The neurobiology of cognition

M. James Nichols and William T. Newsome

Perhaps the deepest mysteries facing the natural sciences concern the higher functions of the central nervous system. Understanding how the brain gives rise to mental experiences looms as one of the central challenges for science in the new millennium.

It is astounding that cognition and emotion — phenomena that cannot be duplicated in our most sophisticated computers — arise naturally from the electrical activity of large systems of neurons within the brain. Scientific investigation of these phenomena is inherently interdisciplinary, drawing strength from fields as diverse as neurophysiology, cognitive psychology and computational theory. Exciting new findings have emerged in recent decades concerning the neural underpinnings of cognitive functions such as perception, learning, memory, attention, decision-making, language and motor planning, as well as the influence of emotion and motivation upon cognition. With very few exceptions, however, our understanding of these phenomena remains rudimentary. We can identify particular locations in the brain where neuronal activity is modulated in concert with particular external or internal stimuli. In some cases we can even artificially manipulate neural activity in a specific brain structure (using electrical or pharmacological techniques) and cause predictable changes in behaviour. But we encounter substantial difficulties in understanding how modulations in neural activity at one point in the nervous system are actually produced by synaptic interactions between neural systems. Thus our current state of knowledge is somewhat akin to looking out the window of an airplane at night. We can see patches of light from cities and towns scattered across the landscape, we know that roads, railways and telephone wires connect those cities, but we gain little sense of the social, political and economic interactions within and between cities that define a functioning society.

To achieve a more sophisticated level of understanding, investigators must develop new experimental techniques for studying functional interactions between neurons and systems of neurons, and new models for understanding the behaviour of complex, dynamic systems like the brain. Whether major breakthroughs occur on the timescale of years or decades depends substantially on success in developing these new techniques. Irrespective of timescale, an increasingly sophisticated understanding of the neural basis of cognition will influence our society profoundly. It will have practical applications such as treatment of mental disease and the design of intelligent machines, and will raise contentious social issues such as the freedom of each individual to choose their behaviour, and the extent to which society can reasonably demand individual responsibility for behaviour.

Figure 1 **Visual scene of apples on a picnic table. Photograph by David G. Muir.**

In what follows, we address three questions. What sorts of cognitive phenomena do neuroscientists seek to explain in terms of brain function? What form do our explanations take, and what technical advances are round the corner to help us? And what are the personal and social implications of understanding the neural underpinnings of consciousness and mental functions?

Visual perception and cognition

The following example of visually based cognition illustrates some of the mental phenomena that neuroscientists seek to understand. Imagine a girl standing before a picnic table (Fig. 1), scrutinizing a jumble of apples in search of the biggest and best granny smith, her favourite kind. She looks, finds one, reaches out and grabs it. The action seems straightforward enough, but this simple description belies the enormous complexity the brain faces in perceiving the scene, deciding on a course of action, and then executing it.

The brain has no direct access to the rich array of objects and surfaces in the three-dimensional visual environment; it must reconstruct the scene from complex, two-dimensional images falling on the two retinae. Consider just three of the many problems the visual system must solve in scene reconstruction. First, although the retinal images are flat, accurate depth perception is critical for assessing the shapes and sizes of the apples, and for reaching accurately to grab the selected apple. The brain reconstructs the third dimension from multiple cues in the retinal image. For example, the brain can infer depth based on the slight disparity between the images an object casts on the two retinae. Second, in order to pick out the largest apple, the girl must accurately assess the size of each apple. Yet the size of an apple's image on her retina depends on its distance from her. Her visual system takes distance into account and automatically estimates the true physical sizes of objects at different distances, a perceptual phenomenon known as size constancy. Third, the apple scene contains a kaleidoscope of contours: texture elements, colour boundaries, shadows, reflections, specks of dirt, and so on. Only a fraction of these delineate the true boundaries of objects, however, and the brain must sort these out (contour extraction) in order to identify objects accurately.

Once the girl's visual system reconstructs an accurate representation of the scene, a higher-level decision process must evaluate the perceptual information and select one apple in particular. More specifically, her brain must categorize the apples (granny smith, red delicious and so on), assign appropriate affective associations (likes and dislikes) to them, deploy spatial attention to salient objects in the scene, and discriminate fine differences in colour, size and shape to select the best of the granny smiths. Thus, the girl's decision is shaped by immediate sensory information, by previously learned categories drawn from visual memory, and by likes and dislikes based on accumulated experience. Other elements in the scene that are totally irrelevant to the decision must be ignored. Finally, once the brain reconstructs the scene and makes a decision, voluntary

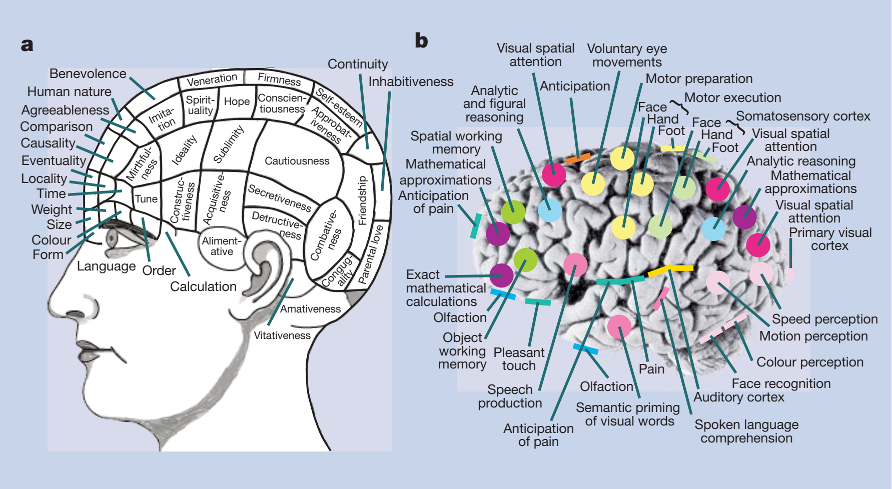

Figure 2 **Localization of mental functions.** a, **According to nineteenth-century phrenologists, lumps on the skull revealed the locations in the brain responsible for mental functions like 'agreeableness' in the frontal lobe and 'cautiousness' in the parietal lobe.** b, **PET and fMRI studies, which measure changes in cerebral blood flow while various perceptual and cognitive tasks are performed, have revealed a different organization. Imaging data assembled by C. J. Doane.**

movement systems must plan and execute the appropriate behavioural response (reaching for the selected apple).

Thus, choosing an apple from an apple cart, like innumerable actions we carry out every day, involves a surprising array of cognitive functions. Scientists would like to understand the neural mechanisms that underlie those functions.

Levels of understanding

Understanding the neural basis of a specific cognitive function typically begins with behavioural observations and hypotheses developed by perceptual and cognitive psychologists. Equipped with sound conceptual frameworks originating in behaviour, neurophysiologists can then study underlying brain function at several levels. We will describe three of the most important levels: localization, representation and microcircuitry.

At the coarsest level, the primary issue is 'localization' of function: identifying, for example, neural systems in the brain that are strongly active in response to visual images or when spatial attention is deployed to different regions of a visual scene. As illustrated in Fig. 2a, localization of function has been a dominant theme in brain science, beginning in the nineteenth century when 'phrenologists' attempted to map mental functions onto the brain by correlating aspects of personality and mental ability with the sizes of bumps at different locations on the skull. More reliable evidence for localization of function emerged in the early part of the twentieth century as neurologists learned to recognize mental deficits (sometimes highly specific) that occurred subsequent to damage in particular regions of the brain.

The most commonly used techniques in modern studies of localization are positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), which generate most of the glossy figures in popular science magazines that depict mental activity as coloured blobs on a picture of the brain. Both PET and fMRI measure changes in blood flow to specific regions of the brain while human subjects perform various cognitive tasks[1]. The blood flow signal is assumed to reflect changes in metabolic demand resulting from altered levels of neural activity. Using PET and fMRI, investigators can study brain activity in humans at the spatial scale of individual brain structures (a few millimetres) and on a timescale of a few seconds.

The recent avalanche of PET and fMRI papers has produced many insights concerning localization of mental functions (Fig. 2b)[2–4]. Although they are substantially more informative than previous findings based on brain damage, distillations of data like the one in Fig. 2b are disconcertingly similar to the one-function-one-brain-area maps of the phrenologists. Will we 'understand' the brain when the map in Fig. 2b is completely filled with blobs? Obviously not; localization data provide little insight into the exact nature of the signals encoded in a given structure, the computations being performed and the interactions between different structures.

Greater insight into how the brain processes and represents information can be obtained by finer-grained studies of individual brain structures using microelectrodes. With these techniques, investigators can measure directly the electrical activity of individual neurons (Fig. 3a), or of small, functionally related clusters of neurons called 'cortical columns'. For example, Fig. 3b schematizes the electrical discharges (black vertical tick marks) of a cortical neuron that responds vigorously to a bar of light sweeping downward across a viewing screen, but not to upward motion of the bar. This neuron, therefore, is 'selective' for downward visual motion. Other neurons are selective for other directions of motion, specific colours, the orientation of line segments, the density of visual texture and many other visual features. Their selectivity allows them to signal or represent the presence or absence of particular features in the visual environment. This 'representational' level of analysis has provided some of the major success stories of systems neuroscience[5,6], including descriptions of early stages of visual processing relevant to the problems of contour extraction and depth perception described in the preceding section.

Investigation at the representational level is particularly powerful when carried out in alert animals trained to perform interesting cognitive tasks. For example, researchers have discovered that certain neurons in the temporal and frontal lobes of the brain are active during short-term memories of specific objects or places[7–9], whereas other neurons in the parietal lobe respond dramatically when the animal deploys attention to one or another region of a visual scene[10–12]. These functional properties can change strikingly from one cortical column to the next within a given brain area (a spatial scale of several tens of micrometres), revealing an intricate level of organization invisible to fMRI and PET.

These discoveries provide researchers with a deeper 'point of entry' for analysis of the neural systems underlying a particular cognitive function. Transformations in the information carried by single neurons can be inferred by recording from successive brain areas in a particular pathway while the animal performs a task. On the basis of such measurements, investigators can develop quantitative models of how these transformations take place. Furthermore, investigators can test hypotheses about the function of a particular pathway by applying electrical or pharmacological techniques to activate or inactivate discrete clusters of neurons. In our laboratory, for example, we have found that electrically activating clusters of direction-selective neurons through a microelectrode (Fig. 3a) can induce rhesus monkeys to report seeing motion in the direction encoded by the activated neurons[13]. Results of this sort, perhaps more decisively than any others, establish a causal link between the activity of particular classes of neurons and specific mental phenomena.

But even this level of analysis begs fundamental questions about how signals are created, encoded and transmitted by single neurons and assemblies of neurons.

Individual cortical neurons (like the one in Fig. 3a) receive input signals from as many as three to ten thousand other neurons. In a typical experiment, however, the neurophysiologist can characterize the responses of only a few neurons at the tip of the electrode. With such a limited data set, it is difficult to determine exactly how the thousands of synaptic inputs to a cell are transformed to create the cell's pattern of output activity.

This third level of analysis — the detailed 'microcircuitry' and dynamic interactions that give rise to the observed activity of single neurons — is the most vexing for systems and cognitive neuroscience. The challenge stems in part from the extreme complexity of even a single cortical neuron, and in part from daunting practical problems of physical accessibility and visualization in an intact, behaving animal. Elegant intracellular recording techniques can be used to study interactions between neurons in simple preparations (for example, thin slices of tissue removed from the brain, or simple invertebrate nervous systems). The sequence of images in Fig. 3c, for example, shows synaptic events within individual compartments of a single cell in the brain of the common housefly. Each compartment receives synaptic signals from a different input neuron. This neuron is selective for downward visual motion, just like the one in Fig. 3b. In the fly brain, however, it is possible to monitor activity sweeping through the individual compartments of the cell as a bar is swept downward through the visual field. Observations of this kind will shed light on how individual neurons transform the input signals they receive from other neurons into an output signal. Even so, understanding the neural microcircuitry underlying complex cognitive phenomena like decision-making in mammals seems a distant hope, although new technical advances could substantially influence this level of analysis.

A look over the horizon

The immediate impediments to progress in systems and cognitive neuroscience are more technical than conceptual. Computational theorists and psychologists have developed plausible models for many cognitive functions; our primary problem lies in acquiring and analysing the neural data needed to evaluate these models. It will be a sign of real progress if this situation reverses over the next two decades, with conceptual issues coming to the fore.

Technical innovation is needed at all levels of analysis. At the level of localization, improvements in fMRI technology could permit study at the spatial resolution of the cortical column, where so much functional specialization resides. This capability is possible in principle and should emerge gradually over the next decade. More problematic, however, is the temporal resolution of fMRI, which is limited by blood-flow dynamics to a scale of seconds. Neural processing, in contrast, occurs on a scale of milliseconds. Our best prospects for recovering temporal resolution lie in combining fMRI with techniques that measure electrical activity directly, such as evoked potential recording[14] and magnetoencephalography[15].

At the level of representational analysis, multi-electrode recording techniques are enabling investigators to obtain data simultaneously from larger populations of neurons, and may provide the most promising approach for analysing dynamic interactions between cortical areas.

Optical techniques are likely to become increasingly important both at the representational level of analysis and at the local circuit level. Currently, CCD cameras are capable of gathering reflected (or fluorescent) light from the brain surface, thereby enabling study of brain function at the level of individual cortical columns[16]. Study of single neuron dynamics will increasingly rely on multi-photon microscopy, which can isolate signals from individual cells *in vivo* hundreds of microns below the cortical surface[17]. Focal stimulation of individual neurons will increasingly involve optical release of caged neurotransmitters[18]. The rapid development of new optical probes, caged neuroactive compounds, and light sources and detectors will extend current resolution limits and may permit the application of optical methods to alert animals.

Implications for society

Progress in understanding higher brain functions, fuelled by technological innovation, will certainly exert a major impact on society in the coming century. New discoveries will be important for the diagnosis and treatment of psychiatric and neurological disease, for improving human learning and communication and for informing the design of intelligent machines. But a scientific understanding of the human mind and human behaviour in terms of brain function could also have a profound impact on how we understand ourselves and our society. Two aspects of mentality deserve special consideration: conscious experience and decision-making.

Consciousness

The relationship between conscious experience and brain function is one of the great remaining mysteries of the natural world. How can three pounds of tissue that scientists study with microelectrodes, microscopes and magnetic resonance machines give rise to conscious experience? Indeed, what exactly do we mean by 'conscious experience'?

A classic thought experiment[19,20] illustrates the problem. Imagine a time in the future when visual neuroscience has come to completion, when everything is known about how the nervous system responds to light and how it produces visually guided behaviour. Now imagine a colour-blind neuroscientist in this golden era who can give a complete account of the neural events underlying a normal observer's ability to discriminate and identify colours. Now suppose our scientist's colour-blindness were miraculously cured and he saw green for the first time. The question is, would he learn anything new about colour perception? Most would answer 'yes': what he learns is precisely the conscious experience of colour. Consciousness is the aspect of our mental life that we can only understand through subjective first-person experience.

Can we study consciousness scientifically? This is controversial among researchers, but we believe that progress will occur only if investigators can develop operational

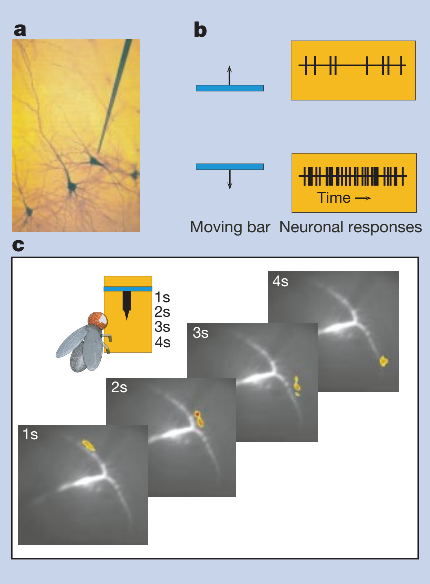

Figure 3 **Illustration of the representational and microcircuit levels of analysis.** a, **Pyramidal neurons in monkey visual cortex, stained by the Golgi method. A tungsten microelectrode typically used for extracellular recording has been superimposed (image from ref. 5).** b, **Schematic of a bar stimulus (blue bar) moving either up or down and the response of a direction-selective neuron. Black tick marks represent the electrical discharges of the neuron. The neuron is direction selective, discharging more vigorously for downward motion than for upward motion.** c, **Tangential cell in a housefly in response to a bar moving downward in the visual field (the preferred direction for that cell). The cell was filled with calcium green, which fluoresces in response to influxes of calcium caused by synaptic inputs from other neurons. The four images represent four successive 1-second intervals as the bar sweeps through the visual field producing electrical changes in different spatial compartments of the cell (S. Single and A. Borst, unpublished data).**

definitions of conscious experience, as has been accomplished for other aspects of cognition. In a classical form of learning, for example, Pavlov's dog generated a measurable dependent behaviour (salivation) in response to controlled manipulation of independent variables (pairing of bell with food). In general, we can hope to understand the neural basis of any cognitive phenomenon that can be defined operationally in a similar manner.

Accomplishing this for consciousness will be difficult, because conscious experience is intrinsically subjective. A person's verbal reports of her conscious experiences may be sufficiently reliable to serve as objective observations for testing hypotheses about consciousness. But the notion that verbal reports (or any other overt behaviour) are an adequate proxy for conscious experiences is itself not scientifically testable. We must therefore acknowledge that this assumption is extra-scientific (as is parsimony, another favourite assumption of scientists). Even if we do accept this assumption in general, verbal reports may fall short under some circumstances. For example, if one person reacts to pain more readily than another, is it because he has a more intense sensation of pain, or because his threshold for reacting to pain is lower? Could any verbal exchange answer such questions? Although animals can be trained to perform simple behavioural tasks, these sorts of questions only become more perplexing without the benefit of language. Therefore, any measure of consciousness in animals is likely to be controversial, despite the common intuition that the family dog enjoys conscious experiences of one sort or another.

Ultimately, no matter how precisely we manage to link verbal reports of conscious experience to brain activity, fundamental mysteries are likely to persist[21]. Exactly how something as ineffable as subjective consciousness can arise from macromolecules, synapses and action potentials may well remain a conundrum.

Decision-making and determinism

An exciting frontier in the study of higher brain function is the attempt to understand mechanistically how decisions are formed. 'Decision processes' are the key cognitive links between perception and action. Perceptual processes carry out the functions of scene reconstruction, contour extraction, and so on, and motor processes implement the planning and execution of a behavioural response. Intermediate levels of the system, however, must evaluate the sensory evidence represented in early cortical areas, and 'decide', for example, which apple is to be the target of a reaching movement.

Neurophysiologists have now begun to investigate the neural underpinnings of decision processes in monkeys trained to perform simple discrimination tasks. By recording at successive stages of known anatomical pathways linking sensory to motor areas of the cortex, investigators have uncovered intriguing evidence for neural systems that represent what the monkey 'decides' about the stimulus as opposed to what the stimulus actually is[22–24]. The neural activity recorded in such experiments can actually predict the outcome of the monkey's decision several seconds before it is revealed by a behavioural response. In addition, neurophysiologists are now taking notice of a fact that psychologists and economists have long recognized: decisions are informed as much by reward expectation and personal history as by the sensory evidence available at any given time. If, for example, an animal knows from past experience that choice A tends to be rewarded more frequently than choice B, it may choose A even if countervailing sensory evidence suggests that choice B will be rewarded on the current trial. Important new discoveries have been made concerning neural signals related to reward expectation and relative reward potency[25–27]; indeed such signals have been found in the inferior parietal cortex of monkeys[28], a cortical region implicated in perceptual decision-making.

Limits to determinism?

The next decade will certainly witness substantial progress in understanding how decisions are formed within the brain, but these studies in particular raise disquieting questions. The crux of the problem lies in the implications of physical determinism for our concepts of personal freedom and moral responsibility. If the most sophisticated aspects of our mental lives, from decisions as trivial as selecting an apple to those as important as choosing a spouse, are determined by the molecular and cellular events that generate electrical activity in the nervous system, what are we to make of our subjective experience of freedom? If our sense of freedom in making moral decisions is illusory, can anyone reasonably be held responsible for his or her actions? Can a criminal reasonably be punished for actions over which he has no meaningful control?

We have no solution to this problem, but several perspectives seem important to bear in mind. The most significant, perhaps, is that both points of view — physical determinism and personal freedom — are utterly essential in humanity's attempt to understand itself. The assumption of cause-and-effect determinism is fundamental not only to science but also to everyday life. For example, every day most of us must negotiate streets full of busy traffic in order to survive until the next day. The basic perceptual, decision-making and motor mechanisms within our brains must be sufficiently deterministic that we take the appropriate action every time. Furthermore, the assumption of determinism underlies our hope of finding cures for devastating psychiatric and neurological diseases. If these conditions do not have physical causes within the brain, there is no reason to hope that they can be cured by physical (that is, medical) interventions.

On the other hand, the assumption of a meaningful degree of personal freedom is essential not only to our personal and social lives, but to science as well. Scientists require the freedom to evaluate data and to reject false hypotheses. But if our mental processes unfold with the physical determinism of a machine, what guarantee do we have that the beliefs the machine generates, by the scientific method or otherwise, are true? And if the machine generates false beliefs (the belief in determinism, perhaps?) how would we discard them? We are left, then, with the paradox that both perspectives seem necessary, for the community of science as well as for everyday life. Perhaps these perspectives will be reconciled at some point in the future. For the present, our only obvious option is to live with both, and accept the paradox as a leavening dose of humility in our intellectual lives.

M. James Nichols and William T. Newsome are at the Howard Hughes Medical Institute and Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA.
e-mail: bill@monkeybiz.stanford.edu

1. Ogawa, S. *et al. Proc. Natl Acad. Sci. USA* **89,** 5951–5955 (1992).
2. Posner, M. I. & Raichle, M. E. *Proc. Natl Acad. Sci. USA* **95,** 763–764 (1998).
3. Sell, L. A. *et al. Eur. J. Neurosci.* **11,** 1042–1048 (1999).
4. Wandell, B. A. *Annu. Rev. Neurosci.* **22,** 145–173 (1999).
5. Hubel, D. H. *Eye, Brain, and Vision* (Scientific American Library, New York, 1995).
6. Parker, A. J. & Newsome, W. T. *Annu. Rev. Neurosci.* **21,** 227–277 (1998).
7. Tanaka, K. *Curr. Opin. Neurobiol.* **2,** 502–505 (1992).
8. Fuster, J. M. *Memory in the Cerebral Cortex* (MIT Press, Cambridge, 1995).
9. Goldman-Rakic, P. S. *Proc. Natl Acad. Sci. USA* **9,** 13473–13480 (1996).
10. Desimone, R. & Duncan, J. *Annu. Rev. Neurosci.* **18,** 193–222 (1995).
11. Maunsell, J. H. *Science* **270,** 764–769 (1995).
12. Colby, C. L. & Goldberg, M. E. *Annu. Rev. Neurosci.* **22,** 319–349 (1999).
13. Salzman, C. D., Murasugi, C. M., Britten, K. H. & Newsome, W. T. *J. Neurosci.* **12,** 2331–2355 (1992).
14. Martinez *et al. Nature Neurosci.* **2,** 364–369 (1999).
15. George, J. S. *et al. J. Clin. Neurophysiol.* **12,** 406–431 (1995).
16. Hubener, M., Shoham, D., Grinvald, A. & Bonhoeffer, T. *J. Neurosci.* **17,** 9270–9284 (1997).
17. Svoboda, K., Helmchen, F., Denk, W. & Tank, D. W. *Nature Neurosci.* **2,** 65–73 (1999).
18. Wang, S. S. & Augustine, G. J. *Neuron* **15,** 755–760 (1995).
19. Jackson, F. *Phil. Quarterly* **32,** 127–136 (1982).
20. Thompson, E. *Phil. Studies* **68,** 321–349 (1992).
21. Nagel, T. *Phil. Rev.* **4,** 435–450 (1974).
22. Shadlen, M. N. & Newsome, W. T. *Proc. Natl Acad. Sci. USA* **93,** 628–633 (1996).
23. Romo, R. & Salinas, E. *Curr. Opin. Neurobiol.* **9,** 487–493 (1999).
24. Schall, J. D. & Thompson, K. G. *Annu. Rev. Neurosci.* **22,** 241–260 (1999).
25. Gallistel, C. R. *Cognition* **50,** 151–170 (1994).
26. Schultz, W., Dayan, P. & Montague, P. R. *Science* **275,** 1593–1599 (1997).
27. Schultz, W. *J. Neurophysiol.* **80,** 1–27 (1998).
28. Platt, M. L. & Glimcher, P. W. *Nature* **400,** 233–238 (1999).

Acknowledgements. We thank M. Schnitzer for useful suggestions during the preparation of the manuscript.