Chapter 5

Synaptic Processes

§ 1. The Signaling Process in the Chemical Synapse

As the principal signaling junction in neural networks, the chemical synapse always involves two neuron cells. The ***presynaptic neuron*** is the cell regarded as the signal source. The ***post-synaptic neuron*** is regarded as the signal destination or "sink." A synapse may not properly be said to "belong" to either neuron by itself. Rather, it is a biological structure one must view as belonging either to both or to neither. Most neuroscientists prefer the former to the latter in thinking about the synapse, and this is the convention used in this textbook.

The structure of the chemical synapse can be regarded as consisting of three distinct parts: (1) ***the presynaptic terminal***, which is regarded as part of the presynaptic neuron; (2) ***the post-synaptic compartment***, which is regarded as part of the postsynaptic neuron: and (3) ***the synaptic cleft***, which is a small extracellular region, typically on the order of 20 to 30 nm in width. Under electron microscope viewing, the cleft is seen to contain filaments of material joining the pre- and post-synaptic cells. These filaments are thought to provide a firm attachment keeping the pre- and post-synaptic regions of the synapse carefully aligned with each other. However, these filaments are absent or largely absent during early synapse formation, during which time a nascent synapse might or might not become established. The formation of a filament "scaffolding" is usually regarded as the sign that a synapse has become firmly established.

Figure 5.1 illustrates the principal signal processing details of the chemical synapse. By far the most interesting dynamic, the ***vesicle cycle***, takes place within the presynaptic terminal. Vesicles are small, spheroidal bodies that serve as containers for neurotransmitter chemicals (NTX). A typical presynaptic terminal contains from 200 to 500 vesicles. The proteins from which vesicles are constructed are originally manufactured in the cell body and transported to the synaptic terminal via a system of microtubules. Once there they are stored in the plasma membrane until needed to manufacture new vesicles or repair old vesicles being recycled after NTX release. Fresh vesicles are manufactured by organelles in the synaptic terminal. This constitutes step 1 in the vesicle cycle.

Fresh vesicles are next filled with their neurotransmitter substance or substances. This process is called neurotransmitter uptake (NTX uptake). It constitutes step 2 in the vesicle cycle. Synaptic vesicles filled with NTX move to the ***active zone*** of the presynaptic terminal during step 3 of the vesicle cycle. They are anchored by actin filaments and form the ***ready pool*** of vesicles.

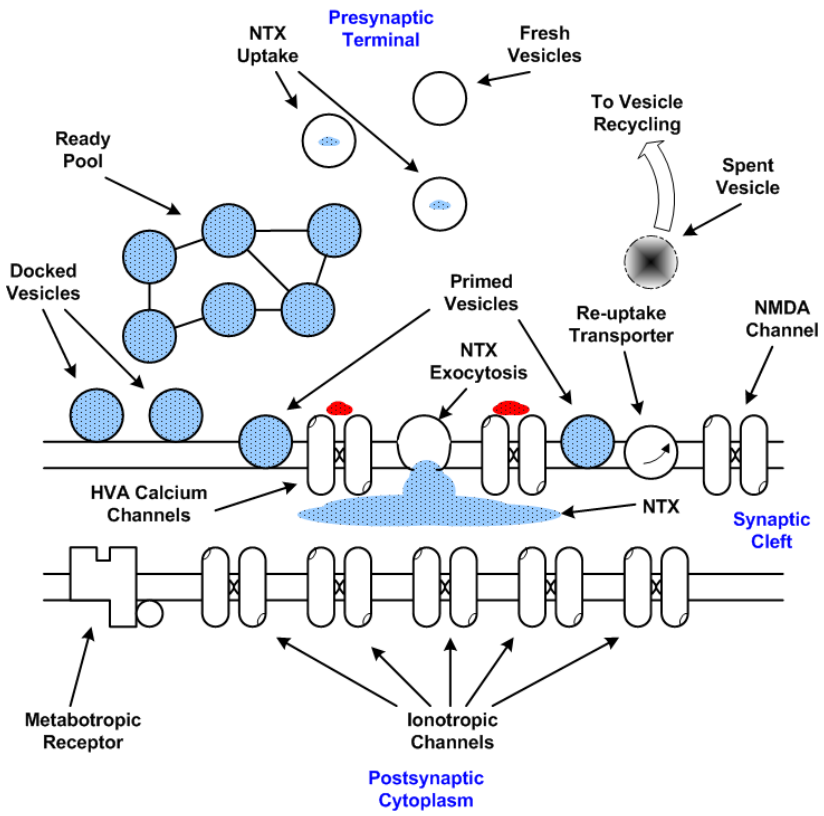

Figure 5.1: Schematic illustration of the principal signaling organization of the chemical synapse. Neurotransmitter molecules (NTX) are represented in blue. Calcium ions are represented in red. The other structures are as labeled in the diagram.

During step 4 some vesicles are taken from the ready pool and moved to the plasma membrane. This step is called ***docking*** the vesicles. Docking occurs only in the active zone of the terminal and involves attachment of the vesicle to the membrane wall. Docking occurs only at neurotransmitter release sites, and these are located in the presynaptic terminal in the region opposite postsynaptic receptors. However, docked vesicles are not yet ready to participate in neurotransmitter exocytosis. They must first undergo a chemical ***prefusion reaction*** that prepares them for calcium-triggered NTX release. This step, step 5 of the vesicle cycle, is called ***priming*** the vesicle and depends on a chemical reaction involving ATP (adenosine tri-phosphate).

The active zone occupies an area of 5-20 μm^2. The number of vesicles attached to the active zone (i.e. either docked or docked-and-primed vesicles) at any one time is variable and typically is on the order of 5 to 10 vesicles. The primed vesicles constitute the ***releasable pool***. A primed vesicles holds its NTX until release is stimulated by reaction with free cytoplasmic Ca^{2+} (step 6 of the vesicle cycle). Calcium enters the presynaptic terminal primarily through HVA calcium channels. As these channels are closed at resting potential levels of the cell membrane, an action potential arriving at the presynaptic terminal is required to open their gates and allow calcium

influx. It is thought that HVA Ca^{2+} channels are located primarily in the active zone. The calcium influx is sometimes colorfully described as a "calcium volcano" erupting amidst the docked and primed vesicles.

Experimental evidence suggests that at least three or four Ca^{2+} ions must act simultaneously (or nearly simultaneously) at the same site in order to trigger NTX exocytosis by the vesicle [SÜDH]. Very high local Ca^{2+} concentrations, greater than 100 mM, are required for triggering the exocytosis. This fact is one reason for making the hypothesis that the HVA channels are located amidst the vesicles in the releasable pool. NTX exocytosis requires only about 0.1 ms, which is far too fast to be explained by enzyme reactions. In addition to directly stimulating NTX exocytosis, the elevated levels of Ca^{2+} are also thought to cause dissolving of actin filaments holding vesicles in the ready pool, thus freeing a new vesicle to take the place of the spent one in the releasable pool [KAND2].

Following exocytosis, the spent vesicle is coated by clathrin and associated proteins and de-attached from the membrane wall. It is transported back into the terminal and undergoes a process of endocytosis in which it is "repaired" and eventually put back into the vesicle cycle at step 1. This process constitutes step 7 of the vesicle cycle. Synaptic vesicles can go through the entire cycle in approximately 60 seconds. Of this time, only 10-20 ms are required for docking and priming, exocytosis requires only 0.1 ms, and removal of the spent vesicle from the membrane wall requires only a few seconds. Thus, a vesicle spends most of its time in the other steps of the vesicle cycle [SÜDH].

NTX is released by the vesicle into the synaptic cleft, where it moves by diffusion to reach the postsynaptic cell wall. Here some of it binds with receptors on ionotropic and/or metabotropic receptor proteins. In the case of ionotropic channel proteins, this binding is followed by the opening of the channel (or, in the case of NMDA receptors, the enabling of the channel). What happens next has been discussed in the previous chapters. In the case of the metabotropic receptor, the immediate result is not the opening of an ionotropic channel, although such a channel opening (or closing) might be a secondary result. The direct result of metabotropic receptor binding is the stimulation of a biochemical cascade reaction inside the postsynaptic cell. The generic name given to this effect is called the ***second messenger signaling*** process. Second messenger processes are very complicated and will be discussed in their own right later in this chapter.

Synaptic vesicles are typically about 35 nm in diameter and are estimated to contain on the order of about 200 protein molecules [SÜDH]. It is thought that one vesicle contains on the order of about 5000 NTX molecules [SCHW1]. Using typical ranges for the active zone area and width

of the synaptic cleft, this puts the concentration of neurotransmitter in the cleft due to exocytosis of one vesicle in the range from about 9 to 20 μM. As the number of NTX molecules in such a concentration is likely to be many times greater than the number of available postsynaptic binding sites, NTX exocytosis results in an excess of NTX molecules in the cleft. The presynaptic neuron helps to clear out this excess by means of ***re-uptake transporters*** located in the terminal. Glial cells also take part in cleansing the extracellular region of excess neurotransmitter. It is known that excessive concentrations of some neurotransmitters, dopamine for instance, are toxic.[1] Thus, this housekeeping chore is essential to the maintenance of the health of the neurons.

As mentioned in chapter 4, it is now known that at least some neurons contain NMDA channels in their presynaptic terminals. This is one putative mechanism by which glial cells might play a role in modulating NTX release activity. At present the location, distribution, and density of presynaptic NMDA channels is not well understood. Similarly, many of the biochemical details governing the vesicle cycle are not fully understood.

§ 2. The Statistics of Neurotransmitter Release

Physiologists first began to suspect neural chemical signaling was quantized as a consequence of the work of Katz and others in the 1950s. Katz was studying the neuromuscular junction, where the connection analogous to the synapse in central systems is called the ***endplate***. Katz noticed the rise in membrane potential in muscle tissue, called the ***endplate potential*** (EPP), appeared to occur in integer multiples of the smallest non-zero response. Close examination of the distribution of "miniature" endplate potentials showed that the "units" of response, n, followed a binomial probability distribution,

$$\Pr[n] = \binom{N}{n} \cdot p^n \cdot (1-p)^{N-n}, \tag{5.1}$$

where p is the probability of an outcome (a "unit endplate potential" in this case), N is the number of trials made, and $\Pr[n]$ is the probability of the outcome occurring n times in N trials. The binomial coefficient is defined as

$$\binom{N}{n} \triangleq \frac{N!}{n!(N-n)!} . \tag{5.2}$$

Historically, this formulation came a bit later, after the ***vesicle hypothesis*** was formulated. In the original studies, Katz and other researchers fit their data to a Poisson process since their

[1] Cocaine has for one of its effects the shutting-down of the dopamine re-uptake process, and thus leads to excessive buildup of extracellular DA. The psychotic effects of cocaine are one consequence of this.

observations were based on the observed amplitudes of the EPSPs and the number of times each amplitude occurred in the course of repeated experimental trials. As it turns out, the Poisson process is the limiting case approached by a binomial process in the limit of large numbers of trials. Outcomes from a binomial process are inherently quantized (either the event does or does not happen), and the pioneering researchers did not know at the start that their experiments would reveal the existence of a quantized process.

After the formulation of the vesicle hypothesis, it became possible to assign physical significance to the variables in (5.1). N is interpreted as the number of primed vesicles in the available pool at the synapse, and n is the number of vesicles undergoing NTX exocytosis in response to an action potential stimulation. The variable p is called the ***release probability***, and it is found experimentally to vary within the range of about $p = 0.1$ to $p = 0.2$ (10% to 20% release probability). The ***expected value*** of n, denoted $\overline{n}$, is the arithmetic mean number of vesicles releasing their NTX per AP stimulation. For a binomial process, $\overline{n} = Np$. The ***variance*** of a binomial process is $\sigma^2 = Np \cdot (1 - p)$.

Figure 5.2 illustrates the binomial distribution for $N = 10$ and $p = 0.1$ and 0.2. One thing we can notice at once from these graphs is the non-zero probability that n will equal zero. This means there is a probability that an action potential will evoke no neurotransmitter release at all. Such an event is called a ***failure***. In general, if p is held constant as N is decreased, the failure probability will go up. For example, $\Pr[n = 0]$ is about 0.35 for $p = 0.1$ and $N = 10$. If N is reduced to $N = 5$, the failure probability goes up to $\Pr[n = 0] = 0.59$ and $\Pr[n = 1]$ drops to 0.328.

The statistical nature of synaptic transmission is often regarded as a sort of "noise" interfering with the otherwise simple (signal-processing-wise) "communication" process between neurons. Some researchers characterize synaptic transmission as "unreliable" and suggest this may be one reason why effective large-scale neural processing seems to be associated with parallel and synchronous AP signaling (spatial summation) or with "burst-like" firing (temporal integration).

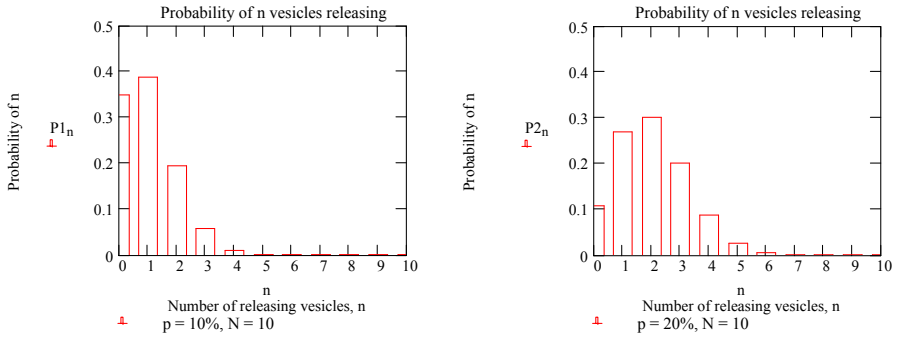

Figure 5.2: Binomial distributions for $p = 0.1$ and $p = 0.2$ with $N = 10$.

Other factors add considerable complications to this so far simple picture of synaptic transmission. First, we may ask: Why is there a stochastic characteristic to synaptic transmission in the first place? It is widely supposed this has to do with the inherently "random walk" nature of calcium diffusion in the presynaptic terminal. Assuming it is true that three or four Ca^{2+} atoms must combine at one release site to produce NTX exocytosis, and further assuming the diffusion process is or is similar to Brownian motion, an element of chance is introduced into the process simply by this property of ion motion in the presynaptic cytoplasm. If all this is so, the amount of Ca^{2+} concentration from the "volcanic plume" of calcium influx through the HVA channels and the distance between these "plumes" and the binding sites for the primed vesicles will be factors in how likely it will be that sufficient calcium binding takes place at any given vesicle. Currently we possess no verified and generally acknowledged *quantitative* model of this qualitative picture, and therefore we are hampered in coming up with definitive experimental tests of the "random walk" hypothesis. Consequently, while few neuroscientists seriously doubt the qualitative model, it remains somewhat speculative at this time.

A second factor coming into play is the fact that cytoplasmic $[Ca^{2+}]$ takes time to be cleared out by the presynaptic terminal's calcium buffering process. Consequently, when AP stimulation takes place through "burst-like" firing, the $[Ca^{2+}]$ factor at work in stimulating NTX release becomes a more complicated function of time and the firing history of the presynaptic neuron. Experiments using "paired pulses" of action potentials separated in time by an interval Δt have shown that the resulting amplitude of excitatory postsynaptic current (EPSC) observed in the postsynaptic cell increases by up to a factor of three when Δt is in the range from 10 to 20 ms. This phenomenon is called ***paired pulse facilitation*** and is one mechanism for what we shall call ***elastic modulation*** of synaptic strength (synaptic weight). Paired pulse facilitation decays more or less exponentially with Δt. For Δt larger than about 20 ms, the ratio of the difference between successive EPSC amplitudes to the amplitude of the first EPSC drops below unity, indicating less than 2:1 facilitation. The time constant for the exponential decay of paired pulse facilitation is on the order of about 200 ms. Results from paired pulse experiments appear to be tracked by estimates of intercellular $[Ca^{2+}]$ vs. time. This is evidence supporting the qualitative model described above [REGE].

Another complicating factor is the decrease in N, the number of primed vesicles in the available pool, following a successful NTX exocytosis event. It takes time for replacement vesicles to be summoned from the ready pool, docked, and primed. While increased $[Ca^{2+}]$ levels favor the likelihood of NTX release (increase of p), depletion of the available pool population, N, works to decrease this likelihood. Long, sustained bursts of action potential stimulation at AP

rates at and above about 50 pulses per second (called a firing ***tetanus***) follow a characteristic course. There is initially some facilitation evident in the amplitudes of the EPSCs seen in the postsynaptic cell, followed by a short term (elastic) ***depression*** of the postsynaptic response. Depletion of the available pool is one hypothetical model that has been proposed as an explanation for this effect, although this hypothesis is not without controversy and at least some experimental studies seem to contradict it. The biophysical and biochemical mechanisms of NTX release are not completely understood at this time, and so the full implications of these experiments are not definitively interpreted yet. This is one area of neural modeling where more and better quantitative models are needed. The Linvill modeling schema introduced in chapter 4 might be able to play an illuminating role in addressing this research question.

Similarly, the precise biophysics and biochemistry involved in the entirety of the vesicle cycle are not well understood as of yet. It seems reasonable to expect that the rate of replenishment of the available pool depends at least in part on the population of vesicles in the ready pool, and that this population in turn depends on the kinetics of the rest of the vesicle cycle. Here is another place where new and better quantitative models of the synaptic processes would be of great use.

To sum this up: There is reason to think both p and N are time-varying functions of the past activity of action potential stimulation of the presynaptic cell. Quantitative mathematical models for these quantities are not yet very far along in their development, and more research work by computational neuroscientists, working with the physiologists, certainly seems called for.

§ 3. Synaptic Arrangements

Stepping back away from the details of the chemical synapse, we next consider the variety of arrangements in which synapses may form. Figure 5.3 illustrates six of the major "themes" by which synapses organize in central systems.

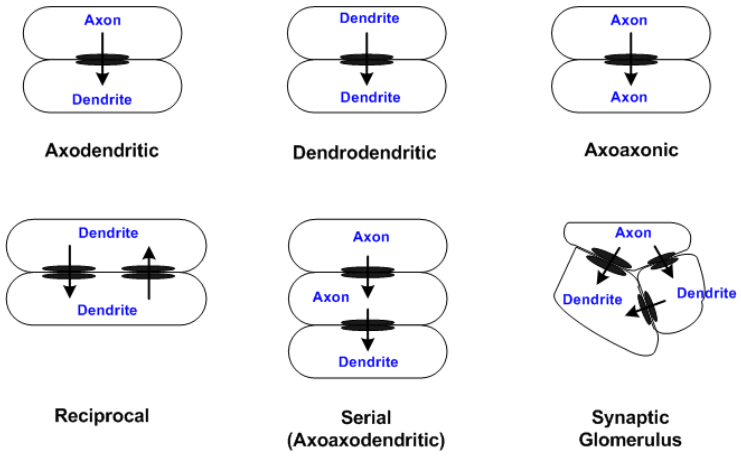

Figure 5.3: Various synaptic arrangements. The six arrangements shown here are not exhaustive.

Two of the most common synapse types are the axodendritic and axosomatic synapses. The axodendritic synapse (shown in Figure 5.3) is a connection from an axon terminal to a dendrite. This connection might or might not be made at a dendritic spine. The axosomatic connection is similar except that in this case the connection is from an axon terminal to the cell body.

Many neurons do not have axons. In their case, the dendrodendritic synapse is a common form of connection. Here the presynaptic structure is found in one of the dendrites and the post-synaptic density (the structure of membrane-spanning receptor proteins) is found in the other. Communication between the neurons is one-way. However, in some cases each dendrite contains both a presynaptic structure and a postsynaptic density, and communication between the neurons is reciprocal (lower left of Figure 5.3). One arrangement of this type frequently occurs where one of the pair of reciprocal synapses is excitatory while the other is inhibitory. An excitatory output by the first neuron also propagates through the dendrite to reach the reciprocal synapse. There it excites a response in the second (inhibitory) neuron. This response, in turn, excites the reciprocal inhibitory synapse, thus inhibiting further output by the first neuron. This type of action is called ***feedback inhibition***.

Axoaxonic connections are common and usually inhibitory. Two important cases are worth considering here. In the first case, the synaptic connection occurs at or near where the axon of the postsynaptic neuron leaves the cell body. Inhibition at this point blocks the generation and spread of the postsynaptic neuron's action potential, and is thus a global inhibition. In the second case, the synapse occurs at or near an axon terminal of the postsynaptic cell. In this case, the inhibition is specific, preventing the signaling by the postsynaptic cell to a third neuron without directly inhibiting that third neuron. This type of connection is often called ***presynaptic feedforward inhibition***.

An illustration of this second (specific) structure is exhibited by the case of serial axo-axo-dendritic (lower center case in Figure 5.3) or axo-axo-somatic connection. Referring to Figure 5.3, the axon shown in the center of this "synaptic sandwich" is excitatory for the dendrite to which it connects. The axoaxonal connection to it, on the other hand, is inhibitory.

Yet another very interesting case, from the signal processing point of view, is exhibited by the synaptic glomerulus (lower right in Figure 5.3). In this case, the signal from the axon is excitatory for both dendrites, but the synaptic connection from the right-most dendrite to the other dendrite is inhibitory. An excitatory signal from the axon produces excitatory responses in both dendrites. However, the excitation of the right-most dendrite produces a follow-on inhibition of the left-most dendrite, thus producing an ***excitatory-inhibitory sequence***. Such a sequence is thought to mediate a kind of ***temporal differentiation*** in the signal processing by this neural "netlet."

Arrangement of this type are common in the thalamus.

Complex synaptic arrangements, like the lower three types shown in Figure 5.3 as well as others, appear to be more the rule than the exception in biological networks. Because the simple types of arrangements illustrated by the three top-most cases in Figure 5.3 were discovered first, they became known as "conventional" synapses. Accordingly, the others became known as "unconventional" synapses. Today, however, it would seem that "unconventional" synapse types are far more numerous than "conventional" ones. The old labels have become misleading.

Because the "unconventional" synapse arrangements are so common throughout the central nervous system, some computational neuroscientists propose that these synaptic clusters be regarded as the basic "computing" units of neural organization. There is much merit in this, because the close interplay among complex synaptic arrangements allows for a greater wealth of signal processing operations to be carried out than can be realized by simple spatial and/or temporal summation of signals from the "conventional" arrangements. However, judging by the current literature on mathematical neural network models, this idea has been slow to take hold among the majority of neural network theorists. The classical neural network models (Adaline, perceptron, Hopfield networks, the "connectionist" models, and so on) were set up based on the "conventional" synapse model, and incorporating some of the more complex dynamics possible with "unconventional" synapses is not easily achieved by these models without making some significant alterations to them. One can, of course, "fake" these dynamics by using a sufficiently more complicated "conventional network" and incorporating feedback into it. This is conceptually no more difficult than constructing computer logic circuits using a sufficiently large number of von Neumann's McCulloch-Pitts "organs" (logic gates). But doing so obviously takes the network model one or more steps away from an easily self-evident biological interpretation.

§ 4. Neuropeptides

In addition to the small vesicles containing small-molecule neurotransmitters discussed earlier, many neurons also contain a smaller number of large vesicles in their presynaptic terminals. These vesicles contain compounds of two or more amino acid residues called ***neuropeptides***. The neuropeptides come in a great variety and produce an even greater variety of effects. All the well-documented effects of neuropeptide transmission are very slow in onset and very long-lasting in duration. So far as is currently known, neuropeptide signaling is exclusively metabotropic. The principal difference between neuropeptide signaling and small-molecule metabotropic signaling lies in the time scales involved, neuropeptides seeming to take significantly longer in onset and having significantly longer-lasting effects, including some that are effectively permanent. (A

number of the neuropeptides are known as ***growth factors***. Obviously, growing new neural structures would be regarded as a permanent or semi-permanent effect).

Unlike the small-molecule neurotransmitters and neuromodulators, production and packaging of neuropeptides in their vesicles does not take place locally at the synaptic terminal. Instead, they are manufactured exclusively (or so is presently believed) in the cell body and transported to the presynaptic terminal using the neuron's internal "transportation network." Far less modeling detail is known about neuropeptide exocytosis than about small vesicle neurotransmitter exocytosis. In at least some cases, neuropeptides are thought to be co-released with the cell's small molecule neurotransmitters. It may be the case that the neuropeptide release probability, p, is smaller than that discussed previously, but this is conjecture. Depending on the species of metabotropic receptors present in the postsynaptic cell, neuropeptide action can be inhibitory or it can be excitatory. The action of the neuropeptide can even be the opposite of the presynaptic cell's small molecule neurotransmitter. In such a case, the difference in time scales required for the onset and cessation of the effects of the two types of substances is significant. So far as is presently known, neuropeptides do not participate in "fast" synaptic communication.

Neuropeptide action has received very little attention by neural network theorists. Primarily this is because so little detail about their signal processing mechanics is known and because their role is primarily regulatory. But another reason may be because the time scales of neurotransmitters and neuropeptide neuromodulators are so different. Many theorists regard the phenomenon of neuropeptide transmission as more properly belonging to the "bias setup" or parametric-determination part of neuronal modeling rather than to the direct information-processing operations of neurons and neural networks. There is much justice in this point of view, at least so far as relatively short-term signal processing operations are concerned. But there is also a certain amount of naivety in this view as well when one turns to the consideration of large-scale integration of neuron system-level activity and considerations dealing with long-term potentiation and long-term depression in the large-scale connectivity of large neural assemblies and systems. Nonetheless, the fact remains that these topics are difficult to model and address on a biologically-sound basis so long as so much about neuropeptide signaling remains undiscovered.

§ 5. Metabotropic Signaling

Metabotropic signals do not directly produce ionotropic currents, but they do indirectly influence such currents. For that reason metabotropic signals are said to be ***modulation*** signals. A ***metabotropic channel*** is determined by both the transmitter substance and the receptor protein to which it binds on the postsynaptic membrane surface. Metabotropic channels are characterized by

TABLE I

RELATIVE TIME SCALES OF VARIOUS SIGNALING AND MODULATION MECHANISMS IN NEURONS. TIMES ARE NORMALIZED WITH ONE UNIT CORRESPONDING TO FAST GLUTAMATE ACTIVATION TIME (APPROX. 1 msec)

Primary First Transmitters	Activation Time Scale	Inactivation Time Scale
Nicotinic acetylcholine, amino acids	1 to 2 or 3 units	10 to 50 units
Muscarinic acetylcholine, catecholamines	10 units or more	as much as 50 kilo-units
Many facilitating or depressing small molecule neurotransmitters	about 100 units	as much as 100 Mega-units
Peptides and hormones	about 10 kilo-units	greater than 1 Giga-unit
Growth factors	about 1 Mega-unit	never; changes are permanent

time scales of onset and inactivation that are slow compared to ionotropic signaling. Table I illustrates the time scales characteristic of different classes of metabotropic channels. The time scale is expressed in "units" of ionotropic action, arbitrarily chosen as 1 ms per "unit."

Metabotropic receptor proteins exist that bind glutamate (MGluRs) and that bind GABA (the $GABA_B$ receptor). Thus, both Glu and GABA are capable of acting as metabotropic signals. Both Glu and GABA are amino acids and so their metabotropic actions belong to the first entry in Table I. Acetylcholine (ACh) is another transmitter substance capable of producing metabotropic actions. The other principal small-molecule metabotropic transmitters are called biogenic amines and include: (1) the catecholamines dopamine (DA), norepinephrine (NE, also known as noradrenaline, NA), and epinephrine; (2) serotonin (5HT); and (3) histamine. All neuropeptides are metabotropic substances.

All these substances are called ***first messenger*** chemicals because their binding to metabotropic receptors in the postsynaptic cell triggers a ***biochemical cascade reaction*** inside the postsynaptic cell. There are many distinct kinds of metabotropic reactions, and this signaling process is one of the most complex found in biological signal processing. Nonetheless, the great majority of these signal processes follow the same general signal process flow, depicted in Figure 5.4.

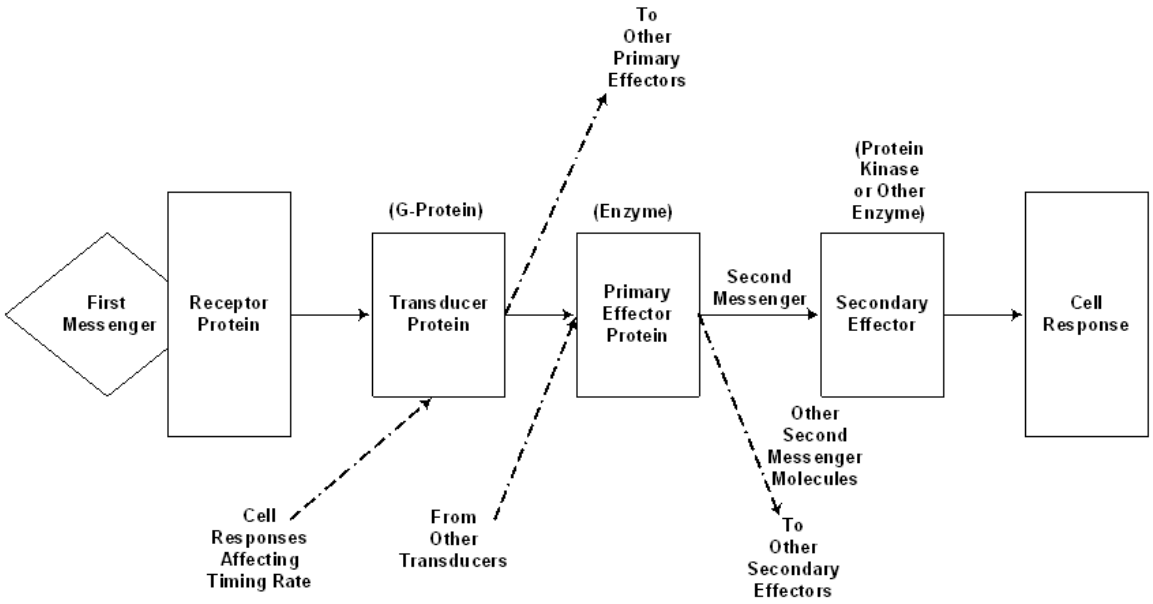

Figure 5.4: General signal flow process for metabotropic "second messenger" cascades.

The process begins with the binding of an extracellular "first messenger" to a metabotropic receptor protein. The receptor responds by activating a ***transducer protein*** located on the cytoplasmic side of the membrane wall of the neuron. In the most extensive class of second messenger systems, the receptor protein activates a ***G-protein*** ("GTP-binding protein"; GTP is guanosine triphosphate). There are, however, other second messenger systems that do not use the G-protein as the transducer. These include the guanylate cyclase receptors, tyrosine kinase receptors, and cytokine receptors. Of these, the G-protein-based second messenger systems are the most studied and we will confine our discussion to them.

The discovery of the G-protein is credited to the work of Alfred Gilman and Martin Rodbell, for which they shared the 1994 Nobel Prize. Speaking in 1994, Gilman tells us,

> It has become abundantly clear, particularly over the past decade, that this relatively large family of GTP-binding and hydrolyzing proteins plays an essential transducing role in linking hundreds of cell surface receptors to effector proteins at the plasma membrane. These systems are widely used in nature, controlling processes ranging from mating in yeast to cognition in man. Receptors that activate G proteins are correspondingly diverse and encompass proteins that interact with hormones, neurotransmitters, autacoids, odorants, tastants, pheromones, and photons. . . Four subfamilies are usually discussed: (1) the small G_s group (G_s and G_{olf}), best recognized as activators of adenylyl cyclases; (2) the large and functionally diverse G_i group, whose members are pertussis toxin substrates with one exception (G_z); (3) the G_q group, activators of several isoforms of phospholipase Cβ; and (4) the most recently recognized G_{12} group, whose functions are unknown. . . If all possible combinations . . . were allowed, we would need to consider at least 600 G protein oligomers [GILM].

G-proteins act as molecular timers and switches, turning on a primary effector protein (which

is typically an enzyme). This primary effector, in turn, causes the production of large numbers of molecules called ***second messengers***. The primary effector can be viewed as a kind of ***signal amplifier*** since one first messenger molecule can lead to the production of an enormous number of second messenger chemicals. The G-protein switches on this process and controls how long the production of second messengers will continue before the primary effector is switched off again. Three of the most important primary effector enzymes are adenylyl cyclase[2] (AC), phospholipase C (PLC), and phospholipase A_2 (PLA). AC converts ATP (adenosine triphosphate) to cAMP (cyclic adenosine monophosphate). PLC produces a pair of second messengers, IP_3 and diacylglycerol (DAG). PLA releases the second messenger arachidonic acid (AA).

The second messengers in turn act upon secondary effectors, which are likewise often enzyme proteins. The secondary effector is so called because it is the efficient cause of the final cell response to the biochemical cascade. In some cases, it should be noted, the G-protein is also capable of exercising direct action on ionotropic channels, and in these cases the G-protein itself acts as an effector. In the other cases, the secondary effector is often a ***protein kinase***[3], the most important of which are protein kinase A (PKA), which reacts to cAMP, protein kinase C (PKC), which reacts to DAG, and the Ca^{2+}/calmodulin-dependent protein kinases (CaM kinases). Other secondary effectors include the arachidonic acid metabolites 5-lipoxygenase, 12-lipoxygenase, and cyclooxygenase, and the receptor tyrosine kinases.

Figure 5.5 illustrates the four principal classes of metabotropic second messenger systems. One of the interesting features of second messenger signaling, which adds to the complex of different effects it exhibits on the cell, is the phenomenon of convergence and divergence in the second messenger pathways. ***Convergence*** occurs when more than one G-protein can act on the same primary effector. Two examples of convergence are shown in Figure 5.5. In the first, both G-protein G_s and G-protein G_i can act on primary effector AC. The action of G_s is to turn on AC, which stimulates production of increased levels of cAMP. The action of G_i is to inhibit AC, leading to a decline in the level of free cytoplasmic cAMP. The second example is the convergence of G_i and G-protein G_q on PLC. In this case, both G-proteins excite the primary effector, resulting in production of two second messengers, IP_3 and DAG.

Divergence occurs when activation of one G-protein or one primary effector produces two or more distinct second messenger cascades. Two examples of this are likewise shown in Figure 5.5.

---

[2] Also known as adenylate cyclase.

[3] A kinase is an enzyme transferring a phosphate group from a high-energy phosphate compound to a recipient molecule. The recipient molecule, which is often an enzyme, is thereby activated and able to perform some function. The action of a kinase is opposed by a ***phosphatase***, which removes the transferred phosphate group.

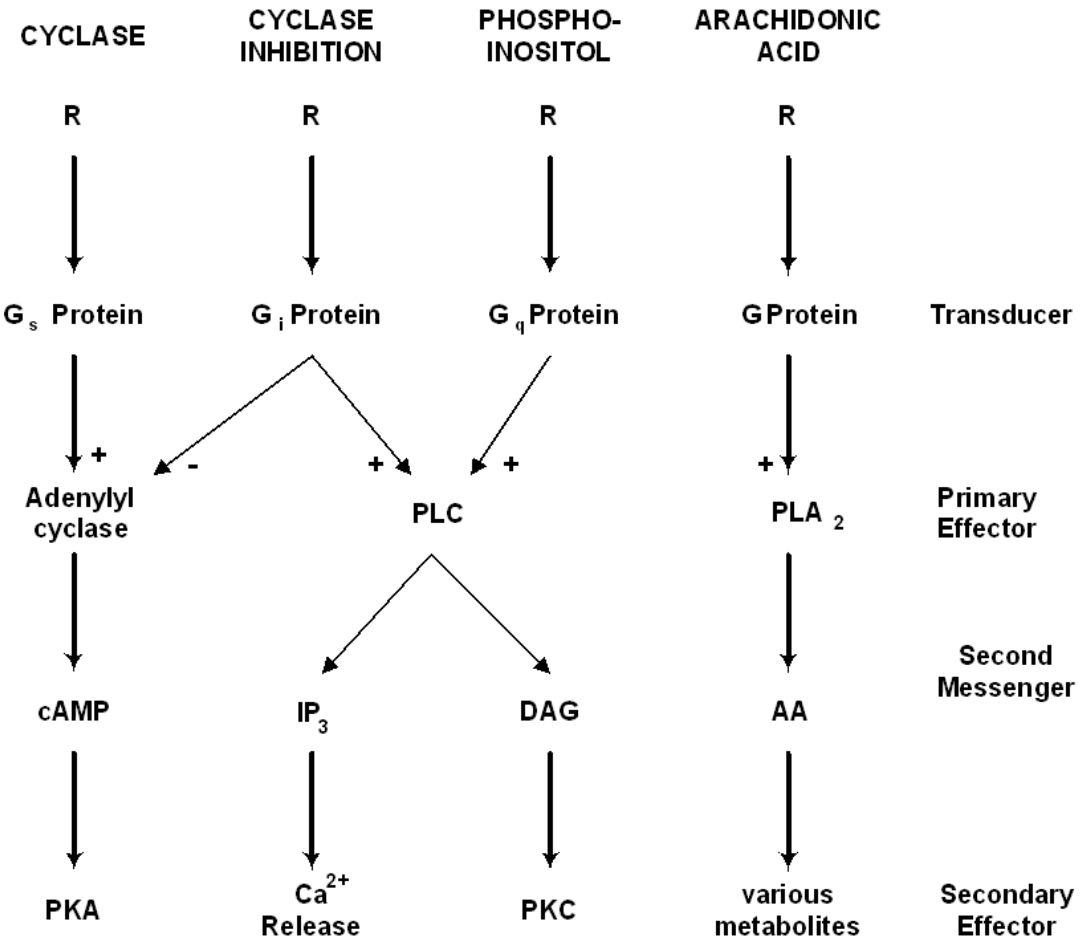

Figure 5.5: Networking pathways of the principal metabotropic second messenger systems. R = receptor. G_s and G_i denote stimulating and inhibitory G proteins, respectively. G_q is another type of G protein. PLC is the enzyme phospholipase C. PLA_2 is the enzyme phospholipase A_2. IP_3 is inositol 1,4,5-triphosphate. DAG is diacylglycerol. AA is arachidonic acid. PKA is the cAMP-dependent protein kinase. PKC is protein kinase C. A "+" sign indicates activation of a primary effector. A "-" sign indicates deactivation. At present not enough is known about the arachidonic acid process to permit a meaningful network description here. However, some of its metabolites are membrane-permeable and are thought to be capable of modulating nearby neurons.

The first we have already discussed, namely the divergent pathways taken from G_i. The second is illustrated by PLC's production of a *pair* of second messengers. IP_3 leads to the release of Ca^{2+} from internal stores in the cell's endoplasmic reticulum. DAG activates PKC, which in turn catalyzes a reaction of ***substrate proteins*** (channel proteins, receptor proteins, enzymes, or cytoskeletal proteins) with ATP to produce a ***phosphoprotein*** (phosphorylation of the target protein). Some PKC enzymes also require Ca^{2+} for activation; thus in many cases the IP_3 cascade must first liberate Ca^{2+} from the ER before the DAG cascade can get fully underway, another example of convergence. Figure 5.5 does not illustrate any "crosstalk" convergence or divergence for the AA cascade. This is merely because we do not yet know what these are for the AA path.

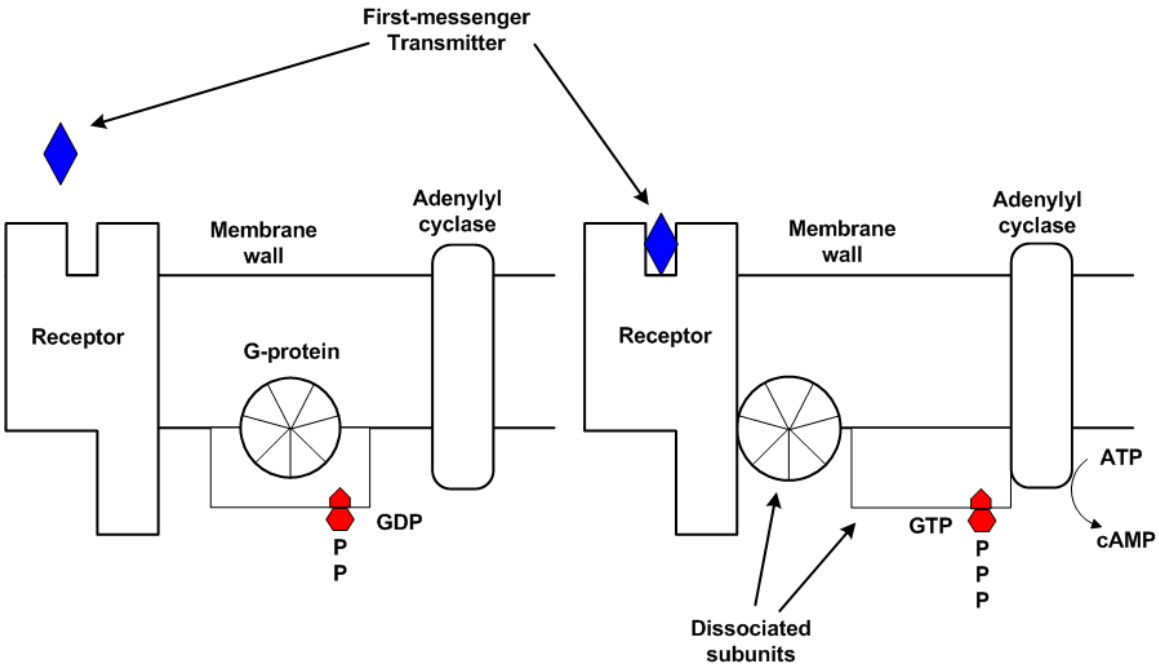

Figure 5.6: Illustration of G_s-protein activation of adenylyl cyclase. Only two steps in the cAMP cycle are illustrated in this figure. The full cycle has seven distinct configurations in it, starting from the figure on the left above, leading to the figure on the right at the fifth step, and returning to the inactive configuration on the right after two more steps in the cycle (re-association of the G-protein subunits, switching AC off again, and ejection of the first messenger neurotransmitter from the receptor protein). GDP = guanosine diphosphate. GTP = guanosine triphosphate. See text for discussion.

Although this text is not concerned with the details of organic chemistry except as this appears at the more abstract level of signal processing theory, it is nonetheless instructive and interesting to briefly look at the biomechanics of the G-protein's ability to act as a molecular timer/switch. A partial illustration in the case of G-protein activation of AC is shown in Figure 5.6. The cAMP cycle, for which Figure 5.6 is a partial illustration, has seven distinctly identifiable stages or "steps." An illustration of the full seven-step cycle is given in [SIEG]. Step 1, the inactive state, is depicted in the left-hand figure of 5.6. We may call this the "resting step" since the receptor protein has not been bound to a first messenger transmitter. In Step 2, a neurotransmitter binds with the receptor protein and the receptor undergoes a change of configuration, exposing a binding site for the G-protein. In Step 3, the G-protein diffuses through the lipid bilayer to bind with the receptor protein. This activates the G-protein for GTP-GDP exchange. In Step 4, GTP causes the G-protein to dissociate into two parts. One remains with the receptor protein. The other moves to an AC protein in the membrane wall. Step 5, illustrated by the right-hand figure in Figure 5.6, is the binding of the second subunit to AC, with subsequent activation of cAMP production. In Step 6, hydrolysis of the GTP by the second subunit returns the G-protein to its original configuration, causing it to dissociate from the AC and re-associate with the other

subunit. Finally, in Step 7 the receptor ejects the transmitter substance and the system returns to its resting configuration. The kinetics of this qualitative model can be given a quantitative and mathematical formulation in the form of a rate process, similar to that which we have seen earlier in the Hodgkin-Huxley model. In this case, the rate process would have seven states rather than only two. The *statistical* form of such a model is called a ***Markov process***, which will be discussed in §7.

§ 6. Phosphorylation and Dephosphorylation

Figures 5.4 and 5.5 might convey the impression that metabotropic signaling is a feedforward type of signal processing. However, this is true only to the extent of the secondary effector producing phosphorylation of a target substrate protein.[4] The overall process is also regulated by a feedback process of dephosphorylation, which is the mechanism by which metabotropic modulation of the cell response is brought to a conclusion.

It has been well established that phosphorylation/dephosphorylation is an important, and perhaps the primary, mechanism for regulating receptor sensitivity and modulating ion channel dynamics [HUGA], [LEVI]. Levitan remarks,

> Modulation of the properties of membrane ion channels is of fundamental importance for the regulation of neuronal electrical activity and of higher neural functions. Among the many potential molecular mechanisms for modulating the activity of membrane proteins such as ion channels, protein phosphorylation has by chosen by cells to play a particularly prominent part. This is not surprising given the central role of protein phosphorylation in a wide variety of cellular, metabolic, and signaling processes. As summarized here, regulation by phosphorylation is not restricted to one or another class of ion channel; rather, many, and perhaps all, ion channels are subject to modulation by phosphorylation. Similarly, a number of different protein kinase signaling pathways can participate in the regulation of ion channel properties, and it is not unusual to find that a particular channel is modulated by several different protein kinases, each influencing channel activity in a unique way. Finally, the biophysical mechanisms of modulation also exhibit a striking diversity that ranges from changes in desensitization rates to shifts in the voltage dependence and kinetics of channel activation and inactivation [LEVI].

While physiologists, molecular biologists, and neurochemists have by now long been aware of the importance and role of phosphorylation/dephosphorylation in the regulation of neuronal activity, this has not yet been widely recognized by neural network theorists in the form of neural network models that include it in their structures. Of the major "schools" of neural network theory, only adaptive resonance theory (ART) has so far incorporated mathematics to give a well-organized account for the modulation phenomena, albeit this accounting is high-level, abstract,

[4] In some cases, the G-protein itself acts in the role of the secondary effector by binding its mobile dissociated subunit to a substrate protein (typically a channel protein or a receptor protein). This action is similar in form to the cAMP cycle illustrated in Figure 5.5 except that the target is not an AC enzyme and the outcome is a direct cell response rather than the production of a second messenger This is sometimes called ***membrane-delimited control*** [HILL6].

and quite indirect. (In ART this type of network modulation is recognized through such mechanisms as the ART "vigilance parameter" and the gain control mechanisms of its "attentional" subsystem [GROSS10], [CARP4-6]). Yet most of the major subsystems in the brain that signal by means of metabotropic messengers have widespread targets across large regions of the neocortex and other cerebral structures. Because these modulating signals, after all is said and done, work through phosphorylation and dephosphorylation, this lack of attention by network theorists is difficult to justify, and perhaps it is impossible to truly justify when one realizes that ART methods are a possible means to take it into account at the network level of the reductionist hierarchy.

Not surprisingly, the kinetics of phosphorylation are different for the many different kinds of kinases. We will not delve into the fine details of the biochemistry involved here, mainly because there are so many but also because this is a topic for specialists. Some excellent reviews are available describing PKA [FRAN], PKC [TANA], the CaM kinases [HANS], and the receptor tyrosine kinases [FANT]. In this text the focus will be given to common features found in the processes of the ***phosphorylation/dephosphorylation cycle***.

A kinase exhibits (at least) two states, called ***active*** and ***inactive***. A kinase in the active state produces phosphorylation of its target substrate protein. Again, phosphorylation is the transfer of a phosphate group by a ***phosphorylase*** to an organic compound. A phosphorylase is any enzyme which catalyzes the addition of phosphate to an organic compound. This compound is then said to be phosphorylated. Phosphorylation is often ATP-dependent, and it produces compounds which are highly reactive in water with other organic molecules in the presence of appropriate enzymes. A kinase is a protein catalyzing the transfer of phosphate. An active kinase is one capable of acting as a catalyst in this way. An inactive kinase is one in a configuration where it does not act as such a catalyst.

Second messengers cause a kinase to enter the active state. Other factors in the cell's chemical milieu, not all of which are particularly well understood, return a kinase to its inactive ("basal") state. In addition to phosphorylating substrate proteins, some kinases have the ability to phosphorylate themselves; this is called ***autophosphorylation***. A kinase in this state remains at least partially active, and is returned to the inactive state through the action of a ***phosphatase***. Let S_0 denote the inactive state of a kinase. Let S_1 represent the normal active state of the kinase, and let S_2 represent the autophosphorylated active state.

By introducing a new element in the Linvill modeling schema, called the ***reactor***, we can represent the process just described. The simplest feasible representation is shown in Figure 5.7. The flux law for a reactor element, R, is identical to that of a transporter, T, except for its physical

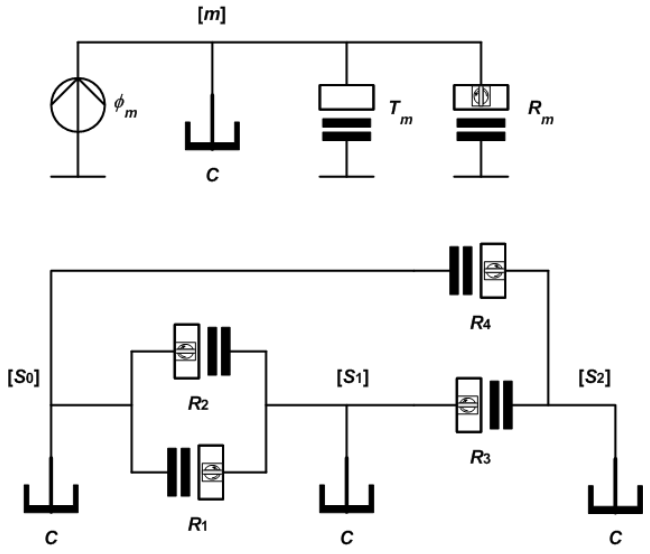

Figure 5.7: A simple Linvill schema for modeling inactive, active, and partially active kinase activated by a second messenger, *m*.

interpretation. Unlike transporter flux, which represents the active transport of a substance from one location to another, reactor "flux" represents the loss of concentration of a substance, [*X*], due to a chemical reaction or a change in chemical state involving substance *X*. The reactor flux is given by

$$\phi_R = R \cdot [X].$$

In general, the value of *R* will be a function of the concentration of some other substance, [*Y*], with which *X* is undergoing chemical reaction or which causes the change in chemical state of *X*. For example, reactor R_2 in figure 5.7 will be a function of the second messenger concentration, [*m*]. Similarly, reactor R_m in the figure will be a function of $[S_0]$, the concentration of inactive secondary effector kinase acted upon by second messenger *m*. Thus, the two Linvill networks in the figure are non-linearly coupled through the co-dependencies of the reactor elements. The sign convention for reactor flux is the same as that of the transporter element. For example, the differential equation describing node $[S_2]$ in the network is

$$C \cdot \frac{d[S_2]}{dt} = R_3 \cdot [S_1] - R_4 \cdot [S_2].$$

Many kinases exhibit a threshold effect for their activation by second messengers. This implies that the dependency of *R* on [*m*] is likely to follow a Boltzmann function or a function similar in general form to a Boltzmann function.

At least some second messengers (such as cAMP) are known to be able to reach any part of a

mammalian cell. In the case of cAMP, this can be accomplished in 100 ms. This omnipresent spread of second messengers might be true for most of the major second messengers in the neuron. It is therefore a reasonable speculation that the rate of activation of the messenger's target kinases is a function of messenger concentration, $[m]$, which is in turn a function of its production rate by the first effector (represented by ϕ_m in figure 5.7). Similarly, it is a reasonable speculation that the rate of phosphorylation of target substrate proteins (e.g. an ionotropic channel protein) is a function of the concentrations of active second effectors ($[S_1]$ and $[S_2]$ in figure 5.7) and of the concentration of non-phosphorylated target proteins (e.g. the number of such proteins in the neuron divided by the aqueous volume of the cell). Modeling this requires a third subnetwork, along the same lines as the lower subnetwork in figure 5.7, be added to the total network model.

The principal protein kinases are widely expressed throughout the central nervous system. Often they are heavily concentrated in the postsynaptic density (the region in and around the synapse where the ionotropic and metabotropic receptors are found). Because these membrane spanning proteins are target substrate proteins for active kinases, this localization is further evidence for how the modulatory role of these kinases is realized in biological signal processing.

At the present state of signal processing modeling theory for the neuron, there are many, many important factors for which we are not yet in possession of quantitative data needed to develop an accurate model. No doubt a large of amount of this information lies buried in the biochemistry and organic chemistry literature, e.g. [BUXB], expressed there in the language of the chemist rather than that of the computational neuroscientist. No doubt, too, some of this information is still entirely unknown, the relevant modeling research question not having yet been posed. This is an under-recognized research field for computational neuroscience and biological signal processing. There is very clear evidence that for many of the important kinases their activity level (which one can assess by examining the rates at which they phosphorylate their targets) is more complex than being merely binary or ternary (as figure 5.7 might suggest). Reports in the bio-chemistry literature indicate that the activity levels of the major kinases is dependent upon levels of ATP concentration (and probably other factors as well). Accounting for this clearly requires a more complex network model than the one given above. Indeed, the need for this accounting is one reason why a method of lumped-element representation of the process is useful and important for understanding the modulation processes of the neuron.

The experimentally observed fact that many of the major protein kinases, especially PKC and CaM, exhibit autophosphorylation, and thus remain partially active even after concentration levels of their second messengers decline, has been a source of great scientific interest. There is a great deal of speculation that autophosphorylation of the kinases is a *memory mechanism* for the

neuron [SCHW2]. This appears to be an undeniable fact in the case of elastic modulation, a form of short-term "memory" for the neuron. Additional evidence suggests that some kinases, particularly CaM and presynaptic PKC, may also support plastic (that is, irreversible) modulation of synaptic efficacy. Presently, this putative role is still somewhat speculative, but it is receiving a great deal of attention and one can reasonably expect the future to bring further clarification to this possibility.

The secondary effectors primarily exert their effects through phosphorylation of their target substrate proteins. The phosphorylation process is accompanied by another regulatory process, namely the process of dephosphorylation. Phosphorylation of the substrate protein can, depending on the kinase involved and the particulars of the target protein, work either to sensitive or to desensitize the response of the channel protein in responding to neurotransmitters. In some cases it can open a normally-closed channel; in others it can close a normally-open channel. One model of this effect, although disputed by some researchers, holds that phosphorylation of channel proteins can be represented by introducing the idea of an ***effective density*** of channel proteins in the synapse. In this phenomenological model, the effect of phosphorylation is looked upon as having the same effect as increasing (or decreasing) the number of receptors on the postsynaptic side of the synapse. AMPA receptors are often treated this way, although there seems to be no strong reason to exclude the other ionotropic receptors from being looked at in this way. However, as already noted, other researchers take issue with this model of cell response to the secondary effectors, and the matter seems to be far from settled on the physiology level. Computational neuroscientists, on the other hand, have been quick to adopt this model of channel sensitization/desensitization. At the modeling level discussed in chapter 4, this effect is taken into account by the setting of the relative synaptic weight, w, or the base g_0 of the channel.

Presynaptic protein kinases might possibly target synaptic vesicle membranes or other structures within the synaptic terminal and thereby effect a change in either p or N or both in equation (5.1). This is presently regarded as rather speculative and is often accompanied by a speculation that membrane-permeable ***retrograde messengers*** such as NO must also be involved in such a process. (NO is a byproduct of the action of the arachidonic acid process; the DAG-PKC second messenger process is thought to produce AA as a divergence byproduct).

Dephosphorylation is the removal of the phosphate group from the target protein. It is also the mechanism by which autophosphorylated protein kinases become inactive once again. Dephosphorylation is produced by enzymes called ***phosphoprotein phosphatases***. Whatever the cell response to the secondary effector was, dephosphorylation terminates the effect. Reactor R_4 in figure 5.7, which represents dephosphorylation of an autophosphorylated kinase, is an example

of how the Linvill modeling schema can represent of this effect. A large value of R_4 represents rapid dephosphorylation; a small value represents slow dephosphorylation.

Dephosphorylation can itself be regulated by metabotropic mechanisms. One well-studied case is that of phosphorylation of K^+ channels by the action of PKA [SIEG]. In this case, a K^+ channel opened by phosphorylation is dephosphorylated and closed by the action of an enzyme called phosphoprotein phosphatase-1. However, the action of this enzyme is regulated by another protein called inhibitor-1. Inhibitor-1 is phosphorylated by the action of PKA, and in this state it inhibits the action of phosphoprotein phosphatase-1. Here is one example of how one metabotropic signaling cascade can diverge (PKA inhibiting dephosphorylation, while at the same time producing it in the K^+ channel). Inhibitor-1 is dephosphorylated by the action of Ca^{2+}, which activates another phosphatase known as calcineurin that in turn dephosphorylates inhibitor-1. A similar process has been reported for regulating phosphorylation/dephosphorylation of NMDA receptors [HALP]. This one involves inactivation of the inhibitor DARPP-32 by CaM and activation of DARPP-32 by PKA produced from elevated cAMP levels stimulated by dopamine signaling.

Each of these actions just described can be approximated by adding additional Linvill networks to those depicted in figure 5.7. In the case of targeted channel proteins, if one adopts the effective-density model the relevant concentrations would be concentrations ν_p of phosphorylated channels and ν_d of dephosphorylated channels. Reactor elements would be used to model the kinetics of the phosphorylation/dephosphorylation process in terms of concentrations of various active secondary effectors. Modeling of dephosphorylation would involve networks similar to the lower network in figure 5.7 modeling the concentrations of inhibited and uninhibited phosphatases, calcineurin, etc. As this model-building is a corollary to the general idea illustrated in figure 5.7, this is left as an exercise at the end of this chapter.

§ 7. Markov Processes

The Hodgkin-Huxley technique for modeling rate processes expressed the rate variable (n, m, h) in terms of quantities with values ranging from 0 to 1. We have previously seen that this can likewise be interpreted in terms of the probabilities of channel deactivation, activation, and inactivation. The Hodgkin-Huxley rate variables, α and β, then take on the significance of being regarded as parameters describing ***conditional transition probabilities***.

When we look at kinetics processes in these statistical terms, what we have is a model that mathematicians and statistical signal processing theorists call a ***Markov process***. Markov processes are widely used in many, many fields. Any physical process that we represent in terms

of a Hodgkin-Huxley-like model, including a Hodgkin-Huxley-like formulation of a Linvill network, can be equivalently expressed in terms of a Markov process model in which the transition probabilities between the ***states*** of the model are functions of the state probabilities. Such a model is a ***time-varying*** Markov process said to be ***linear in parameters*** because the next state of the model is a linear function of the present state (linear in the transition probability parameters), but the parameters themselves are different after this transition (thus making the overall model time-varying).

How does one take a model representation from the "physical" representation of a Hodgkin-Huxley or a Linvill model to the more abstract representation of a Markov process? This can be easily demonstrated by showing how this is done for the case of the Hodgkin-Huxley rate process. We begin with a diagrammatic representation of the H-H rate process as shown in Figure 5.8(A). For simplicity we pick the activation variable, n, which represents the channel gate as having two states, open and closed. We will let S_0 represent the closed (deactivated) state and S_1 represent the open (activated) state. From our previous rate equations (chapter 3), we obtain the diagram of figure 5.8(A). The corresponding pair of differential equation is

$$\frac{dS_0}{dt} = -\alpha \cdot S_0 + \beta \cdot S_1$$
$$\frac{dS_1}{dt} = \alpha \cdot S_0 - \beta \cdot S_1$$

By interpreting states S_0 and S_1 as the fraction of gates closed or open and then imposing the constraint $S_0 = 1 - S_1$ we obtain the original Hodgkin-Huxley rate equation with S_1 interpreted as

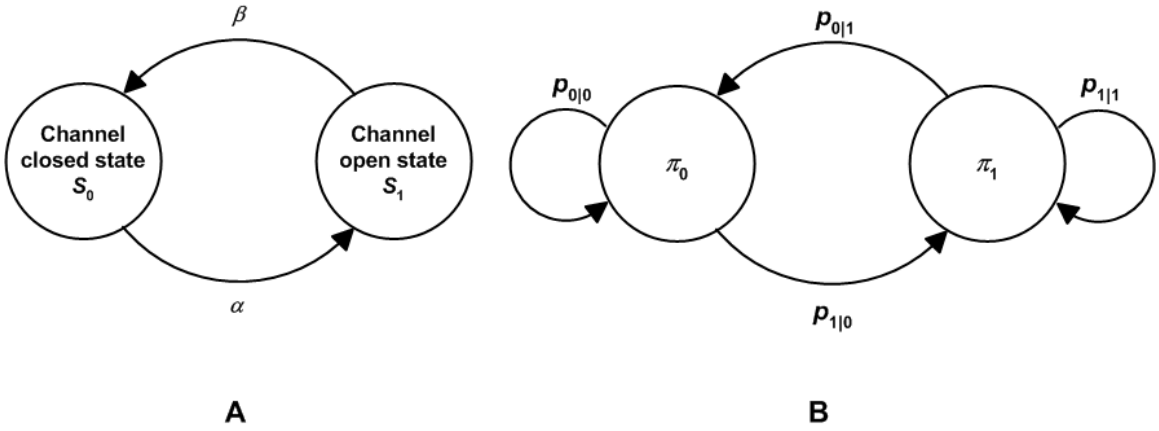

Figure 5.8: Rate process and equivalent Markov process diagrams. (A) The Hodgkin-Huxley first order rate process in diagram form. This diagram represents the process in continuous time in the form of a pair of differential equations. By letting S_0 and S_1 denote the fractions of the populations closed or open, respectively, and setting $S_0 = 1 - S_1$ we obtain the original Hodgkin-Huxley equation. (B) Equivalent Markov process diagram. This form is obtained from converting the pair of differential equations describing (A) to a pair of difference equations.

activation variable. Converting the pair of differential equations to difference equations using Euler's method gives us

$$\begin{bmatrix} S_0(t+\Delta t) \\ S_1(t+\Delta t) \end{bmatrix} = \begin{bmatrix} (1-\alpha\cdot\Delta t) & \beta\cdot\Delta t \\ \alpha\cdot\Delta t & (1-\beta\cdot\Delta t) \end{bmatrix} \begin{bmatrix} S_0(t) \\ S_1(t) \end{bmatrix} . \tag{5.3}$$

Now, when one interprets $S_0(t)$ and $S_1(t)$ as the fraction of gates closed or opened, and if one further assumes (as Hodgkin and Huxley did) that the closing and opening of any particular gate is statistically independent of the closing or opening of any other gate, then S_0 and S_1 can be equally well interpreted as representing the probabilities, π_0 and π_1 of any particular gate being closed or open. We may then replace the state variables in (5.3) by their corresponding probabilities. The state matrix on the right-hand side of (5.3) then has the interpretation of being a matrix of ***state transition probabilities***. These are conditional probabilities, e.g. $p_{0|1}$ is the probability of making a transition to state S_0 given that the current state is S_1.

Making these changes of variables in (5.3) results in the expression

$$\begin{bmatrix} \pi_0(t+\Delta t) \\ \pi_1(t+\Delta t) \end{bmatrix} = \begin{bmatrix} p_{0|0} & p_{0|1} \\ p_{1|0} & p_{1|1} \end{bmatrix} \begin{bmatrix} \pi_0(t) \\ \pi_1(t) \end{bmatrix} . \tag{5.4}$$

This is the system of equations represented in diagram form by Figure 5.8(B). This diagram is called a ***Markov process diagram***. The entries in the state transition matrix of (5.4) correspond one-to-one to the entries in the state matrix of (5.3) and give them a statistical interpretation. (5.4) is written in more concise form by using standard vector-matrix notation, $\Pi(t + \Delta t) = \mathbf{P}\ \Pi(t)$.

The obvious disadvantage of using expressions (5.3) or (5.4) is that they require two equations to be solved, whereas the original Hodgkin-Huxley formulation requires only one. The rows of the system of equations (5.3) or (5.4) are not linearly independent. Note that if we sum the column entries in the state matrix of (5.3), both columns sum to unity. The same is true for the state transition matrix in (5.4), and this is a general feature of all Markov processes. The proper interpretation of this mathematical property is simply this: Given whatever state the system is in at time t, it is certain to be in some state at time $t + \Delta t$. Furthermore, the sum of all the state probabilities at any time t is always unity. Thus, the state matrices of (5.3) and (5.4) are not full rank. The expression for the steady state in (5.4) is simply $\Pi(t + \Delta t) = \Pi(t)$. If we make this substitution into (5.4), we find that the solution for $\Pi(t)$ is indeterminate. In order to solve (5.4) for the steady state, one must replace one of the rows (which one doesn't matter) by the constraint equation $1 = \pi_0 + \pi_1$. This is, in effect, how Hodgkin and Huxley were able to reduce (5.3) to a single difference equation.

This disadvantage of the Markov process representation becomes less of a disadvantage as the number of states in the system increases. If **P** is an $n \times n$ matrix, its rank will always be $n - 1$. For any column of **P**, the sum of the column entries will always equal 1 because the sum of all the state probabilities must always add up to unity. The column-sum property is a mathematical constraint placed on *any* probabilistic model of *any* rate process kinetics.

As an example, let us derive the Markov process model for the kinase process illustrated in the lower part of Figure 5.7. Here we have three state variables. Summing the effluxes from each node to obtain the system of differential equations, converting these to difference equation form by Euler's method, and making the substitution of the state probability variables, we obtain

$$\begin{bmatrix} \pi_0(t+\Delta t) \\ \pi_1(t+\Delta t) \\ \pi_2(t+\Delta t) \end{bmatrix} = \begin{bmatrix} (1-\Delta t \cdot R_2/C) & (\Delta t \cdot R_1/C) & (\Delta t \cdot R_4/C) \\ (\Delta t \cdot R_2/C) & (1-\Delta t \cdot (R_1+R_3)/C) & 0 \\ 0 & (\Delta t \cdot R_3/C) & (1-\Delta t \cdot R_4/C) \end{bmatrix} \begin{bmatrix} \pi_0(t) \\ \pi_1(t) \\ \pi_2(t) \end{bmatrix} . \qquad (5.5)$$

The derivation of (5.5) is left as an exercise at the end of the chapter. Note that for each column of the state transition matrix, the sum of the column elements adds up to unity, as it must for any correct model of this class. The entries in the state matrix of (5.5) are interpreted as the transition probabilities, $p_{n|m}$, for making a transition to state n from an initial state m.

The combination of a network representation using the Linvill schema with the probabilistic representation of a Markov process provides a powerful tool for general modeling of the dynamical processes in a neuron. Furthermore, the mathematical constraint that the columns of transition probabilities must add to 1 provides a check on the correctness of the kinetics network represented by the Linvill model. Note that the three equations of (5.5) can be combined with the single difference equation for [m] in Figure 5.7. The equation for [m] will be a difference equation for the second messenger concentration rather than for a probability, and this equation will be uncoupled from the three state probability equations. (The coupling in the physical process being represented is indirect; it occurs through co-dependencies of the process parameters in just the same manner as membrane voltage is coupled to the rate constants in the Hodgkin-Huxley model). The constraint on the sum of column elements applies only to the submatrix representing the state transition probabilities because the terms relating to [m] are not interpreted as state transition probabilities. The equations for the model of Figure 5.7 are easily shown to be

$$\begin{bmatrix} [m](t+\Delta t) \\ \pi_0(t+\Delta t) \\ \pi_1(t+\Delta t) \\ \pi_2(t+\Delta t) \end{bmatrix} = \begin{bmatrix} (1-\Delta t \cdot (T_m+R_m)/C) & 0 & 0 & 0 \\ 0 & p_{0|0} & p_{0|1} & p_{0|2} \\ 0 & p_{1|0} & p_{1|1} & p_{1|2} \\ 0 & p_{2|0} & p_{2|1} & p_{2|2} \end{bmatrix} \begin{bmatrix} [m](t) \\ \pi_0(t) \\ \pi_1(t) \\ \pi_2(t) \end{bmatrix} + \begin{bmatrix} \Delta t/C \\ 0 \\ 0 \\ 0 \end{bmatrix} \phi_m(t)$$

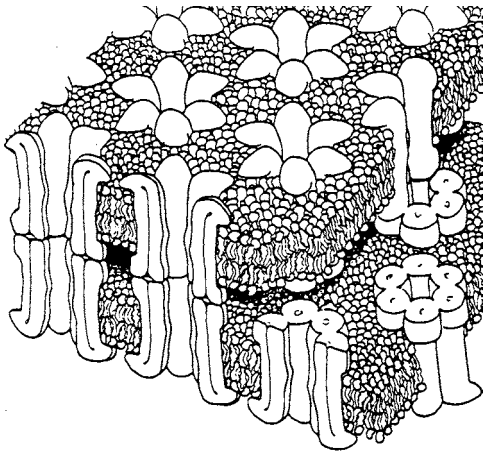

Figure 5.9: Illustration of the gap junction synapse. The spacing of the synaptic cleft at a gap junction narrows to around 3.5 nm (compared to 20 nm for a chemical synapse). The channel is formed by a pair of six-sided membrane-spanning structures called connexons. The spacing between connexons is about 8.7 nm, giving a channel density of around 26,400 pores per μm^2. The pore conductance is on the order of 100 to 120 pS/connexon.

where the $p_{n|m}$ entries are obtained from (5.5).

§ 8. Gap Junction Synapses

Although chemical synapses are the most common form of synapse found in the neurons of the central nervous system, a second type, the ***gap junction synapse*** is also found to occur between some neurons. Gap junctions also are the principal synapse connecting glia to one another. Figure 5.9 illustrates the arrangement of a gap junction synapse. At the gap junction, the spacing between cell membrane walls narrows to about 3.5 nm. Gap junctions range in diameter from about 0.1 to about 10 μm. The channel is formed by a six-sided structure called a ***connexon***. The channel diameter is around 1.5 nm and connexons are spaced at intervals of approximately 8.7 nm. This gives a channel pore density of approximately 26,400 pores/μm^2. Channel conductance is large, around 100 to 120 pS/pore. Thus, the over conductance of the junction, G_{gj}, ranges from around 20 nS to 200 μS, depending on the diameter of the junction.

Gap junctions come in two different types, non-rectifying and rectifying. The non-rectifying junction is the most common in the central nervous system. It's model is quite simple, merely an electrical conductance, G_{gj}, connecting two neurons. A simplified circuit model of the gap junction connection is shown in Figure 5.10. There is very little voltage drop across the gap junction, and in equilibrium the two cells will maintain the same membrane voltage. The gap junction causes a larger overall conductance to be presented to the chemical synapse circuit, which means it is more difficult for the chemical synapse to stimulate the neuron into firing. When one neuron in a gap-junction-connected network is stimulated into firing, the other neurons

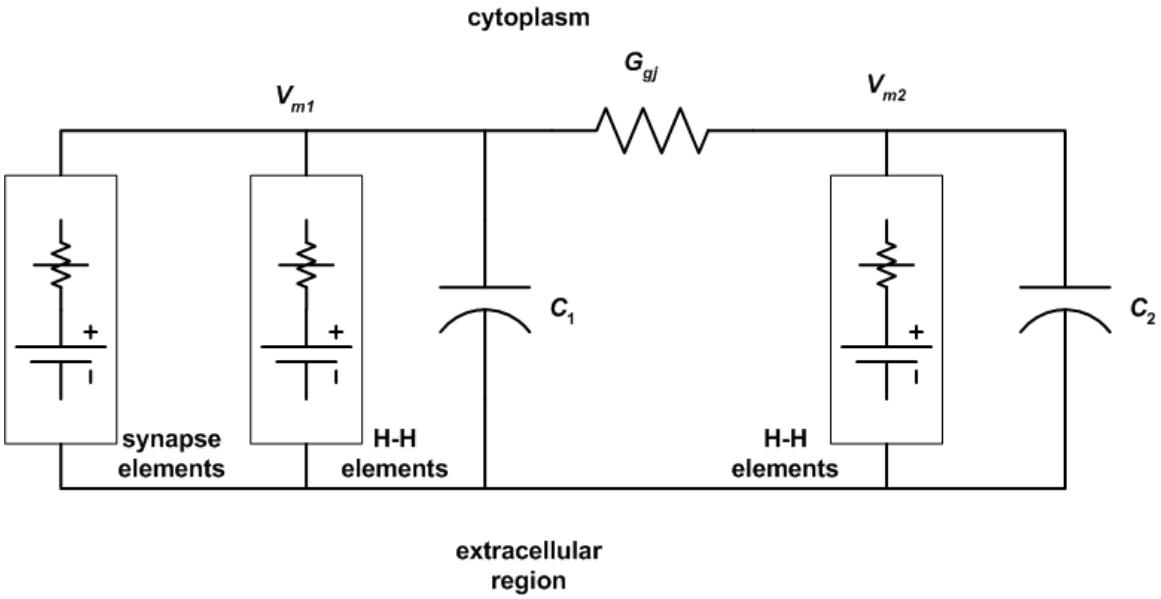

Figure 5.10: Simplified circuit model of two neurons connected by a gap junction synapse. The synaptic and Hodgkin-Huxley circuit elements have been simplified (cell-level Thévenin equivalent circuits) in this figure. Current flow from neuron 1 to neuron 2 is simply $I = G_{gj} \cdot (V_{m1} - V_{m2})$.

will likewise be stimulated into firing. Thus, one can regard a network of gap-junction-connected neurons as, in effect, constituting a single gigantic neuron with multiple inputs and outputs. It has been proposed that gap junctions are a mechanism for promoting synchronized firing by groups of neurons.

The second, less common, form of gap junction is the rectifying gap junction. This type of junction easily conducts current in one direction, say from cell 1 to cell 2 in figure 5.10, but not in the other. Its conductance is well approximated using the Heaviside function,

$$G = G_{gj} \cdot \mathbf{H}\left(V_{m1} - V_{m2}\right).$$

Why some gap junctions rectify is not entirely understood. Some have proposed there may be a minute voltage-dependency for the opening of the channel pore that is more sensitive on one side of the junction than on the other, but this hypothesis is at present somewhat speculative.

It is known that non-rectifying gap junctions can be modulated by chemical messengers within the cell. Neuronal gap junctions have been found to close in response to lowered pH levels, and in response to elevated Ca^{2+} levels. Other neurotransmitters stimulating metabotropic signaling have also been found to alter gap junction conductance. One function such modulation might perform is to disconnect gap-junction-connected neurons. Gap junctions are have been found in the retina, between certain types of inhibitory interneurons in neocortex, and in the neural circuitry of the brainstem.

Because their membrane voltage is more or less inactive, in networks of gap-junction-connected glial cells ion transport seems likely to be primarily a matter of diffusion rather than ion drift caused by electric potential differences. If ion diffusion current dominates electrical current in glia, the Linvill model is a more appropriate model for transport of chemicals via the glial network. Because glia are known to transport cytoplasmic calcium waves, this suggests that gap junctions in glial cells do not share the propensity seen in neurons for elevated Ca^{2+} to close gap junctions. This might be merely a matter of the possible existence of some threshold phenomenon that depends on levels of $[Ca^{2+}]$, this perhaps being higher in glia than in neuronal cytoplasm, or it might hint at a quite different set of biophysical conditions for glia. This, however, is speculation and merits its own investigation into whether gap junctions behave differently for glia vs. neurons.

Exercises

1. Calculate and plot the probabilities Pr[n] for exocytosis by n vesicles as a function of n for a releasable pool of $N = 5$ vesicles and release probabilities p of 0.1 and 0.2. Calculate the variance for each of the two release probabilities.
2. For $p = 0.1$ and $N = 5$, what is the average number of action potentials required to stimulate exocytosis by at least 1 vesicle assuming that the action potentials are spaced far enough apart in time for Ca^{2+} buffering to return presynaptic Ca^{2+} concentration to its basal level? How might one expect faster AP rates, relative to the Ca^{2+} buffering rates, to affect neurotransmitter release?
3. Develop a first-order Linvill model for presynaptic calcium concentration in response to the occurrence of an action potential. Assume a basal $[Ca^{2+}]$ level of 100 nM and that action potential induced calcium influx into the presynaptic terminal is by means of HVA calcium channels. What physical parameters are required for this model to make quantitative predictions of calcium concentrations?
4. Why would synchronous AP signaling at many synapses at the same time improve the reliability of neuronal signaling in a neural network? Why would high-rate action potential signaling at one synapse improve the reliability of neuronal signaling in a neural network? If on the average $p = 0.1$ and $N = 5$, how many synapses would need to receive action potentials in order to activate signaling for an average of 50 synapses in the postsynaptic neuron?
5. The binomial distribution model of synaptic exocytosis makes the implicit assumption that the measured level of post synaptic current is proportional to the number of vesicles undergoing exocytosis in response to an AP. What does this assumption imply for NTX binding on the postsynaptic side of the cleft?
6. Assume the active area of a synapse contains $N = 5$ primed vesicles arranged in the configuration illustrated in Figure E.6. Pick out 5 reasonable geometric centers for a 5-compartment Linvill model of NTX diffusion into the cleft and propose such a model network. Assume that outside the active area [NTX] is effectively zero and ignore NTX re-uptake. Model NTX exocytosis as an impulsive flux event and make reasonable scaling approximations for compartment volumes (use the center compartment as reference C) and make reasonable scaling approximations for the relative diffusance values (taking those attached to the centermost compartment as reference value D). Your compartments do not need to have equal volumes. Assume that within the synaptic cleft D is proportional to the cross sectional area between compartments and inversely proportional to the straight-line distance from the vesicle center to the compartment center. For diffusion out of the cleft assume D is inversely proportional to the mean distance from the center of the compartment to the periphery of the cleft region. Make whatever other

assumptions you deem reasonable, but explain them.

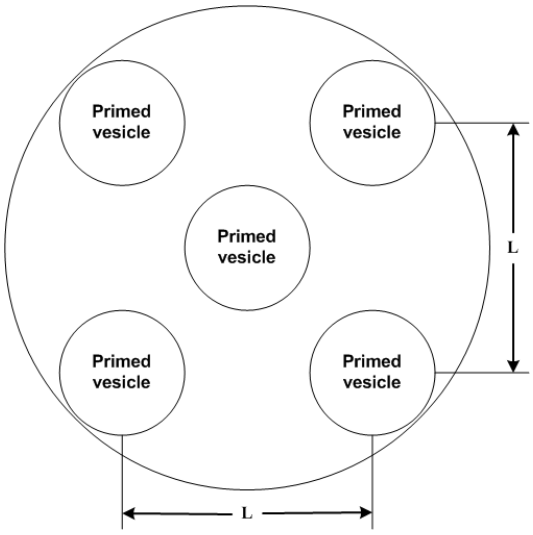

Figure E.6

7. Sketch the Hodgkin-Huxley circuit model for a two-neuron network connected by a reciprocal synapse arrangement as shown in Figure 5.3. In representing each neuron, it is sufficient to represent the neuron's voltage-gated channels using a single Thévenin equivalent circuit segment. Clearly label what the activating signal is for each of the two synapses.

8. Repeat exercise 7 for the case of the synaptic glomerulus of Figure 5.3.

9. The simplest kinetic model one might propose for the action of a G-protein is the second-order statistical model shown in Figure E.9. Models of this type are called Markov process models. Let π_0 denote the probability that the G-protein has not switched on its primary effector. Let π_1 denote the probability the G-protein has switched on its primary effector. Let k denote a time index in units of Δt. The terms $p_{n|m}$ denote the conditional probabilities for entering state n given that the present state is m. Assume the rate of second messenger production is proportional to π_1. Let $\pi_0(0) = 1$ and $\pi_1(0) = 0$. Show that for $k \geq 0$ the state probabilities are given by

$$\begin{bmatrix} \pi_0(k) \\ \pi_1(k) \end{bmatrix} = \begin{bmatrix} p_{0|0} & p_{0|1} \\ p_{1|0} & p_{1|1} \end{bmatrix}^k \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Plot the two state probabilities vs. k for $p_{0|0} = 0.8$, $p_{0|1} = 0.01$, $p_{1|0} = 0.2$, $p_{1|1} = 0.99$. Do you think this model properly accounts for the eventual turning off of the G-protein? Why or why not? Based on your answer, is this kinetics model a correct model for G-protein action?

10. The modeling error for the model in exercise 9 is improper treatment of the π_0 state. When a G-protein returns to the inactive condition, its receptor protein ejects the first messenger substance and so the conditions are different from those represented by the π_0 state. This can be accounted for by adding a *trap state* to the model as shown in Figure E.10. In this model $p_{0|1}$ is set to zero, denoting that the system cannot return to the π_0 state directly from the π_1 state. Transition probability $p_{2|1}$ replaces the old $p_{0|1}$ transition probability, and $p_{2|2} = 1$ denotes that the G-protein remains inactive until a new first messenger re-starts the cycle. Derive the new system of difference equations describing the model of Figure E.10 and repeat the rest of exercise 9 for the new system. Except as noted here, use the same initial conditions and state transition probabilities as before.

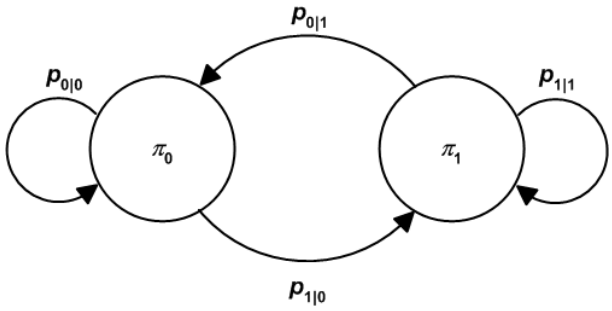

Figure E.9.

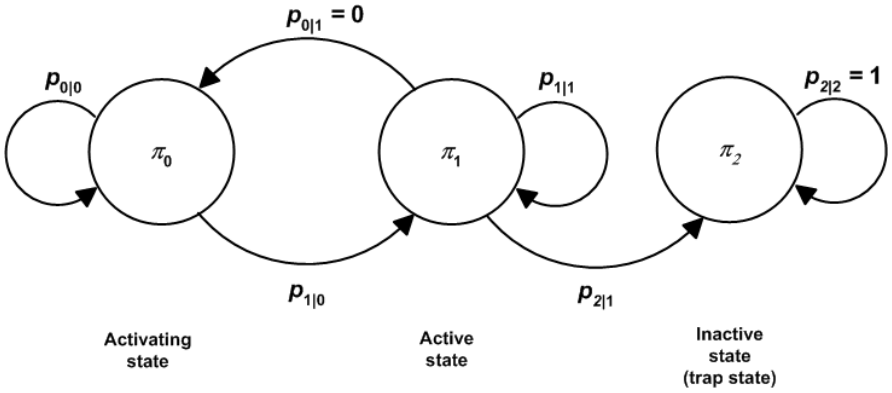

Figure E.10

11. Derive the full set of four differential equations describing Figure 5.7.
12. Deduce equation (5.4) directly from the diagram of Figure 5.8(B). How does this deduction differ from the process by which differential equations are derived from rate process diagrams such as Figure 5.8(A)? (The reason the processes differ is because one describes a *differential* equation at a single instant in time, t, whereas the other describes a *difference* equation giving the values of the variables at the *next* time step).
13. Derive equation (5.5).