Protein Science (1992), *1*, 777–785. Cambridge University Press. Printed in the USA.

The disulfide bond pairing of the pheromones E*r*-1 and E*r*-2 of the ciliated protozoan *Euplotes raikovi*

ALBERT E. STEWART,[1] SIMONA RAFFIONI,[1] TANUJA CHAUDHARY,[2] BRIAN T. CHAIT,[2] PIERANGELO LUPORINI,[3] AND RALPH A. BRADSHAW[1]

[1] Department of Biological Chemistry, College of Medicine, University of California, Irvine, California 92717
[2] Mass Spectrometry Laboratory, The Rockefeller University, New York, New York 10021
[3] Department of Cell Biology, University of Camerino, 62032 Camerino (MC), Italy

(RECEIVED December 27, 1991; REVISED MANUSCRIPT RECEIVED February 13, 1992)

Abstract

The disulfide pairings of the two *Euplotes raikovi* pheromones E*r*-1 and E*r*-2 have been determined by chemical and mass spectrometric analyses. Cystine-linked peptides from thermolytic digestions of the native molecules were purified by reverse-phase high performance liquid chromatography and identified in the known sequences to make the assignments. The same pairing, Cys(I)–Cys(IV), Cys(II)–Cys(VI), and Cys(III)–Cys(V), was found in both pheromones, suggesting that this pattern occurs commonly throughout this family of molecules. This arrangement of disulfides indicates that the three-dimensional structure is defined by three loops, which can vary in size and charge distribution from one pheromone to another.

Keywords: evolution; mass spectroscopy; polypeptide hormones; protein conformation

Proteins involved in intercellular communication and recognition are constitutively secreted by some species of ciliated protozoa. These molecules, known as pheromones, distinguish different intraspecific classes of cells, commonly referred to as "mating types." Once secreted, they apparently interact with cell surface receptors (Ortenzi et al., 1990). Significant molecular evidence has accumulated to support the hypothesis that ciliate mating types and their products evolved as a self-recognition system (Luporini & Miceli, 1986), similar to the autocrine mechanisms found in higher vertebrate cells. In this model, the cell producing the pheromone also produces a receptor for that pheromone. The identification of a putative membrane-bound form of the precursor to a *Euplotes raikovi* pheromone, which arises from alternate splicing, may indicate that the ligand and its receptor are actually derived from the same gene (Miceli et al., 1992).

Pheromones have been isolated and partially characterized from *Blepharisma japonicum* (Kubota et al., 1973; Miyake & Beyer, 1974), *Euplotes raikovi* (Miceli et al., 1983; Concetti et al., 1986; Luporini et al., 1986; Raffioni et al., 1987), *Euplotes octocarinatus* (Weischer et al., 1985; Schulze Dieckhoff et al., 1987; Meyer et al., 1991), *Euplotes patella* (Akada, 1986), and *Dileptus anser* (Parfenova et al., 1989). In *E. raikovi*, each pheromone (termed euplomone and abbreviated E*r*, where *r* stands for the specific name *raikovi*) (Miceli et al., 1983) is controlled by one of a multiple series of codominant alleles at the Mendelian locus *mat* (Luporini & Miceli, 1986). These pheromones constitute a family of small homodimeric proteins, which possess an acidic isoelectric point and an average molecular weight of 9,000. The complete amino acid sequence has been determined for seven members of the family; however, only five were found to be unique (Raffioni et al., 1988, 1989, 1992). Each protein contains three intrachain disulfide bonds formed by half-cystine residues in similar but not identical positions. The precursor structures, determined from the corresponding cDNA sequences, have been elucidated for three members of the family (Miceli et al., 1989, 1991).

In this report, the disulfide bond pairings for two members of the *E. raikovi* pheromone family, E*r*-1 and E*r*-2, have been elucidated, completing the determination of the covalent structure of the mature pheromone in each case. Given the degree of similarity between E*r*-1, E*r*-2, and the other members of the family thus far identified (Raffioni et al., 1992), it is probable that this pattern will be generally conserved throughout.

Reprint requests to: Ralph A. Bradshaw, Department of Biological Chemistry, College of Medicine, University of California, Irvine, California 92717-1700.

Results

Thermolytic digestion of Er-2

The extent of thermolytic digestion of E*r*-2 was assessed by microbore reverse-phase high performance liquid chromatographic (HPLC) analysis of aliquots at 6, 12, and 24 h. Nearly complete fragmentation was observed after 24 h, as judged by the disappearance of the peak that corresponded to native E*r*-2 (data not shown). The material digested for 24 h was used for all subsequent analyses (Fig. 1A).

Identification of cystine-containing peptides

An aliquot of the thermolytic digest of E*r*-2 was treated with mercaptoethanol at alkaline pH to reduce the cystine-containing peptides into their reduced cysteine peptide constituents; the remainder of the reaction mixture was unmodified. The resulting samples were fractionated by microbore reverse-phase HPLC. A comparison of the HPLC profiles of the reduced (Fig. 1B) and unreduced (Fig. 1A) samples showed a marked decrease in the intensity of peaks P1, P2, and P3, indicating that they were the predominant cystine-containing peptides. Each peak in the chromatogram of the nonreduced sample was subjected to amino acid analysis; cystine was indeed observed only in the fractions corresponding to peaks P1, P2, and P3. These samples were further characterized by sequence analysis by automated Edman degradation, and all were found to possess multiple amino termini, which suggested each was comprised of more than one peptide component (data not shown). Pools P1, P2, and P3 were further fractionated by microbore reverse-phase HPLC utilizing triethylamine-acetate at pH 6.0 as the ion pairing agent. As shown in Figure 2, each sample was resolved into a major component (peaks P1a, P2a, and P3a) and several minor components, consistent with the initial sequencing results, which indicated the three samples were not homogeneous.

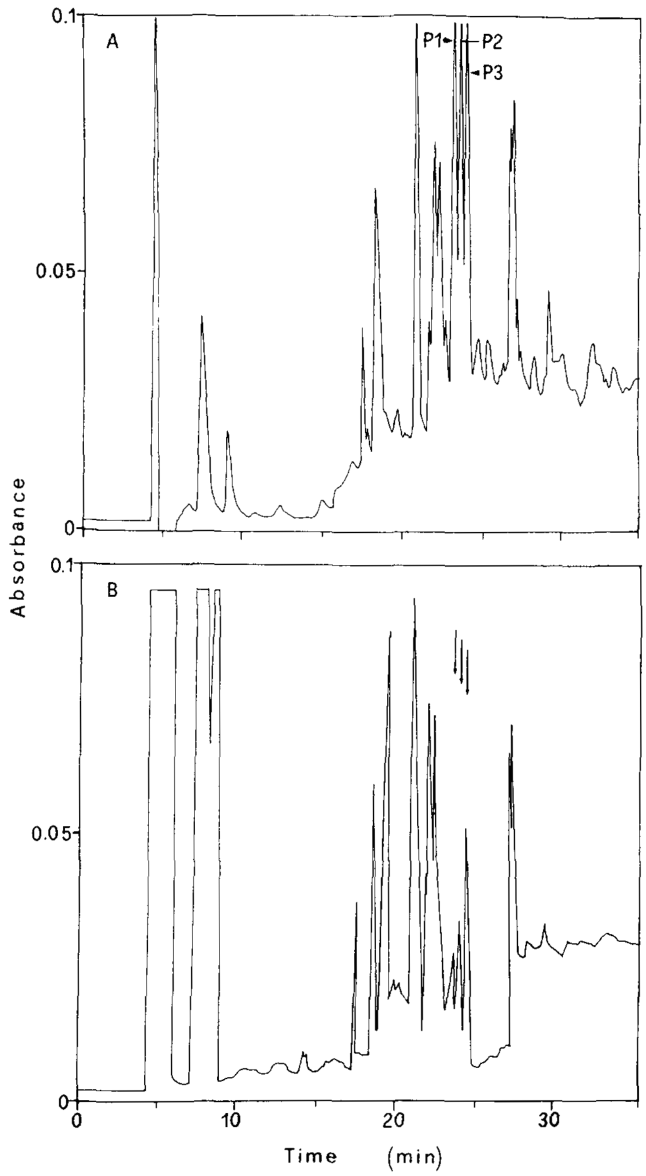

Fig. 1. Reverse-phase HPLC profiles of unreduced and reduced aliquots of thermolysin-digested E*r*-2. Native E*r*-2 was digested with thermolysin for 24 h. Two equivalent aliquots ($\frac{1}{25}$ of the total digest) were removed. One (**A**) was untreated and the other (**B**) was treated with 0.1% β-mercaptoethanol in 0.2 M Tris, pH 8.5, for 5 h. Both samples were analyzed by microbore reverse-phase HPLC on a C18 column (2.1 × 220 mm) utilizing $CH_3CN:H_2O$ (4:1) containing 0.08% TFA as the mobile phase. The peptides were eluted with a linear gradient and detected by their absorbance at 214 nm as described under Materials and methods. P1, P2, and P3 were the only peaks that contained detectable cystine after acid hydrolysis.

Characterization of cystine-containing peptides

Amino acid analysis of each component obtained by refractionation of pools P1, P2, and P3 indicated cystine was present only in peaks 1a, 2a, and 3a (data not shown). The results of peptide sequence analysis of P1a, P2a, and P3a are given in Table 1. Two sequences, as would be anticipated for peptides linked by a disulfide bond, were observed for each sample. Single residue sequencing cycles immediately preceded by a cycle in which two residues were identified, or a 50% decrease from the previous cycle, indicated a probable half-cystine residue at that position in one of the two cystine-linked peptides. This was observed in cycle two of peptides P1a and P2a and in cycle three of peptide P3a. In each case, the sequences that were consistent with the known primary structure of E*r*-2 (Raffioni et al., 1992) could be assigned (Table 2). The peptide sequencing data were in good agreement with results obtained from amino acid composition analyses.

To confirm the assignments made by amino acid composition and sequence analysis, peptides P1a, P2a, and P3a were subjected to mass spectrometric analysis (Fig. 3;

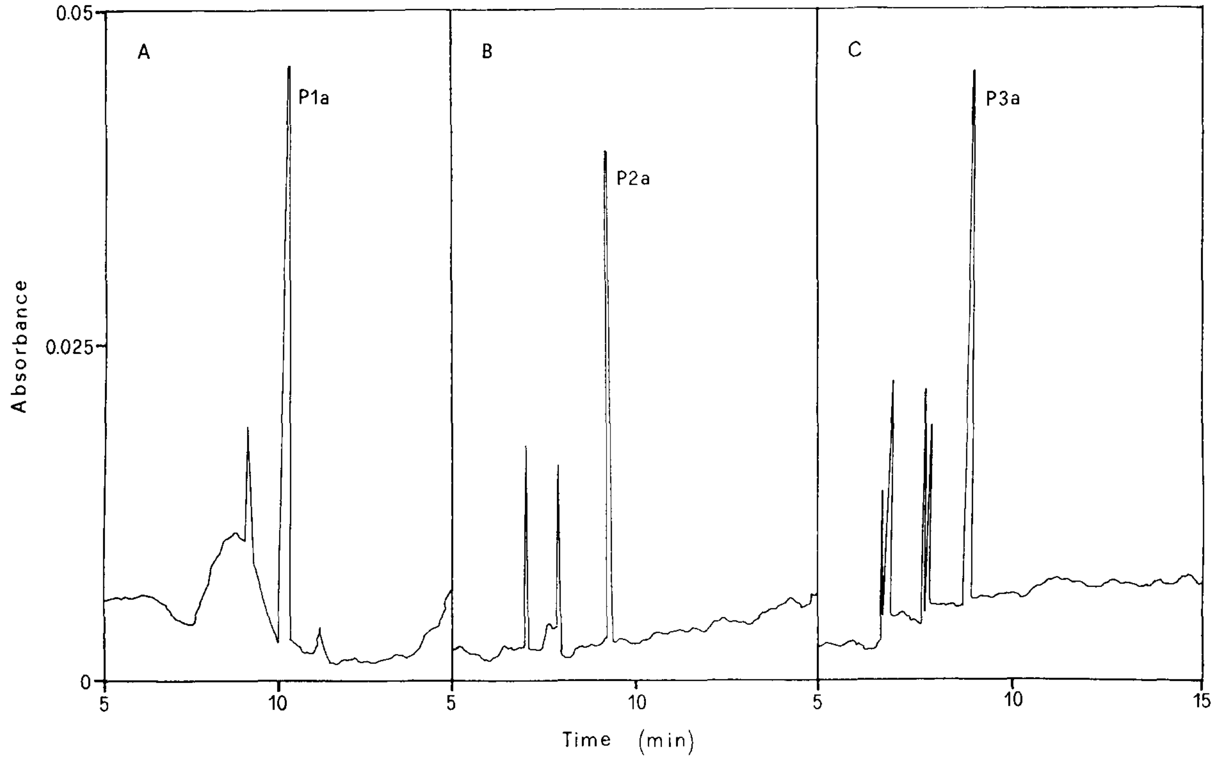

Fig. 2. Purification of cystine-linked peptides by microbore reverse-phase HPLC. Samples of P1, P2, and P3 (containing approximately 150 pmol each of cystine) (Fig. 1A) were further fractionated by microbore reverse-phase HPLC on a C18 column (2.1 × 250 mm) equilibrated in 20 mM TEA-acetate, pH 6.0. The peptides were eluted by a linear increase in the concentration of CH_3CN:20 mM TEA-acetate (1:1) and detected by their absorbance at 214 nm as described under Materials and methods.

Table 2). The electrospray mass spectrum of P1a yielded a singly protonated molecule ion $(M + H)^+$ at m/z 661.2, which agreed with the isotopically averaged mass of 660.8 (data not shown). Peptide P2a showed a diprotonated $(M + 2H)^{2+}$ species at m/z 588.5 in the electrospray mass spectrum. This corresponds to a mass of 1,175.1 as compared with the calculated mass of 1,175.3. A singly protonated molecule ion species was not observed however. In order to verify the mass of P2a, a matrix-assisted laser desorption mass spectrum was obtained (Fig. 3A). An $(M + H)^+$ ion at m/z 1,176.5 was observed, as well as ions corresponding to $(M + Na)^+$, $(M + K)^+$, and

Table 1. *Amino acid sequence analysis of thermolytic peptides of Er-2 and Er-1*

	PTH-amino acid[a]					
	E*r*-2[b]			E*r*-1[b]		
Cycle	P1a	P2a	P3a	P′1a	P′2a	P′3a
1	Met 97	Met 94, Tyr 49	Ala 121, Ile 149	Ala 96, Glu 50	Asp 48, Leu 159	Ile 257
2	Thr 44	Thr 63	Ser 34, Thr 65	Gly 37	Ala 171	Gln 71, Tyr 113
3	Gly 61	Gln 64	Thr 71	Glu 44	Thr 113	Ser 20
4		Glu 34, Gly 47	Glu 11, Asp 20	Asp 26		Asn 52
5		Gln 24, Pro 21	His 9, Pro 53	Arg 21		–
6			Thr 8, Glu 28	Thr 16		Pro 35
7				Gly 33		Pro 57
8				–		Tyr 14
9				Tyr 15		Val 10

[a] Values are yields reported in pmol; – indicates a blank cycle.
[b] Estimated amount loaded in pmol: P1a, 100; P2a, 130; P3a, 180; P′1a, 150; P′2a, 200; P′3a, 230.

Table 2. *Calculated and determined molecular masses of the cystine peptides of Er-2*

Peptide	Proposed structure[a]	Calculated mass	Determined mass[b]
P1a	16 18 M-C-G | M-T-C 26 28	660.8	661.2
P2a	3 7 M-T-C-E-Q | Y-C-Q-G-P 19 23	1,175.3	1,175.1,[d] 1,175.5[e]
P3a	10 15 A-S-C-E-H-T | I-T-T-D-P-E-C 31 37	1,422.5 (1,055.2)[c]	1,422.0,[f] 1,422.9,[f] 1,055.7

[a] Residue numbers are for the E*r*-2 sequence (Raffioni et al., 1992) (see Fig. 4).
[b] Molecular masses determined by mass spectroscopy.
[c] Molecular weight for the peptide lacking the E-H-T sequence.
[d] Calculated from the diprotonated species (m/z = 588.5) observed in the electrospray spectrum.
[e] Determined from matrix-assisted laser desorption mass spectrum.
[f] Calculated from the di- and triprotonated species (712.0 and 475.3) observed in the electrospray spectrum.

$(M + Cu)^+$. This corresponded to a mass of 1,175.5, which was in good accord with the calculated value. A trace amount of P3a (incompletely removed by the HPLC purification step) can also be seen at an m/z of 1,422.5. The electrospray mass spectrum of P3a showed a very weak monoprotonated ion $(M + H)^+$ at m/z 1,055.7 (Fig. 3B) (visible only on an expanded scale), and a more intense diprotonated ion $(M + 2H)^{2+}$ at m/z 528.7 (Fig. 3B), which corresponded to isotopically averaged masses of 1,054.7 and 1,055.4, respectively. These values do not agree with the calculated mass of 1,422.5 for P3a (Table 2). However, they do correspond well to a truncated form of P3a (calculated mass 1,055.1) that would result from the loss of the Glu-His-Thr tripeptide. In addition, diprotonated $(M + 2H)^{2+}$ and triprotonated $(M + 3H)^{3+}$ species were observed at m/z values of 712.0 and 475.3, respectively (Fig. 3B). These species correspond to masses of 1,422.0 and 1,422.9, respectively, which are in good agreement with the calculated mass of the untruncated peptide. The data suggest that the sample is composed of two peptide species. The carboxy-terminal truncated form of P3a was not anticipated, as fractionation of P3a by a reverse-phase HPLC utilizing two distinct mobile phases resulted in a single, well-resolved peak (Figs. 1, 2). However, peptide sequencing of P3a does suggest the presence of the truncated form, since the yield of Glu, His, and Thr in the fourth, fifth, and sixth cycles is clearly lower than the second sequences. Unfortunately the yield of phenylthiohydantoin (PTH)-His and PTH-Thr is usually low, which makes this interpretation less secure. Nonetheless, the molecular mass measurements are entirely consistent with the proposed sequences.

Assignment of disulfide bond pairing in Er-2

The sequence and molecular mass analyses identify three unique cystine peptides that allow the direct assignment of the disulfide pairs as Cys(I)–Cys(IV), Cys(II)–Cys(VI), and Cys(III)–Cys(V). The detection of only three cystine-containing peptides and their recovery in good yield (54, 48, and 61% for P1a, P2a, and P3a, respectively) strongly support the view that these pairings represent the alignments in the native molecule. In addition, as shown in Figure 4, four non-cystine-containing peptides were isolated and sequenced (data not shown) that account for all of the molecule but the amino-terminal dipeptide. Thus, despite the extensive digestion employed, the pattern remained focused and consistent with thermolysin specificity (Bradshaw, 1969).

Assignment of disulfide bond pairing in Er-1

An identical approach was used to determine the disulfide bond pattern of E*r*-1. The native molecule was subjected to thermolytic digestion for 24 h, and the resulting mixture was fractionated in two steps utilizing identical reverse-phase columns and buffer systems as described for E*r*-2 (data not shown). The cystine-containing peptides subsequently identified were characterized by amino

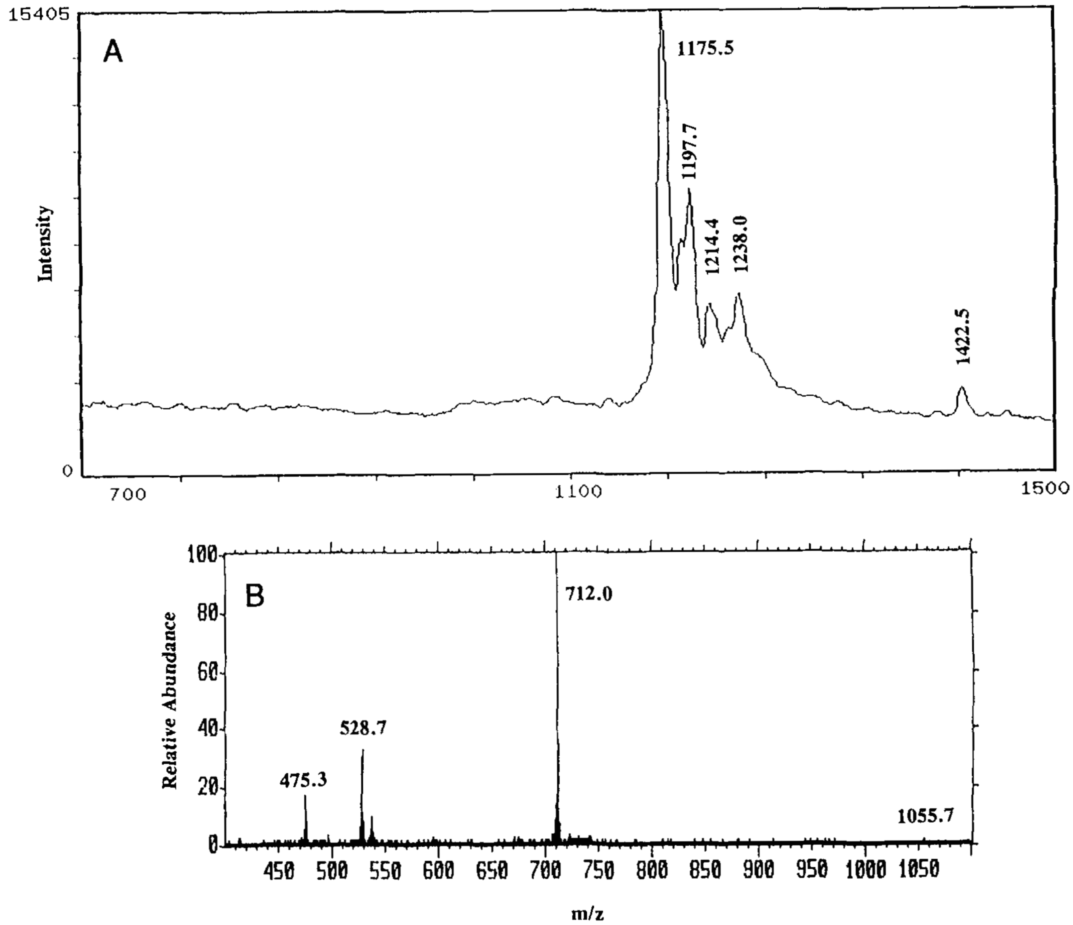

Fig. 3. Mass spectra of cystine peptides P2a and P3a obtained from thermolytic digestion of E*r*-2. Peptide P2a (**A**) was analyzed by matrix-assisted laser desorption mass spectrometry on a time-of-flight mass spectrometer as described under Materials and methods. Peptide P3a (**B**) was brought to a concentration of 5 pmol/μL with H_2O:CH_3OH:acetic acid (45:50:5) and analyzed by electrospray mass spectrometry.

acid composition (data not shown), peptide sequence (Table 1), and mass spectrometric analysis (Table 3). The calculated molecular weights of the cystine peptides were found to be in excellent agreement with the corresponding observed masses. As with E*r*-2, the disulfide bond assignments of Cys(I)-Cys(IV), Cys(II)-Cys(VI), and Cys(III)-Cys(IV) could be made directly from these peptides. The recovery of peptides P′1a, P′2a, and P′3a in yields of 63, 41, and 57%, respectively, and the isolation of several additional non-cystine peptides (Fig. 4) indicated that the cleavages in the E*r*-1 molecule by thermolysin were also consistent with the proposed assignment. Thus, the disulfide bond pairing determined for E*r*-1 was identical to that of E*r*-2.

Discussion

Assignment of disulfide bonds requires the characterization of unique cystine-containing peptides, which must be generated by introducing cleavages in the protein molecule between each half-cystine residue (except when sequentially adjacent half-cystines form a closed loop). In E*r*-2, the four Met residues and the acid-labile Asp-Pro imide bond (Piszkiewicz et al., 1970) at residues 34–35 suggested a possible route through cleavage by CNBr in 70% formic acid. This procedure in fact did allow the tentative assignment of Cys(II)-Cys(VI) by fast atom bombardment mass spectrometric (FABMS) analysis of an unfractionated mixture (Bradshaw et al., 1990). However, the lack of a cleavage point between Cys(III) and Cys(IV) ultimately prevented any further assignments. Thus, thermolysin was selected as an alternative, based on its specificity (Matsubara et al., 1965; Bradshaw, 1969) and the absence of cysteine/cystine residues in the primary structure (Titani et al., 1972), the presence of which could obscure the analysis of pheromone-derived cystine-linked peptides. Hydrophobic residues, the primary targets for thermolysin, are found between each half-cystine residue in E*r*-2. In addition, cleavage at Glu was observed in peptides P3a, P′1a, and P′2a. Although

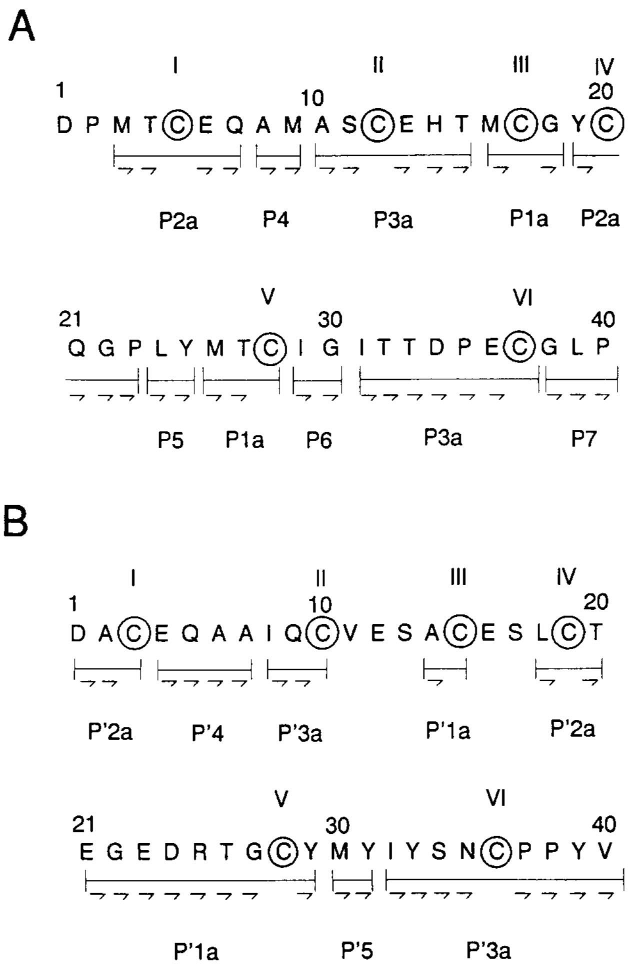

Fig. 4. Amino acid sequences of E*r*-1 and E*r*-2 and their peptides obtained from thermolysin digestion. The primary sequences of E*r*-2 (**A**) and E*r*-1 (**B**) were aligned to place the half-cystines, indicated by Roman numerals, in equivalent positions. The thermolytic peptides were characterized by amino acid composition analysis and automated Edman degradation (arrows).

Glu residues generally do not represent a favorable cleavage site for thermolysin, cleavage at this site at pH 6.5 has been reported (Guy et al., 1971). Minimization of disulfide bond interchange to maintain the integrity of the native pairings was also a major concern. Low pH and temperature restrict such exchanges (Spackman et al., 1960), and the digests were accordingly performed at pH 6.0 and 37 °C (although they are not the optimal conditions for the enzyme).

The amino acid sequences of E*r*-2 (Raffioni et al., 1992) and E*r*-1 (Raffioni et al., 1988) with the proposed disulfide bridge structures are shown in Figure 5. Clearly, the disulfide bond patterns suggest a three-dimensional organization that can be defined by three loops. Loop I, formed by the link between Cys(I) and Cys(IV), involves most of the amino-terminal part of the protein (including two internal half-cystines). Loop II, defined by the Cys(III)–Cys(V) bond, overlaps part of Loop I and includes one internal half-cystine. Loop III is bounded by Cys(III)–Cys(V) and Cys(II)–Cys(VI). The variability in these loops among the five unique *E. raikovi* pheromones is summarized in Figure 6 and Table 4. These comparisons are based on the assumption that E*r*-10, E*r*-11, and E*r*-20 (Raffioni et al., 1992) possess disulfide bond pairings equivalent to those determined in this study (Fig. 6). Loop I displays the narrowest range in size, being 16 or 17 residues in all molecules. Loop II shows the greatest variability, being comprised of 10–14 residues. Loop III is about the same size as Loop I and is slightly more variable (15–17 residues). The conservation in size of Loop I reflects the invariability of segments A and B (Fig. 6), which may reflect a common property of these molecules such as dimer formation. The greater variability observed in Loops II and III, particularly as manifested in the lengths of segments C, D, and E, suggests that they may be involved in functions more unique to each molecule, such as receptor recognition.

Secondary structure evaluations by Chou–Fasman analyses of E*r*-1 and E*r*-2 (Bradshaw et al., 1990) suggested an alpha-helical segment at the amino-terminus, with little secondary structure other than beta-turns predicted for the remainder of the molecule. This helical region is not, however, expected to comprise many residues, since the Cys(I)–Cys(IV) and Cys(II)–Cys(VI) bonds and the size of Loop I might be expected to restrict the propagation of any secondary structure in that particular region of the molecule. Similarly, hydropathy profiles have in-

Table 3. *Calculated and determined molecular masses of the cystine peptides of Er-1*

Peptide	Proposed structure[a]	Calculated mass	Determined mass[b]
P′1a	14 15 A-C | E-G-E-D-R-T-G-C-Y 21 29	1,219.3	1,219.2
P′2a	1 3 D-A-C | L-C-T 18 20	640.7	640.3
P′3a	8 10 I-Q-C | I-Y-S-N-C-P-P-Y-V 32 40	1,415.4	1,415.6

[a] Residue numbers are for the E*r*-1 sequence (Raffioni et al., 1988) (see Fig. 4).
[b] Molecular masses determined by mass spectroscopy.

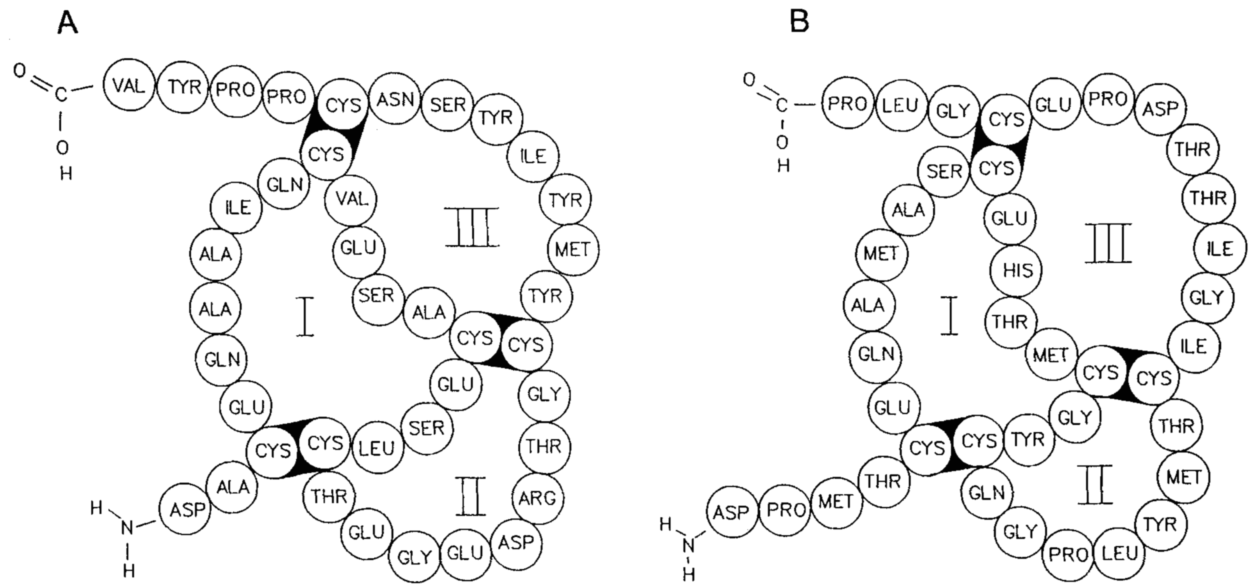

Fig. 5. Schematic diagram of the amino acid sequences of the pheromones E*r*-1 (**A**) and E*r*-2 (**B**) that emphasizes the basic three-loop structure of the *E. raikovi* pheromones.

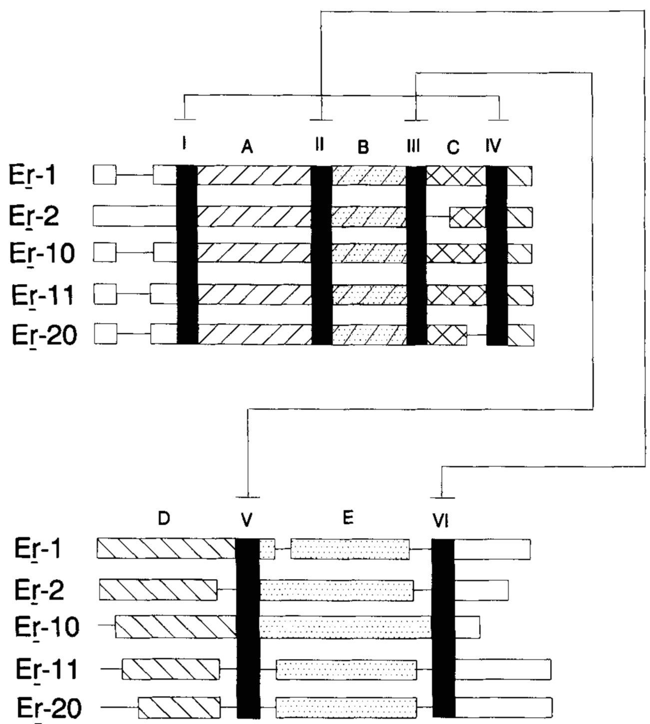

Fig. 6. Schematic presentation of the covalent structures of five members of the *E. raikovi* pheromone family. The proposed conserved disulfide bond pairings are indicated by lines connecting the half-cystines (Roman numerals). Boxes indicate sequences aligned by the insertion of gaps (solid lines) (Raffioni et al., 1992). The five segments between the half-cystines are designated A–E. Loop I: A + B + C . Loop II: C + D . Loop III: B + E . Residues shared by Loops I and III are indicated by ; residues shared by Loops I and II are defined as .

dicated dramatic variations in the distribution of hydrophobic residues in these molecules (Raffioni et al., 1992). These are also manifested in the loop structures, particularly in Loops II and III. This is well illustrated by the comparison of E*r*-1 and E*r*-2 (Fig. 5). E*r*-1 (as well as E*r*-10, which has a very similar hydropathy profile) is characterized by a pronounced hydrophilic peak corresponding to segment D (Fig. 6) that makes up most of Loop II; E*r*-2 has, in contrast, a markedly hydrophobic sequence for segments C and D, the two components of Loop II. The opposite is true of Loop III. Segment E of E*r*-1 is quite hydrophobic, whereas in E*r*-2 it is mildly hydrophilic. These differences are underscored by the charge distribution in the loops (Table 4). In E*r*-1, Loop II has five charged residues as compared to none for E*r*-2. Conversely, Loop III has only one charge in E*r*-1; there are four charges in Loop III of E*r*-2. Both have three charges in Loop I. Interestingly, E*r*-11 and E*r*-20 show a low charge distribution in all of the loops, reflecting the much lower overall charge density on these molecules. This is reflected in the decidedly more hydrophobic profiles observed in both molecules (Raffioni et al., 1992).

The assignment of the disulfide bond pairs of two members in the *E. raikovi* pheromone family, E*r*-1 and E*r*-2, is part of our continuing efforts to elucidate the structure–function relationships in this class of extracellular signalling molecules. The identical pairing of disulfide bonds found in these two molecules provides additional evidence for similar overall structures that may be shared in this family of pheromones. The functional role of the three proposed loop structures with regard to receptor in-

Table 4. *Number of residues in loop structures of E. raikovi pheromones*

	Loop I (A + B + C)[a]		Loop II (C + D)[a]		Loop III (B + E)[a]	
	Residues	Charges[b]	Residues	Charges[b]	Residues	Charges[b]
E*r*-1	17	3/10	14	5/10	15	1/10
E*r*-2	16	3/8	12	0/8	16	4/8
E*r*-10	17	3/10	13	4/10	17	2/10
E*r*-11	17	0/5	12	1/5	15	1/5
E*r*-20	16	2/5	10	1/5	15	1/5

[a] Loops are defined as including all half-cystine residues that provide covalent bonds necessary to give a closed structure (see Figs. 5, 6).

[b] Positive and negative charges are summed (His is treated as charged) and the value listed relative to the total side-chain charges.

teraction and protein dimer formation can best be probed through the combined use of chemical modification and site-directed mutagenesis. Efforts, which are in progress, to determine the three-dimensional crystal structure of E*r*-1 by X-ray crystallography (Anderson et al., 1990) and the solution structure by two-dimensional NMR spectroscopy should also be facilitated by the disulfide bond assignments elucidated in this study.

Materials and methods

Materials

E*r*-1 and E*r*-2 were obtained from the clones 1aF13 (*mat-1/mat-1*) and 1bF13 (*mat-2/mat-2*) of *E. raikovi*, respectively, and prepared as previously described (Concetti et al., 1986; Raffioni et al., 1987). Chemicals and reagents were purchased as follows: thermolysin from Boehringer Mannheim; 2-[*N*-morpholino]ethanesulfonic acid (MES), trifluoroacetic acid (TFA), mercaptoethanol, acetic acid, triethylamine (TEA), phenylisothiocyanate (PITC), and 6 N hydrochloric acid from Pierce; dithiothreitol from Calbiochem; HPLC-grade CH_3CN and CH_3OH from Burdick and Jackson; HPLC grade water from a Millipore Milli-Q purification system; and protein sequencer chemicals from Applied Biosystems. All other chemicals were of reagent grade.

Thermolysin digestion

Eight nanomoles of lyophilized protein was dissolved in 50 μL of 0.1 M MES, pH 6.0, 0.1 mM $CaCl_2$, and 0.002% NaN_3. Thermolysin was added at an enzyme:substrate ratio of 1:30 and the mixture incubated under nitrogen at 37 °C. At 6, 12, and 24 h, 2-μL aliquots were removed and brought to 20 μL with 0.1% TFA to determine the extent of digestion. After 24 h, the reaction was terminated by adding 5 μL of 1 M TFA.

Reduction of cystine-linked peptides

Six microliters of the 24-h time point sample was dried under vacuum, brought to 50 μL with 0.2 M Tris, pH 8.5, and sparged with nitrogen for 5 min. Undiluted β-mercaptoethanol (14 M) was added to a concentration of 0.1% (v/v), and the reaction was carried out under nitrogen at 37 °C for 5 h. One-third of the reduced material was analyzed by microbore reverse-phase HPLC using a CH_3CN:H_2O:TFA buffer system.

Peptide separation

The 6-, 12-, and 24-h aliquots of the thermolysin digest were fractionated by microbore reverse-phase HPLC on an Applied Biosystems model 140 system equipped with an Isco CV4 UV-VIS micro flow-cell detector and an Applied Biosystems C18 PTH column (2.1 × 220 mm). A linear gradient from 2 to 50% buffer B was run over a period of 30 min at a flow rate of 200 μL/min. Buffer A was 0.1% TFA and buffer B was CH_3CN:H_2O (4:1) containing 0.08% TFA. The absorbance at 214 nm was monitored. The remainder of the digest was fractionated on a Hitachi Instruments model L-6200 HPLC system using a Vydac C18 column (4.6 × 250 mm) with a linear gradient from 2 to 50% buffer B developed over 40 min at a flow rate of 0.7 mL/min. Cystine peptides, identified by reduction and amino acid composition analysis, were collected, dried under vacuum, and fractionated by microbore reverse-phase HPLC using a linear gradient from 2 to 50% buffer D developed over 30 min at a flow rate of 200 μL/min. Buffer C was 20 mM TEA-acetate; buffer D was CH_3CN:20 mM TEA-acetate (1:1). Absorbance was monitored at 214 nm.

Amino acid analysis

Samples were hydrolyzed in the vapor phase with 6 N HCl at 110 °C for 16 h, dried, and derivatized using

CH_3OH:TEA:H_2O:PITC in a ratio of 9:2:1:1 (v/v) for 15 min at room temperature. The phenylthiocarbamyl derivatives were analyzed by reverse-phase HPLC on a Hewlett-Packard model 1090 utilizing 150 mM sodium acetate/CH_3CH, pH 6.0, as the mobile phase (Heinrikson & Meredith, 1984). The calculated yields of cystine peptides obtained by thermolytic digestion were based on the recovery of cystine from 500 pmol each of E*r*-1 and E*r*-2, which had been hydrolyzed for 20 h.

Amino acid sequence determination

Sequence analysis was performed using an Applied Biosystems pulsed-liquid Sequencer equipped with an on-line model 120 PTH analyzer.

Mass spectrometry

Mass spectrometric analyses were performed by both matrix-assisted laser desorption and electrospray mass spectrometry using time-of-flight mass spectrometers constructed at the Rockefeller University (Beavis & Chait, 1989) as previously described (Beavis & Chait, 1990; Chowdhury et al., 1990).

Acknowledgments

We thank Steven Disper, Richard Kendall, Jose Sy, and Kevin Burke for helpful discussions. This work was supported by a Ministero dell'Universita e della Ricerca Scientifica e Tecnologica e CNR grant to P.L. and U.S.P.H.S. research grants RR00862 and GM38274 to B.T.C. and DK32465 to R.A.B.

References

Akada, R. (1986). Partial characterization of the mating-inducing factors (gamones) produced by mating type IV cells in the ciliate *Euplotes patela* syngen 2. *J. Exp. Zool. 237*, 287–290.

Anderson, D., Raffioni, S., Luporini, P., Bradshaw, R.A., & Eisenberg, D. (1990). Crystallization of the *Euplotes raikovi* mating pheromone E*r*-1. *J. Mol. Biol. 216*, 1–2.

Beavis, R.C. & Chait, B.T. (1989). Factors affecting the ultraviolet laser desorption of proteins. *Rapid Commun. Mass Spectrom. 3*, 233–237.

Beavis, R.C. & Chait, B.T. (1990). High accuracy molecular mass determination of proteins using matrix-assisted laser desorption mass spectrometry. *Anal. Chem. 62*, 1836–1840.

Bradshaw, R.A. (1969). The amino acid sequence of bovine carboxypeptidase A. III. Specificity of peptide bond cleavage by thermolysin and the complete sequence of cyanogen bromide fragment F_{III}. *Biochemistry 8*, 3871–3877.

Bradshaw, R.A., Raffioni, S., Luporini, P., Chait, B.T., Lee, T., & Shively, J. (1990). Amino acid sequence-mass spectrometric analyses of mating pheromones of the ciliate *Euplotes raikovi*. In *Analytical Biotechnology* (Horvath, C. & Nikelly, J.G., Eds.), pp. 154–161. American Chemical Society, Washington, D.C.

Chowdhury, S.K., Katta, V., & Chait, B.T. (1990). An electrospray ionization mass spectometer with new features. *Rapid Commun. Mass Spectrom. 4*, 81–87.

Concetti, T., Raffioni, S., Miceli, C., Barra, D., & Luporini, P. (1986). Purification to apparent homogeneity of the mating pheromone of *mat-1* homozygous *Euplotes raikovi. J. Biol. Chem. 261*, 10582–10586.

Guy, O., Shapanka, R., & Greene, L.J. (1971). The structure of the bovine pancreatic secretory trypsin inhibitor — Kazal's inhibitor. *J. Biol. Chem. 246*, 7740–7747.

Heinrikson, R.L. & Meredith, S.C. (1984). Amino acid analysis of reverse-phase high-performance liquid chromatography: Precolumn derivatization with phenylisothiocyanate. *Anal. Biochem. 136*, 65–74.

Kubota, T., Tokoroyama, T., Tsukuda, Y., Koama, H., & Miyake, A. (1973). Isolation and structure determination of Blepharismin, a conjugation initiating gamone in the ciliate *Blepharisma. Science 179*, 400–402.

Luporini, P. & Miceli, C. (1986). Mating pheromones. In *Molecular Biology of Ciliated Protozoa* (Gall, J.G., Ed.), pp. 263–299. Academic Press, Inc., New York.

Luporini, P., Raffioni, S., Concetti, A., & Miceli, C. (1986). The ciliate *Euplotes raikovi* heterozygous at the *mat* genetic locus coreleases two individual species of mating pheromone: Genetic and biochemical evidence. *Proc. Natl. Acad. Sci. USA 83*, 2889–2893.

Matsubara, H., Singer, A., Sasaki, R., & Jukes, T.H. (1965). Observations on the specificity of a thermostable bacterial protease "thermolysin." *Biochem. Biophys. Res. Commun. 21*, 242–247.

Meyer, F., Schmidt, H.J., Plumper, E., Hasalik, A., Mersmann, G., Meyer, H.E., Engstrom, A., & Heckman, K. (1991). UGA is translated as cysteine in pheromone 3 of *Euplotes octocarinatus. Proc. Natl. Acad. Sci. USA 88*, 3758–3761.

Miceli, C., Concetti, A., & Luporini, P. (1983). Isolation of the mating inducing factor of the ciliate *Euplotes. Exp. Cell Res. 149*, 593–598.

Miceli, C., La Terza, A., Bradshaw, R.A., & Luporini, P. (1991). Structural characterization of mating pheromone precursors of the ciliate *Euplotes raikovi. Eur. J. Biochem. 202*, 759–764.

Miceli, C., La Terza, A., Bradshaw, R.A., & Luporini, P. (1992). Identification and structural characterization of a cDNA clone encoding a membrane-bound protein form of E*r*-1 in the ciliate protozoan *Euplotes raikovi. Proc. Natl. Acad. Sci. USA 89*, 1988–1992.

Miceli, C., La Terza, A., & Melli, M. (1989). Isolation and structural characterization of cDNA clones encoding the mating pheromone E*r*-1 secreted by the ciliate *Euplotes raikovi. Proc. Natl. Acad. Sci. USA 86*, 3016–3020.

Miyake, A. & Beyer, J. (1974). Blepharmone: A conjugation-inducing glycoprotein in the ciliate *Blepharisma. Science 185*, 621–623.

Ortenzi, C., Miceli, C., Bradshaw, R.A., & Luporini, P. (1990). Identification and characterization of an autocrine pheromone receptor in the protozoan ciliate *Euplotes raikovi. J. Cell Biol. 111*, 607–614.

Parfenova, E.V., Afon'kin, S.Y., Yudin, A.L., & Etingof, R.N. (1989). Characterization and partial purification of mating pheromone excreted by mating type II cells of the ciliate *Dileptus anser. Acta Protozool. 28*, 11–21.

Piszkiewicz, D., Landon, M., & Smith, E.L. (1970). Anomalous cleavage of aspartyl-proline peptide bonds during amino acid sequence determinations. *Biochem. Biophys. Res. Commun. 40*, 1173–1178.

Raffioni, S., Luporini, P., & Bradshaw, R.A. (1989). Purification, characterization and amino-acid sequence of the mating pheromone E*r*-10 of the ciliate *Euplotes raikovi. Biochemistry 28*, 5250–5256.

Raffioni, S., Luporini, P., Chait, B.T., Disper, S.S., & Bradshaw, R.A. (1988). Primary structure of the mating pheromone E*r*-1 of the ciliate *Euplotes raikovi. J. Biol. Chem. 263*, 18152–18159.

Raffioni, S., Miceli, C., Concetti, A., Barra, D., & Luporini, P. (1987). Purification and characterization of new mating pheromones of the ciliate *Euplotes raikovi. Exp. Cell Res. 172*, 417–424.

Raffioni, S., Miceli, C., Vallesi, A., Chowdhury, S.K., Chait, B.T., Luporini, P., & Bradshaw, R.A. (1992). The primary structure of *E. raikovi* pheromones: A comparison of five sequences of pheromones from cells with variable mating interactions. *Proc. Natl. Acad. Sci. USA 89*, 2071–2075.

Schulze Dieckhoff, H.S., Freiburg, M., & Heckmann, K. (1987). The isolation of gamones 3 and 4 of *Euplotes octocarinatus. Eur. J. Biochem. 168*, 89–94.

Spackman, D.H., Stein, W.H., & Moore, S. (1960). The disulfide bonds of ribonuclease. *J. Biol. Chem. 235*, 648–659.

Titani, K., Hermodson, M.A., Ericsson, L.H., Walsh, K.A., & Neurath, H. (1972). Amino acid sequence of thermolysin. *Nature New Biol. 238*, 35–37.

Weischer, A., Freiburg, M., & Heckmann, K. (1985). Isolation and characterization of gamone 1 of *Euplotes octocarinatus. FEBS Lett. 191*, 176–180.