

NEURONES in higher visual motion areas in the superior temporal sulcus (STS) of the macaque monkey respond to abstract random dot optic flow stimuli. Higher motion areas may not only represent, but in a next computational stage also analyse the flow field to determine, for instance, the direction of heading for navigation purposes. Real world visual scenes differ in several aspects from these abstract optic flow stimuli. We tested the neuronal response to naturalistic optic flow stimuli which simulated egomotion in different virtual environments and contained different numbers of visual cues. Neuronal activity depended mainly on the position of the focus of expansion rather than on other visual cues. This finding supports the hypothesis that higher motion areas within the STS analyse optic flow in natural scenes and can thus signal the direction of heading.

Key Words: Optic flow; Monkey; Extrastriate cortex; Heading; Area MST

Neuronal responses in the motion pathway of the macaque monkey to natural optic flow stimuli

Martin Pekel,^{CA} Markus Lappe,
Frank Bremmer, Alexander Thiele and
Klaus-Peter Hoffmann

Allgemeine Zoologie und Neurobiologie, Ruhr
Universität Bochum, D-44780 Bochum, Germany

^{CA}Corresponding Author

Introduction

Optic flow motion patterns occur whenever an observer is moving through the world. To avoid collisions or to navigate through a 3D environment containing stationary or moving obstacles, correct interpretation of the pattern of motion on the retina that is produced by self-motion is crucial. Humans can use this visual motion to accurately detect the direction of heading from optic flow stimuli.¹ The medial superior temporal area (area MST) in the parietal part of the superior temporal sulcus (STS) is thought to be involved in the analysis of optic flow.^{2–5} All studies so far have used abstract dot or stripe patterns, which are highly artificial stimuli compared with the natural environment of primates. The visual signals arising on the retina when an observer moves through a natural 3D environment differ from these stimuli in several aspects (motion parallax, colour, texture, object sizes, complex forms). To determine heading, the visual system has to perform at least two successive steps: the first involves the computation of the optic flow field itself from the retinal illumination patterns. The second step is the analysis of the optic flow to determine the egomotion parameters.^{6,7} When the direction of heading has to be determined from natural visual scenes, some difficulties arise. For instance, the aperture problem occurs in the first step of visual analysis.⁸ Local motion detectors in V1 cannot discriminate the true 2D-motion of an extended edge crossing the receptive field of a neurone. Instead they respond to the resulting vector perpendicular to the orientation of the edge. Thus, the signalled motion is dependent on the form of the

object. Pooling the information of several motion detectors can resolve this problem but leads to additional problems in the detection of motion discontinuities or transparent motion.⁹ A natural environment gives richer information to the visual system such as colour or texture which can be used to discriminate objects and, in turn, to overcome the aperture problem while simultaneously detecting motion boundaries or transparent motion.

A hierarchical processing of visual motion starts in V1 with small receptive fields and leads to area MT (V5) with medium-sized receptive fields. Area MT is believed to represent the flow field and to solve the aperture problem.¹⁰ MT neurones in turn project to area MST, which is thought to perform the optic flow analysis.^{2,4,6} From this point of view, optic flow analysis should be invariant towards the details of the environment. Neurones involved in optic flow analysis should be able to respond to naturalistic moving scenes in much the same way as to abstract random dot stimuli, provided both represent the same egomotion.

Materials and Methods

Animal preparation: Animal preparation followed standard procedures.¹¹ Briefly, two monkeys were trained to fixate a stationary spot of red light (0.8° diameter) during presentation of a stimulus. Under sterile conditions a head-holder, two scleral search-coils and two recording chambers were chronically implanted. The experiments started some weeks after the implantation. During an experimental session the

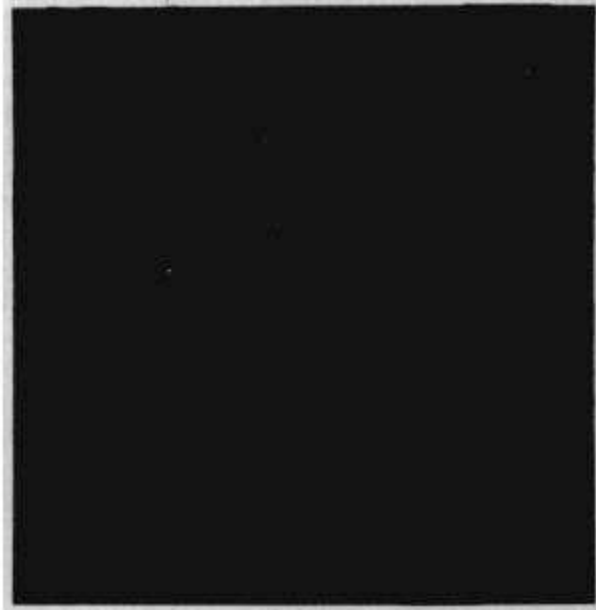


FIG. 1. Single frame of the natural scene stimulus. The naturalistic environment consisted of a ground plane on which a number of trees and stones were randomly placed. Single frame images were produced with a ray tracing program and later joined into a movie sequence. The viewpoint in the successive frames was varied in order to give the desired egomotion trajectory. The direction of view was varied in different movies in order to test different positions of the focus of expansion.

monkey was seated in a primate chair and his head was fixed. Eye position and neuronal data were recorded by a CED 1401 Interface (Cambridge Electronic Design LTD) and stored on a 486-PC. Area MST was located by physiological criteria. In the first monkey, histochemical markers were applied during the last recording session and the injection sites were determined using standard histological procedures.¹² The recording sites were clustered in a part of area MST that was adjacent to MT and extending down to the fundus of the superior temporal sulcus (STS). Histology of monkey 2 is not yet available. Experimental procedures are in accord with the published guidelines on the use of animal research (European Communities Council Directive 86/609/ECC).

Paradigm and stimulus: During fixation, a visual stimulus of $90 \times 90^\circ$ size was presented for 1.3 s on a flat screen 36 cm in front of the animal. The optic flow stimuli consisted of computer-generated movie sequences which simulated movement of an observer through a virtual environment. Three different types were used. The control stimulus stimulated movement towards a black wall covered with white dots. The neuronal response to the control stimulus was compared to the responses to stimuli in which motion parallax, colour and complex object forms were present. They simulated movement through a cloud of dots (parallax stimulus), or over a naturalistic

ground plane covered with stones and trees (naturalistic scene, Fig. 1).

The focus of expansion was presented at nine different locations on the screen: central and at eight different positions 40° eccentric. This was achieved by adjusting the simulated direction of gaze of the virtual observer with respect to his direction of movement. These conditions are similar to a paradigm used recently by Duffy and Wurtz.¹³ A period of 650 ms of simulated forward motion was followed by 650 ms of backward motion. Forward motion resulted in an expansional flow pattern, backward motion yielded a contractional flow pattern. The simulated speed of the observer was 3 m/s^{-1} . The distance to the wall in the control stimulus was 4.0 m. Ninety dots with a diameter of 20 μm were visible. This resulted in an average visual dot size of 3.8° . The cloud of dots in the parallax stimuli extended from 2 to 40 m in depth and contained 90 dots with the same absolute size of 20 μm . Visual dot size in the cloud stimulus depended on the distance of the individual dot from the observer. The luminance contrast of dots versus background was $>99\%$. The height above the ground plane in the naturalistic scene was 1.3 m. The naturalistic scene extended from 2 to 40 m in depth.

Data analysis: We analysed the neuronal data by taking the mean responses over fixed time intervals of 250–650 ms for expansion and 900–1300 ms for contraction. The delay of 250 ms between the onset of the stimulus and the measures start time was introduced to avoid including cell activity within the latency period. To determine whether a significant response to the visual stimulus occurred for an individual cell we performed a Kruskal–Wallis ANOVA on ranks over 18 stimulus conditions and the background activity. Only neurones that showed a significant response modulation by our visual stimuli ($p < 0.05$) were included in the data analysis. Several indices are introduced below for a quantitative comparison of the data. They are based on the mean values and on the PSTH data.

Results

All neurones recorded showed the typical properties of cells in area MST. They had large receptive fields (often covering the whole $90 \times 90^\circ$ screen, sometimes with smaller hot spots, and were highly motion sensitive to full-field unidirectional moving stimuli. Fifty-four neurones were recorded with optic flow stimuli of at least two stimulus types. Forty neurones could be recorded with all nine focus positions. All showed a significant response to at least one optic flow stimulus which was shown by an ANOVA on rank test. The response strength depended strongly on the position of the focus of

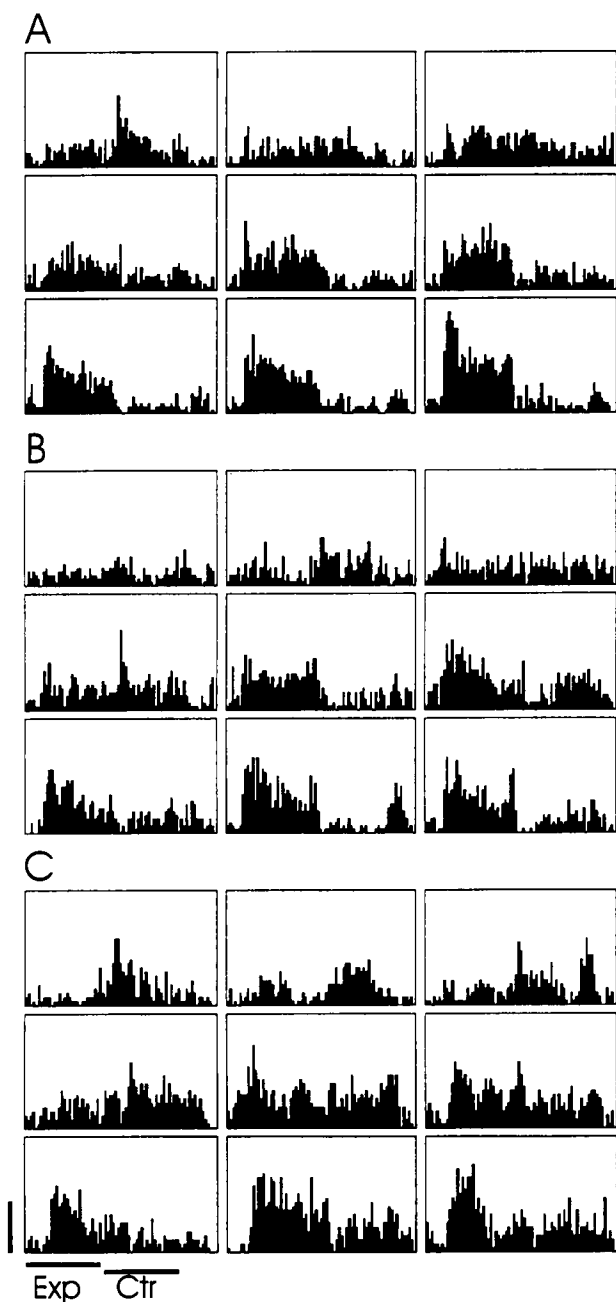


FIG. 2. Typical response of a neuron to the three stimulus types for all nine focus positions. Forward motion occurred during 0–650 ms, backward motion during 650–1300 ms. The response to the random dot stimulus without motion parallax (A) can be directly compared with the response to a random dot stimulus with parallax (B) and the naturalistic scene (C). The arrangement of the nine PSTHs corresponds to the position of the focus of expansion, which was centred or 40° eccentric.

expansion. The neurones showed qualitatively similar responses in the different environments. They were more or less independent of the number of visual cues present in the simulation but much more dependent on the simulated movement direction.

The neurone shown in Figure 2 is a typical example. The vertical bar on the lower left peristimulus–time

histogram (PSTH) indicates a spike rate of 100 spikes s^{-1} . The response strength varied little with the different stimulus types, but strongly with the position of the focus of expansion. The best response to expansion (0–650 ms) was observed for the lower right (control and parallax) or lower central (natural scene) focus position. The time course was more variable for the naturalistic scene than for the control. The upper-left, right and lower-right PSTHs show this clearly.

To quantify differences in the response strength, we calculated two indices. The first, the relative response strength index (RS) compared the responses to all nine egomotion stimuli in the two conditions for the single neurones. The second, the relative maximum response index (RM) compared only the maximum responses of all nine focus positions for two stimulus conditions.

$$RS = \frac{1}{9} \sum \frac{r_1 - r_2}{r_1 + r_2}$$

$$RM = \frac{r_{\max 1} - r_{\max 2}}{r_{\max 1} + r_{\max 2}}$$

where r_1 = response to test stimulus (parallax or naturalistic scene), r_2 = response to control stimulus (stimulus without parallax), $r_{\max 1}$ = maximal response to test stimulus of all nine focus positions, $r_{\max 2}$ = maximal response to control stimulus of all nine focus positions.

In a few cases RS compared responses for different focus positions. However a possible difference in the location of the maximum response is separately considered in the response gradient parameter below. Figure 3A shows the mean over all neurones of both indices for parallax and naturalistic scene versus control. Both indices can produce values in the range from –1.0 to 1.0. A value of zero indicates equal responses in both conditions. For both complex stimuli RS and RM are close to zero, indicating little difference in the response strength of the neurones for different environments.

Of particular interest for the determination of egomotion parameters is the ‘response gradient’ of the neurone, i.e. the direction from the fovea along which the greatest selectivity for the location of the focus of expansion (or the direction of heading in our stimuli) exists.^{6,13} We determined this direction by taking the response difference between expansion and contraction for each of the nine focus positions tested, performing a 2D multiple linear regression on the nine values, and determining the gradient of the regression plane. Only neurones for which the regression for all stimulus conditions was significant (F-test, $p < 0.05$) and which revealed a significant gradient along the horizontal or vertical axis (t -statistic, $p < 0.05$) were used for the analysis. The

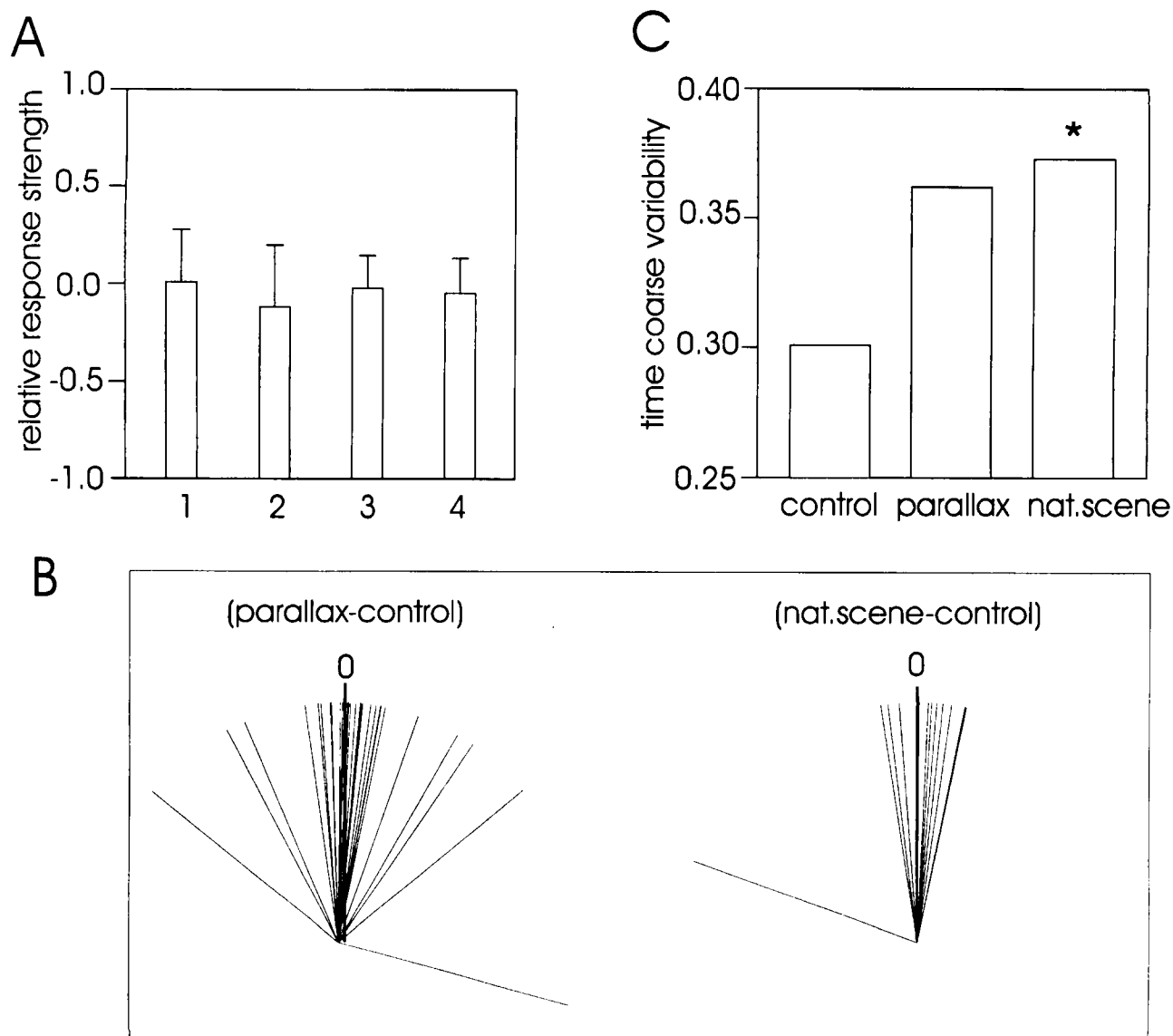


FIG. 3. (A) Average relative response strength for parallax vs control stimulus and for naturalistic scene vs control over all cells were calculated as indices RS and RM. RS includes all nine focus positions while RM gives the relative response strength for the maximum responses to the nine focus positions. 1: RS, parallax vs control; 2: RS, naturalistic scene vs control; 3: RM, parallax vs control; 4: RM, naturalistic scene vs control. The mean response of the cells is not greatly modulated if visual cues such as parallax and colour change. (B) Differences between the response gradient for parallax vs control and for natural scene vs control. The response gradient is the direction from the fovea along which the greatest selectivity for the location of the focus of expansion exists. The cells often show small changes of the response gradient in different environments while a few cells have larger differences for both the parallax and naturalistic scene paradigm compared to control. (C) The median of the 'relative variability over time' for the control stimulus, for the parallax stimulus and for the naturalistic scene. With increasing visual stimulus complexity the time course variability of the neuronal responses increases. Differences are significant for the natural scene compared to the control stimulus.

differences between the response gradients for the parallax paradigm vs control and for the naturalistic scene vs control are plotted in Figure 3B. For both stimulus types the angular differences were usually low. This means that the neuronal dependency on the position of the focus of expansion is similar in the different environments. Some neurones, however, show larger differences for the response gradients, indicating some influence of motion parallax or other cues present in the natural scenes. The mean of the direction difference over all neurones was close to zero for both complex stimulus types compared with

control (6.5° for parallax-control and -3.7 deg for natural scene-control). To quantify the observed differences in the variability of the time course of the response we calculated the standard deviations of the spike rates within 20 ms bins ('time course variability'). We only used the standard deviation from the PSTH showing the maximum response of all nine focus positions. The 'relative variability over time' (quotient of standard deviation and mean response) was calculated for each cell and each stimulus type separately. The median of the resulting values for all neurones is shown in Figure 3C. There

was an increase in the median time course variability when using more complex stimulus types which was significant for the natural scene against the control stimulus (U-test, $p = 0.017$).

Discussion

We compared the responses of single neurones in area MST of the macaque monkey to optic flow stimuli simulating egomotion in different virtual environments. Neuronal responses were strongly modulated by the position of the focus of expansion, and the mean response strength was largely invariant to the environmental layout. Variability of the response over time was greater in more complex environments.

The mean activities of the cells depended less on the visual details of the scene than on the simulated direction of movement of the observer. The response modulation agrees with recent findings by Duffy and Wurtz.¹³ Cells which modulate strongly in relation to specific visual properties of the scene (colour, contrast, motion parallax, etc.) would not be expected to analyse global optic flow and signal the direction of heading.

The influence of motion parallax was investigated explicitly in psychophysical studies of heading detection.¹ Motion parallax is a necessary cue for correct heading detection in conditions with simulated eye movements. When pure observer translations without eye movements were simulated, absence of motion parallax caused no difference for the detection of the focus of expansion. In our experiments, the monkey was always fixating. We found that neuronal responses changed little in the presence of motion parallax. Whether neurones behave differently during eye movements or whether the small differences observed are sufficient cannot be answered by the present experiments. On the other hand, during eye movements the system also receives extraretinal input,¹⁴ which could also be used to overcome ambiguity of the visual information.^{1,15}

One major difference between the naturalistic scenes and the random dot stimuli was the presence of colour. Several studies suggest that the macaque visual system is divided into a colour-sensitive object or form pathway and a colour-insensitive motion pathway.¹⁶ Dobkins and Albright found that neuronal responses in MT and oculomotor behaviour show the same preference to isoluminant heterochromatic (red/green) moving gratings.¹⁷ Gegenfurtner *et al*

showed that one-third of the neurones in area MT was sensitive to colour-modulated movement stimuli.¹⁸ Because of a high threshold, however, they concluded that MT neurones are unlikely to be the source of behaviourally relevant chromatic motion signals. This is supported by our findings because area MT provides the major source of input for area MST.¹⁹ In our study we found few differences in neuronal responses to black and white stimuli (random dots) and the coloured naturalistic scenes. We found some increase in variability of the neuronal response overtime when coloured stimuli with large objects were applied or when motion parallax was present. This could be due to the fact that the distribution of objects in these stimuli was less homogeneous than in the control stimulus. Moreover, for the parallax and naturalistic scene paradigms the angular object size increased with the distance between the object and the observer.

Conclusions

The recorded neurones are modulated mainly by the position of the focus of expansion. They are little influenced by other visual cues, such as colour and object size. Our results support the view that neurones in area MST are involved in the computation of self-motion.

References

- Warren WH and Hannon DJ. *J Ophthalmol Soc Am* **A7**, 160–169 (1990).
- Duffy CJ and Wurtz RH. *J Neurophysiol* **65**, 1329–1345 (1991).
- Tanaka K and Saito H. *J Neurophysiol* **62**, 626–641 (1989).
- Lagae L, Maes H, Raiguel S *et al.* *J Neurophysiol* **71**, 1597–1626 (1994).
- Graziano MSA, Andersen RA and Snowden RJ. *J Neurosci* **14**, 54–67 (1994).
- Lappe M and Rauschecker JP. *Neural Comput* **5**, 374–391 (1993).
- Lappe M and Rauschecker JP. *Nature* **369**, 712–713 (1994).
- Marr D and Ullman S. *Proc R Soc Lond B* **211**, 151–180 (1981).
- Hildreth EC and Koch C. *Annu Rev Neurosci* **10**, 477–533 (1987).
- Movshon JA, Adelson EH, Gizzi MS *et al.* In: Chagas C, Gattass R and Gross C, eds, *Pattern Recognition Mechanisms*. New York: Springer, 1985: 117–151.
- Bremmer F, Ilg UJ, Thiele A *et al.* *J Neurophysiol* In Press (1995).
- Distler C, Boussaoud D, Desimone R *et al.* *J Comp Neurol* **334**, 125–150 (1993).
- Duffy CJ and Wurtz RH. *J Neurosci* **15**, 5192–5208 (1995).
- Newsome WT, Wurtz RH and Komatsu H. *J Neurophysiol* **60**, 604–620 (1988).
- Royden CS. *Vision Res* **34**, 3215–3222 (1994).
- Ungerleider LG and Mishkin M. In: Ingle DJ, Goodale MA and Mansfield RJW, eds, *Analysis of Visual Behavior*. Cambridge, MA: MIT Press, 1982: 549–586.
- Dobkins KR and Albright TD. *Vis Neurosci* **12**, 321–332 (1995).
- Gegenfurtner KR, Kiper DC, Beusmans JMH *et al.* *Vis Neurosci* **11**, 455–466 (1994).
- Ungerleider LG and Desimone R. *J Comp Neurol* **248**, 190–222 (1986).

ACKNOWLEDGEMENTS: This work was supported from the Deutsche Forschungsgemeinschaft and ESPRIT.

Received 8 January 1996;
accepted 9 February 1996