Journal of Biomolecular NMR, **19:** 75–78, 2001.
KLUWER/ESCOM

Letter to the Editor: NMR structure of the pheromone E*r*-22 from *Euplotes raikovi*

Aizhuo Liu[a,*], Peter Luginbühl[a], Oliver Zerbe[a,**], Claudio Ortenzi[b], Pierangelo Luporini[b,***] & Kurt Wüthrich[a,***]

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

Received 27 September 2000; Accepted 6 October 2000

Key words: ciliate pheromone E*r*-22, *Euplotes raikovi*, NMR structure

Biological context

Cellular recognition and signal transduction processes mediated by polypeptides or small proteins, e.g., growth factors and hormones, are an important feature of intercellular communication in higher eukaryotes (e.g., Bradshaw, 1996). The pheromone regulatory system in *Euplotes raikovi (E. raikovi)* represents a similar, simpler system in a unicellular eukaryote, which may represent an evolutionary predecessor of the more sophisticated regulatory mechanisms in multicellular species (Luporini et al., 1996). This makes the *E. raikovi* system an attractive paradigm for investigating the structural basis of 'self' and 'nonself' recognition processes that elicit different cellular responses. NMR structures of the pheromones E*r*-1 (Mronga et al., 1994), E*r*-2 (Ottiger et al., 1994), E*r*-10 (Brown et al., 1993) and E*r*-11 (Luginbühl et al., 1996b), as well as an X-ray crystal structure of E*r*-1 (Weiss et al., 1995) have previously been determined. These four pheromones are all members of the 'PR group' (Luporini et al., 1995) and share a common architecture. Local structure variations confer specificity in receptor association (Luginbühl et al., 1994) without interfering with the ability to compete for each other's cell receptors in homologous (autocrine) reactions for cell growth stimulation, and in heterologous (paracrine) reactions for mating induction (e.g., Ortenzi et al., 2000). Here we describe the NMR structure of the 37-residue pheromone E*r*-22, which belongs to the 'GA group' of *E. raikovi* pheromones (Luporini et al., 1995). Since the PR and GA groups represent mating-incompatible *E. raikovi* strains (Vallesi et al., 1996), the E*r*-22 structure should provide further insight into the molecular basis of signaling mechanisms in this organism.

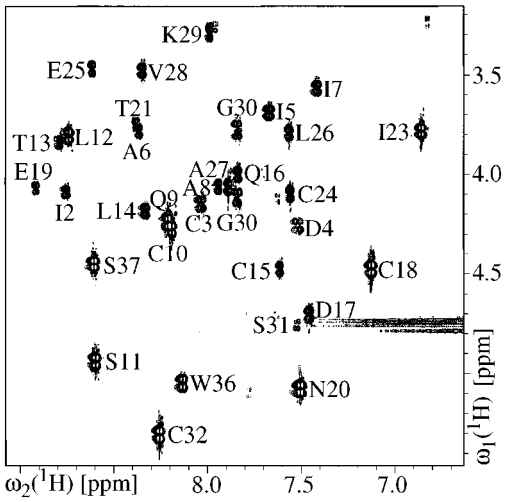

Figure 1. Fingerprint region of the 600 MHz phase-sensitive [^{1}H,^{1}H]-2QF-COSY spectrum of E*r*-22. Resonance assignments are given by the one-letter amino acid code and the sequence position.

*Present address: Tularik Inc., Two Corporate Drive, South San Francisco, CA 94080, U.S.A.
**Present address: Institut für Pharmazeutische Wissenschaften, Winterthurerstr. 190, Eidgenössische Technische Hochschule, CH-8057 Zürich, Switzerland.
***To whom correspondence should be addressed. Fax: P.L. +39 0737 636216; K.W. +41 1 633 11 51.

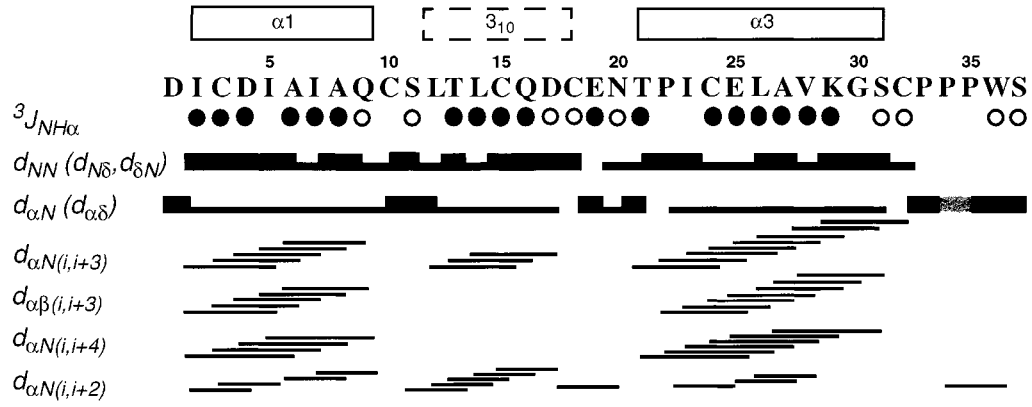

Figure 2. Resonance assignment and secondary structure of E*r*-22. Open and filled circles identify residues with ${}^3J_{HN\alpha} > 8.0$ Hz, and ${}^3J_{HN\alpha} < 6.0$ Hz, respectively. For the sequential NOEs $d_{\alpha N}$, d_{NN} and $d_{\beta N}$ ($d_{N\delta}$, $d_{\alpha\delta}$ and $d_{\beta\delta}$ for Xxx-Pro, $d_{\delta N}$ for Pro-Xxx), thick and thin bars indicate strong and weak intensities. Medium-range NOEs (Wüthrich, 1986) are indicated by lines connecting the two related residues. The crosshatched bar in $d_{\alpha N}$ between Pro 34 and Pro 35 indicates a sequential $C^{\alpha}H$-$C^{\alpha}H$ NOE, which shows that this Pro-Pro dipeptide bond is in the *cis*-conformation. The numeration of the three helices at the top corresponds to that used for previously studied pheromones, with the 3_{10}-helix in the place of α2 (Luginbühl et al., 1994).

Methods and results

The pheromone E*r*-22 at natural isotope distribution was isolated from natural sources (Vallesi et al., 1996). The structure determination was performed at pH 5.0 and 23 °C. Homonuclear ^{1}H NMR spectra and a [^{13}C,^{1}H]-COSY experiment were recorded at 600 MHz on a Bruker AMX spectrometer and at 750 MHz on a Varian Unity-plus spectrometer. The spectral analysis was supported by the program XEASY (Bartels et al., 1995), and the structure calculation and refinement were performed with the programs DIANA (Güntert et al., 1991) and OPAL (Luginbühl et al., 1996a).

The [^{1}H,^{1}H]-COSY spectrum of Figure 1 illustrates the high quality of the NMR data obtained for this small protein. Sequence-specific ^{1}H resonance assignments were obtained using sequential NOEs and ^{1}H–^{1}H scalar couplings (Wüthrich, 1986) (Figure 1). The Pro34–Pro35 dipeptide showed a sequential $C^{\alpha}H$–$C^{\alpha}H$ NOE and thus is in the *cis*-conformation (Wüthrich, 1986). ^{1}H–^{1}H two-quantum spectroscopy and [^{13}C,^{1}H]-COSY at natural ^{13}C abundance were used to establish the chemical shifts for geminal methylene protons and $^{13}C^{\alpha}$, and the Asn and Gln side chain amide protons were individually assigned from intraresidual NOEs (Sevilla-Sierra et al., 1987). The complete ^{1}H and $^{13}C^{\alpha}$ assignments have been deposited in the BioMagResBank (accession code BMRB-4820).

Table 1. Characterization of the 20 energy-refined DIANA conformers representing the NMR structure of E*r*-22[a]

Diana target function (Å^2)[b]	0.73 ± 0.14 (0.51. . . 0.98)
NOE violations > 0.1 Å	0.1 ± 0.2 (0. . . 1)
AMBER energy (kcal/mol)	-1006 ± 24 (-1058. . . -940)
RMSD, N, C^{α}, C′ (Å)[c]	0.47 ± 0.10 (0.28. . . 0.62)
RMSD, all heavy atoms (Å)[c]	0.71 ± 0.07 (0.56. . . 0.85)

[a] Average values ± standard deviations for the 20 DIANA conformers, minimum and maximum values for individual conformers in parentheses.
[b] Before energy minimization.
[c] RMSD values relative to the mean coordinates for residues 1–37.

Patterns of successive d_{NN} connectivities and small ${}^3J_{HN\alpha}$ coupling constants, and nearly complete sets of $d_{\alpha N}(i, i+3)$, $d_{\alpha\beta}(i, i+3)$, and $d_{\alpha N}(i, i+4)$ NOE connectivities indicate two regular α-helices from residues 2–9 and 21–31, and a distorted helical structure is indicated for the residues 12–18 (Wüthrich, 1986) (Figure 2). The secondary structures are also clearly reflected by the $^{13}C^{\alpha}$ shifts (data not shown) (Spera and Bax, 1991).

A total of 598 NOESY cross peaks obtained at 750 MHz with a mixing time of 45 ms were assigned and integrated, which yielded 497 NOE upper limit distance constraints. The combined information from the intraresidual and sequential NOEs (Figure 2), and from measurement of 27 ${}^3J_{HN\alpha}$ and 24 ${}^3J_{\alpha\beta}$ scalar coupling constants yielded 32, 32 and 26 constraints on dihedral angles φ, ψ and χ^1, respectively (Güntert et al., 1991). Stereospecific assignments were

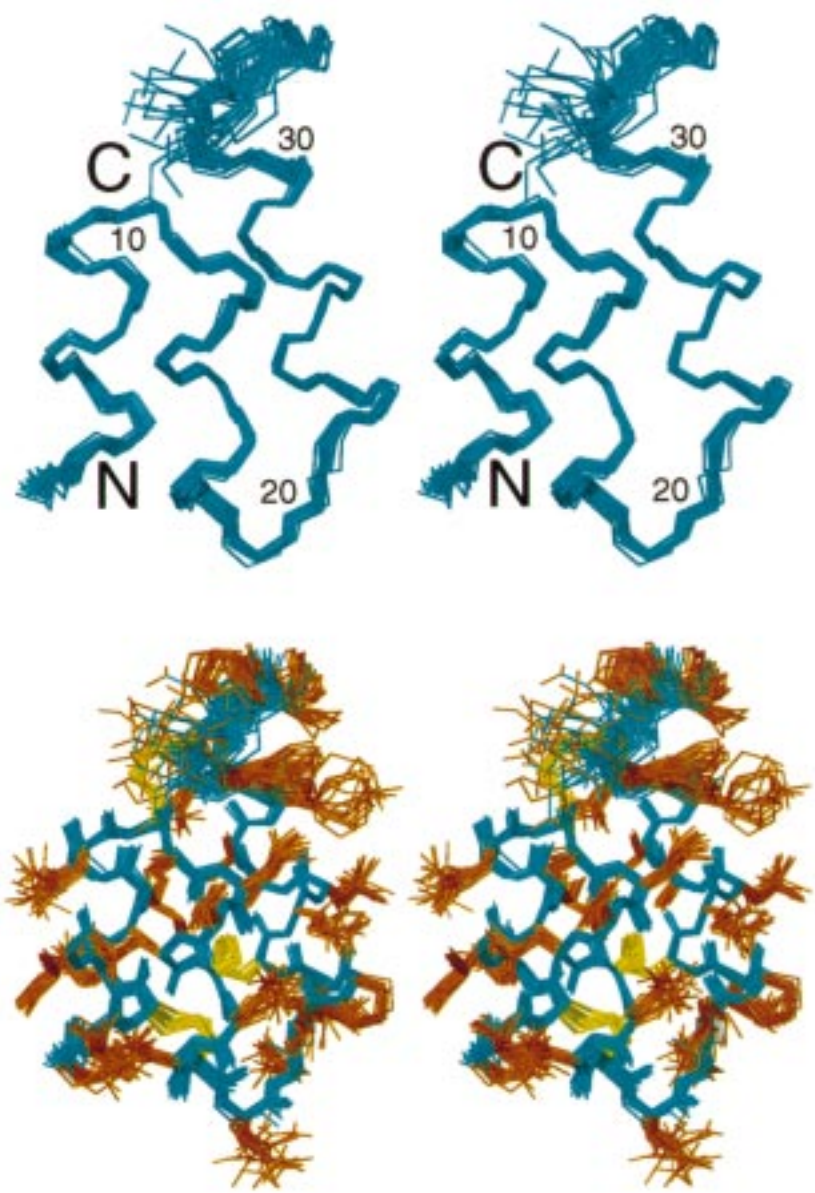

Figure 3. Stereoviews of the NMR structure of E*r*-22. Top: backbone of the 20 energy-refined DIANA conformers, with best fit for N, C^{α} and C' of residues 2–33. Bottom: all-heavy-atom representation of the same 20 E*r*-22 conformers. The backbone is blue, the side-chains are red and the disulfide bonds yellow.

obtained for 16 diastereotopic pairs of methylene protons and isopropyl methyl groups. The three disulfide bonds were enforced by 9 upper and 9 lower distance constraints (Williamson et al., 1985). Thereby we started with the assumption that the disulfide bonds correspond to those determined previously by chemical methods in E*r*-1 and E*r*-2, i.e., with the Cys combinations 3–18, 10–32 and 15–24 (Stewart et al., 1992). Subsequently this assumption was validated by structure calculations with alternative Cys–Cys combinations in the disulfides, which invariably resulted in higher residual values of the DIANA error function. The statistics on the 20 conformers used to characterize the NMR structure, which resulted from a final DIANA calculation with 50 randomized starting conformations, are given in Table 1. Figure 3 affords a visual impression of the high quality of the E*r*-22 structure determination.The atom coordinates have

been deposited in the Protein Data Bank (PDB ID code 1hd6).

E*r*-22 consists of an antiparallel bundle of three helices comprising residues 2–9, 12–18 and 21–31 (Figures 2 and 3). The three helix axes are nearly parallel, with an up-down-up topology in the orientation of Figure 3. As was already indicated by the data of Figure 2, helix 1 and helix 3 are regular α-helices, whereas helix 2 is a continuous stretch of distorted 3_{10}-helix turns. The distortion of helix 2 ensures that Cys15 and Cys18 are properly oriented to form disulfide bridges to the other two helices (Figure 3b). The dipeptide segment Cys10-Ser11 is stabilized by O′-HN hydrogen bonds with Ile7 and Ala6. The dipeptide Glu19-Asn20 forms a hydrogen bond with Thr21 O′, and is further stabilized by a Cys18 O′-Thr21 γOH hydrogen bond. The polypeptide segment following helix 3 adopts an extended conformation, with a kink at the *cis* peptide bond linking Pro34 and Pro35. The disulfide bridge Cys3-Cys18 connects the helices 1 and 2, Cys15-Cys24 connects the helices 2 and 3, and Cys10-Cys32 ties the C-terminal 'tail' of residues 32 to 37 to the dipeptide link between the helices 1 and 2 (Figure 3b).

Discussion and conclusions

From sequence alignment, Vallesi et al. (1996) observed that in the pheromones of the GA group the polypeptide segment from the third to the fifth Cys residue is at least two residues shorter than in the pheromones of the PR group, and they hypothesized that this might generate the different group specificities. In E*r*-22 one of these deletions leads to the 3_{10}-type helix 2. Since this coincides with the structure of E*r*-2 (Ottiger et al., 1994), this feature does not appear to discriminate between the PR and GA groups. The second deletion shortens the loop between the helices 2 and 3, which is a unique feature of the presently known GA pheromones. The spatial arrangement of the C-terminal segment in E*r*-22 is very similar to E*r*-11, with the C-terminal residue on top of helix 2, whereas in E*r*-1, E*r*-2 and E*r*-10 it is on top of helix 3 (Luginbühl et al., 1994). The amino acid sequences from the end of helix 3 to the C-terminus are very similar in all GA pheromones (Vallesi et al., 1996), and hence the conformation type of the C-terminal 'tail' from the end of helix 3 to the C-terminus might be a critical feature for discrimination between the GA and PR groups of *E. raikovi* pheromones. This conclusion is in line with previous suggestions that more subtle variations of the arrangement of the C-terminal polypeptide segment further confer specificity to the *E. raikovi* pheromones within the GA and PR groups (Luginbühl et al., 1994), and would also provide a rationale for the 'intermediate behaviour' of cells producing E*r*-11 relative to mating induction with GA pheromones (Luporini et al., 1995; Vallesi et al., 1996).

Acknowledgements

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 C.N.R.

References

Bartels, C., Xia, T., Billeter, M., Güntert, P. and Wüthrich, K. (1995) *J. Biomol. NMR*, **6**, 1–10.

Bradshaw, R.A. (1996) In *Comprehensive Human Physiology: From Cellular Mechanism to Integration* (Greger, R. and Windhorst, U., Eds.), Vol. 1, Springer, Berlin, pp. 431–450.

Brown, L.R., Mronga, S., Bradshaw, R.A., Ortenzi, C., Luporini, P. and Wüthrich, K. (1993) *J. Mol. Biol.*, **231**, 800–816.

Güntert, P., Braun, W. and Wüthrich, K. (1991) *J. Mol. Biol.*, **217**, 517–530.

Luginbühl, P., Ottiger, M., Mronga, S. and Wüthrich, K. (1994) *Protein Sci.*, **3**, 1537–1546.

Luginbühl, P., Güntert, P., Billeter, M. and Wüthrich, K. (1996a) *J. Biomol. NMR*, **8**, 136–146.

Luginbühl, P., Wu, J., Zerbe, O., Ortenzi, C., Luporini, P. and Wüthrich, K. (1996b) *Protein Sci.*, **5**, 1512–1522.

Luporini, P., Vallesi, A., Miceli, C. and Bradshaw, R.A. (1995) *J. Euk. Microbiol.*, **42**, 208–212.

Luporini, P., Vallesi, A., Miceli, C. and Bradshaw, R.A. (1996) In *New Directions in Invertebrate Immunology* (Soderhall, K., Iwanaga, S. and Vasta, G.R., Eds.), SOS Publications, Fair Haven, NJ, USA, pp. 143–154.

Mronga, S., Luginbühl, P., Brown, L.R., Ortenzi, C., Luporini, P., Bradshaw, R.A. and Wüthrich, K. (1994) *Protein Sci.*, **3**, 1527–1536.

Ortenzi, C., Alimenti, C., Vallesi, A., Di Pretoro, B., La Terza, A. and Luporini, P. (2000) *Mol. Biol. Cell*, **11**, 1445–1455.

Ottiger, M., Szyperski, T., Luginbühl, P., Ortenzi, C., Luporini, P., Bradshaw, R.A and Wüthrich, K. (1994) *Protein Sci.*, **3**, 1515–1526.

Sevilla-Sierra, P., Otting, G. and Wüthrich, K. (1994) *J. Mol. Biol.*, **235**, 1003–1020.

Spera, S. and Bax, A. (1991) *J. Am. Chem. Soc.*, **113**, 5490–5492.

Stewart, A.E., Raffioni, S., Chaudhary, T., Chait, B.T., Luporini, P. and Bradshaw, R.A. (1992) *Protein Sci.*, **1**, 777–785.

Vallesi, A., La Terza, A., Miceli, C. and Luporini, P. (1996) *Eur. J. Protistol*, **32**, Suppl. I, 170–176.

Weiss, M.S., Anderson, D.H., Raffioni, S., Bradshaw, R.A., Ortenzi, C., Luporini, P. and Eisenberg, D. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 10172–10176.

Williamson, M.P., Havel, T.F. and Wüthrich, K. (1985) *J. Mol. Biol.*, **182**, 295–315.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.