Part I

An overview of functional magnetic resonance imaging

Part IA Introduction to the physiological basis of functional neuroimaging

The subject to be observed lay on a delicately balanced table which could tip downwards either at the head or at the foot if the weight of either end were increased. The moment emotional or intellectual activity began in the subject, down went the balance at the head-end, in consequence of the redistribution of blood in his system …

… We must suppose a very delicate adjustment whereby the circulation follows the needs of the cerebral activity. Blood very likely may rush to each region of the cortex according as it is most active, but of this we know nothing.

William James (1890) The Principles of Psychology

Chapter 1

Neural activity and energy metabolism

Introduction

Metabolic activity accompanies neural activity

The goal of understanding the functional organization of the human brain has motivated neuroscientists for well over 100 years, but the experimental tools to measure and map brain activity have been slow to develop. Neural activity is difficult to localize without placing electrodes directly in the brain. Fluctuating electric and magnetic fields measured at the scalp or near the head provide information on electrical events within the brain, and from these data the location of a few sources of activity can be estimated, but the information is not

sufficient to produce a detailed map of the pattern of activation. However, precise localization of the *metabolic* activity and blood flow changes that follow neural activity are much more feasible and form the basis for most of the functional neuroimaging techniques in use today, including positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). Although comparatively new, fMRI techniques are now a primary tool for basic studies of the organization of the working human brain, and clinical applications are growing.

In 1890, William James published *The Principles of Psychology*, a landmark in the development of psychology as a science grounded in physiology. The possibility of measuring changes in brain blood flow associated with mental activity clearly lay behind the experiment performed by Angelo Mosso and recounted by James in the quotation at the beginning of this section. By current standards of blood flow measurement, this experiment is quaintly crude, but it indicates that the idea of inferring neural activity in the brain from a measurement of changes in local blood flow long preceded the ability to do such measurements (Raichle 1998).

In fact, this experiment is unlikely to have worked reliably for an important reason. The motivation for this experiment may have been an analogy with muscle activity. Vigorous exercise produces substantial muscle swelling through increased blood volume, and thus a redistribution in weight. But the brain is surrounded by fluid and encased in a hard shell, so the overall fluid volume within the cranium must remain nearly constant, a principle often referred to as the Munro–Kellie doctrine. Blood volume changes do occur in the brain, and the brain does move with cardiac pulsations, but these changes most likely involve shifts of cerebrospinal fluid (CSF) as well. As a result, the weight of the head should remain approximately constant.

Furthermore, this experiment depends on a change in blood volume, rather than blood flow, and blood flow and blood volume are distinct quantities. *Blood flow* refers to the volume per minute moving through the vessels, while *blood volume* is the volume occupied by the vessels. In principle, there need be no fixed relation between blood flow and blood volume; flow through a set of pipes can be increased by increasing the driving pressure without changing the volume of the plumbing. Physiologically, however, experiments typically show a strong correlation between cerebral blood flow (CBF) and cerebral blood volume (CBV), and functional neuroimaging techniques are now available for measuring both of these quantities.

The working brain requires a continuous supply of glucose and oxygen (O_2), which must be supplied by CBF. The human brain receives approximately 15% of the total cardiac output of blood (approximately 700 mL/min) and yet accounts for only 2% of the total body weight. Within the brain, the distribution of blood flow is heterogeneous, with gray matter receiving several times more flow per gram of tissue than white matter. Indeed, the flow per gram of tissue to gray matter is comparable to that in heart muscle, the most energetic organ in the body. The activity of the brain generates approximately 11 W/kg of heat, and glucose and O_2 provide the fuel for this energy generation. Yet the brain has virtually no reserve store of O_2, and is therefore dependent on continuous delivery by CBF. If the supply of O_2 to the brain is cut off, loss of consciousness quickly follows.

Table 1.1 lists the primary physiological variables associated with brain energy metabolism and blood flow, along with approximate values for the resting human brain. With brain activation, glucose metabolism, O_2 metabolism, blood flow and blood volume all increase in the active area. Unexpectedly, however, the oxygen extraction fraction (OEF) – the fraction of the delivered O_2 that leaves the blood and is metabolized in the cells – *decreases* with activation, and this phenomenon is exploited in fMRI.

Table 1.1. Typical energy metabolism and blood flow variables for the resting human brain

Physiological variable	Abbreviation	Typical value
Cerebral blood flow	CBF	0.5 mL/(g·min)
Cerebral blood volume	CBV	4%
Cerebral metabolic rate of glucose	CMRGlc	8.5 μmol/(g·min)
Cerebral metabolic rate of oxygen	$CMRO_2$	1.6 μmol/(g·min)
Oxygen extraction fraction	OEF (or E)	40%
Arterial oxygen content	$[O_2]_a$	8 μmol/mL

Functional MRI

Positron emission tomography provided the first technology for mapping patterns of activation in the human brain with high spatial resolution by measuring changes in blood flow and energy metabolism, and these methods are described later in this chapter. More recently, fMRI methods have dominated the field of functional neuroimaging, primarily based on a phenomenon called the blood oxygenation level dependent (BOLD) effect. This effect arises because of two distinct phenomena. The first is that when hemoglobin, the molecule in blood that carries O_2, loses that O_2 to become deoxyhemoglobin, the magnetic properties change in a subtle way. The effect of this is that the MR signal changes slightly, increasing when the blood becomes more oxygenated. This phenomenon alone, while interesting from a biophysical point of view, only becomes useful when combined with a second, physiological phenomenon: when an area of brain is activated, the blood flow increases much more than the O_2 metabolic rate ($CMRO_2$). This leads to a reduction of the OEF, a seemingly paradoxical scenario in which the venous blood is more oxygenated – despite the increase in $CMRO_2$ – because the blood flow has increased more. Taken together, these two phenomena produce the BOLD effect, a local increase of the MR signal owing to a reduction of the OEF during increased neural activity.

Functional MRI based on the BOLD effect is the most widely used method for exploring brain function in human subjects, but MRI offers several additional techniques as well. Although the term fMRI is often taken in a narrow sense to mean BOLD imaging, in this book the term is taken in a broader sense to mean any MRI technique that moves beyond anatomical imaging and provides information on physiological function. Chapters 1 and 2 provide an introduction to the physiological basis of these methods, as well as background on other techniques such as PET. The connection between neural activity, energy metabolism, and blood flow is the foundation of functional neuroimaging, yet this area of physiology is still not well understood. Recent research has emphasized the key role played by astrocytes, cells that have processes projecting to both neurons and blood vessels. This has led to the concept of the *neurovascular unit*, a close interaction between neurons, astrocytes, and blood vessels. The first two chapters describe these current results and ideas.

Neural signaling

Like all organs, energy metabolism in the brain is necessary for the basic processes of cellular work, such as chemical synthesis and chemical transport. But the particular work done by the brain, which requires the high level of energy metabolism, is the generation of electrical

activity required for neuronal signaling. We begin by reviewing the basic processes involved in neural activity from the perspective of thermodynamics, in order to emphasize the essential role of energy metabolism. A more complete description can be found in standard neuroscience texts (Nicholls *et al.* 1992).

Neural activity

Neurons have a complex structure, with an intricate tree of fine *dendrites* extending outward from the cell body, and a single *axon* that carries outgoing signals (Fig. 1.1A). The axon divides into many branches, and a branch makes contact with a dendrite or cell body of another neuron at a *synapse*. In the primate brain, a neuron may have on the order of 10 000 synapses where it can receive signals from other neurons. Most of these connections are with nearby neurons (within a few millimeters), but some connections are much longer.

A key aspect of all cells, but in particular for neurons, is the *membrane potential*. Electrodes placed inside and outside the cell record an electric potential difference across the cell membrane of approximately −70 mV, with the potential more negative inside. One can think of the membrane potential as the medium of neuronal signaling: it is the physiological property altered in one neuron when it receives a signal from another neuron. An *action potential* or *spike* is a transient disturbance of that potential, a rapid depolarization of the membrane near the origin of the axon, initiated by a partial depolarization of the membrane potential. For example, if positive current is injected into the cell, the membrane potential will slowly increase (depolarize, approaching zero) until a threshold is reached that

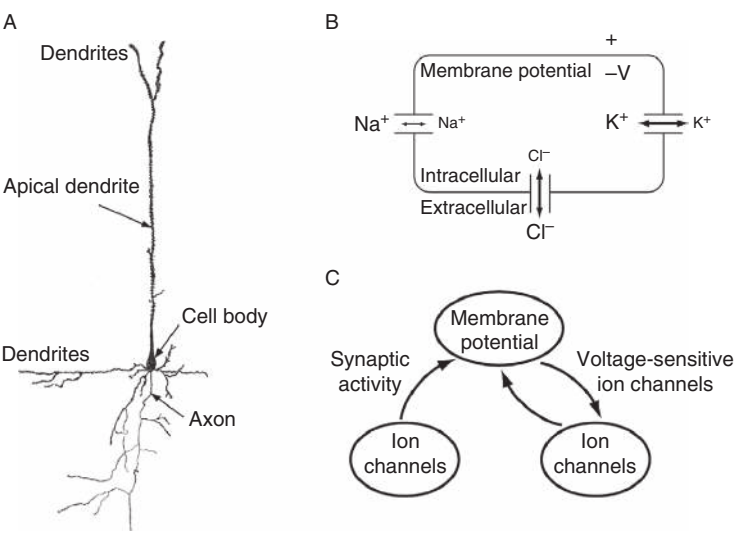

Fig. 1.1. Neuronal signaling. (A) The schematic diagram shows the dendrites, cell body, and axon for a pyramidal cell, the principal neuron of the cortex. Axons from other neurons make contact at synapses on the dendrites and cell body, causing ion channels to open and current to flow into or out of the cell. If a sufficient depolarizing current reaches the cell body, an action potential is generated that travels along the axon to signal other neurons. (B) Ions have different extracellular and intracellular concentrations. The membrane potential depends on these distributions plus the permeability of the cell membrane to each ion. Synaptic activity leads to transient opening or closing of specific ion channels, producing fluctuations of the membrane potential. (C) A source of the non-linear complexity of neuronal signaling, including the generation of an action potential, is that other ion channels are voltage dependent and so are affected by the membrane potential, creating a feedback loop.

triggers the action potential, an abrupt further depolarization that quickly recovers. This rapid depolarization triggers an action potential in a nearby section of the axon, and in this way the action potential propagates down the axon until it reaches a junction with another neuron at a synapse.

The arrival of the action potential then influences the firing of the second neuron by creating a local fluctuation in the membrane potential of the post-synaptic neuron. With an *excitatory post-synaptic potential* (EPSP), the potential inside is raised, creating a slight depolarization, and for an *inhibitory post-synaptic potential* (IPSP) the potential inside is decreased, creating a slight hyperpolarization. Each neuron thus has the capacity to integrate the inputs from many other neurons through their cumulative effect on the post-synaptic potential. A new action potential is generated by the post-synaptic neuron if the membrane potential in the region where the axon originates becomes sufficiently depolarized. From an electrical viewpoint, the working neuron is an intricate pattern of continuously fluctuating membrane potentials caused by synaptic activity punctuated by occasional sharp action potentials.

The membrane potential

The membrane potential depends on two factors: the extracellular/intracellular distribution of ions across the cell membrane, and the permeability of the membrane to each of those ions (Fig. 1.1 B, C). In general, ions cannot freely diffuse across the membrane but instead must pass through *ion channels* created by specialized proteins embedded in the membrane. Ion channels can be very specific and can be modulated by chemical messengers such as neurotransmitters, or by the membrane potential itself. The interesting complexity of neuronal signaling is that permeability to specific ions determines the membrane potential, but the membrane potential, in turn, affects the permeability of voltage-sensitive ion channels. For example, a voltage-sensitive sodium (Na^+) channel will open if the membrane in its vicinity becomes depolarized, and this will lead to an additional Na^+ influx that will further depolarize the membrane. This non-linear cooperative behavior is a key part of the generation of an action potential.

In addition to Na^+, the other principal ions involved in neuronal signaling are potassium (K^+), calcium (Ca^{2+}) and chloride (Cl^-), which are distributed with higher concentrations of Na^+, Ca^{2+} and Cl^- outside the cell and a higher concentration of K^+ inside the cell. Whenever there is a concentration difference across a membrane, there is a tendency for the ions to diffuse from the side with the higher concentration to the side with the lower concentration. However, this tendency is offset by the membrane potential; a negative potential inside the cell will favor a higher intracellular concentration of positive ions. For a given intracellular/extracellular distribution of an ion, the *equilibrium potential* is the membrane potential that would balance that distribution so that there is no tendency for a net movement across the membrane. For example, a typical distribution of Cl^- is an extracellular/intracellular concentration ratio of approximately 12:1, a ratio that is in equilibrium with a membrane potential of approximately −70 mV. Because this is close to what is observed as the existing membrane potential in resting neurons, we conclude that even if the membrane is highly permeable to Cl^- there will nevertheless be no tendency for a net flux of ions through the membrane.

Sodium has a similarly high extracellular/intracellular concentration ratio, approximately 10:1, but because Na^+ is positively charged the equilibrium potential is approximately +60 mV, substantially different from Cl^-. At rest, the permeability of the membrane to

Na^+ is low, so there is little flux of Na^+ down its gradient (although there is a slow leak). If membrane ion channels open to allow passage of Na^+, making the membrane permeable to Na^+, there is a strong inward current because both the concentration gradient and the negative potential drive a Na^+ flux into the cell. Potassium has an even more asymmetric distribution, with an extracellular/intracellular concentration ratio of approximately 1:40, and the corresponding equilibrium potential is approximately −95 mV. Opening a K^+ channel will lead to an outward flux of K^+ down its gradient. In short, relative to the resting membrane potential, the Cl^- distribution is near equilibrium; the K^+ distribution is somewhat out of equilibrium; and the Na^+ distribution is far from equilibrium.

Given this complex non-equilibrium system with multiple ions in different distributions, what actually determines the membrane potential? Ultimately, the membrane potential depends on a very slight imbalance of charge inside and outside the cell, but the amount of charge involved is much smaller than the fluxes of charge across the membrane through ion channels. One can think of the neuron as being in a *steady state* in which the net flux of charge across the membrane is zero. That is, positive and negative charges are moving back and forth through membrane channels, and for any particular ion there may be a steady flux in one direction, such as a slow but steady leak of Na^+ into the cell. Overall, however, there is no net charge transfer across the membrane. Because the membrane potential is highly sensitive to any imbalance of charge across the membrane, any departure from this steady state, leading to a net flux of charge, will quickly alter the membrane potential. Then, for any combination of ion distributions and membrane permeabilities, the membrane potential takes on the value that will create this steady state with no net flux of charge.

For example, if the membrane is highly permeable to one ion, but only weakly permeable to the others, the membrane potential will approach the equilibrium potential of that ion, because that ion alone determines the net flux of charge. However, if the membrane also is permeable to another ion with a different equilibrium potential, the resulting membrane potential will be intermediate between the two equilibrium potentials, weighted by the relative permeability to each of the different ions. That is, as the permeability to the second ion increases, the membrane potential will shift toward the equilibrium potential of that ion. In this case, because the membrane is permeable to both ions but the membrane potential does not match either equilibrium potential, there will be a steady leak of ions that tends to degrade the ionic distributions across the membrane, even though there is no net flux of charge. The stability of the cell depends on maintaining these ionic distributions, so homeostasis requires that the ions be pumped back against their gradient, requiring energy metabolism (see below). In short, the membrane potential depends on the ion distributions across the membrane and the permeability of the membrane to each ion, and these ion permeabilities are altered in neuronal signaling.

Synaptic activity

An action potential arriving at a synapse with another neuron initiates a process that causes ion channels on the post-synaptic neuron to open or close (Fig. 1.2). This process begins on the pre-synaptic side when the incoming action potential triggers an increase of the membrane permeability to Ca^{2+}, allowing Ca^{2+} entry into the pre-synaptic terminal. Within the pre-synaptic terminal, *neurotransmitter* molecules are concentrated in small packages called *vesicles*, and the influx of Ca^{2+} triggers these vesicles to merge with the cell membrane and spill their contents into the synaptic cleft separating the pre- and post-synaptic membranes. The neurotransmitter molecules diffuse across this thin (20–40 nm) gap and bind to receptor

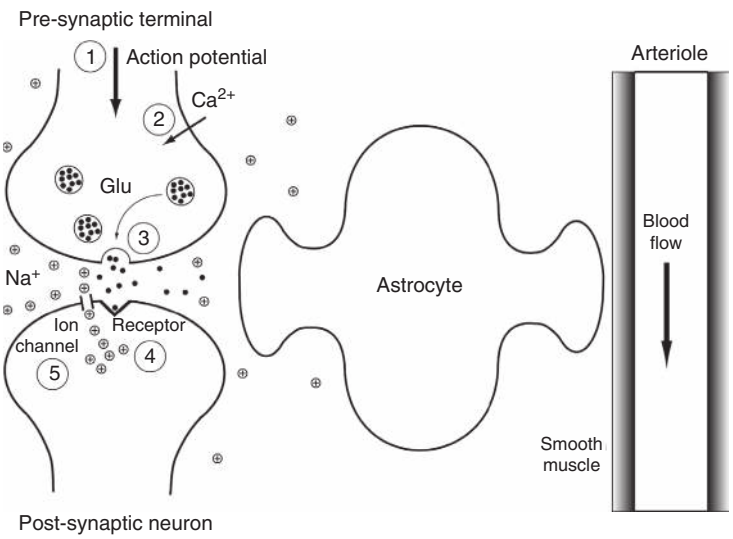

Fig. 1.2. Synaptic signaling. An action potential arriving at the pre-synaptic terminal (1) initiates Ca^{2+} influx (2), triggering neurotransmitter (glutamate [Glu]) release into the synaptic cleft (3), binding of neurotransmitter to receptors on the post-synaptic neuron (4), opening of Na^+ channels (5), and creation of a strong influx of Na^+ into the post-synaptic neuron. Astrocytes have processes that project to neuronal synapses and also to blood vessels, and are involved in clearing neurotransmitter and signaling blood flow changes (illustrated in Figs. 1.4 and 2.4).

sites on the post-synaptic membrane. At each synapse, the neurotransmitter released is characteristic of the pre-synaptic neuron. Glutamate is the most common excitatory neurotransmitter in mammalian cortex, and gamma-aminobutyric acid (GABA) is a common inhibitory neurotransmitter (Erecinska and Silver 1990). Glutamatergic neurons include the *pyramidal cells*, the principal neurons of the cortex. GABAergic neurons include a diverse class of *interneurons* that are important for controlling the net activity of ensembles of neurons.

There are two general classes of receptor. *Ionotropic* receptors are proteins that are themselves ion channels, so binding of neurotransmitter leads to a conformational change of the receptor that opens the ion channel, and the channel remains open while the neurotransmitter is bound. These receptors are very fast, operating on a time scale of milliseconds, and many glutamate and GABA receptors operate in this way. In contrast, at *metabotropic* receptors, binding of the neurotransmitter initiates a chemical cascade that changes the concentration of intracellular *second messengers*, such as cyclic adenosine monophosphate (cAMP), cyclic guanosine monophosphate (cGMP), and Ca^{2+}. These receptors are also called *G-protein-coupled receptors*, because the initiating step involves guanosine triphosphate (GTP), an energy-rich molecule that is a close relative of ATP, which is discussed later in this chapter. These signaling molecules then gate ion channels or exert other modulatory effects on the post-synaptic neuron, and the time scale for these effects can be much slower and longer lasting (seconds to minutes or more) compared with ionotropic receptors. For this reason, activation of these receptors is often described as having a *neuromodulatory* role, affecting different aspects of neuronal signaling from neurotransmitter release to post-synaptic effects, and these synapses are thought to play an important role in learning and

memory. Examples include different types of glutamate and GABA receptor as well as a wide range of other receptors including those for serotonin (5-hydroxytryptamine [5-HT]), histamine, and dopamine.

In summary, binding of neurotransmitter to a receptor on the post-synaptic membrane initiates a process that opens or modulates particular ion channels either directly or through the action of second messengers. Opening Na^+ channels creates a strong inward positive current, which will tend to depolarize the membrane potential in the vicinity of the synapse, creating an EPSP. If, however, the action at the synapse is to open K^+ channels, the effect is to make the potential more negative – hyperpolarizing the membrane – because the equilibrium potential for K^+ is more negative, creating an IPSP. If Cl^- channels are opened, there may be no change in the membrane potential itself, because the resting membrane potential is already close to the Cl^- equilibrium potential. However, this also has an inhibitory effect on the post-synaptic neuron, because now opening the same Na^+ channels will have less of an effect on the membrane potential. Opening Cl^- channels pulls the potential more strongly toward the Cl^- equilibrium potential, which happens to be approximately the resting potential. For this reason, more Na^+ channels would need to open to achieve the same depolarization as when the Cl^- channels are not open. In short, opening many Na^+ channels will depolarize the membrane; opening many K^+ channels will hyperpolarize the membrane; and opening many Cl^- channels will tend to peg the membrane potential near the resting value.

At any moment, the neuron may be receiving signals from many other neurons, creating a complex pattern of flickering fluctuations of the membrane potential on the dendritic tree. Along a dendrite, the currents and associated membrane potentials combine; and if there is a sufficient net current that reaches the cell body, the membrane potential there will be depolarized, generating an action potential and sending a new signal to other neurons. Note, though, that there is likely to be a great deal of sub-threshold synaptic activity that may not lead to spiking. Excitatory synapses generally are located on the dendrites, while inhibitory synapses are generally closer to the cell body. Because of the complex geometry of a neuron, incoming excitatory currents down a dendrite toward the cell body can be dissipated by IPSPs nearer the cell body. For this reason, the output of the post-synaptic cell is a complex combination of the inputs, and it is important to keep in mind the distinction between synaptic activity and spiking activity. As discussed below, much of the energy cost of neural signaling is thought to lie in the recovery from synaptic activity.

Electrophysiology measurements

The opening and closing of ion channels creates currents across the cell membrane and also within the extracellular space. Because of Ohm's law, these fluctuating currents are associated with fluctuating electric potentials (Fig. 1.3). These *extracellular potentials* can be measured invasively with an electrode with high temporal resolution. For example, opening a Na^+ channel at a synapse near the electrode, creating a positive current into the cell, creates a negative deflection of the potential at the location of the electrode. Extracellular potentials are often analyzed by dividing the signal into low- and high-frequency components. The low-frequency components, called *local field potentials* (LFPs), primarily reflect synaptic activity, while the high-frequency activity, called *multi-unit activity* (MUA), primarily reflects spiking activity (Logothetis 2002). While electrode studies are primarily carried out in animal models, implanted electrodes in patients with epilepsy have made possible human recordings as well.

In addition, some components of extracellular potentials are detectable outside the head with non-invasive techniques. The *electroencephalogram* (EEG) is produced from