Protein Science (1996), 5:1512–1522. Cambridge University Press. Printed in the USA.

The NMR solution structure of the pheromone Er-11 from the ciliated protozoan *Euplotes raikovi*

PETER LUGINBÜHL,[1] JIHUI WU,[1] OLIVER ZERBE,[1] CLAUDIO ORTENZI,[2] PIERANGELO LUPORINI,[2] AND KURT WÜTHRICH[1]

[1] Institut für Molekularbiologie und Biophysik Eidgenössische Technische Hochschule, CH-8093 Zürich, Switzerland
[2] Department of Molecular, Cellular and Animal Biology, University of Camerino, 62032 Camerino (MC), Italy

(RECEIVED March 28, 1996; ACCEPTED May 6, 1996)

Abstract

The NMR solution structure of the pheromone Er-11, a 39-residue protein from the ciliated protozoan *Euplotes raikovi*, was calculated with the distance geometry program DIANA from 449 NOE upper distance constraints and 97 dihedral angle constraints, and the program OPAL was employed for structure refinement by molecular mechanics energy minimization in a water bath. For a group of 20 conformers used to characterize the solution structure, the average of the pairwise RMS deviations from the mean structure calculated for the backbone heavy atoms N, C^{α}, and C′ of residues 2–38 was 0.30 Å. The molecular architecture is dominated by an up-down-up bundle of three short helices with residues 2–9, 12–19, and 22–32, which is closely similar to the previously determined structures of the homologous pheromones Er-1, Er-2, and Er-10. This finding provides structural evidence for the capability shown by these pheromones to compete with each other in binding reactions to their cell-surface receptors.

Keywords: cellular recognition; NMR structure; pheromone Er-11 from *Euplotes raikovi*

Families of polypeptide growth factors play a key role in the control of cell survival, proliferation, and differentiation. Insights into their structure and regulatory activity are mostly derived from studies of cell systems provided by advanced vertebrates (James & Bradshaw, 1984; Sporn & Roberts, 1990). However, single-cell eukaryotes such as the protozoan ciliate *Euplotes raikovi* have already evolved signaling systems that may be regarded as attractive paradigms for investigating signal transduction mechanisms at a primordial, apparently simpler level (Luporini et al., 1995).

Like numerous other species of ciliates (Dini & Nyberg, 1983), *E. raikovi* is represented in nature by a virtually unlimited number of morphologically similar but chemically distinct cell (mating) types. The chemical markers of these cell types are polypeptides of 37–40 amino acid residues, usually referred to as pheromones, and denoted Er-1, Er-2, and so forth (Luporini et al., 1995). Each one is encoded by one of a series of alleles at the highly polymorphic *mat* (or mating type) locus, and acts as either a mitogenic or a mating-inducing factor, according to whether it binds to cells in autocrine or paracrine fashion (Ortenzi & Luporini, 1995).

To establish a structural basis for systematic studies of the evolutionary and functional relationships of this pheromone system, NMR and crystallographic analyses of a series of pheromones have been undertaken. Previously, the conformations of Er-1 (Mronga et al., 1994; Weiss et al., 1995), Er-10 (Brown et al., 1993), and Er-2 (Ottiger et al., 1994) have been determined, and a detailed comparison of these structures has been reported (Luginbühl et al., 1994). In this article, we describe the NMR solution structure of Er-11 and show that it fully conforms with the three-helix-bundle architecture of the three other pheromone structures, although the only invariant amino acid residues among these four species are the N-terminal aspartic acid and six cysteines that are involved in three intrachain disulfide bonds (Stewart et al., 1992).

Results

The determination of the Er-11 structure was performed on preparations with natural isotope distribution. The aqueous solution conditions for the structure determination were: protein concentration = 2.5 mM, pH 5.0, $T = 27$ °C. The amide proton exchange measurements were performed at 13 °C.

Reprint requests to: Kurt Wüthrich, Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule, CH-8093 Zürich, Switzerland.

Abbreviations: 2QF, two-quantum-filtered; TPPI, time-proportional phase incrementation; COSY, 2D correlation spectroscopy; TOSCY, 2D total correlation spectroscopy; E. COSY, 2D exclusive correlation spectroscopy; NOESY, 2D NOE spectroscopy; TMS, tetramethylsilane; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate; REDAC, redundant dihedral angle constraint, Er-x, *Euplotes raikovi* pheromone of mating type x.

^{1}H NMR assignments and identification of regular secondary structure elements

The resonance assignments were based on ^{1}H spin system identification with homonuclear and ^{13}C-^{1}H heteronuclear correlation spectroscopy, and on sequential NOE connectivities (Billeter et al., 1982; Wagner & Wüthrich, 1982; Wider et al., 1982). The Pro spin systems were inserted into the correct sequence positions on the basis of $d_{\alpha\delta}$, $d_{N\delta}$, $d_{\alpha N}$, and $d_{\delta N}$ cross peaks observed using 2D NOESY experiments (Wüthrich, 1986). The Pro 36–Pro 37 peptide bond was found to be in the *cis*-conformation based on the observation of the sequential $C^\alpha H$-$C^\alpha H$ NOE cross peak. [^{13}C,^{1}H]-COSY at natural abundance was also used to identify the $^{13}C^\alpha$ chemical shifts. The degenerate methylene proton chemical shifts of $C^\beta H_2$ of Asn 5, Leu 22, Glu 24, and Met 31, and of $C^\gamma H_2$ of Glu 2, Gln 9, Glu 24, Glu 35, Pro 36, and Pro 37, were established by the fact that only one cross peak was observed in [^{13}C,^{1}H]-COSY (Table 1). The pairs of side-chain amide protons of Asn were assigned individually from the relative intensity of the intraresidual NOEs to the individual β-methylene protons (Otting, 1987; Sevilla-Sierra et al., 1994). Overall, complete assignments for the nonlabile protons and the backbone and side-chain amide protons were obtained (Table 1).

Figure 1 affords a survey of the NMR parameters used for the sequential assignments and the initial secondary structure determination. The patterns of successive d_{NN} connectivities and

Table 1. *$^{13}C^\alpha$ and ^1H chemical shifts for Er-11 at pH 5.0 and 27 °C*

	Chemical shift (ppm)				
Residue	$^{13}C^\alpha$	NH	$C^\alpha H$[a]	$C^\beta H$[a]	Others[a]
Asp 1	52.61		4.16	2.98, 2.84	
Glu 2	59.06	9.17	3.88	1.43, 1.01	$C^\gamma H_2$ 1.93, 1.93
Cys 3	57.41	8.18	4.20	2.99, 3.25	
Ala 4	54.81	7.95	4.07	1.38	
Asn 5	55.06	8.00	4.35	2.76, 2.76	$N^\delta HE$ 7.60, $N^\delta HZ$ 6.81
Ala 6	55.28	7.83	3.86	1.63	
Ala 7	53.76	6.97	3.90	1.52	
Ala 8	54.14	7.67	4.08	1.49	
Gln 9	55.21	7.60	4.25	2.34, 1.92	$C^\gamma H_2$ 2.41, 2.41
Cys 10	55.08	8.15	4.25	3.60, 3.28	
Ser 11	55.00	7.49	4.98	3.48, 3.79	
Ile 12	63.10	8.30	3.46	2.11	$C^\gamma H_2$ 1.59, 1.35; $C^\gamma H_3$ 0.83; $C^\delta H_3$ 0.59
Thr 13	66.73	8.96	3.86	3.90	$C^\gamma H_3$ 1.13
Leu 14	56.83	8.40	4.08	1.83, 1.53	$C^\gamma H$ 1.81; $C^\delta H_3$ 0.95, 0.85
Cys 15	56.81	8.37	4.52	3.32, 2.87	
Asn 16	55.92	7.70	4.23	2.82, 2.54	$N^\delta HE$ 7.29, $N^\delta HZ$ 6.87
Leu 17	59.86	7.56	4.05	1.23, 0.96	$C^\gamma H$ 1.07; $C^\delta H_3$ 0.69, 0.53
Tyr 18	58.61	7.57	4.97	2.53, 3.06	$C^\delta H$ 7.10, 7.10; $C^\epsilon H$ 6.79, 6.79
Cys 19	58.00	7.68	4.60	3.23, 2.69	
Gly 20	45.48	9.42	3.61, 4.19		
Pro 21	64.11		4.54	2.41, 2.06	$C^\gamma H_2$ 2.08, 2.03; $C^\delta H_2$ 4.09, 3.81
Leu 22	53.79	7.93	4.73	1.87, 1.87	$C^\gamma H$ 1.60; $C^\delta H_3$ 0.95, 0.83
Ile 23	67.29	7.39	3.12	1.64	$C^\gamma H_2$ 1.94, 1.09; $C^\gamma H_3$ 0.72; $C^\delta H_3$ 0.96
Glu 24	58.31	8.65	3.98	2.02, 2.02	$C^\gamma H_2$ 2.30, 2.30
Ile 25	63.09	7.42	3.81	1.86	$C^\gamma H_2$ 1.48, 1.28; $C^\gamma H_3$ 0.88; $C^\delta H_3$ 0.82
Cys 26	57.64	7.94	4.10	2.90, 3.27	
Glu 27	59.88	8.61	3.52	2.74, 2.19	$C^\gamma H_2$ 1.98, 1.88
Leu 28	57.83	8.34	3.98	1.74, 1.56	$C^\gamma H$ 1.68; $C^\delta H_3$ 0.82, 0.81
Thr 29	67.06	7.98	3.75	4.41	$C^\gamma H_3$ 1.15
Val 30	66.32	8.08	3.64	1.95	$C^\gamma H_3$ 0.98, 0.14
Met 31	58.07	8.02	3.81	2.02, 2.02	$C^\gamma H_2$ 2.52, 2.31; $C^\epsilon H_3$ 2.06
Gln 32	60.52	7.91	4.19	2.11, 2.01	$C^\gamma H_2$ 2.49, 2.34
Asn 33	52.93	7.68	4.88	2.87, 2.83	$N^\delta HE$ 7.56, $N^\delta HZ$ 7.12
Cys 34	56.09	8.22	5.08	3.68, 3.21	
Glu 35	54.07	8.74	4.88	2.05, 1.90	$C^\gamma H_2$ 2.46, 2.46
Pro 36	61.28		3.32	1.53, 1.43	$C^\gamma H_2$ 1.83, 1.83; $C^\delta H_2$ 3.77, 3.70
Pro 37	61.52		4.09	2.21, 2.02	$C^\gamma H_2$ 1.84, 1.84; $C^\delta H_2$ 3.52, 3.37
Phe 38	54.94	8.33	5.17	3.47, 2.96	$C^\delta H$ 7.10, 7.10; $C^\epsilon H$ 7.35, 7.35
Ser 39	59.76	8.30	4.48	4.00, 3.96	

[a] Chemical shifts of hydrogen atoms for which stereospecific assignments were obtained are underlined, with the atom with the lower branch number shown first, e.g., β^2H or δ^1CH_3. Methylene hydrogens with degenerate chemical shifts for which direct evidence of the degeneracy was obtained from a heteronuclear [^{13}C,^{1}H]-COSY spectrum are listed with two identical chemical shifts. The reference for the ^{13}C and ^{1}H chemical shifts is internal DSS.

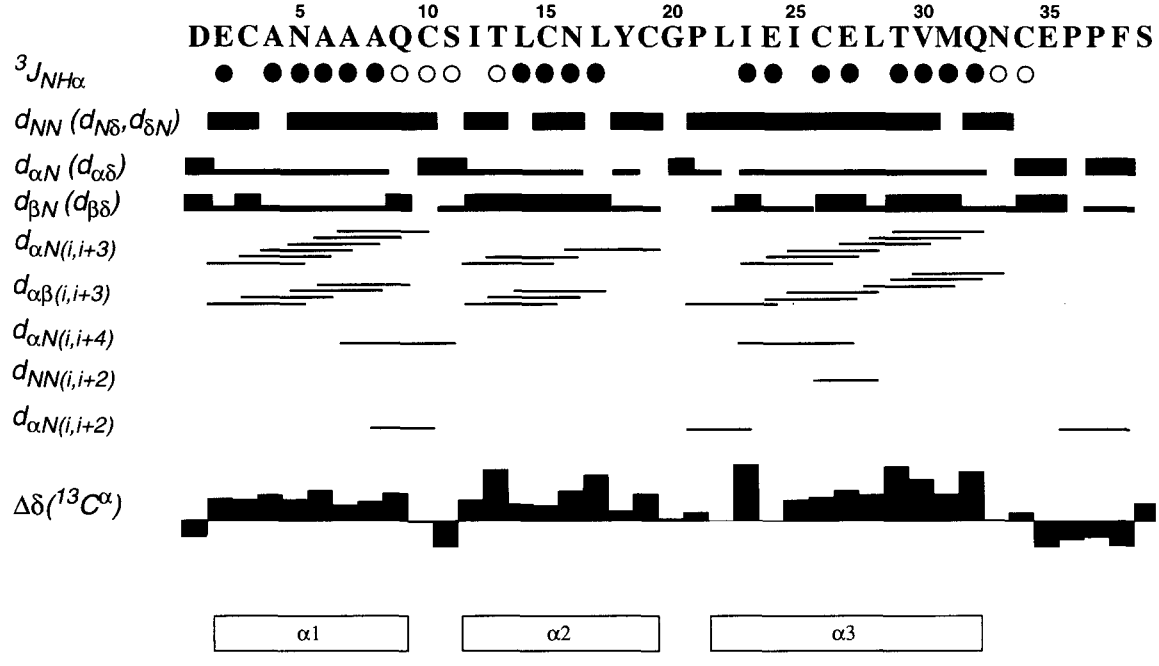

Fig. 1. Amino acid sequence of E*r*-11 and survey of the sequential and medium-range NOEs, spin-spin coupling constants $^3J_{HN\alpha}$, and conformation-dependent $^{13}C^{\alpha}$ chemical shifts. Below the sequence, open and filled circles identify residues with $^3J_{HN\alpha} > 8.0$ Hz, and $^3J_{HN\alpha} < 6.0$ Hz, respectively. For the sequential NOE connectivities $d_{\alpha N}$, d_{NN}, and $d_{\beta N}$ ($d_{N\delta}$, $d_{\alpha\delta}$, and $d_{\beta\delta}$ for Xxx-Pro dipeptides, $d_{\delta N}$ for Pro-Xxx), thick and thin bars indicate strong and weak NOE intensities. Medium-range NOEs are indicated by lines connecting the two residues that are related by the NOE. $^{13}C^{\alpha}$ chemical shifts relative to the random coil values (Richarz & Wüthrich, 1978), $\Delta\delta^{13}C^{\alpha}$, are plotted at the bottom of the figure, where positive values are shifts to lower field, and the largest shift is about 6 ppm. (The shifts have been corrected by +1.43 ppm to account for the difference between the two calibration methods used, i.e., internal DSS in the present study [Wishart et al., 1995], and external TMS by Richarz and Wüthrich [1978].) Downfield shifts of the size observed here are indicative of helical secondary structure (Spera & Bax, 1991). The locations of the three helices are identified at the bottom.

small $^3J_{HN\alpha}$ coupling constants indicate that E*r*-11 contains three helices extending approximately from residues 2–9, 12–19, and 22–32. In the first and third helical regions, nearly complete sets of $d_{\alpha N}(i, i+3)$ and $d_{\alpha\beta}(i, i+3)$ connectivities suggest the presence of regular α-helices (Wüthrich et al., 1984), whereas for the second helix, only a subset of the expected medium-range NOEs was observed, indicating that this secondary structure is not a regular α-helix. There is also a good correlation between the presence of successive residues with conformation-dependent downfield $^{13}C^{\alpha}$ chemical shifts, $\Delta\delta^{13}C^{\alpha}$, exceeding 1.5 ppm and their location in the regular α-helices 1 and 3 (Spera & Bax, 1991; Wishart et al., 1991; Wishart & Sykes, 1994; Luginbühl et al., 1995), and irregularities in helix 2 are reflected by the fact that some of the $^{13}C^{\alpha}$ shifts are close to the random coil values (Richarz & Wüthrich, 1978).

Collection of structural constraints and structure calculation

A total of 577 cross peaks was assigned and integrated in two 600 MHz [^{1}H,^{1}H]-NOESY spectra measured with a mixing time of 70 ms in H_2O and ^2H_2O solution, respectively, using several cycles of data collection and structure calculation. This involved checking constraints with consistent violations in the calculated structures for possible incorrect NOESY cross peak assignments, assignment of further NOESY cross peaks by reference to the preliminary structures with the program ASNO (Güntert et al., 1993), and repetition of the structure calculation with the new, expanded data sets. Calibration with the program CALIBA and an initial screening with the program DIANA (Güntert et al., 1991a) yielded 449 meaningful NOE upper-limit distance constraints for the final structure calculation (Fig. 2). For the three disulfide bonds 3–19, 10–34, and 15–26, an additional nine upper and nine lower distance constraints were added (Williamson et al., 1985). Constraints are scarce for residues 20 and 21, which are located in the loop between the helices 2 and 3. In addition, values for 27 $^3J_{HN\alpha}$ scalar coupling constants were obtained by inverse Fourier transformation from the NOESY spectrum in H_2O solution (Szyperski et al., 1992). Thirty $^3J_{\alpha\beta}$ scalar coupling constants were measured in an E. COSY spectrum (Griesinger et al., 1985). Using the programs HABAS (Güntert et al., 1989) and GLOMSA, these data resulted in 35, 35, and 27 constraints on dihedral angles ϕ, ψ, and χ^1, respectively, and in stereospecific assignments for 15 diastereotopic pairs of methylene protons or isopropyl methyl groups (Table 1).

The final DIANA calculation started with 50 randomized conformations. The REDAC procedure (Güntert & Wüthrich, 1991) was applied once, followed by a DIANA computation at full target level using the original 97 dihedral angle constraints obtained with HABAS. Of the 50 computed conformers, the 20 conformers with the lowest residual target function values had small numbers and magnitudes of residual constraint violations. The mean energy of these 20 conformers could be reduced by about 450 kcal/mol through energy-minimization with the program

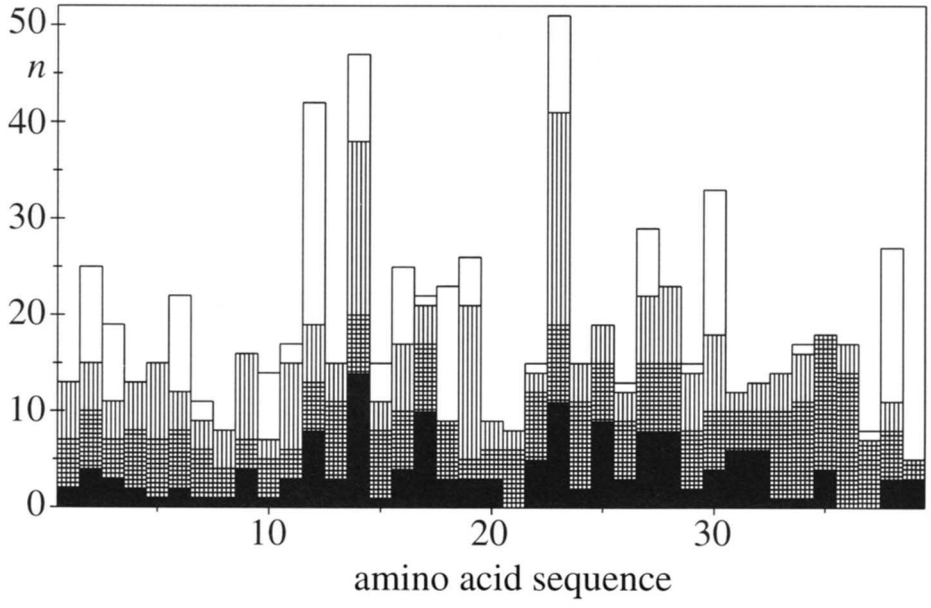

Fig. 2. Plot of the number of NOE constraints per residue used in the calculation of the E*r*-11 structure versus the sequence of E*r*-11. Filled, crosshatched, vertically hatched, and open bars represent, respectively, intraresidual, sequential, medium-range, and long-range NOEs (for the notation used, see Wüthrich, 1986).

OPAL (P. Luginbühl, M. Billeter, P. Güntert, & K. Wüthrich, unpubl.), with only minor changes in the residual constraint violations and the RMS deviations (RMSDs) relative to the mean coordinates (Table 2). For the backbone atoms N, C^{α}, and C′ of the 20 energy-refined conformers, the global RMSDs are 0.39 Å for all 39 residues and 0.30 Å for residues 2–38. A visual impression of the high quality of the structure determination is afforded by Figure 3. The global displacements calculated for the backbone atoms N, C^{α}, and C′ of the individual residues, D_{glob}^{bb} (Billeter et al., 1989), are below 0.7 Å for residues 2–38. For 16 "best-defined" residues (see Table 2) the global displacement of the side chains, D_{glob}^{sc}, is smaller than 0.7 Å (see also Fig. 4).

NMR solution structure of the pheromone Er-11

Figure 4 presents a summary of the information available on the NMR structure of E*r*-11, which consists of an antiparallel bundle of three helices comprising residues 2–9, 12–19, and 22–32 (Fig. 3). The axes of the three helices are almost parallel, with an up-down-up topology in the orientation of Figure 3A. The first and second helices are connected by a dipeptide segment of residues Cys 10 and Ser 11 (Fig. 3B), stabilized by an O′-HN hydrogen bond between residues 7 and 10 (Table 3). The second and third helices are connected by the dipeptide segment of residues 20–21, which forms a 3_{10}-helical turn (Fig. 3A) as indicated by the three O_i'–HN_{i+3} hydrogen bonds connecting residues 19–22, 20–23, and 21–24 (Table 3). The C-terminal polypeptide segment following helix 3 adopts an extended conformation that is only interrupted by the *cis* peptide bond Pro 36–Pro 37. The disulfide bridge 3–19 in E*r*-11 connects the first and second helices, the disulfide bridge 15–26 connects the second and third helices, and the disulfide bridge 10–34 ties the C-terminal tail of the molecule to the dipeptide link between the first and second helices (Fig. 4). The chain terminal residues Asp 1 and Ser 39 have few nonbonding contacts with the rest of the molecular structure, and their spatial arrangement is less well determined by the NMR data than that of most of the rest of the polypeptide chain (Fig. 3).

All three helices in E*r*-11 are quite short. As was already indicated by the data in Figure 1, helices 1 and 3 are regular α-helices, involving a complete set of O_i'–HN_{i+4} hydrogen bonds. Helix 1 ends with the C-terminal residue Gln 9 in a 3_{10}-helical conformation and an Ala 7 O′–Cys 10 HN hydrogen bond involving the amide proton of the first residue following the helix (see also Table 3). It is well known that regular α-helices in globular proteins tend to end with one or two residues in a local 3_{10}-helix conformation (Schulz & Schirmer, 1979). Helix 2 contains α-type O_i'–HN_{i+4} hydrogen bonds linking the residue pairs 12–16, 13–17, and 15–19, but there is also an O_i'–HN_{i+3} hydrogen bond linking residues 15–18 (Table 3). The distortion of the helix near residue 18 is also reflected in the experimental data by scarcity of medium-range NOEs and small conformation-dependent $^{13}C^{\alpha}$ chemical shifts (Fig. 1). This distortion of the helix ensures that the Cys residues 15 and 19 are oriented properly to form the disulfide bridges to the other two helices (Fig. 4). E*r*-11 thus shares characteristic traits of other highly stable, "small disulfide-rich" proteins, in which regular secondary structure elements are distorted by the high density of disulfide bonds (Richardson, 1981).

The C-terminal region of E*r*-11 possesses two segments that adopt extended conformations with residues 33–35 and 37–39, respectively, and that are separated only by the *cis* Pro 36–Pro 37 dipeptide segment (Figs. 3A, 4). The first one of these extended segments corresponds closely to the C-terminal structure of E*r*-1 (Mronga et al., 1994). However, E*r*-1 forms an extended loop on top of helix 3, whereas in E*r*-11, the kink at the *cis* Pro–Pro dipeptide segment causes the polypeptide chain to end on top

Table 2. *Quantitative characterization of the 20 DIANA conformers used to represent the solution structure of Er-11 before and after energy minimization with the program OPAL*

Parameter	DIANA[a,c]	OPAL[b,c]
DIANA target function (Å²)[d]	0.78 ± 0.23 (0.47 to 1.22)	
AMBER energy (kcal/mol)	−576 ± 36 (−640 to −516)	−1,038 ± 20 (−1,078 to −1,006)
van der Waals energy	186 ± 28 (136 to 235)	−107 ± 4 (−115 to −97)
Electrostatic energy	−978 ± 25 (−1,026 to −928)	−1,080 ± 20 (−1,126 to −1,039)
Residual NOE distance constraint violations (Å)		
Number > 0.1 Å	12.0 ± 2.8 (6 to 17)	1.1 ± 0.9 (0 to 6)
Sum	3.5 ± 0.5 (2.7 to 4.7)	3.8 ± 0.3 (3.3 to 4.4)
Maximum	0.25 ± 0.08 (0.17 to 0.38)	0.10 ± 0.01 (0.09 to 0.11)
Residual dihedral angle constraint violations (deg)		
Number > 2.5°	0.4 ± 0.6 (0 to 2)	0
Sum	5.9 ± 1.7 (3.1 to 9.2)	7.0 ± 2.2 (4.8 to 12.1)
Maximum	2.1 ± 0.8 (0.9 to 3.6)	1.8 ± 0.2 (1.4 to 2.1)
RMSD values (Å)[e]		
Backbone N, C^{α}, C′ (1-39)	0.36 ± 0.10 (0.19 to 0.53)	0.39 ± 0.07 (0.29 to 0.53)
All heavy atoms (1-39)	0.78 ± 0.08 (0.65 to 0.91)	0.79 ± 0.08 (0.67 to 0.91)
Backbone N, C^{α}, C′ (2-38)[f]	0.30 ± 0.10 (0.16 to 0.49)	0.30 ± 0.05 (0.20 to 0.40)
Same + best-defined side chains[f]	0.31 ± 0.10 (0.16 to 0.50)	0.31 ± 0.05 (0.21 to 0.39)

[a] DIANA conformers before energy minimization.
[b] DIANA conformers after restrained energy refinement with the program OPAL.
[c] The numbers given are the average ± standard deviation calculated for the group of 20 DIANA conformers, with the minimum and maximum values for individual conformers given in parentheses.
[d] The values for the DIANA target function (see text) are given with respect to all distance constraints and all dihedral angle constraints obtained from the HABAS treatment of the experimental data. For all other parameters in the table, the values are given with respect to a reduced data set from which were omitted the disulfide bridge constraints, the constraints for all ψ angles, and those for the ϕ and χ^1 angles for which there is no direct experimental evidence from ^3J coupling constants. The target function is not given for the energy-minimized conformers, because these do not have the ECEPP/2 standard covalent geometry (Momany et al., 1975; Némethy et al., 1983).
[e] The RMSD values are relative to the mean coordinates.
[f] For all residues 2-38,the global displacement for the backbone atoms N, C^{α}, C′, D_{glob}^{bb}, was < 0.7 Å. Among these, the following "best-defined" 16 residues have global displacements of the side-chain heavy atoms, D_{glob}^{sc}, < 0.7 Å: Cys 3, Ala 4, Ala 6, Ala 7, Ala 8, Ser 11, Ile 12, Cys 15, Cys 19, Gly 20, Pro 21, Ile 23, Cys 26, Thr 29, Val 30, Pro 36.

of helix 2, i.e., in front of the protein in the orientation of Figure 3. Long-range NOE constraints in this region were found between the residue pairs Cys 34/Cys 10 and Pro 37/Ile 12, and between Phe 38 and the residues 10–12, 14, and 30. The long-range NOEs with Phe 38 indicate a novel feature of the C-terminal conformation of E*r*-11 when compared with E*r*-1, E*r*-2, and E*r*-10 (Luginbühl et al., 1994). This new feature includes nearly complete solvent protection of residues 10–12 by the aromatic ring of Phe 38 with, for example, a solvent accessibility of only 1% of the total surface area for Ser 11 (Fig. 6). There is no hydrogen bond in E*r*-11 beyond Cys 34 (Table 3), which contrasts with the other pheromone structures, where the C-terminal segment forms one or two turns that are stabilized by backbone–backbone hydrogen bonds.

The amide proton exchange rates (Fig. 5) provide direct evidence for the relative solvent accessibilities of different polypeptide segments in E*r*-11. Overall, there is a good correlation between hydrogen bond formation and slowed exchange of the backbone amide protons. With the exceptions of residues 4, 11, 12, 13, 25, and 39, exchange rates slower than 10^{-1} min^{-1} were found only for hydrogen bonded protons (Fig. 5; Table 3). An Asp 1 δO–Ala 4 NH hydrogen bond was found in 5 of the 20 conformers (not listed in Table 3), which explains the slightly slowed amide proton exchange of Ala 4. The small solvent accessibility of residues 11 and 12 may explain the high protection factor of the amide protons in these cases. Conversely, no residue with $k_{NH} > 10^{-1}$ min^{-1} forms a hydrogen bond identified in Table 3 (the amide proton exchange of Cys 19 could not be determined due to spectral overlap).

Apart from the three disulfide bridges, there are six other residues for which the solvent-accessible surface is below 15% (Fig. 6). These are Ala 6, Ala 7, and Val 30, which lie in the interfacial region between the three helices, the aforementioned Ser 11 and Ile 12, which are protected from the solvent by the C-terminal peptide segment (see above), and Pro 36, which is shielded from the solvent by the side chains of Ile 12, Met 31, Glu 35, and Pro 37 (Fig. 4).

Discussion

The conclusions from a previous detailed comparison of the NMR structures of the pheromones E*r*-1, E*r*-2, and E*r*-10 from *E. raikovi* (Luginbühl et al., 1994) receive further support by the present work. Residues Cys 3, Ala 6, Cys 15, Cys 19, Cys 26, and Val 30 in E*r*-11 act as anchor residues for three helices, and coincide in the structural positions with corresponding anchor residues in E*r*-1, E*r*-2, and E*r*-10 (Fig. 6). Helices 1 and 3 are regular α-helices, whereas helix 2 is distorted due to the two anchoring Cys residues. At the start of helix 2, an N-cap formed by hydrogen bonds Asn 11 δO–Gly 14 NH in E*r*-10 and Ser 11

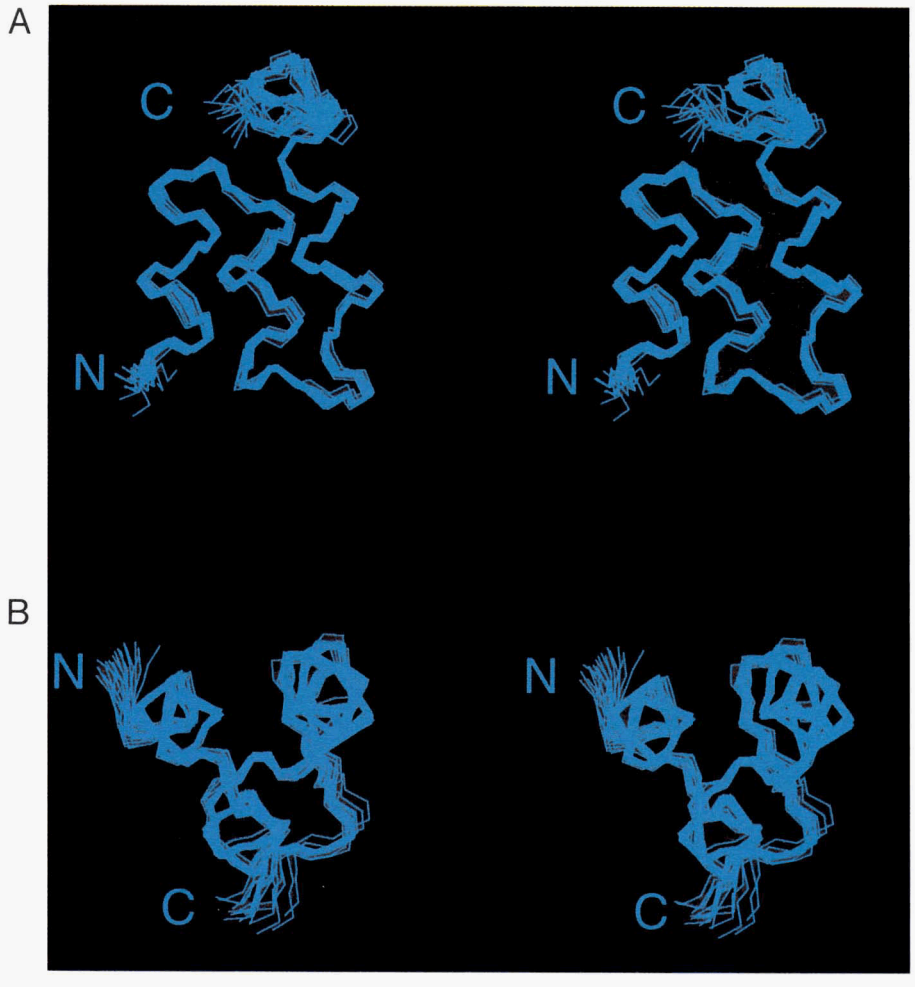

Fig. 3. Two stereo views of the polypeptide backbone of the 20 energy-refined DIANA conformers of E*r*-11 used to represent the solution structure, with superposition of the backbone atoms N, C^{α}, and C′ of residues 2–38. (A) Front view. (B) Top view obtained after rotation by 90° about a horizontal axis in the projection plane.

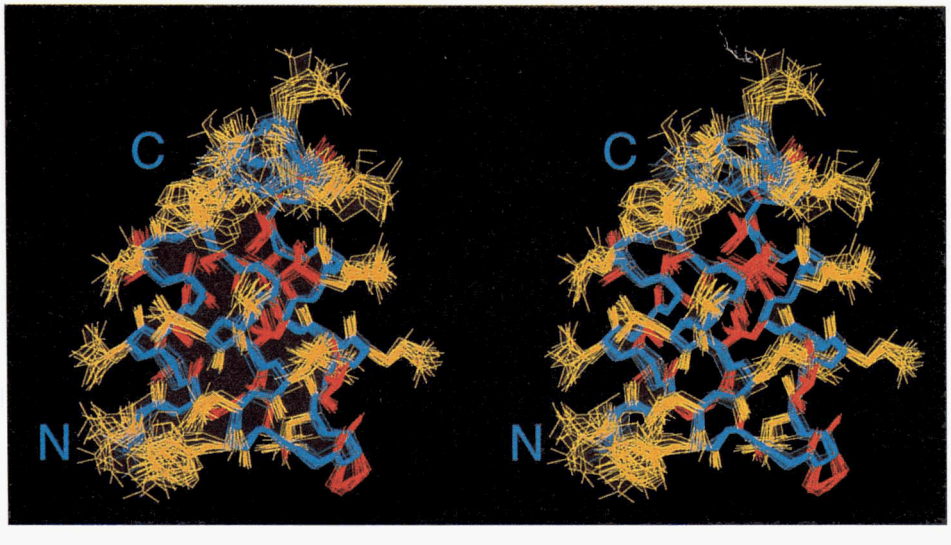

Fig. 4. Stereo view of the all-heavy-atom representation of the 20 E*r*-11 conformers of Fig. 3 in the same orientation as in Fig. 3A. The backbone is colored cyan, the "best-defined" side chains (Table 2) red, and the remaining side chains yellow.

Table 3. *Hydrogen bonds observed in the NMR solution structure of Er-11*

Donor[a]	Acceptor[a]	Number[b]
5 Asn NH	1 Asp O′	12
6 Ala NH	2 Glu O′	20
7 Ala NH	3 Cys O′	18
8 Ala NH	4 Ala O′	20
9 Gln NH	5 Asn O′	20
10 Cys NH	7 Ala O′	16
14 Leu NH	11 Ser γO	19
15 Cys NH	11 Ser O′	18
16 Asn NH	12 Ile O′	20
17 Leu NH	13 Thr O′	16
18 Tyr NH	15 Cys O′	11
19 Cys NH[c]	15 Cys O′	19
22 Leu NH	19 Cys O′	10
23 Ile NH	20 Gly O′	19
24 Glu NH	21 Pro O′	12
26 Cys NH	22 Leu O′	20
27 Glu NH	23 Ile O′	20
28 Leu NH	24 Glu O′	20
29 Thr NH	25 Ile O′	20
29 Thr γOH	25 Ile O′	16
30 Val NH	26 Cys O′	20
31 Met NH	27 Glu O′	19
32 Gln NH	28 Leu O′	19
33 Asn NH	29 Thr O′	11
34 Cys NH	30 Val O′	18

[a] Hydrogen bonds are listed if they are present in at least 10 of the 20 energy-refined DIANA conformers used to characterize the solution structure. The criterion for identification of a hydrogen bond was a maximal proton–acceptor distance of 2.4 Å and a maximal angle of 35° between the donor–proton bond and the line connecting the donor and acceptor heavy atoms (Levitt, 1983). The donor and acceptor atoms are identified by the sequence positions, the amino acid type, and the atom position, where NH and O′ stand for the backbone amide proton and the backbone carbonyl oxygen, respectively.

[b] The number of occurrences among the 20 energy-minimized DIANA conformers is given.

[c] For Cys 19, no amide proton exchange measurements were possible (see text). All other amide protons listed in this table exchange slowly. Slightly slowed exchange was also found for residues 4, 11, 12, 13, 25, and 39 (Fig. 5), of which Ala 4 appears to be involved in a weak hydrogen bond, and Ser 11 and Ile 12 are otherwise completely protected from the solvent (see text).

γO–Leu 14 NH in Er-11, respectively, is a local detail preserved between these two proteins. In Er-1 and Er-2, this hydrogen bond is replaced by a backbone–backbone hydrogen bond between the corresponding residues. The largest differences between Er-11 and the previously determined structures of Er-1, Er-2, and Er-10 are, on the one hand, that Er-11 contains a 3_{10}-helical turn in the place of an extended loop between helices 2 and 3, which can be related to an additional deletion in this segment, and, on the other hand, the presence of a novel conformation of the C-terminal segment, as described in the preceding section.

The NMR determination of the Er-11 conformation shows that the extensive variations that distinguish the amino acid sequence of this pheromone from its homologues Er-1, Er-2, and Er-10 do not overcome the functional constraints that impose strict conservation of a common three-helix-bundle molecular architecture. This finding accounts for the similar capability shown by all these pheromones to compete for binding to each other's receptors, and thus in eliciting a variety of different cellular responses (Ortenzi & Luporini, 1995). A cell may be stimulated to reproduce (proliferate), or to arrest temporarily its reproduction and form mating pairs, depending on whether the "own" pheromone or a "foreign" homologue is prevalently bound by its pheromone receptors.

E. raikovi pheromones have been divided into two separate groups on the basis of the spectrum of cell types that they can induce to form mating pairs (Luporini et al., 1995). The pheromones Er-1, Er-2, Er-7, Er-10, and Er-11 are members of the PR group, and Er-20, Er-21, and Er-22 belong to the GA group. Within a given group, each pheromone can induce cells that secrete another pheromone to form mating pairs, whereas intergroup pheromone/cell combinations are ineffective except in combinations with cells that secrete Er-11 (Luporini et al., 1995). However, the Er-11 cells show unusually weak restrictions to forming mating pairs, also in response to pheromone preparations from other *Euplotes* species. It is of obvious interest to investigate whether this functional classification is also supported by the molecular structures of the pheromones.

The pheromone-binding unit of the cellular receptors appears to consist of a membrane-anchored pheromone isoform of which the extracellular C-terminal domain has an identical amino acid sequence as the secreted pheromone. In fact, it arises from a mechanism of alternative splicing of the primary transcript of the same gene, which contains the information for the sequence of the secreted pheromone (Miceli et al., 1992). In this context, the availability of a reference topological picture of unique and common structural traits that make a set of homologous pheromones equally able to function as self or non-self cell ligands is also of interest with regard to a search of molecular features that distinguish homologous versus heterologous (or autocrine versus paracrine) pheromone/pheromone-receptor interactions and consistently modulate different transduction pathways. These interactions may, in fact, be mimicked by the study of interactions between identical or different pheromone molecules, which have recently been shown by Er-1 crystallographic analyses to involve at least two types of cooperative intermolecular helix–helix association (Weiss et al., 1995). Here, the novel C-terminal conformation found in Er-11 may have an important role. The triangular shape of Er-11 gives rise to three surfaces formed by two helices each (Fig. 3). Whereas in Er-1, Er-2, and Er-10 the C-terminus is located on top of the face formed by helices 1 and 3, in Er-11 it is located on top of the face formed by helices 2 and 3. Because the C-terminal amino acid sequence of Er-11 is very similar to the C-terminal sequences of Er-20, Er-21, and Er-22 (Luporini et al., 1995), one would predict that the C-terminal conformation of these latter pheromones should resemble that of Er-11. This would go along with the facts that Er-11 shows higher sequence identity to pheromones of the GA group (between 51% and 59% for Er-20, Er-21, and Er-22) than to the other members of the PR group (between 26% and 34% for Er-1, Er-2, Er-7, and Er-10), and that, accordingly, the sequence identity between Er-11 and the GA group is about twice as high as between any of the other pheromones from the PR group and any member of the GA group (Luporini et al., 1995). Overall, the data on sequence and conformation thus provide a rationale for the weak restrictions of the Er-11 cells to forming mating pairs also in response to pheromones from the GA group.

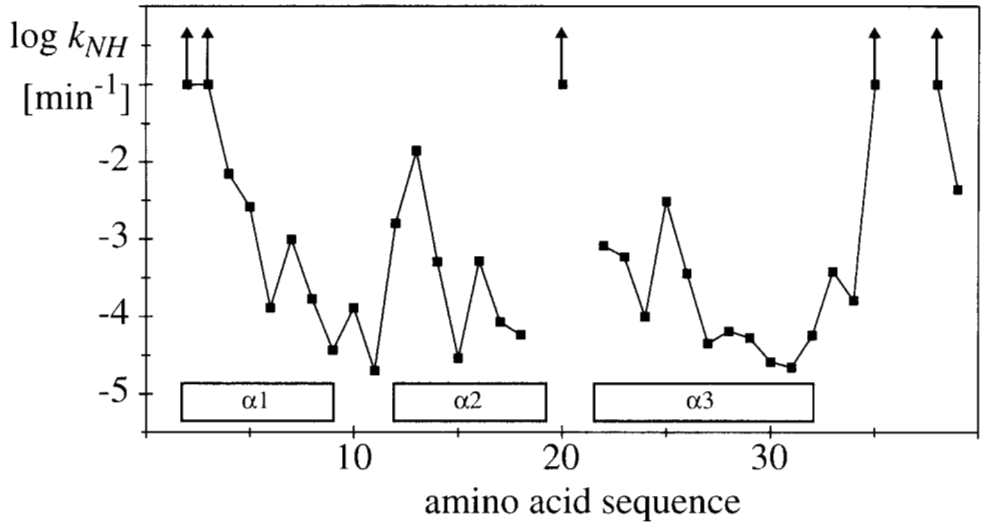

Fig. 5. Plot of the rates of exchange with solvent ^2H_2O for the backbone amide protons of E*r*-11 at pH 5.0 and $T = 13$ °C versus the amino acid sequence. Arrows indicate that the exchange was too fast for the amide proton resonance to be observed with the method used, and only a lower limit for the exchange rate constant could be established. For Cys 19, the amide proton exchange rate was not evaluated because of spectral overlap. The locations of the three helices of E*r*-11 are given at the bottom.

Materials and methods

The pheromone E*r*-11 was purified from supernatants of *E. raikovi* cultures of offspring clones homozygous for the E*r*-11 coding gene following published procedures (Concetti et al., 1986). The cell cultures were derived from mating pairs induced to form within another clone denoted "1aF$_1$1N," which was used originally for the determination of the E*r*-11 primary structure (Raffioni et al., 1992). The protein sample was dialyzed against H_2O, using a SPECTRUM™ spectra/por membrane (mwco: 1,000). For the NMR experiments, 2.5 mM solutions were prepared in ^2H_2O or in a mixed solvent of 90% H_2O and 10% ^2H_2O. For the ^2H_2O samples, complete exchange of the labile protons was obtained by keeping the sample at 40 °C in ^2H_2O for 10 h, followed by lyophilization and redissolving in ^2H_2O. All NMR experiments were recorded at 27 °C and at pH 5.0, with the exception that the amide proton exchange measurements were performed at 13 °C.

^{1}H NMR spectra were obtained on Bruker AMX 500 and AMX 600 spectrometers in the pure phase absorption mode

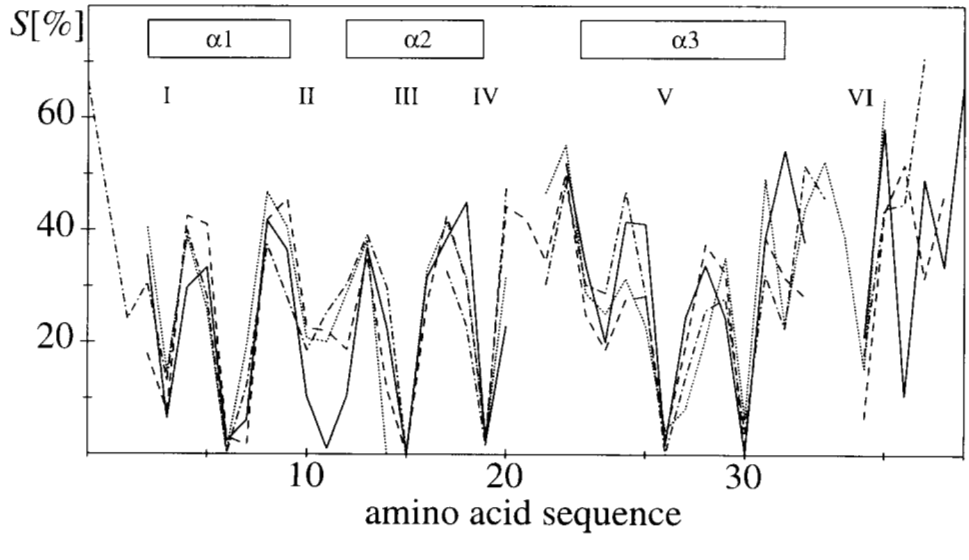

Fig. 6. Comparison of the average solvent-accessible surface area (Richmond, 1984; Billeter et al., 1990b) relative to the total surface area for individual residues versus the amino acid sequence. Numeration at the bottom is for E*r*-11. Boxes indicate the locations of the helices in E*r*-11, and Roman numerals identify the positions of the Cys residues. Solid line, E*r*-11; broken line, E*r*-1; dotted line, E*r*-10, dotted-broken line, E*r*-2.

using the States-TPPI method (Marion et al., 1989b). To obtain ^{1}H resonance assignments, a 2QF-COSY spectrum (Rance et al., 1983), a clean-TOCSY spectrum (Griesinger et al., 1988) with a mixing time of 100 ms, and a NOESY spectrum (Anil-Kumar et al., 1980) with a mixing time of 120 ms were recorded in H_2O solution and in ^2H_2O solution. In addition, an E. COSY spectrum (Griesinger et al., 1985) and a [^{13}C,^{1}H]-COSY spectrum (Bodenhausen & Ruben, 1980) were recorded in ^2H_2O. Typical data sizes were 1,024 and 512 complex data points in t_2 and t_1, respectively. The digital resolution after zero-filling was about 3 Hz/point along ω_2 and 6 Hz/point along ω_1. For the collection of upper bound ^{1}H-^{1}H distances, a NOESY spectrum in H_2O solution and a zero-quantum-suppressed NOESY spectrum (Otting et al., 1990) in ^2H_2O solution were recorded at 600 MHz, with a mixing time of 70 ms and the following experimental conditions: recorded data size 700 complex points in t_1 and 1,024 points in t_2, $t_{1max} = 60.3$ ms, $t_{2max} = 164$ ms, digital resolution after zero-filling 3.1 Hz/point along ω_2 and 5.7 Hz/point along ω_1.

For measurements of amide proton exchange rates, a sample of 500 μL of a fully protonated 6 mM protein solution at pH 5.0 was lyophilized. Subsequently, the protein was redissolved in the same amount of ^2H_2O. We began to record a series of 10 1D ^{1}H NMR spectra 15 minutes after the preparation of the ^2H_2O solution. Immediately thereafter, a series of 30 2D NOESY spectra with mixing times of 100 ms was recorded. The individual 2D experiments were started between 40 and 56,379 minutes after preparation of the ^2H_2O solution, and the total measuring time per spectrum was between 0.5 h for the first spectrum and 4 h for the last recording. The rate constants, k_{NH}, were obtained from a nonlinear least-squares fit of a single exponential to the experimental data.

For all 2D NMR spectra, the residual water signal after preirradiation was further reduced using the convolution method of Marion et al. (1989a). Prior to Fourier transformation, the time domain data were multiplied with a sine bell window function shifted by $\pi/2$ (DeMarco & Wüthrich, 1976). Baseline distortions were corrected using the FLATT procedure (Güntert & Wüthrich, 1992). Processing of the spectra was performed with the program PROSA (Güntert et al., 1992). Peak picking in the NMR spectra, spin system identification, and volume integration of the NOESY cross peaks were performed with the interactive program XEASY (Bartels et al., 1995). Once an initial structure had been obtained, the program ASNO (Güntert et al., 1993) was used for computer-supported assignments of additional NOESY cross peaks by reference to this structure. Values of the $^3J_{HN\alpha}$ scalar coupling constants were obtained by inverse Fourier transformation and time-domain fitting of cross sections taken through in-phase NOESY cross peaks (Szyperski et al., 1992), so that this information came from the same data set as the NOE distance constraints measured in H_2O solution. The $^3J_{\alpha\beta}$ coupling constants for residues with $C^\beta H_2$-groups were determined from an E. COSY spectrum (Griesinger et al., 1985).

Upper distance constraints were obtained from NOESY cross peak volumes using the program CALIBA (Güntert et al., 1991a, 1991b), and dihedral angle constraints were determined from the spin-spin coupling constants and the intraresidual and sequential NOEs using the program HABAS (Güntert et al., 1989). At early stages of the structure determination, dihedral angle constraints derived from conformation-dependent $^{13}C^\alpha$ chemical shifts were used to improve convergence (Luginbühl et al., 1995). Upper and lower distance constraints for disulfide bridges were introduced as described previously (Williamson et al., 1985). HABAS also provided a number of stereospecific assignments for β-methylene groups. At later stages of the structure refinement, additional stereospecific assignments of $C^\beta H_2$ and of other pairs of diastereotopic substituents were obtained with the program GLOMSA (Güntert et al., 1991a, 1991b). The structure calculations were performed with the distance geometry program DIANA (Güntert et al., 1991a, 1991b), using the REDAC procedure (Güntert & Wüthrich, 1991).

Energy minimization using the AMBER all-atom force field (Weiner et al., 1986) was performed with the program OPAL (P. Luginbühl, P. Güntert, M. Billeter, & K. Wüthrich, unpubl.), which includes pseudoenergy terms for distance constraints and dihedral angle constraints (Widmer et al., 1989; Billeter et al., 1990b). The pseudoenergy was adjusted such that violations of 0.10 Å for distance constraints and 2.5 degrees for dihedral angle constraints corresponded to $k_BT/2$ at room temperature. The distance constraints for the chemically determined disulfide bridges (Williamson et al., 1985) and those dihedral angle constraints from HABAS for which no direct spin-spin coupling information was available were removed from the set of constraints used during energy refinement. The dihedral angles of the peptide bonds were constrained to 180 ± 20 degrees, with the sole exception of the *cis* peptide bond linking Pro 36 and Pro 37, which was constrained to 0 ± 20 degrees. The energy minimization was conducted after immersing each DIANA conformer in a 6-Å-thick shell of water molecules, and using a constant dielectric permitivity for the electrostatic interactions. For each conformer, 1,500 steps of conjugate gradient minimization were performed, after which a small total gradient was attained. The analysis of the structures in terms of RMSD values (McLachlan, 1979), atom displacements (Billeter et al., 1989), identification of hydrogen bonds (Levitt, 1983), and solvent-accessible surface (Richmond, 1984; Billeter et al., 1990a) were performed with the program MOLMOL (Koradi et al., 1996), which was also used to prepare the stereo drawings of the E*r*-11 structure.

Acknowledgments

Financial support was obtained from the Schweizerischer Nationalfonds (project 31.32033.91), the Italian Ministero dell' Università e della Ricerca Scientifica e Tecnologica, and the Italian C.N.R. We thank R. Marani for careful processing of the manuscript.

References

Anil-Kumar, Ernst RR, Wüthrich K. 1980. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for elucidation of complete proton-proton cross relaxation networks in biological macromolecules. *Biochem Biophys Res Commun 95*:1-6.

Bartels C, Xia T, Billeter M, Güntert P, Wüthrich K. 1995. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J Biomol NMR 5*:1-10.

Billeter M, Braun W, Wüthrich K. 1982. Sequential resonance assignments in protein ^{1}H nuclear magnetic resonance spectra. Computation of sterically allowed proton-proton distances and statistical analysis of proton-proton distances in single crystal protein conformations. *J Mol Biol 155*:321-346.

Billeter M, Kline AD, Braun W, Huber R, Wüthrich K. 1989. Comparison

of the high-resolution structures of the α-amylase inhibitor tendamistat determined by nuclear magnetic resonance in solution and by X-ray diffraction in single crystals. *J Mol Biol 206*:677–687.

Billeter M, Qian YQ, Otting G, Müller M, Gehring W, Wüthrich K. 1990a. Determination of the three-dimensional structure of the *Antennapedia* homeodomain from *Drosophila* in solution by ^{1}H nuclear magnetic resonance spectroscopy. *J Mol Biol 214*:183–197.

Billeter M, Schaumann Th, Braun W, Wüthrich K. 1990b. Restrained energy refinement with two different algorithms and force fields of the structure of the α-amylase inhibitor tendamistat determined by NMR in solution. *Biopolymers 29*:695–706.

Bodenhausen G, Ruben DJ. 1980. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem Phys Lett 69*:185–188.

Brown LR, Mronga S, Bradshaw RA, Ortenzi C, Luporini P, Wüthrich K. 1993. Nuclear magnetic resonance solution structure of the pheromone E*r*-10 from the ciliated protozoan *Euplotes raikovi*. *J Mol Biol 231*:800–816.

Concetti A, 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.

DeMarco A, Wüthrich K. 1976. Digital filtering with a sinusoidal window function: An alternative technique for resolution enhancement in FT NMR. *J Magn Reson 24*:201–204.

Dini F, Nyberg D. 1983. Sex in ciliates. *Adv Microbial Ecol 13*:85–153.

Griesinger C, Otting G, Wüthrich K, Ernst RR. 1988. Clean TOCSY for ^{1}H spin system identification in macromolecules. *J Am Chem Soc 110*: 7870–7872.

Griesinger C, Sørensen OW, Ernst RR. 1985. Two dimensional correlation of connected NMR transitions. *J Am Chem Soc 107*:6394–6396.

Güntert P, Berndt KD, Wüthrich K. 1993. The program ASNO for computer-supported collection of NOE upper distance constraints as input for protein structure determination. *J Biomol NMR 3*:601–606.

Güntert P, Braun W, Billeter M, Wüthrich K. 1989. Automated stereospecific ^{1}H NMR assignment and their impact on the precision of protein structure determinations in solution. *J Am Chem Soc 111*:3997–4004.

Güntert P, Braun W, Wüthrich K. 1991a. Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J Mol Biol 217*:517–530.

Güntert P, Dötsch V, Wider G, Wüthrich K. 1992. Processing of multi-dimensional NMR data with the new software PROSA. *J Biomol NMR 2*:619–629.

Güntert P, Qian YQ, Otting G, Müller M, Gehring W, Wüthrich K. 1991b. Structure determination of the Antp (C39S) homeodomain from nuclear magnetic resonance data in solution using a novel strategy for the structure calculations with the programs DIANA, CALIBA, HABAS and GLOMSA. *J Mol Biol 217*:531–540.

Güntert P, Wüthrich K. 1991. Improved efficiency of protein structure calculations from NMR data using the program DIANA with redundant dihedral angle constraints. *J Biomol NMR 1*:447–456.

Güntert P, Wüthrich K. 1992. FLATT – A new procedure for high-quality baseline correction of multidimensional NMR spectra. *J Magn Reson 96*:403–407.

James R, Bradshaw RA. 1984. Polypeptide growth factors. *Annu Rev Biochem 53*:259–292.

Koradi R, Billeter M, Wüthrich K. 1996. MOLMOL: A program for display and analysis of macromolecular structures. *J Mol Graph 27*. Forthcoming.

Levitt M. 1983. Protein folding by restrained energy minimization and molecular dynamics. *J Mol Biol 170*:723–764.

Luginbühl P, Ottiger M, Mronga S, Wüthrich K. 1994. Structure comparison of the NMR structures of the pheromones E*r*-1, E*r*-10, and E*r*-2 from *Euplotes raikovi*. *Protein Sci 3*:1537–1546.

Luginbühl P, Szyperski T, Wüthrich K. 1995. Statistical basis for the use of $^{13}C^{\alpha}$ chemical shifts in protein structure determination. *J Magn Reson B 109*:229–233.

Luporini P, Vallesi A, Miceli C, Bradshaw RA. 1995. Chemical signaling in ciliates. *J Euk Microbiol 42*:208–212.

Marion D, Ikura K, Bax A. 1989a. Improved solvent suppression in one- and two-dimensional NMR spectra by convolution of time-domain data. *J Magn Reson 84*:425–430.

Marion D, Ikura K, Tschudin R, Bax A. 1989b. Rapid recording of 2D NMR spectra without phase cycling: Application to the study of hydrogen exchange in proteins. *J Magn Reson 85*:393–399.

McLachlan AD. 1979. Gene duplication in the structural evolution of chymotrypsin. *J Mol Biol 128*:49–79.

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

Momany FA, McGuire RF, Burgess AW, Scheraga HA. 1975. Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *J Phys Chem 79*:2361–2381.

Mronga S, Luginbühl P, Brown LR, Ortenzi C, Luporini P, Bradshaw RA, Wüthrich K. 1994. The NMR solution structure of pheromone E*r*-1 from the ciliated protozoan *Euplotes raikovi*. *Protein Sci 3*:1527–1536.

Némethy G, Pottle MS, Scheraga HA. 1983. Energy parameters in polypeptides. 9. Updating of geometrical parameters, nonbonded interactions, and hydrogen bond interactions for the naturally occurring amino acids. *J Phys Chem 87*:1883–1887.

Ortenzi C, Luporini P. 1995. Competition among homologous polypeptide pheromones of the ciliate *Euplotes raikovi*. *J Euk Microbiol 42*:242–248.

Ottiger M, Szyperski T, Luginbühl P, Ortenzi C, Luporini P, Bradshaw RA, Wüthrich K. 1994. The NMR solution structure of the pheromone E*r*-2 from the ciliated protozoan *Euplotes raikovi*. *Protein Sci 3*:1515–1526.

Otting G. 1987. Strukturermittlung an kleinen Proteinen mit n.m.r.: neue Methoden und Anwendungen am Beispiel der Cardiotoxine von *Naja Mossambica Mossambica* und des P22 c2 Repressors [thesis no. 8314]. Zürich, Switzerland: ETH.

Otting G, Orbons LPM, Wüthrich K. 1990. Suppression of zero-quantum coherence in NOESY and soft-NOESY. *J Magn Reson 86*:496–508.

Raffioni S, Miceli C, Vallesi A, Chowdhury SK, Chait BT, Luporini P, Bradshaw RA. 1992. Primary structure of *Euplotes raikovi* pheromones: Comparison of five sequences of pheromones from cells with variable mating interactions. *Proc Natl Acad Sci USA 89*:2071–2075.

Rance M, Sørensen OW, Bodenhausen G, Wagner G, Ernst RR, Wüthrich K. 1983. Improved spectral resolution in COSY proton NMR spectra of proteins via double quantum filtering. *Biochem Biophys Res Commun 117*: 479–485.

Richardson JS. 1981. The anatomy and taxonomy of protein structure. *Adv Protein Chem 34*:167–339.

Richarz R, Wüthrich K. 1978. Carbon-13 NMR chemical shifts of the common amino acid residues measured in aqueous solutions of the linear tetrapeptides H-Gly-Gly-X-L-Ala-OH. *Biopolymers 17*:2133–2141.

Richmond TJ. 1984. Solvent-accessible surface area and excluded volume in proteins. Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J Mol Biol 178*:63–89.

Schulz GE, Schirmer RH. 1979. *Principles of protein structure*. Berlin: Springer.

Sevilla-Sierra P, Otting G. Wüthrich K. 1994. Determination of the nuclear magnetic resonance structure of the DNA-binding domain of the *P22* c2 repressor (1 to 76) in solution and comparison with the DNA-binding domain of the *434* repressor. *J Mol Biol 235*:1003–1020.

Spera S, Bax A. 1991. Empirical correlation between protein backbone conformation and C^{α} and C^{β} ^{13}C nuclear magnetic resonance shifts. *J Am Chem Soc 113*:5490–5492.

Sporn MB, Roberts AB. 1990. *Peptide growth factors and their receptors, I and II*. Berlin: Springer.

Stewart AE, Raffioni S, Chaudhary T, Chait BT, Luporini P, Bradshaw RA. 1992. The disulfide bond pairing of the pheromones E*r*-1 and E*r*-2 of the ciliated protozoan *Euplotes raikovi*. *Protein Sci 1*:777–785.

Szyperski T, Güntert P, Otting G, Wüthrich K. 1992. Determination of scalar coupling constants by inverse Fourier transformation of in-phase multiplets. *J Magn Reson 99*:552–560.

Wagner G, Wüthrich K. 1982. Sequential resonance assignments in protein ^{1}H nuclear magnetic resonance spectra. Basic pancreatic trypsin inhibitor. *J Mol Biol 155*:347–366.

Weiner SJ, Kollman PA, Nguyen DT, Case DA. 1986. An all atom force field for simulations of proteins and nucleic acids. *J Comput Chem 7*:230–252.

Weiss MS, Anderson DH, Raffioni S, Bradshaw RA, Ortenzi C, Luporini P, Eisenberg D. 1995. A cooperative model for receptor recognition and cell adhesion: Evidence from the molecular packing in the 1.6-Å crystal structure of the pheromone E*r*-1 from the ciliated protozoan *Euplotes raikovi*. *Proc Natl Acad Sci USA 92*:10172–10176.

Wider G, Lee KH, Wüthrich K. 1982. Sequential resonance assignments in protein ^{1}H nuclear magnetic resonance spectra. Glucagon bound to perdeuterated dodecylphospho-choline micelles. *J Mol Biol 155*:367–388.

Widmer H, Billeter M, Wüthrich K. 1989. Three dimensional structure of the neurotoxin ATX Ia from *Anemonia sulcata* in aqueous solution determined by nuclear magnetic resonance spectroscopy. *Proteins Struct Funct Genet 6*:357–371.

Williamson MP, Havel TF, Wüthrich K. 1985. Solution conformation of pro-

teinase inhibitor IIa from bull seminal plasma by ^{1}H nuclear magnetic resonance and distance geometry. *J Mol Biol 182*:295–315.

Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD. 1995. ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J Biomol NMR 6*:135–140.

Wishart DS, Sykes BD. 1994. The ^{13}C chemical-shift index: A simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J Biomol NMR 4*:171–180.

Wishart DS, Sykes BD, Richards FM. 1991. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol 222*:311–333.

Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: Wiley.

Wüthrich K, Billeter M, Braun W. 1984. Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton distances. *J Mol Biol 180*:715–740.