The human stretch reflex and the motor cortex

Peter B. C. Matthews

The spinal stretch reflex, exemplified by the tendon jerk, appears to be less important in humans than a delayed 'long-latency' response. This is easily observed when muscles of the hand are stretched while they are already contracting voluntarily. On limited evidence, many have long held that the delayed response is a transcortical reflex and have tended to neglect alternative possibilities, particularly that it might be a spinal reflex dependent upon slow afferents. New experiments have now eliminated the alternatives, leaving the transcortical hypothesis in command of the field.

Peter B. C. Matthews is at the University Laboratory of Physiology, Oxford OX1 3PT, UK.

Liddell and Sherrington's description in 1924 of the stretch reflex in the decerebrate cat stands as a pillar of classical neurophysiology[1]; the long-known tendon jerk was relegated to being defined as a 'fractional manifestation' of a reflex that normally acted tonically (Fig. 1). Their successors enthusiastically analysed the spinal mechanisms of this 'simplest' of reflexes in the cat, while for the textbooks it became generally accepted that 'our posture is largely maintained by its action'. Such restriction of vision and simplistic generalization received a rude shock some 30 years later when in 1960 Hammond studied the actual effect of stretching a human muscle during its voluntary contraction[2] (see Ref. 3 for review). It did, of course, respond but the major action occurred with a latency that was too long for the classical, monosynaptic spinal stretch reflex to be responsible and too short for the reaction to be 'voluntary'. Hammond's experiment has since been repeated endlessly. There has been much quibbling over detail, but it has come to be generally accepted that the human stretch reflex has at least two quite separate components: the classical short-latency spinal reflex (M1), followed by a long-latency component (M2) of more complex origin[4,5] (Fig. 2A). These are not always well demarcated from each other, so it is hardly surprising that there is little agreement over their precise relative timing and magnitude. They also vary from muscle to muscle, with the long-latency response being particularly well developed for muscles of the hand.

Hammond suggested two ways in which the extra delay of the second response might be produced: either 'it is carried by slower afferents' or 'it takes a longer route in the central nervous system' (Fig. 2B). He left it at that, and for the next 10 years his work was largely neglected. Then, in 1970, Phillips raised the specific question: 'is the CM (cortico-motoneuronal) projection itself the efferent limb of a transcortical spindle circuit?'[6]; this query was a result of reviewing his extensive work on the baboon, and without reference to Hammond. In favour of the answer 'yes', Phillips cited two convincing findings: (1) Neurones in area 4 of the motor cortex project monosynaptically to spinal motoneurones. (2) Fast muscle-spindle afferents project to area 3a of the sensory cortex. However, two years later he returned the answer 'no' when he failed to find a direct connection from 3a to 4 (Ref. 7). Nevertheless, his original argument had by then been enthusiastically accepted, and for many years those favouring it routinely cited Phillips' 'yes', using his authority to support their case, and simply ignored his subsequent 'no'. Moreover, any 'long-latency' response was often automatically taken to be a 'long-loop' reflex, without considering other possibilities for the delay, and mistakenly equating the two conceptually distinct terms. However, Phillips' first paper was rightly influential, since the essence of science lies in asking the right question at the right time; it is now just a detail that the routing of afferent activity to area 4 remains under debate[8,9], since its neurones do respond to muscle stretch[10]. It remains disturbing, however, that too many blindly followed fashion to the detriment of critical analysis. This helped trigger revolt, as described below, by those who recognized that the emerging consensus was based more on enthusiastic belief than on the elimination of alternative hypotheses.

Accumulating support

Despite Phillips' negative conclusion, solid evidence continued to accumulate in favour of the transcortical hypothesis.

(1) The minimum time needed to traverse the potential pathway was assessed ever more accurately in both animals and humans. This was done by breaking into the pathway at various points, either to record the activity elicited by physiological stimuli or to activate the loop artificially by electrical stimulation. It was proved that there is ample time available for the transcortical transmission of the later part of the reflex. Indeed, in relation to the peripheral transit time, the human cortex is remarkably close to the spinal cord. (Conduction time from the motor cortex to cervical motoneurones is 5–6 ms, from these to the hand it is 15 ms, and afferent timing is similar. Corresponding reflex latencies: M1, 30 ms; M2, 50–60 ms.)

(2) Lesions along the central course of the pathway in humans were frequently found to have a much greater effect on the later 'M2' component of the response than on the initial M1 component[11]. However, such observations cannot prove that the reflex pathway itself has been interrupted (the excitability of crucial spinal interneurones might simply have changed on interfering with their descending inputs).

(3) Evarts[12,13] was the first to make recordings from single cortical neurones in conscious animals while the reflex was being elicited; appropriate increases in firing were then found in the monkey but not in the cat[14]. This approach culminated in the experiments of Cheney and Fetz, who studied cortical neurones that had been shown to project monosynaptically to the motoneurones of the stretched muscle[15]. This paper confirmed beyond doubt that the cortex was contributing to the reflex

0166 - 2236/91/$02.00

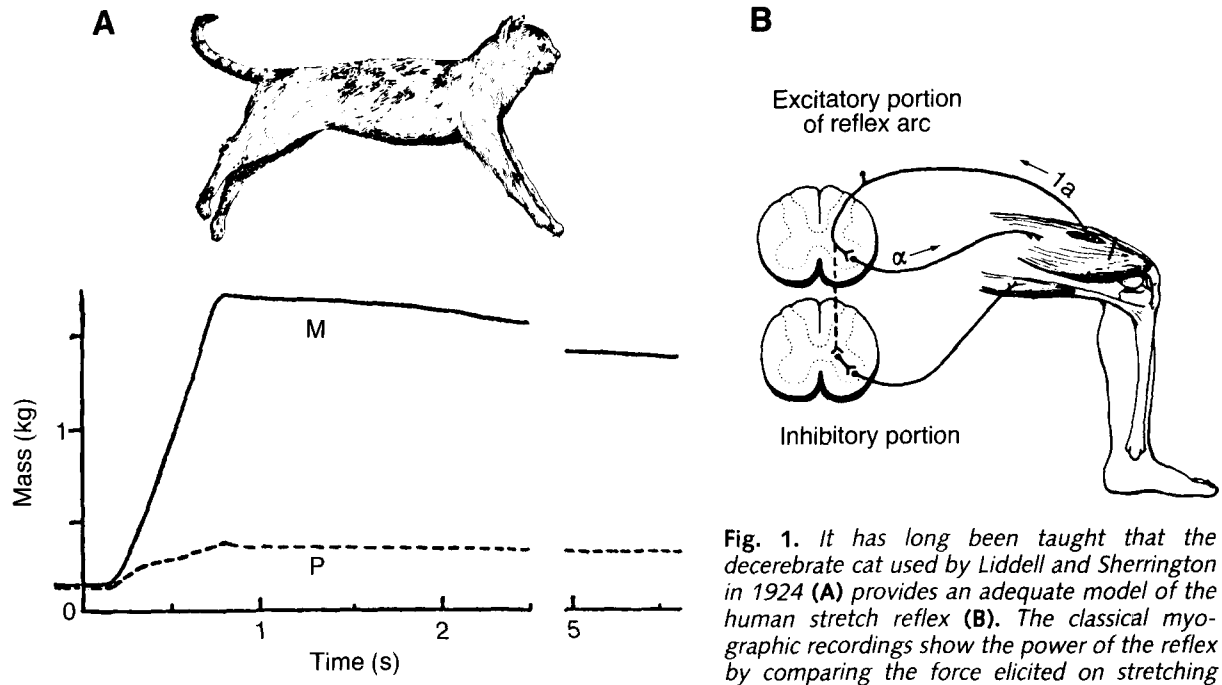


Fig. 1. *It has long been taught that the decerebrate cat used by Liddell and Sherrington in 1924* **(A)** *provides an adequate model of the human stretch reflex* **(B).** *The classical myographic recordings show the power of the reflex by comparing the force elicited on stretching quadriceps in the decerebrate cat with (M) and without (P) the reflex in action. (B) A typical diagram from modern teaching shows how the reflex arc mediates the human knee jerk. More prolonged, physiologically significant stretch reflex action is commonly supposed to depend solely upon the same neural circuitry.*

under study, without showing whether or not it was entirely responsible. However, it can be questioned whether the reflex then studied was indeed the homologue of the human stretch reflex. Monkeys responded to the disturbance by co-contracting agonists and antagonists, so that a given muscle gave the same long-latency response irrespective of whether it was being stretched or released. In contrast, the human stretch reflex behaves symmetrically so that unloading a contracting muscle reduces its contracting activity (Fig. 3A). Separate short- and long-latency components of action can again be recognized, both of which are representative of a disfacilitation of the pathways on shortening of the agonist, rather than of a reflex inhibition from the already stretched antagonist (Fig. 3B)[16].

An element of confusion was introduced by the suggestion that cutaneous afferents were entirely responsible for the standard M2 response, so that muscle afferents were presumed to be responsible for only the M1 response[17]. (In human and whole-animal experiments, muscle stretch is routinely accompanied by joint rotation and cutaneous stimulation.) However, the M2 responses that are elicited by moving a digit can survive after the digit has been locally anaesthetized, suggesting that in fact muscle afferents must be involved[18]. Cutaneous afferents probably also contribute on occasion, since long-latency responses of comparable latency are found when cutaneous nerves are stimulated[19]. In other words, there is more than one reflex of 'M2' latency[16]. All in all, in spite of some loose ends, the varied experiments discussed above combine to make a persuasive case for the human stretch reflex containing a transcortical component – but they failed to close the issue.

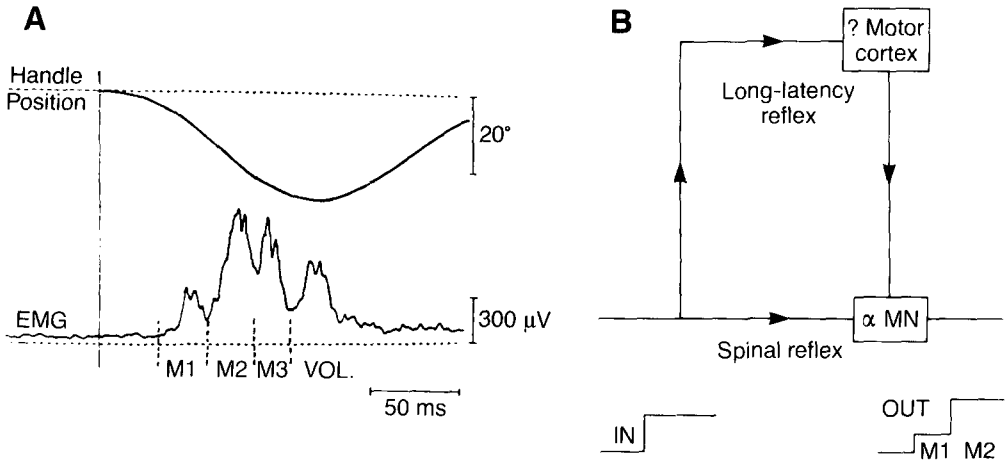


Fig. 2. (A) *The complex segmented response seen on stretching a human muscle during its voluntary contraction. The top trace shows the stimulus applied to the extensor muscles of the wrist (forcible flexion of the wrist on moving a handle that the subject was grasping). The bottom trace shows their electromyographic response with separable M1 and M2 components, preceding any voluntary response*[4]*. M3 is no longer widely recognized. EMG is rectified and averaged, and zero is indicated by a dashed line.* **(B)** *How the delay of the M2 response might arise.* (Taken, with permission, from Ref. 25.)

Counter-attacks

The counter-attacks on the transcortical hypothesis were mounted independently on three separate fronts and they occurred at the same time as the data in its favour were still accumulating.

(1) Lesion experiments on both cats and monkeys showed that responses resembling the M2 responses of intact animals could survive interruption of the putative transcortical pathway[20,21]. This demonstrated that the cortex could not be uniquely responsible for the appearance of later waves in the EMG response – these are also explicable in various other ways.

(2) A specific alternative mechanism was proposed for M2 in the intact state on the basis of recording from single afferent fibres, especially in humans. The recordings emphasized that, as expected, stretch afferents continued to discharge throughout the course of the stimulus. Thus, the later components of a response might in principle be entirely due to continuing short-latency action. More significantly, the afferents tended to fire in synchronized bursts, associated with mechanical oscillations of the muscle, each of which would seem capable of eliciting a short-latency response[22]. Thus, 'segmentation' of the reflex into a series of components cannot be taken as an unequivocal sign of long-latency action, since it might be due simply to a succession of short-latency responses. It seems inevitable that such behaviour must contribute to the complex segmented pattern of EMG activity recorded when a muscle is stretched. However, it cannot be the whole story, and it is incapable of explaining the occurrence of a long-latency response in the absence of a short-latency response; this has now been well documented for various muscles of the hand on both stretch and release (Fig. 4)[5,16,23].

(3) I recently re-examined[24,25] Hammond's original alternative mechanism, which stated that the delayed reflex is mediated by slow afferents. The essential new observation was that vibration fails to elicit an appreciable long-latency response that is comparable to that elicited by stretch of the muscle. Likewise, cessation of vibration fails to give the long-latency reduction of EMG activity produced by releasing the stretch. This is equally so in parkinsonian patients in whom the delayed response to stretch tends to be exaggerated[4,26]. Yet vibration (about 1 mm at 100 Hz to tendon or to digit) is known to excite fast spindle (Ia) afferents at least as powerfully as stretch. It was suggested that the failure of vibration to elicit the reflex arose because of its known failure to excite the slower spindle afferents (group II) effectively, whereas stretch does so. These afferents were thus suggested to be crucially involved in mediating the reflex, with the excess delay arising peripherally through their low conduction velocity (assumed to be about half the Ia value).

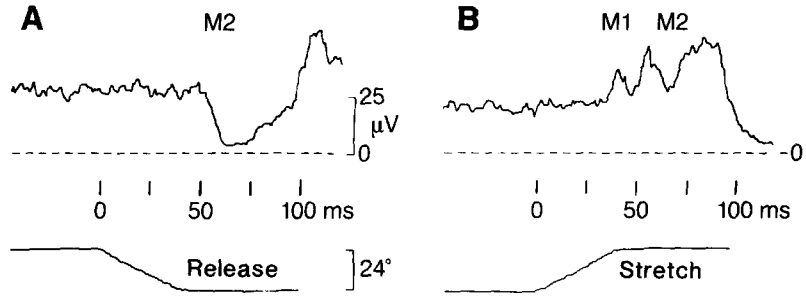


Fig. 3. (A) *The long-latency reduction of EMG activity found on unloading a contracting finger flexor. The appropriate finger was initially flexed against a stop, which was then gradually released (as shown).* **(B)** *The corresponding stretch response elicited by extending the flexor. This muscle does not show an initial short-latency unloading effect, even though stretch elicits an M1 response (i.e. its M1 pathway does not act tonically).* (Taken, with permission, from Ref. 16.)

Recent consolidation

Having revived the slow afferent hypothesis, I proceeded to abandon it on the basis of further experiments with cooling, which were initiated in the expectation of actually consolidating the idea[18,23]! By cooling the human arm, an increase of over 50% can readily be produced in the time taken for an afferent volley to travel from the hand to the spinal cord. The absolute increase, in milliseconds, varies inversely with the size of the fibres concerned and their initial conduction velocity. Work on animals has shown that the percentage change in conduction velocity on cooling is the same for fibres of all sizes. Thus, cooling will have more effect in slowing a reflex with a given 'long' latency if its excess delay (over that of the tendon jerk) arises from slow peripheral transmission by small afferents rather than from long-loop central transmission. The unequivocal experimental finding was that cooling slowed the long-latency reflex by the amount appropriate for fast afferents, showing that the excess delay arises centrally rather than peripherally. Control experiments established that the reflex being studied was mediated by muscle rather than cutaneous afferents, and the cutaneous long-latency reflex elicited by stimulation of digital nerves was also confirmed to depend upon fast afferents. Thus, slow peripheral transmission cannot be responsible for the 'long' latency of either muscular or cutaneous responses, at least for the intrinsic muscles of the hand. However, slow spindle afferents might possibly contribute to the main body of the stretch response. In 1989, with the elimination

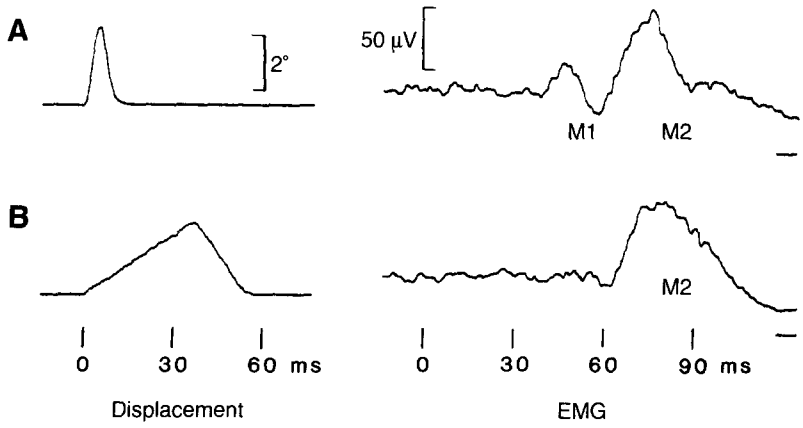


Fig. 4. *Isolation of long-latency excitation by reducing the velocity of stretch.* **(A)** *Segmented M1/M2 EMG response elicited by a tap.* **(B)** *Pure, late (M2) response produced by a slower stretch. (The muscle tested was the first dorsal interosseus of the hand. Zero of EMG is shown by horizontal bars.)* (Taken, with permission, from Ref. 23.)

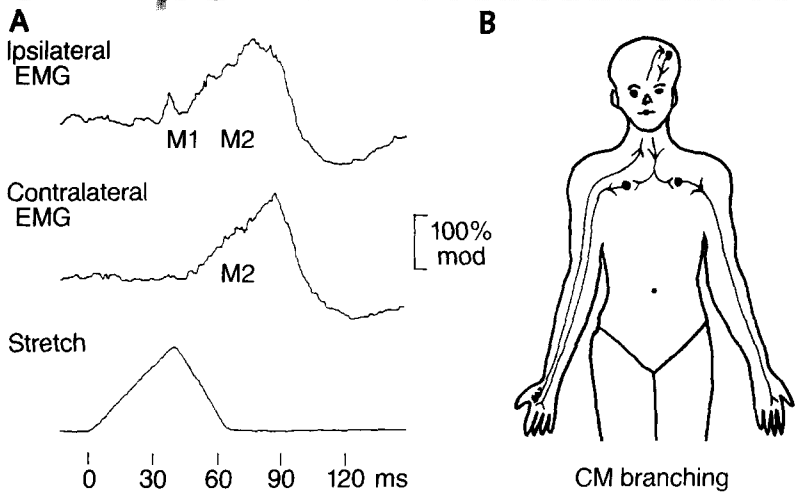


Fig. 5. (A) *The bilateral occurrence of the long-latency EMG response of the first dorsal interosseus in a patient with mirror movements[28]. Note that the stretched muscle also gave a small M1 response.* **(B)** *The suggested underlying neural circuitry. The descending CM fibres of the pyramidal tract branch abnormally to supply corresponding motoneurones on both sides of the spinal cord.* (Taken in part, with permission, from Ref. 28.)

of its principal rival, the transcortical hypothesis was therefore left in command of the field.

Further positive support followed rapidly from studies of patients with the Klippel–Feil syndrome (a congenital dysplasia of the upper cervical vertebrae) who displayed characteristic 'mirror movements'[27,28]. These consist of an automatic involuntary movement of one hand matching an intentional movement of the other hand. In such subjects, typical ipsilateral stretch responses were accompanied by a long-latency response of the equivalent contralateral muscle; the short-latency M1 response, however, was restricted to the stretched muscle (Fig. 5A). The mirror movements probably arise through many corticospinal axons branching abnormally to supply homologous motoneurones bilaterally (unilateral percutaneous cortical stimulation elicits responses bilaterally, and the short-term synchronization of single motor units in equivalent muscles of the two hands is similar to that of units within the same muscle) (Fig. 5B). Accepting this, the long-latency stretch reflex seems almost certainly to involve corticospinal neurones, since the spinal circuitry appears to be normal. Cogent evidence for the transcortical hypothesis has thus been obtained, unexpectedly and rather late in the day, by the classical method of analysing patients with neurological defects.

Nonetheless, the vindication of the transcortical hypothesis does not automatically exclude all other possibilities, particularly for muscles less dominated by the cortex than those controlling the digits; as in all debates, the alternatives too easily appear black and white and mutually exclusive, rather than being taken as potentially complementary mechanisms. The reflex action, if any, of the discharges elicited by stretch from muscle spindle secondary endings remains unknown. Moreover, the enigma persists that Ia discharges appear to have different effects, depending upon whether they are elicited by stretch or by vibration, with vibration failing to activate the transcortical pathway as effectively. Perhaps there is some trivial explanation for this, or perhaps human spindle secondary endings are supplied by much faster afferents than in the cat. Alternatively, we might be failing to read an important clue to the working of the central machinery: conceivably, the cortex is more discriminating than the spinal cord and demands an appropriate pattern of activity in more than one type of afferent (as from both spindles and tendon organs) before it is prepared to respond. With the triumph of the transcortical hypothesis, the detailed mechanism of the stretch reflex is likely to prove more difficult to unravel than if it had been a purely spinal reaction.

Functional 'set'

The involvement of the cortex in such an apparently simple reaction indicates that there must be rather more to the 'stretch reflex' than is commonly supposed; a simple servo control of individual muscles would seem well within the capabilities of the spinal cord. Hammond suggested, at the very outset, that the reflex becomes larger if the subject is instructed to resist the stretch, and smaller if he is told to give way[2]. How far such changes represent a simple voluntary response to the disturbance is still under debate, along with the question of the mechanical effectiveness of the stretch reflex itself[29].

There is widespread support for the view that cortical mediation allows for the better adjustment of the reflex to prevailing conditions. A particular instance is provided by the spread of the response to muscles other than the one being stretched; in the decerebrate cat the reflex is strictly localized, but in humans the response can spread widely with little extra delay. This was first noted in the context of posture where a variety of muscles can be activated without being directly stretched, and contribute to the maintenance of the stability of the body[30]. Such responses depend crucially upon neural 'set', and disappear when a muscle acts in a different context. However, multiple innervation is not simply a means of maintaining posture, but can also occur between muscles that happen to be acting together in movement. When the thumb and index finger are opposed, then interference with movement of the thumb elicits a response from muscles acting on both of the digits[31]. Yet more remarkably, in contrast to the spinal M1 reflex, the long-latency response can be routed to an apparent antagonist if its contraction were mechanically advantageous. Thus, stretch of the biceps muscle when it is acting to supinate the arm elicits long-latency excitation of the triceps, as well as of the biceps itself. The triceps then counteracts the unwanted component of flexion produced by the reflex activation of the biceps[32]. Such initially unexpected synergies also occur during voluntary contraction.

Perhaps part of the rationale for routing the 'stretch reflex' via the cortex is to make use of its machinery to establish complex and shifting patterns of connectivity. Cutaneous inputs to the cortex can also elicit reflex responses that vary with the behavioural context; for example, the 'micro-slip' of an object that is being gripped between finger and

thumb reflexly elicits a tightening of the grip[33]. All in all, the classical neurophysiological analysis of stereotyped reflex action in the cat, decerebrate or otherwise, failed to alert us to the complexities and interesting properties of the human stretch reflex.

Selected references

1 Liddell, E. G. T. and Sherrington, C. S. (1924) *Proc. R. Soc. London Ser. B* 96, 212–242
2 Hammond, P. H. (1960) *Proc. III Int. Conf. Med. Electron.* 190–199
3 Matthews, P. B. C. (1990) *Can. J. Physiol. Pharmacol.* 68, 347–354
4 Lee, R. G. and Tatton, W. G. (1978) *Prog. Clin. Neurophysiol.* 8, 167–177
5 Marsden, C. D., Merton, P. A. and Morton, H. B. (1976) *J. Physiol.* 257, 1–44
6 Phillips, C. G. (1969) *Proc. R. Soc. London Ser. B* 173, 141–174
7 Phillips, C. G., Powell, T. P. S. and Wiesendanger, M. (1971) *J. Physiol.* 217, 419–446
8 Wiesendanger, M. (1986) *Prog. Brain Res.* 64, 67–74
9 Wiesendanger, M. and Miles, T. S. (1982) *Physiol. Rev.* 62, 1234–1270
10 Wolpaw, J. R. (1980) *J. Neurophysiol.* 49, 1122–1138
11 Marsden, C. D., Merton, P. A., Morton, H. B. and Adam, J. (1977) *Brain* 100, 503–526
12 Evarts, E. V. (1973) *Science* 179, 501–503
13 Evarts, E. V. and Tanji, J. (1976) *J. Neurophysiol.* 39, 1069–1080
14 Tatton, W. G., North, A. G. E., Bruce, I. C. and Bedingham, W. (1983) *J. Neurosci.* 3, 1807–1817
15 Cheney, P. D. and Fetz, E. B. (1984) *J. Physiol.* 349, 249–272
16 Matthews, P. B. C. and Miles, T. S. (1988) *J. Physiol.* 404, 515–534
17 Darton, K., Lippold, O. C. J., Shahani, M. and Shahani, U. (1985) *J. Neurophysiol.* 53, 1064–1618
18 Matthews, P. B. C. (1989) *Prog. Brain Res.* 80, 103–112
19 Jenner, J. R. and Stephens, J. A. (1982) *J. Physiol.* 333, 405–419
20 Ghez, C. and Shinoda, Y. (1978) *Exp. Brain Res.* 32, 55–68
21 Miller, A. D. and Brooks, V. B. (1981) *Exp. Brain Res.* 41, 146–158
22 Eklund, G., Hagbarth, K-E., Hagglund, J. V. and Wallin, E. U. (1982) *J. Physiol.* 326, 79–90
23 Matthews, P. B. C. (1989) *J. Physiol.* 419, 519–538
24 Matthews, P. B. C. (1984) *J. Physiol.* 348, 383–415
25 Matthews, P. B. C. (1985) in *Feedback and Motor Control in Invertebrates and Vertebrates* (Barnes, W. J. P. and Gladden, M. H., eds), pp. 431–449, Croom Helm
26 Cody, F. W. J., MacDermott, N., Matthews, P. B. C. and Richardson, H. C. (1986) *Brain* 109, 229–249
27 Capaday, C., Fraser, R., Forget, R. and Lamarre, Y. (1989) *Soc. Neurosci. Abstr.* 15, 74.13
28 Matthews, P. B. C., Farmer, S. F. and Ingram, D. A. (1990) *J. Physiol.* 428, 561–577
29 Marsden, C. D., Merton, P. A., Morton, H. B., Rothwell, J. C. and Traub, M. M. (1981) *J. Physiol.* 316, 47–60
30 Marsden, C. D., Merton, P. A. and Morton, H. B. (1981) *Brain* 104, 513–534
31 Cole, K. J., Gracco, V. L. and Abbs, J. H. (1984) *Exp. Brain Res.* 56, 582–585
32 Gielen, C. C. A. M., Ramaekers, L. and van Zuylen, E. J. (1988) *J. Physiol.* 407, 275–292
33 Johansson, R. S. and Westling, G. (1984) *Exp. Brain Res.* 56, 550–564

perspectives

Neuroscience in Hungary

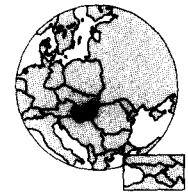
J. Hámori

A strong foundation in neuroscience in Hungary is threatened with erosion if the financial climate of general science funding cannot be improved through international cooperation and collaboration.

If one wished to be successful in Hungary in the 19th century, one became a poet or a lawyer. Although poetry and literature are still highly esteemed in the 20th century (for example, the present president of the republic, A. Göncz, is a writer), science – including biochemistry (with the Nobel prize winner A. Szentgyörgyi its most famous representative) and, particularly, neurobiology – emerged as the second most popular vocation of many young hopefuls. This is not unexpected, since Hungary's contribution to neuroscience has very rich traditional roots. In the last part of the 19th century and at the beginning of the 20th century, the 'founding fathers' of neuroscience in Hungary were neuroanatomists (for example, I. Apathy or his scientific opponent M. von Lenhossek), neuropathologists (for example, K. Schaffer and D. Miskolczi) and neurologists (for example, E. Högyes).

Lenhossék, who contributed to the verification of Ramón y Cajal's neurone concept, was professor of anatomy at the Budapest University (now named after Semmelweis) and was the founder of the Hungarian Neuroanatomy School, which is still active today. The internationally recognized neuroanatomist J. Szentágothai was Lenhossék's pupil, and now works at the Semmelweis University. The pupils of Szentágothai himself include G. Székely, J. Hámori and P. Somogyi (neuroanatomy), and B. Flerko, B. Halász and M. Palkovits (neuroendocrinology), among others. Of those who have not emigrated, most lead their own research teams by now (see Table I).

K. Schaffer, who first described the histopathology of Tay–Sachs disease (and whose name is attributed to the hippocampus 'Schaffer collaterals') was also succeeded by eminent neuropathologists, such as D. Miskolczi, S. Környei and, particularly, K. Sántha. Unfortunately, Sántha was a victim of the communist regime in the early 1950s. E. Högyes, who published his original research on the labyrinthine reflex between 1875 and 1880, is perhaps not as well recognized internationally. However, his observations greatly influenced the work of G. Bekessy, who won the Nobel prize for his results on hearing, and also influenced the work of Szentágothai (1947–1950) on the interactions between the labyrinthine reflex and the optokinetic system. Furthermore, Högyes can be considered as the pioneer of functional studies, which were continued in the 1950s by K. Lissák, E. Grastyán and G. Ádám, who established and initiated modern neurophysiology in Hungary. Single-cell neurophysiology



József Hámori is the chairman of the Neurobiology Committee and a Professor of Anatomy at the Semmelweis University, Budapest, Hungary.

0166 - 2236/91/$02.00