of *Euplotes*, it has also been shown that the many different mating pheromones produced—supposedly hundreds in *E. raikovi* (as in other species of *Euplotes* (10, 11))—are proteins controlled by a series of alleles co-dominant at the Mendelian mating type (*mat*)[1] locus (6, 12). Each mating pheromone appears to be inherited in association with one *mat*-allele: thus, the pheromone Er-1 is inherited with the allele *mat*-1, Er-2 with *mat*-2, and so forth. (The designation Er is the abbreviation of "E" for "euplomone," a general term proposed to indicate *Euplotes* mating pheromones, and *r* for *raikovi* (4).)

The mating type mechanism in ciliates is not related to sex. Each one of the often multiple intraspecific mating types is comprised of diploid "hermaphrodite" cells and the mating type mechanism is functionally better compared with devices used by, for example, flowering plants to ensure self-recognition (10, 13-15). This process requires specific pheromone recognition and presumably involves receptor-pheromone interactions at the cell surface.

In this report, we present the complete amino acid sequence of a ciliate mating pheromone, which is secreted by *E. raikovi* of mating type I homozygous for the allele *mat*-1. This pheromone, Er-1, was initially isolated together with the pheromone Er-2 from the supernatant of heterozygous *mat-1/mat-2* cells of the wild type strain 13 (4). Purification to apparent homogeneity of Er-1 was carried out by a three-stage procedure that involved reverse-phase chromatography on Sep-Pak C_{18} cartridges, gel filtration on Sephadex G-50, and ion exchange chromatography on a Mono Q column (5). The homogeneous material was reported (5) to be a protein with a molecular mass of 12,000 daltons and an isoelectric point of 3.7. It was active at concentrations of 10^{-12} M when *mat-1* homozygous clones were obtained as sexual offspring from strain 13 and used as a source of Er-1 (6). The covalent structure determined, which is consistent with mass spectrometer measurements, suggests the native molecule occurs as a dimer (or larger aggregate).

EXPERIMENTAL PROCEDURES

Materials—Er-1 was obtained from the clone $1aF_113$ (*mat-1/mat-1*) of *E. raikovi* and prepared as previously described by Concetti *et al.* (5). Chemicals and reagents were purchased as follows: cyanogen bromide, guanidine hydrochloride, iodoacetic acid (recrystallized before use), *N*-tosyl-L-phenylalanine chloromethyl ketone-treated trypsin, chymotrypsin, and carboxypeptidase Y from Sigma; *Staphylococcus aureus* V8 protease and trifluoroacetic acid from Pierce; formic acid from Mallinkrodt Chemicals; iodo[1-^{14}C]acetic acid, 17.9 mCi/

* This work was supported by United States Public Health Service Research Grants DK32465 (to R. A. B.) and RR00862 (to B. T. C.) and the Italian Ministero della Pubblica Istruzione. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence and reprint requests should be addressed.

[1] The abbreviations used are: *mat*, mating type locus designation; Er-1, euplomone *r*-1; CM-Er-1, *S*-^{14}C-carboxymethylated Er-1; HPLC, high performance liquid chromatography; PTH, phenylthiohydantoin.

mmol, from Du Pont-New England Nuclear; HPLC-grade solvents from Fisher; gas-phase Sequencer chemicals from Applied Biosystems. All other chemicals were of reagent grade.

S-Carboxymethylation—*S*-Carboxymethylation was performed essentially according to Angeletti *et al.* (16). Er-1 was dissolved in 0.5 M Tris and 6 M guanidine hydrochloride at pH 8.0 (1 mg/ml), reduced with 0.01 M dithiothreitol under nitrogen for 3–4 h at 37 °C, and carboxymethylated with a 3-fold excess of iodoacetic acid neutralized with NaOH and containing 0.05 mCi of iodo[1-^{14}C]acetic acid. The reaction was allowed to proceed at room temperature for 5 min and stopped by the addition of 2-mercaptoethanol. The labeled protein was eventually dialyzed and lyophilized.

Protein Cleavage—Cyanogen bromide cleavage of *S*-CM-Er-1 (0.5 mg) was performed in 70% formic acid with a 50-fold molar excess of CNBr over methionine residues. The mixture was incubated under nitrogen at room temperature for 24 h and then diluted with 15 volumes of distilled water prior to lyophilization (17).

The insoluble material remaining after cyanogen bromide cleavage of CM-Er-1 was digested with trypsin (1% w/w) in 0.1 M NH_4HCO_3, 0.1 mM $CaCl_2$, at pH 8.1 (18). After incubation for 10 h at 37 °C, the reaction was terminated by acidification.

CM-Er-1 (0.5 mg) was digested with *S. aureus* V8 protease (2% w/w) in 0.05 M NH_4HCO_3, at pH 7.8 (19). After incubation for 18 h at room temperature, the reaction was terminated by the addition of excess phenylmethylsulfonyl fluoride.

The conditions for the digestion of CM-Er-1 (0.5 mg) with chymotrypsin are those described above for trypsin. After 8 h, the reaction was stopped by acidification with acetic acid.

For the determination of the carboxyl-terminal residues, CM-Er-1 was dissolved in 0.02 M sodium phosphate at pH 6.0 and carboxypeptidase Y was added at a final ratio 1:100 (20). Aliquots were removed at designated times, acidified with 1 N HCl, and centrifuged, and the supernatant was dried prior to derivatization and amino acid analysis.

Peptide Separation—All CM-Er-1 peptides were separated by reverse-phase HPLC on a Hewlett-Packard model 1090 chromatograph. A Vydac C18 column (4.6 × 250 mm) was equilibrated with 5% acetonitrile and 0.1% trifluoroacetic acid. Radioactive peaks were identified on all by liquid scintillation counting of aliquots of the fractions.

Amino Acid Analysis—Samples (0.5 nmol of Er-1 or 10% of peptide fraction from reverse-phase chromatography) were hydrolyzed in 6 N HCl for 22 h, dried, and derivatized using ethanol:triethylamine:water:phenylisothiocyanate in a ratio of 7:1:1:1 (v/v) for 15 min at room temperature. The phenylthiocarbamyl derivatives were analyzed by reverse-phase HPLC on a Hewlett-Packard model 1090 chromatograph (21).

Amino Acid Sequence Determination—Sequence analysis was performed using either an Applied Biosystems 470A gas-phase or 477A pulsed-liquid Sequencer. PTH-derivatives (from the 470A) were identified by reverse-phase HPLC with a Hewlett-Packard model 1084B chromatograph equipped with an Altex Ultrasphere ODS column (4.6 × 250 mm). The aqueous phase was 7.5 mM sodium phosphate, pH 5.5, and the PTH-amino acids were eluted with a gradient of methanol:acetonitrile in 17:3 ratio (v/v). The flow rate was 1.65 ml/min. PTH-cysteine was detected as *S*-[^{14}C]carboxymethylcysteine and confirmed by liquid scintillation counting.

Mass Spectrometry—Er-1 (2 nmol in 20 μl of 0.19% trifluoroacetic acid) was applied to a thin layer of nitrocellulose, prepared by electrospraying 50 μg of nitrocellulose onto an aluminized polyester support with 1 cm^2 of surface area (22). The sample was allowed to adsorb onto the nitrocellulose surface for 3 min, whereupon the surface was washed with 0.1% trifluoroacetic acid. The sample was then thoroughly dried on the nitrocellulose under reduced pressure, and its mass spectrum was obtained using a ^{252}Cf fission fragment ionization time-of-flight mass spectrometer constructed at the Rockefeller University (23). The molecular mass of the sample was determined with an accuracy of better than 1 dalton.

Nomenclature of Peptides—Peptides were designated CNBr, T, V8, and CT for cleavage by cyanogen bromide, trypsin, *S. aureus* V8 protease, and chymotrypsin, respectively. Peptides are numbered in the order expected for the anticipated (or observed) cleavage sites.

Most of the supporting data for the separations, amino acid compositions, and quantitative sequencing are included in the Miniprint Supplement[2] following the text. Tables and figures are numbered in the order in which they occur in the text, but are followed by a superscript asterisk if they appear in the Miniprint Supplement.

RESULTS

The complete amino acid sequence of Er-1 and the strategy used for its determination are shown in Fig. 1. The sequence was established on the basis of automated Edman degradations of the whole Er-1, previously *S*-carboxymethylated and radiolabeled with ^{14}C, and of peptides produced by cleavage with cyanogen bromide, trypsin, *S. aureus* V8 protease, and chymotrypsin. The proposed structure accounts for the amino acid composition of Er-1 (shown in Table I), both as determined in this study and as reported by Concetti *et al.* (5), in a single unambiguous sequence of 40 amino acids.

Automated Edman degradation of an undigested 2-nmol sample of CM-Er-1 allowed identification of the first 20 amino acid residues (Table II*). The remainder of the sequence was determined from a combination of peptides derived from four different cleavage reactions. This seemingly exhaustive proof was deemed necessary because of the initial discrepancy in the expected molecular weight (~12,000) and that finally determined (4,410).

Cyanogen Bromide Cleavage—The CNBr cleavage of CM-Er-1 produced only one soluble peptide, CNBr-2, which was entirely sequenced after purification by reverse-phase chromatography (Fig. 2*). It consisted of 10 residues (Table III*) and was assumed to constitute the carboxyl-terminal portion of Er-1 as it did not contain homoserine (Table IV*).

Trypsin Digestion—The insoluble fraction remaining after the cleavage of CM-Er-1 with CNBr (containing both peptide CNBr-1 and residual undigested CM-Er-1) was recovered by centrifugation and further digested with trypsin. Reverse phase chromatography of the digest produced numerous peaks, only four of which (designated *1* to *4* in Fig. 3*) contained only a single peptide in amounts sufficient to be analyzed (Table IV*). Peak 1 corresponded to peptide T-3, a pentapeptide corresponding to residues 26–30 (Table V*). It also contained homoserine, suggesting that it preceded CNBr-2. Peptide T-3-4, found in peak 4, confirmed this assignment. It was generated by cleavage at a tyrosine residue in the residual undigested CM-Er-1 as judged by the intact methionine residue. It encompassed peptide T-3 and the first 3 residues of peptide CNBr-2 (Table V*).

Peaks 2 and 3 included peptides T-2 and T-1, respectively, the former derived by cleavage at the single arginine in Er-1 and the latter by another nonspecific cleavage at Gln-9. Both these peptides were partially sequenced, and their alignment was established by reference to the 20-residue amino-terminal sequence earlier determined on intact Er-1 (Table VI*). The remaining two tryptic peptides were not further characterized.

Staphylococcus aureus V8 Protease Digestion—The reverse-phase HPLC separation of the *S. aureus* V8 protease digest of CM-Er-1 gave seven major peaks (designated *1* to *7* in Fig. 4*) and yielded five of the six peptides predicted by the presence of 5 glutamic acid residues in Er-1 (Table VII*). Only the dipeptide, V8-5, was not recovered.

Peak 1 contained two tetrapeptides, V8-1 and V8-3, which were sequenced directly without further separation and assigned to positions 1–4 and 13–16 on the basis of the previous sequence information (see above) (Table VIII*). Peaks 2 to 5 each contained one peptide: V8-4, V8-3-4, V8-2, and V8-1-2, respectively, the compositions of which are given in Table VII*. Peptides V8-2 and V8-4 were also sequenced (Table VIII*) and juxtaposed to V8-1 and V8-3 in positions 5–12 and 17–21, respectively; they collectively confirmed the 20 assignments from the analysis of the whole protein and extended it by a residue (Glu-21). Peptides V8-1-2 and V8-3-4 appear to

[2] Tables II*–XIII* and Figs. 2*–6* are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

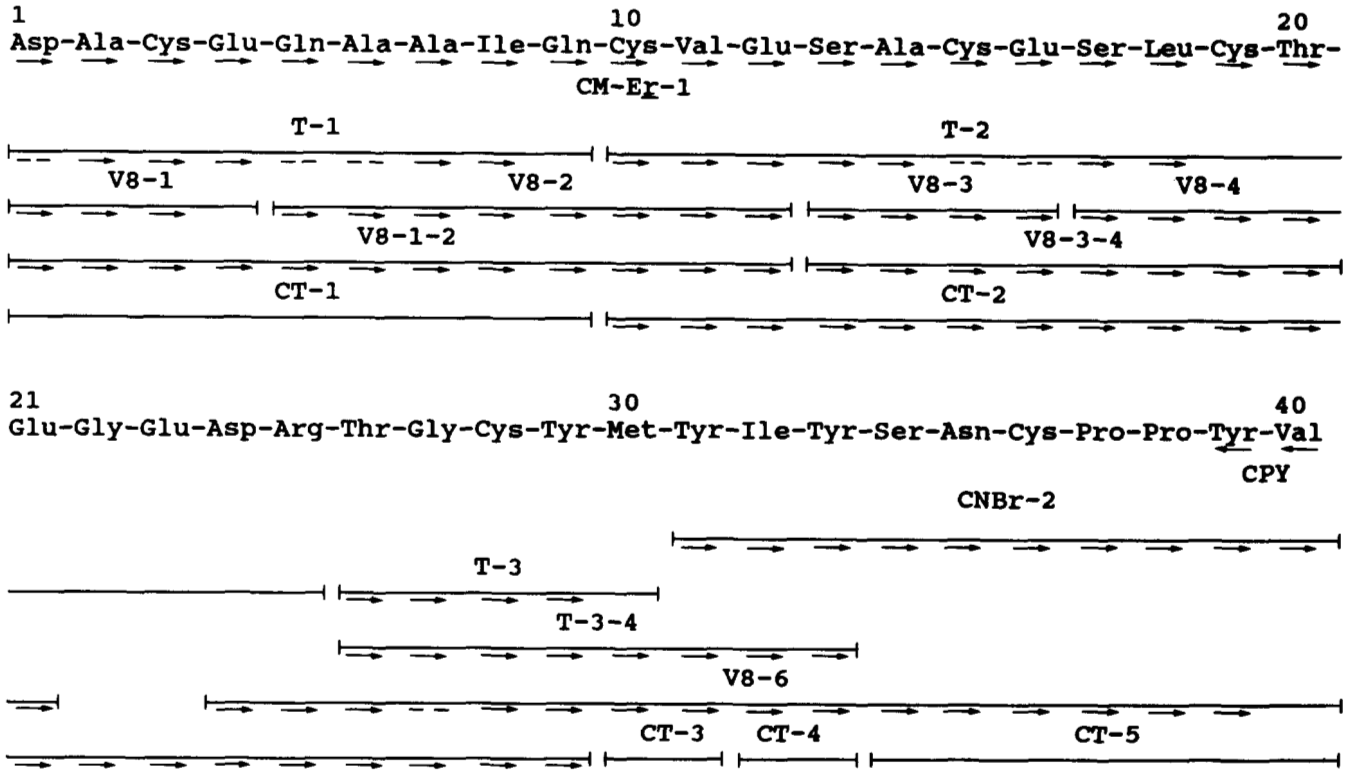

FIG. 1. **Summary of the data used to establish the complete amino acid sequence of Er-1 mating pheromone.** The peptides have been designated and numbered according to the type of digest and the theoretical order in which they appear in the sequence. Designations are: *CNBr*, cyanogen bromide; *T*, trypsin; *V8*, *S. aureus* V8 protease; *CT*, chymotrypsin. Peptides indicated by 2 *numbers* connected with a *hyphen* result from partial cleavage. Residues directly identified by automated Edman degradation and carboxypeptidase Y digestion (*CP-Y*) are marked by *right* and *left arrows*, respectively. Residues identified by amino acid composition are indicated by *dashed lines*.

TABLE I

Amino acid composition of E. raikovi mating pheromone Er-1

Amino acid	Residues/molecule[a]		
	From acid hydrolysates[b]		From amino acid sequence
	1	2	
Aspartic acid	2.7	3.2	3
Glutamic acid	6.9	7.1	7
Serine	2.7	2.9	3
Glycine	2.4	3.3	2
Arginine	0.9	0.3	1
Threonine	1.9	2.0	2
Alanine	4.4	4.4	4
Proline	2.3	2.0	2
Tyrosine	3.6	2.3	4
Valine	2.1	1.8	2
Methionine	0.8	0.6	1
Half-cystine[c]	5.9	6.1	6
Isoleucine	2.1	1.9	2
Leucine	1.3	1.3	1
Lysine		0.3	
Tryptophan	ND[d]	0.1	
Total	40.0	39.6	40

[a] Assuming M_r = 4,400.

[b] Column 1, this study; column 2, Concetti *et al.* (5), recalculated from original M_r = 12,000.

[c] Determined as cysteic acid.

[d] Not determined.

have been generated by incomplete cleavages at Glu-4 and Glu-16 (Table IX*). Peak 7 contained a single peptide sufficiently pure to be sequenced for 15 of its 17 residues (Table X*). This peptide, V8-6, contained the carboxyl terminus of the protein and provided the assignment of residues 24 and 25.

Chymotrypsin Digestion—To obtain the requisite data to connect the amino- and carboxyl-terminal fragments and complete the Er-1 sequence, CM-Er-1 was treated with chymotrypsin and the digest was separated by reverse-phase HPLC (Fig. 5*). Only peptide CT-2, eluted in peak 5 and shown by amino acid analysis (Table XI*) to contain 1 arginine residue, was sequenced. It consisted of 20 residues, spanning residues 10 to 29, and provided the necessary overlap as well as the identification of residues 22 and 23 (Table XII*).

Carboxypeptidase Y Digestion—To confirm the carboxyl-terminal sequence of Er-1, previously determined through the sequence of peptide CNBr-2, a carboxypeptidase Y digestion of the entire CM-Er-1 was performed. A timed analysis of the released amino acids permitted the quantitative confirmation of the last 2 residues, Tyr-Val (Table XIII*).

Mass Spectrometric Analysis—A sample of native Er-1 was subjected to Cf-252 fission fragment ionization mass spectrometry. As shown in Fig. 6*, a strong protonated quasimolecular ion $(M + H)^+$ appears in the positive spectrum at m/z 4411.2, corresponding to an isotopically averaged mass of 4410.2. In addition, a weaker doubly protonated quasimolecular ion $(M + 2H)^{2+}$ appears at m/z 2206.2, corresponding to an isotopically averaged mass of 4410.4. The spectrum of Er-1 also contains a secondary peak at m/z 8819, which might arise from the presence of homodimers of noncovalently linked subunits in the analyzed sample.

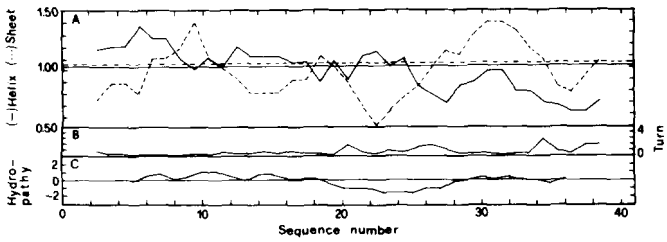

FIG. 7. **Secondary structure prediction and hydropathy profile of Er-1 mating pheromone.** Profiles in *A* and *B* were derived according to Chou and Fasman (24). The hydropathy profile in *C* was obtained using the scale of Kyte and Doolittle (25) at a span setting of 9 residues; values *above* and *below* the *horizontal line* indicate hydrophobic and hydrophilic regions, respectively.

DISCUSSION

The mating pheromone Er-1 of *E. raikovi* was determined to be a single chain polypeptide of 40 residues. The complete amino acid sequence was determined by automated Edman degradation of intact CM-Er-1, and its fragments were derived by digestion with CNBr, trypsin, *S. aureus* V8 protease, and chymotrypsin. The dipeptide sequence Tyr-Val at the carboxyl terminus was confirmed by the results of CM-Er-1 treatment with carboxypeptidase Y. The Er-1 amino acid composition deduced from the sequence analysis is the same as that determined after hydrolysis.

The 6 half-cystine residues are unreactive in the native molecule (data not shown) and presumably exist, as expected for an extracellular protein, as either intra- or interchain disulfide bonds. The even number of half-cystines would require a minimum of two interchain bridges. Molecular weight measurements of unreduced and reduced samples should resolve this issue. However, previous determinations of the native and subunit molecular weights of Er-1 have been difficult to interpret because they varied over a wide range. In urea/sodium dodecyl sulfate electrophoresis, reduced Er-1 migrated as a broad band corresponding to a molecular weight of 29,000 (5). This value was considered to be anomalous and was attributed to intrinsic structural properties of Er-1, such as the strongly negative charge and the high content of half-cystine residues. Molecular weight determinations of native Er-1 samples on gel filtration, which gave values of 12,000, were thought to be more reliable. The Er-1 amino acid composition subsequently reported was calculated based on this value (5). In more recent studies,[3] Er-1 samples have been found to elute with a calculated molecular weight of 9,000 on a Superose-12 column (equilibrated with 0.1 M Tris-HCl and 0.15 M KCl at pH 7.5). Furthermore, 20% sodium dodecyl sulfate-polyacrylamide gel electrophoresis after carboxymethylation gave a diffuse band corresponding to a molecular weight of 4000–5000.

These later results, taken with the molecular weight determined from the sequence, suggest a homodimeric structure (or possibly larger aggregate) for native Er-1 with only intrachain disulfide bonds. Consistent with this interpretation, the mass spectral analysis showed primarily a single species of mass 4411.2. The calculated mass from the proposed sequence (assuming three disulfide bonds) is 4411.0. A minor peak (~3%) at *m/z* 8819 in the mass spectrum may represent the homodimer and suggests that the subunits are associated by tight, noncovalent forces.

An evaluation of the secondary structure of Er-1, based on the parameters proposed by Chou and Fasman (24), is illustrated in Fig. 7 together with the hydropathy profile delineated according to Kyte and Doolittle (25). The amino terminus may contain a helical segment, but the remainder of the molecule probably contains little secondary structure other than β-turns. Residues 20–27 are the most polar region and probably occur on the surface. As such, it is a possible candidate to function as a site of molecular recognition. It may also contain a β-turn. Such areas often serve as loci for receptor binding and antibody recognition (26).

All ciliate mating pheromones isolated so far are polypeptides with the exception of "gamone 2" of *B. japonicum*, which has been identified as calcium-3-(2′-formylamino-5′-hydroxybenzoyl)lactate (1). Amino acid sequence information has not yet been determined for the slightly basic gamone 1 of this species (M_r = 20,000) (2), and for the acidic gamones 1, 2, 3, and 4 of *E. octocarinatus* with molecular masses between 18,500 and 23,500 Da (8, 9). Only the amino termini of three other *E. raikovi* mating pheromones, Er-2, Er-3, and Er-9, have been analyzed and reported (27), in comparison to Er-1, to be:

```
     1                                  10
Er-1: Asp-Ala-Cys-Glu-Gln-Ala-Ala-Ile-Gln-Cys-Val-Glu-Ser-
Er-3: Asp-Ala-Cys-Glu-Gln-Ala-Ala-Ile-Gln-Cys-Val-Glu-Val-
Er-2: Asp-Pro-Met-Thr-Cys-Glu-Gln-Ala-Met-Ala-Thr-Cys-Glu-
Er-9: Asp-Pro-Met-Gln-Cys-Glu-Gln-Ala-Met-Ala-Ser-Cys-Glu-
```

These sequences all show aspartic acid as a common amino-terminal residue and that, with a 2-residue shift for Er-2 and Er-9, the motif Cys-Glu-Gln-Ala-*X*-*X*-*X*-Cys is found in each mating pheromone. This suggests that the pheromone family may be homologous, and/or this sequence may be important for a functionality common to *E. raikovi* mating pheromones. As proposed for other protein families characterized by a conserved amino-terminal region (28, 29), one such property might be the interaction with membranes and other molecules, especially small hydrophobic compounds. Clearly, extensive similarity exists between Er-1 and Er-3 and between Er-2 and Er-9.

It has been proposed (10) that, as in autocrine secretion systems, the primary target cell of a ciliate mating pheromone is the same as that which actually secretes the pheromone. The completion of the amino acid sequence of Er-1 is the first step in defining the molecular basis of pheromone-cell interactions and thus testing this innovative concept, which contradicts the traditional view that ciliate mating pheromones are involved in a process of recognition of cells characterized by molecular differences that are "complementary" for mating (12, 30).

Acknowledgments—We wish to thank Owen J. Bates and Kevin P. Burke for their assistance with various analyses. We also thank Debby Thyssen and Linda Kishimoto for expert assistance with the manuscript.

REFERENCES

1. Kubota, T., Tokoroyama, T., Tsukuda, Y., Koama, H., and Miyake, A. (1973) *Science* **179,** 400–402
2. Miyake, A., and Beyer, J. (1974) *Science* **185,** 621–623
3. Braun, V., and Miyake, A. (1975) *FEBS Lett.* **53,** 131–134
4. Miceli, C., Concetti, A., and Luporini, P. (1983) *Exp. Cell Res.* **149,** 593–598
5. Concetti, A., Raffioni, S., Miceli, C., Barra, D., and Luporini, P. (1986) *J. Biol. Chem.* **261,** 10582–10586
6. Luporini, P., Raffioni, S., Concetti, A., and Miceli, C. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 2889–2893
7. Raffioni, S., Miceli, C., Concetti, A., Barra, D., and Luporini, P. (1987) *Exp. Cell Res.* **172,** 417–424

[3] S. Raffioni and P. Luporini, unpublished observations.

8. Weischer, A., Freiburg, M., and Heckmann, K. (1985) *FEBS Lett.* **191,** 176–180
9. Dieckhoff, H. S., Freiburg, M., and Heckmann, K. (1987) *Eur. J. Biochem.* **168,** 89–94
10. Luporini, P., and Miceli, C. (1986) in *Molecular Biology of Ciliated Protozoa* (Gall, J., ed) pp. 263–299, Academic Press, Orlando, FL
11. Nobili, R., Luporini, P., and Esposito, F. (1987) in *Invertebrate Models: Cell Receptors and Cell Communication* (Greenberg, A. H., ed) pp. 6–28, S Karger AG, Basel
12. Heckmann, K., and Kuhlmann, H. W. (1986) *J. Exp. Zool.* **237,** 87–96
13. Sleigh, M. A. (1973) *The Biology of Protozoa,* Arnold, London
14. Ammermann, D. (1982) *Arch. Protistenkd.* **126,** 373–381
15. Dini, F., and Luporini, P. (1985) *J. Protozool.* **32,** 111–117
16. Angeletti, R. H., Bradshaw, R. A., and Wade, R. D. (1971) *Biochemistry* **10,** 463–470
17. Gross, E., and Witkop, B. (1961) *J. Am. Chem. Soc.* **83,** 1510–1511
18. Allen, G. (1981) *Sequencing of Proteins and Peptides* (Work, T. S., and Burdon, R. H., eds) Elsevier Scientific Publishing Co., Amsterdam
19. Houmard, J., and Drapeau, G. R. (1972) *Proc. Natl. Acad. Sci. U. S. A.* **69,** 3506–3509
20. Hayashi, R., Moore, S., and Merrifield, R. B. (1973) *J. Biol. Chem.* **248,** 3889–3892
21. Heinrikson, R. L., Mora, R., and Maraganore, J. M. (1987) in *Modern Methods in Protein Chemistry* (L'Italien, J., ed) Pergamon Press, New York, in press
22. Chait, B. T., and Field, F. H. (1986) *Biochem. Biophys. Res. Commun.* **134,** 420–426
23. Chait, B. T., Agosta, W. C., and Field, F. H. (1981) *Int. J. Mass Spectrom. Ion Phys.* **39,** 339–366
24. Chou, P. Y., and Fasman, G. D. (1974) *Biochemistry* **13,** 222–245
25. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132
26. Rose, G. D., Gierasch, L. M., and Smith, J. A. (1985) in *Advances in Protein Chemistry* (Anfinsen, C. B., Edsall, J. T., and Richards, F. M., eds) Vol. 37, pp. 1–109, Academic Press, Orlando, FL
27. Disper, S. S., Bradshaw, R. A., Raffioni, S., Micelli, C., and Luporini, P. (1986) *Eur. J. Cell Biol.* **42** (Suppl. 16), 40
28. Sundelin, J., Das, S. R., Eriksson, U., Rask, L., and Peterson, P. A. (1985) *J. Biol. Chem.* **260,** 6494–6499
29. Böhmer, F.-D., Kraft, R., Otto, A., Wernstedt, C., Hellman, U., Kurtz, A., Müller, T., Rohde, K., Etzold, G., Lehmann, W., Langen, P., Heldin, C.-H., and Grosse, R. (1987) *J. Biol. Chem.* **262,** 15137–15143
30. Mikaye, A. (1981) in *Biochemistry and Physiology of Protozoa* (Levandowsky, M., and Hutner, S. H., eds) pp. 125–198, Academic Press, Orlando, FL

SUPPLEMENTAL MATERIAL TO

The Primary Structure of Er-1 Mating Pheromone of the ciliate Euplotes raikovi

Simona Raffioni, Pierangelo Luporini, Brian T. Chait, Steven S. Disper and Ralph A. Bradshaw

The amino acid composition values are followed in parentheses by the integer number of each residue as determined from the sequence. PTH-amino acid values were not corrected for carryover or background because of the relatively short length of each run.

TABLE II*

Amino acid sequence analysis of S-[^{14}C]-carboxymethyl Er-1

CYCLE	PTH-Amino Acid	Yield (pmol)[a]
1	Aspartic Acid	2360
2	Alanine	-[b]
3	Cysteine (CM)	-[c]
4	Glutamic Acid	1200
5	Glutamine	1050
6	Alanine	700
7	Alanine	930
8	Isoleucine	380
9	Glutamine	700
10	Cysteine (CM)	-[c]
11	Valine	640
12	Glutamic Acid	1100
13	Serine	240
14	Alanine	450
15	Cysteine (CM)	-[c]
16	Glutamic acid	860
17	Serine	120
18	Leucine	240
19	Cysteine (CM)	-[c]
20	Threonine	-[d]

[a] Estimated amount loaded: 2700 pmol

[b] Integrator failure prevented quantitation

[c] Identified on HPLC as PTH-glutamic acid and confirmed by radioactive analysis

[d] Identified qualitatively

TABLE III*

Amino acid sequence analysis of fragment CNBr-2 from CM-Er-1

CYCLE	PTH-Amino Acid	Yield (pmol)[a]
1	Tyrosine	1430
2	Isoleucine	1190
3	Tyrosine	970
4	Serine	200
5	Asparagine	740
6	Cysteine (CM)	-[b]
7	Proline	330
8	Proline	450
9	Tyrosine	190
10	Valine	180

[a] Estimated amount loaded: 2000 pmol

[b] Identified on HPLC as PTH-glutamic acid and confirmed by radioactive analysis

TABLE IV*

Amino acid composition of CNBr and tryptic peptides of Er-1[a]

Amino Acid	CNBr-2	T-1	T-2	T-3	T-3-4
Aspartic acid	0.9 (1)	0.7 (1)	0.7 (1)		
Glutamic acid		2.8 (3)	3.4 (4)		
Serine	0.7 (1)		1.3 (2)		
Glycine			1.2 (1)	1.2 (1)	1.3 (1)
Arginine			0.6 (1)		
Threonine			0.6 (1)	0.7 (1)	0.7 (1)
Alanine		2.7 (3)	1.1 (1)		
Proline	2.3 (2)				
Tyrosine	2.6 (3)			0.6 (1)	2.5 (3)
Valine	1.2 (1)[b]		0.8 (1)		
Methionine				0.2[c](1)	0.6 (1)
Half-cystine[d]	0.8 (1)	0.8 (1)	2.5 (3)	0.8 (1)	0.8 (1)
Isoleucine	1.0 (1)[b]	1.0 (1)			1.0 (1)
Leucine			1.0 (1)		
Total	10	9	16	5	8

[a] Values in parentheses are the integral values determined from the amino acid sequence

[b] Valine and Isoleucine determined after 96 h hydrolysis

[c] Methionine detected as homoserine

[d] Half-cystine determined as S-(carboxymethyl) cysteine

TABLE V*

Amino acid sequence analysis of tryptic peptides T-3 and T-3-4 of CM-Er-1

CYCLE	PTH-Amino Acid	Yield (pmol)[a]
Peptide T-3:		
1	Threonine	310
2	Glycine	560
3	Cysteine (CM)	-[b]
4	Tyrosine	150
Peptide T-3-4:		
1	Threonine	210
2	Glycine	270
3	Cysteine (CM)	-[b]
4	Tyrosine	320
5	Methionine	240
6	Tyrosine	430
7	Isoleucine	150
8	Tyrosine	260

[a] Estimated amount loaded: T-3: ~ 500 pmol; T-3-4: ~ 500 pmol

[b] Identified on HPLC as PTH-glutamic acid and confirmed by radioactive analysis

TABLE VI*

Amino acid sequence analysis of tryptic peptides T-1 and T-2 of CM-Er-1

CYCLE	PTH-Amino Acid	Yield (pmol)[d]
Peptide T-1:		
1	(Aspartic Acid)[b]	-[c]
2	Alanine	330
3	Cysteine (CM)	-[d]
4	(Glutamic Acid)[b]	-[c]
5	(Glutamine)[b]	-[c]
6	Alanine	330
7	Alanine	880
8	Isoleucine	190
9	Glutamine	90
Peptide T-2:		
1	Cysteine (CM)	-[d]
2	Valine	560
3	Glutamic Acid	530
4	Serine	70
5	Alanine	390
6	(Cysteine (CM))[b]	-[c]
7	(Glutamic acid)[b]	-[c]
8	Serine	70
9	Leucine	120

[a] Estimated amount loaded: T-1: ~ 500 pmol; T-2: ~ 750 pmol

[b] Tentative identification

[c] Not quantitated

[d] Identified on HPLC as PTH-glutamic acid and confirmed by radioactive analysis

TABLE VII*

Amino acid composition of S. aureus V8 protease peptides from Er-1[a]

Amino acid	V8-1+V8-3	V8-2	V8-4	V8-6	V8-1-2	V8-3-4
Aspartic acid	1.1 (1)			1.6 (2)	0.9 (1)	
Glutamic acid	2.3 (2)	2.7 (3)	0.9 (1)	0.8 (1)	3.5 (4)	2.1 (2)
Serine	0.8 (1)		0.7 (1)	0.9 (1)		1.2 (2)
Glycine				2.2 (2)		
Arginine				0.7 (1)		
Threonine			0.7 (1)	0.9 (1)		0.7 (1)
Alanine	2.5 (2)	2.4 (2)			3.6 (3)	1.2 (1)
Proline				2.4 (2)		
Tyrosine				3.2 (4)		
Valine		1.2 (1)		1.3 (1)	1.2 (1)	
Methionine				0.8 (1)		
Half-cystine	2.0 (2)	0.9 (1)	0.8 (1)	1.9 (2)	1.9 (2)	1.8 (2)
Isoleucine		1.0 (1)		1.0 (1)	1.0 (1)	
Leucine			1.0 (1)			1.0 (1)
Total	8	8	5	19	12	9

[a] Values in parentheses are the integral values determined from the amino acid sequence

[b] Half-cystine determined as S-(carboxymethyl) cysteine

TABLE VIII*

Amino acid sequence analysis of S. aureus V8 protease peptides V8-1/V8-3, V8-2 and V8-4 of CM-Er-1

CYCLE	PTH-Amino Acid	Yield (pmol)[a]
Peptide V8-1/V8-3		
1	Aspartic Acid/Serine	1240/850
2	Alanine	5420
3	Cysteine (CM)	-[b]
4	Glutamic Acid	950
Peptide V8-2		
1	Glutamine	930
2	Alanine	1300
3	Alanine	1220
4	Isoleucine	880
5	Glutamine	750
6	Cysteine (CM)	-[b]
7	Valine	800
8	Glutamic Acid	230
Peptide V8-4:		
1	Serine	1070
2	Leucine	3330
3	(Cysteine (CM))	-[b]
4	Threonine	780
5	Glutamic acid	520

[a] Estimated amount loaded: V8-1/V8-3: ~ 7500 pmol; V8-2: ~ 1500 pmol V8-4: ~ 4000 pmol

[b] Identified on HPLC as PTH-glutamic acid and confirmed by radioactive analysis

TABLE IX*

Amino acid sequence analysis of S. aureus V8 protease peptides V8-1-2 and V8-3-4 of CM-Er-1

CYCLE	PTH-Amino Acid	Yield (pmol)[a]
Peptide V8-1-2		
1	Aspartic Acid	430
2	Alanine	1660
3	Cysteine (CM)	-[b]
4	Glutamic Acid	1400
5	Glutamine	1360
6	Alanine	1290
7	Alanine	1720
8	Isoleucine	860
9	Glutamine	910
10	Cysteine (CM)	-[b]
11	Valine	850
12	Glutamic Acid	340
Peptide V8-3-4:		
1	Serine	170
2	Alanine	1180
3	Cysteine (CM)	-[b]
4	Glutamic Acid	900
5	Serine	80
6	Leucine	370
7	Cysteine (CM)	-[b]
8	Threonine	60
9	Glutamic Acid	80

[a] Estimated amount loaded: V8-1-2: ~2000 pmol; V8-3-4: ~1500 pmol

[b] Identified on HPLC as PTH-glutamic acid and confirmed by radioactive analysis

TABLE X*

Amino acid sequence analysis of S. aureus V8 protease peptide V8-6 of CM-Er-1

CYCLE	PTH-Amino Acid	Yield (pmol)[a]
1	Aspartic Acid	630
2	Arginine	420
3	Threonine	210
4	Glycine	220
5	Cysteine (CM)	-[b]
6	Tyrosine	430
7	Methionine	330
8	Tyrosine	680
9	Isoleucine	260
10	Tyrosine	930
11	Serine	150
12	Asparagine	410
13	Cysteine (CM)	-[b]
14	Proline	90
15	Proline	210
16	Tyrosine	310

[a] Estimated amount loaded: ~ 1000 pmol

[b] Identified on HPLC as PTH-glutamic acid and confirmed by radioactive analysis

TABLE XI*

Amino acid composition of chymotryptic peptides from Er-1[a]

Amino Acid	CT-1	CT-2	CT-3	CT-4	CT-5
Aspartic acid	0.9 (1)	0.8 (1)			0.9 (1)
Glutamic acid	2.9 (3)	3.4 (4)			
Serine		1.6 (2)			0.7 (1)
Glycine		2.3 (2)			
Arginine		0.8 (1)			
Threonine		1.7 (2)			
Alanine	3.4 (3)	1.3 (1)			
Proline					2.4 (2)
Tyrosine		0.6 (1)	0.6 (1)	0.7 (1)	0.6 (1)
Valine		1.1 (1)			1.0 (1)
Methionine			1.0 (1)		
Half-cystine[b]	0.8 (1)	3.5 (4)			0.8 (1)
Isoleucine	1.0 (1)			1.0 (1)	
Leucine		1.0 (1)			
Total	9	20	2	2	7

[a] Values in parentheses are the integral values determined from the amino acid sequence

[b] Half-cystine was determined as S-(carboxymethyl) cysteine

TABLE XII*

Amino acid sequence analysis of chymotryptic peptide CT-2 of CM-Er-1

CYCLE	PTH-Amino Acid	Yield (pmol)[a]
1	Cysteine (CM)	-[b]
2	Valine	940
3	Glutamic Acid	440
4	Serine	200
5	Alanine	940
6	Cysteine (CM))	-[b]
7	Glutamic acid	180
8	Serine	130
9	Leucine	380
10	Cysteine (CM)	-[b]
11	Threonine	130
12	Glutamic Acid	270
13	Glycine	170
14	Glutamic Acid	300
15	Aspartic Acid	120
16	Arginine	80
17	Threonine	100
18	Glycine	130
19	Cysteine (CM)	-[d]
20	Tyrosine	50

[a] Estimated amount loaded: ~ 1300 pmol

[b] Identified on HPLC as PTH-glutamic acid and confirmed by radioactive analysis

TABLE XIII*

Release of amino acids from Er-1 by carboxypeptidase Y

Time	Amino Acid[a]	
min	Valine	Tyrosine
5	172	72
30	225	105
60	246	156
120	280	215
240	424	338

[a] Quantity, reported in pmols, of amino acids detected following carboxypeptidase Y digestion of 5 nmols of CM-Er-1

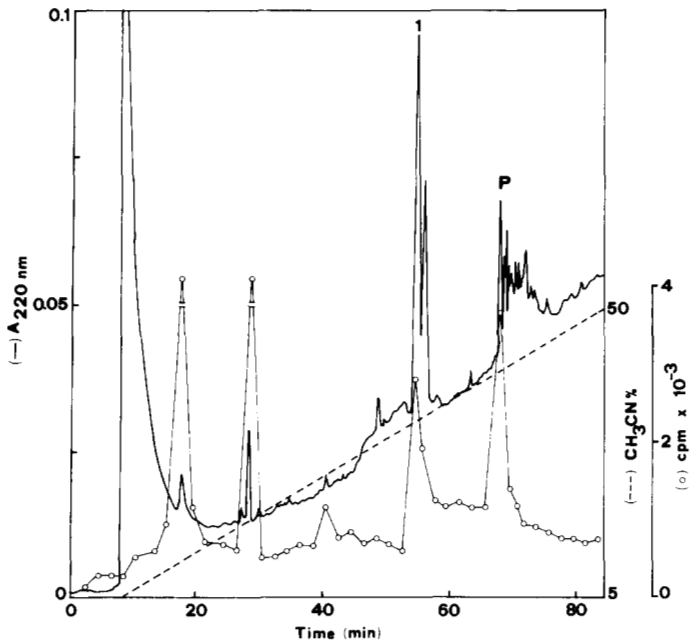

Fig. 2*. Separation of soluble CNBr peptides of CM-Er-1 by HPLC. The peptides were eluted with a linear gradient of acetonitrile and detected by their absorbance at 220 nm. Radioactivity was counted on 50 µl aliquots taken from every other fraction. The only soluble peptide, CNBr-2, is contained in peak 1. Peak P is undigested protein; the labeled peaks at 19 and 28 min are non-peptidic by-products of the carboxymethylation modification.

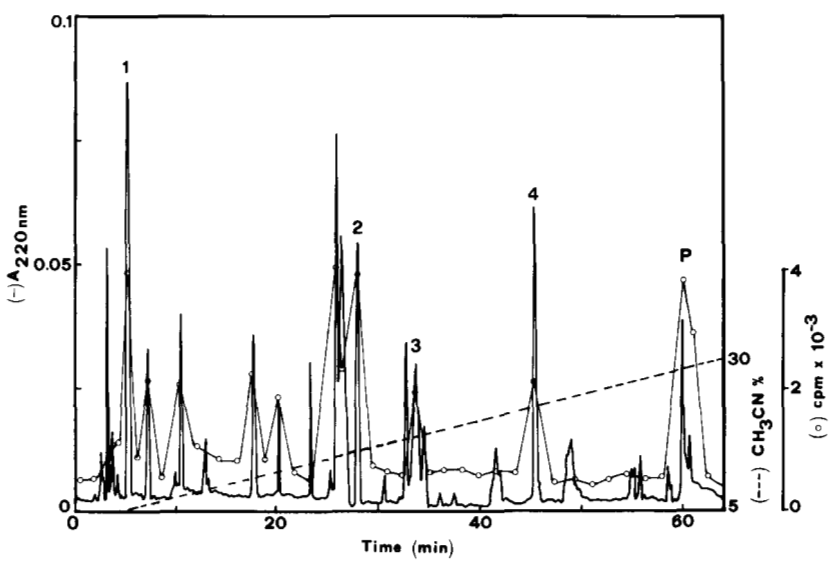

Fig. 3*. Separation by HPLC of tryptic peptides from insoluble CNBr cleavage of CM-Er-1. The same conditions as described in Fig. 2* were used.

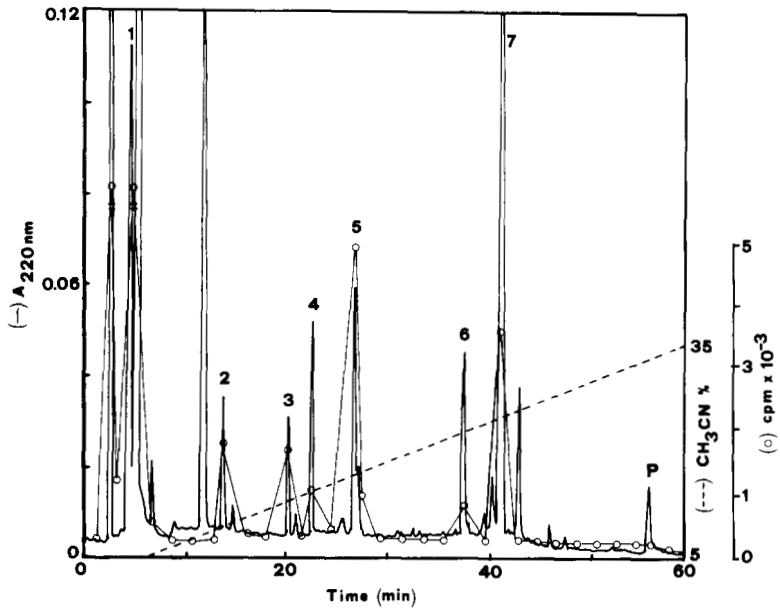

Fig. 4*. Separation by HPLC of S. aureus V8 peptides of CM-Er-1. The same conditions as described in Fig. 2* were used.

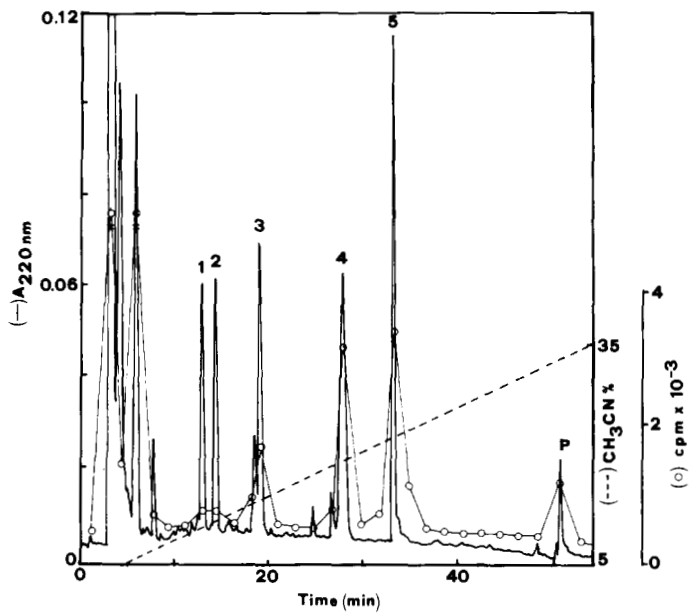

Fig. 5*. Separation by HPLC of chymotryptic peptides of CM-Er-1. The same conditions as described in Fig. 2* were used.

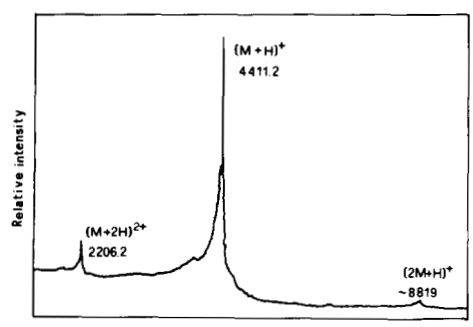

Fig. 6*. Cf-252 fission fragment ionization time-of-flight mass spectrum of Er-1. The region between m/z 1600-10000 is shown. M designates the intact Er-1 molecule.