CHAPTER 8

Release of Neurotransmitters

Robert S. Zucker, Dimitri M. Kullmann, and Thomas L. Schwarz

The synapse is the point of functional contact between one neuron and another. It is the primary place at which information is transmitted from neuron to neuron in the central nervous system or from neuron to target (gland or muscle) in the periphery. The simplest way for one cell to inform another of its activity is by direct electrical interaction, in which the current generated extracellularly from the action potential in the first cell passes through neighboring cells. Owing to the shunting of current by the highly conductive extracellular fluid, a 100 mV action potential may generate only 10–100 µV in a neighboring neuron. This coupling can be improved if neighboring cells are joined by a specialized conductive pathway through gap junctions (see Chapter 15); even then, a presynaptic spike is not likely to generate more than about 1 mV postsynaptically, unless the presynaptic process is nearly as large or larger than the postsynaptic process. This biophysical constraint limits the number of presynaptic cells that can converge on and influence a postsynaptic cell, and such electrical connections normally can be only excitatory and short lasting, are bidirectional in transmission, and show little plasticity or modifiability. They have limited potential for complex computation but can be useful when a postsynaptic neuron must be activated with high reliability and speed or when concurrent activity in a large number of presynaptic afferents must be signaled.

ORGANIZATION OF THE CHEMICAL SYNAPSE

Most interneuronal communication relies on the use of a chemical intermediary, or transmitter, secreted subsequent to action potentials by presynaptic cells to influence the activity of postsynaptic cells. In chemical transmission, a single action potential in a small presynaptic terminal can generate a large postsynaptic potential (PSP) (as large as tens of millivolts). This is accomplished by the release of thousands to hundreds of thousands of molecules of transmitter that can bind to postsynaptic receptor molecules and open (or close) as many ion channels in about 1 ms. There is room for many afferents (often thousands) to interact and influence a postsynaptic neuron, and the effect can be either excitatory or inhibitory, depending on the ions that permeate the channels operated by the receptor. The resulting responses are either excitatory postsynaptic potentials (EPSPs) or inhibitory postsynaptic potentials (IPSPs), depending on whether they drive the cell toward a point above or below its firing threshold. Different afferents can have different effects, with different strengths and kinetics, on each other as well as on postsynaptic cells. These differences depend on the identity of the transmitter(s) released and the receptors present (see Chapters 9, 11, and 16). Chemical synapses are often modified by prior activity in the presynaptic neuron.
Chemical synapses are also particularly subject to modulation of presynaptic ion channels by substances released by the postsynaptic or neighboring neurons. This flexibility is essential for the complex processing of information that neural circuits must accomplish, and it provides an important locus for modifiability of neural circuits underlying adaptive processes such as learning (Chapter 19).

Transmitter Release Is Quantal

One of the first applications of the microelectrode was the discovery that transmitter release is quantal in nature (Katz, 1969). Transmitter is released spontaneously in multimolecular packets called quanta in the absence of presynaptic electrical activity. Each packet generates a small postsynaptic signal—either a miniature excitatory or a miniature inhibitory postsynaptic potential (mEPSP or mIPSP, respectively, or just mPSP); under voltage clamp, a miniature excitatory or a miniature inhibitory postsynaptic current (mEPSC or mIPSC, respectively, or just mPSC) is generated. An action potential tremendously, but very briefly, accelerates the rate of secretion of quanta and synchronizes them to evoke a PSP. Vertebrate skeletal neuromuscular junctions are frequently used as model synapses, because both receptors and nerve terminals are relatively accessible for anatomical, electrophysiological, and biochemical studies. At the neuromuscular junction, the motor nerve forms a cluster of small unmyelinated processes that lie in shallow gutters in the muscle to form a structure called an end plate, and PSPs, PSCs, mPSPs, and mPSCs are called end-plate potentials (EPPs), end-plate currents (EPCs), miniature end-plate potentials (mEPPs), and miniature end-plate currents (mEPCs), respectively.

Why is transmission quantized? Neural circuits must process complex and quickly changing information fast enough to generate timely appropriate responses. This requires rapid transmission across synapses. Fast-acting chemical synapses accomplish this by concentrating transmitter in membrane-bound structures, ~50 nm in diameter, called synaptic vesicles and docking these vesicles at specialized sites called active zones along the presynaptic membrane (Fig. 8.1A). Vesicles not docked at the membrane are clustered behind it and associated with cytoskeletal elements (Heuser, 1977). Action potentials release transmitter by depolarizing the presynaptic membrane and opening Ca²⁺ channels that are strategically colocalized with the synaptic vesicles in the active zone (Robitaille et al., 1990). The local intense rise in Ca²⁺ concentration triggers the fusion of docked vesicles with the plasma

![Ultrastructural images of synaptic vesicle exocytosis. Synapses from frog sartorius neuromuscular junctions were quick-frozen milliseconds after stimulation in 4-aminopyridine to broaden action potentials and enhance transmission.](image-url)
membrane (called exocytosis; Figs. 8.1B and C) (Heuser and Reese, 1981) and the release of their contents into the narrow synaptic cleft (about 100 nm wide) separating the presynaptic terminal from high concentrations of postsynaptic receptors. The fusion of one vesicle releases about 5000 transmitter molecules within a millisecond (Fletcher and Forrester, 1975; Kuffler and Yoshikami, 1975; Whittaker, 1988) and generates the quantal response recorded postsynaptically. No membrane carrier can release so much transmitter this fast, nor can a pore or channel unless some mechanism exists to concentrate the transmitter behind the pore, which may be regarded as the function of synaptic vesicles. Evidence that transmitter is released from vesicles and that one quantum is due to exocytosis of a vesicle is summarized in Box 8.1 (Betz and Bewick, 1993; Eccles, 1964; Edwards, 1992; Heuser and Reese, 1973; Heuser et al., 1979; Hurlbut et al., 1990; Large and Rang, 1978; Murthy and Stevens, 1998; Prior, 1994; Ryan et al., 1997; Schiavo et al., 1994; Searl et al., 1991; Torri-Tarelli et al., 1985, 1992; Van der Kloo, 1988, 1991; Van der Kloo and Molgó, 1994; von Wedel et al., 1981).

At neuromuscular junctions, transmitter from one vesicle diffuses across the synaptic cleft in 2 μs and

**BOX 8.1**

**EVIDENCE THAT A QUANTUM IS A VESICLE**

**Transmitter Is Released from Vesicles**

1. All chemically transmitting synaptic terminals contain presynaptic vesicles (Eccles, 1964).
3. Rapid freezing of neuromuscular junctions during stimulation shows vesicle exocytosis occurring at the moment of transmitter release (Torri-Tarelli et al., 1985).
5. Retarding the filling of vesicles by using transport inhibitors (e.g., vesamicol for acetylcholine) or by reducing the transvesicular pH gradient generates a class of small mEPSPs that probably represent partially filled vesicles; drugs that enhance vesicle loading increase mEPSP size (Prior, 1994; Searl et al., 1991; Van der Kloo, 1991).
6. Quantal size is independent of membrane potential or cytoplasmic acetylcholine concentration altered osmotically (Van der Kloo, 1988).
7. Synaptic vesicles formed by endocytosis load with extracellular electron-dense and fluorescent dyes (horseradish peroxidase and FM1–43, respectively) after nerve stimulation; the dye is released by subsequent stimulation (Betz and Bewick, 1993; Heuser and Reese, 1973).
8. False transmitters synthesized from choline derivatives load slowly into cholinergic vesicles; they are co-released with acetylcholine in proportion to their concentrations in vesicles (Large and Rang, 1978).
9. Clostridial toxins that interfere with the synaptic vesicle–plasma membrane interaction block neurosecretion (Schiavo et al., 1994).

**One Quantum Is One Vesicle**

1. The number of acetylcholine molecules in isolated vesicles corresponds to the number of molecules released in a quantum (Fletcher and Forrester, 1975; Kuffler and Yoshikami, 1975; Whittaker, 1988).
2. When release is enhanced and the collapse of vesicle fusion images is prolonged by treatment with the potassium channel blocker 4-aminopyridine to broaden action potentials, the number of vesicle fusions observed corresponds to the number of quanta released by an action potential (Heuser et al., 1979). In these special circumstances, several vesicles are released at each active zone (Fig. 8.1C).
3. The number of vesicles present in nerve terminals corresponds to the total store of releasable quanta. When endocytosis is blocked by the temperature-sensitive Drosophila mutant sliche (Van der Kloo and Molgó, 1994) or pharmacologically (Hurlbut et al., 1990) and the motor nerve is stimulated to exhaustion, the number of quanta released corresponds to the original number of presynaptic vesicles.
4. The statistical variations in quantal release match the statistical variations in vesicle release and recovery measured with staining and destaining of lipophilic dyes (see Box 8.3 and section titled “Quantal Analysis”) (Ryan et al., 1997; Murthy and Stevens, 1988).
reaches a concentration of about 1 mM at the postsynaptic receptors (Matthews-Bellinger and Salpeter, 1973). These receptors bind transmitter rapidly, opening from 1000 to 2000 postsynaptic ion channels (Van der Kloot et al., 1994) (two molecules of transmitter must bind simultaneously to receptors to open each channel; see Chapters 11 and 16). Each channel has a 25 pS conductance and remains open for about 1.5 ms, admitting a net inflow of 35,000 positive ions. A single action potential in a motor neuron can release 300 quanta within about 1.5 ms along a junction that contains about 1000 active zones. The resulting postsynaptic depolarization, which begins after a synaptic delay of about 0.5 ms and reaches a peak of tens of millivolts, is typically sufficient to generate an action potential in the muscle fiber.

At fast central synapses, postsynaptic cells make contact with presynaptic axon swellings called varicosities when they occur along fine axons and boutons when they are located at the tips of terminals. Each varicosity or bouton contains one active zone or a few of them. The postsynaptic process is often on a fine dendritic branch or tiny spine with a length of a few micrometers, having a very high input resistance and capable of generating active propagating responses (Chapter 17). At inhibitory GABAergic synapses and excitatory glutamatergic synapses (Edwards et al., 1990; Jonas et al., 1993), each action potential releases from 5 to 10 quanta, and each quantum released elevates the transmitter concentration (Clemens et al., 1992; Tang et al., 1994; Tong and Jahr, 1994) in the cleft to about 1 mM and activates about 30 ion channels. At excitatory synapses, this release may be sufficient to generate EPSPs of 1 mV or less in amplitude, clearly subthreshold for generating action potentials. But central neurons often receive thousands of inputs, each of which has a “vote” on how the cell should respond (see Chapter 16). No input has absolute, or even majority, control over postsynaptic cell activity, but the matching of quantal size to input resistance ensures that inputs are reasonably effective. Consequently, at synapses onto larger central neurons with lower input resistances, quanta open between 100 and 1000 postsynaptic channels.

**Synaptic Vesicles Are Recycled**

A constant supply of vesicles filled with transmitter must be available for release from the nerve terminal at all times. Maintaining this supply requires the efficient recycling of synaptic vesicles. For this purpose, two partly overlapping cycles are used: one for the components of the synaptic vesicle membrane and another for the vesicle contents (transmitter substances). The cycles overlap from the time of transmitter packaging into vesicles until exocytosis. The cycles are distinct during the stages in which vesicle membrane and transmitter are recovered for reuse. The various steps of these cycles are common to all chemical synapses and are summarized in Fig. 8.2.

**Vesicle Membrane Cycle**

The components of the synaptic vesicle membrane are initially synthesized in the cell body before being transported to nerve terminals by fast axoplasmic transport (Bennett and Scheller, 1994; Jahn and Südhof, 1994) (see Chapter 2). Within the nerve terminal, the synaptic vesicles are loaded with transmitter and either anchored to each other and actin filaments (McGuiness et al., 1989) or targeted to plasma membrane docking sites at active zones. These docking sites are also rich in clusters of high-voltage-activated Ca²⁺ channels (Haydon et al., 1994; Robitaille et al., 1990), mainly N- and P/Q-type Ca²⁺ channels, depending on the synapse (Dunlap et al., 1995; Wheeler et al., 1994; see Chapters 5 and 6). Depolarization of the plasma membrane by an invading action potential opens these voltage-dependent Ca²⁺ channels to admit Ca²⁺ ions in the neighborhood of docked vesicles. The local high concentration of Ca²⁺ resulting from the opening of multiple Ca²⁺ channels triggers exocytosis. After exocytosis, some vesicles may rapidly reclose, but most fuse fully with the plasma membrane (Granstedt et al., 2006). The latter are recovered by endocytosis, a budding off of the vesicular membrane to form a new “coated” vesicle covered by the protein clathrin. Endocytosis may also be regulated by presynaptic [Ca²⁺] (Schweizer and Ryan, 2006). Recovered vesicular membrane often fuses to form large membranous sacs, called endosomes or cisternae, from which new synaptic vesicles are formed. The molecular mechanisms of the vesicle cycle of exo- and endocytosis are discussed later in this chapter.

**Transmitter Cycle**

The steps of the transmitter cycle vary with the type of transmitter (for additional details see Chapters 9 and 10). Some transmitters are synthesized from precursors in the cytoplasm before transport into synaptic vesicles, whereas other transmitters are synthesized in synaptic vesicles from transported precursors. Peptide transmitters are synthesized exclusively in the cell body and not locally recycled. At most synapses, a transporter that harnesses the energy in the proton gradient across the vesicular membrane functions to concentrate transmitter (or
FIGURE 8.2  Steps in the life cycle of synaptic vesicles: (1) Na\(^+\)-dependent uptake of transmitter (XMTR) or XMTR precursors into the cytoplasm, (2) synthesis of XMTR, (3) delivery of vesicle membrane containing specialized transmembrane proteins by axoplasmic transport on microtubules, (4) production of transvesicular H\(^+\) gradient by vacuolar ATPase, (5) concentration of XMTR in vesicles by H\(^+\)/XMTR antipporter, (6) synapsin I-dependent anchoring of vesicles to actin filaments near active zones, (7) releasable vesicles docked in active zones near Ca\(^{2+}\) channels, (8) depolarization of nerve terminal and presynaptic bouton by action potential, (9) opening of Ca\(^{2+}\) channels and formation of regions of local high [Ca\(^{2+}\)] (“Ca\(^{2+}\) microdomains”) in active zones, (10) triggering of exocytosis of docked vesicles comprising quantal units of XMTR released by overlapping Ca\(^{2+}\) microdomains, (11) nonquantal leakage of XMTR through vesicle membrane fused with plasma membrane and exposure of vesicle proteins to synaptic cleft, (12) recovery of vesicle membrane by dynamin-dependent endocytosis of clathrin-coated vesicles, (13) fusion of coated vesicles with endosomal cisternae, (14) formation of synaptic vesicles from endosomes. Also shown are postsynaptic receptors with multiple XMTR binding sites and extracellular XMTR-degradative enzymes in synaptic cleft.
transmitter precursors) in vesicles (Edwards, 1992). The pH gradient arises from the action of a vacuolar proton ATPase that uses the energy of ATP hydrolysis to transport protons into vesicles. After exocytosis, released transmitter diffuses across the synaptic cleft and rapidly binds to receptors. As transmitter falls off receptors, it is typically recovered from the synaptic cleft by sodium-dependent uptake transporters (see Chapter 9). At cholinergic synapses, acetylcholine is hydrolyzed to acetate and choline by the enzyme acetylcholinesterase present in the synaptic cleft. This enzyme is saturated by the initial gush of transmitter following exocytosis but can keep up with its subsequent slower release from receptors. The choline so produced is recovered by a presynaptic choline transporter and made available for the synthesis of new transmitter. Much of the evidence for the steps outlined in Fig. 8.2 comes from ultrastructural and pharmacological experiments. Some of this evidence is outlined in Box 8.2 (Chan et al., 1991; Cooper et al., 1991; Edwards et al., 1985; Elmquist and Quastel, 1965; Llina’s et al., 1991; Parsons et al., 1993; Van der Bliek and Meyerowitz, 1991) and Box 8.3 (Ryan et al., 1993).

**Summary**

Chemical synapses are ideally suited to permit one neuron to rapidly and effectively excite or inhibit the activity of another cell. A diversity of transmitters and receptors guarantees a multiplicity of postsynaptic responses. The opportunity for presynaptic and postsynaptic interactions between inputs provides for marvelously complex computational capabilities.

**BOX 8.2**

**EVIDENCE FOR SOME OF THE EVENTS IN THE LIFE HISTORY OF VESICLES**

Numbers refer to the steps in Figure 8.2.

1. Uptake of transmitter or transmitter precursors is prevented by specific inhibitors, such as hemicholinium-3 block of choline uptake at cholinergic synapses, ultimately leading to failure of synaptic transmission (Elmgvist and Quastel, 1965).

2. Cholinergic synapses can be identified by the presence of the synthetic enzyme choline acetyltransferase, GABAergic synapses by the enzyme glutamic acid decarboxylase, adrenergic synapses by the enzyme dopamine β-hydroxylase, and so forth (Cooper et al., 1991).

3. Vesicular transport into nerve terminals is blocked by inhibitors of axoplasmic transport such as antibodies to the microtubule motor protein kinesin (Bennett and Scheller, 1994; Jahn and Sündhof, 1994).

4. The storage of transmitter in vesicles can be blocked by inhibitors of vacuolar ATPase, such as bafilomycin A2, or of an H+-dependent transporter, such as vesamicol for acetylcholine (Parsons et al., 1993).

5. Dephosphorylation of synapsin I inhibits vesicle movements and transmission, whereas its phosphorylation by Ca2+-calmodulin-dependent kinase II protects against this inhibition (Llina’s et al., 1991; McGuiness et al., 1989).

6. Toxins from Clostridium bacteria, which proteolyze the vesicular protein synaptobrevin or plasma membrane proteins SNAP-25 and syntaxin, block exocytosis, whereas mutants deficient in the vesicle protein synaptotagmin and injection of peptides derived from synaptotagmin show defects in evoked transmitter release (more details later in this chapter).

7. Block of action potential propagation by local application of tetrodotoxin prevents transmission, and depolarization by elevating potassium in the bath accelerates mPSP frequency, as long as Ca2+ is present in the medium (Katz, 1969).

8. N- and P/Q-type calcium channel antagonists, such as α-conotoxin and α-agatoxin IVA, prevent Ca2+ influx and block transmission at many synapses (Wheeler et al., 1994; Dunlap et al., 1995).

9. Cholinergic synapses show a nonquantal leak of acetylcholine that is enhanced after stimulation; it is blocked by vesamicol, the vesicular acetylcholine transport inhibitor, indicating that the leak is due to transport through vesicular membrane fused with the plasma membrane (Edwards et al., 1985).

10. Cholinergic synapses show a nonquantal leak of acetylcholine that is enhanced after stimulation; it is blocked by vesamicol, the vesicular acetylcholine transport inhibitor, indicating that the leak is due to transport through vesicular membrane fused with the plasma membrane (Edwards et al., 1985).

11. Endocytosis is blocked at high temperature in the shibire mutant of Drosophila, which affects the protein dynamin in endocytosis of coated vesicles (Chan et al., 1991; Van der Bliek and Meyerowitz, 1991).

Morphological evidence for steps 13 and 14 in Figure 8.2 is given in Box 8.3.
The packaging of transmitter into vesicles and its release in quanta enable a single action potential to secrete hundreds of thousands of molecules of transmitter almost instantaneously at a synapse onto another cell. Neurochemical and ultrastructural studies have provided a rich picture of the life cycle of synaptic vesicles from their exocytosis at active zones to their recovery by endocytosis, their refilling with transmitter, and redocking at release sites.

**EXCITATION–SECRETION COUPLING**

Shortly after an action potential invades presynaptic terminals at fast synapses, the synchronous release of many quanta of transmitter generates the postsynaptic potential. Since the work of Locke in 1894, the presence of calcium in the external medium has been known to be a requirement for transmission. What is the central role of Ca$^{2+}$ in triggering neurosecretion?

**Calcium Triggers Release of Transmitters at Internal Sites**

Calcium was originally believed to act at an external site to enable neurons to release transmitter. The pioneering work of Bernard Katz (1969) and his co-workers showed that Ca$^{2+}$ acts intracellularly. This conclusion is based on many lines of evidence:

1. Calcium must be present only at the moment of invasion of the nerve terminal by an action potential for transmitter to be released.
2. Calcium entry is retarded by a large presynaptic depolarization, and transmitter release is delayed until the voltage gradient is reversed at the end of the pulse, whereupon Ca$^{2+}$ enters and release occurs as an off-EPSP until Ca$^{2+}$ channels close. Sodium influx is not necessary for secretion, and K$^+$ ions also play no role.
3. Elevation of intracellular [Ca$^{2+}$] accelerates the spontaneous release of quanta of transmitter (Steinbach and Stevens, 1976; Rahamimoff et al.,

---

**BOX 8.3**

**HISTOLOGICAL TRACERS CAN BE USED TO FOLLOW VESICLE RECYCLING**

An elegant picture of the life history of synaptic vesicles comes from studies using electron-dense or fluorescent markers of intracellular regions that have been in contact with the extracellular space. Horseradish peroxidase (HRP) is an enzyme that catalyzes the oxidation of diaminobenzidine, forming an electron-dense product that can easily be identified in tissues fixed with osmium tetroxide for electron microscopy; FM1–43 is an amphiphatic styryl dye that becomes highly fluorescent on partitioning into cell membranes. When frog muscles were soaked in HRP and the motor neurons were stimulated at 10 Hz for 1 min, the enzyme appeared in coated vesicles in nerve terminals in regions outside active zones. After more prolonged stimulation, most of the HRP collected in endosomal cisternae, owing to the fusion of endocytotic vesicles with these organelles. When the HRP was washed out and the neurons were rested for an hour before fixation, HRP appeared in small clear synaptic vesicles in active zones. When rested neurons were stimulated again before fixation, this time in the absence of HRP, the filled vesicles gradually disappeared owing to their release by exocytosis (Heuser and Reese, 1973).

Another study traced the uptake of FM1–43 into living motor nerve terminals with the use of confocal fluorescence microscopy. High-frequency stimulation for just 15 s in FM1–43 was marked by uptake of dye into nerve terminals. More prolonged stimulation followed by a period of rest without the dye in the bath resulted in the persistent staining of synaptic vesicles in active zones. Subsequent stimulation at 10 Hz gradually destained the terminals in minutes; destaining required the presence of Ca$^{2+}$ in the medium and represented exocytosis of stained vesicles. After about 1 min, the rate of destaining decreased as the vesicle pool began to be diluted with unstained vesicles newly recovered by endocytosis (Betz and Bewick, 1993). Exposing dissociated hippocampal neurons to FM1–43 at various times after stimulation showed that endocytosis proceeded for about 1 min after exocytosis. Cells loaded with dye and then restimulated began to destain about 30 s after endocytosis, which is a measure of the time needed for recycling of recovered vesicles into the pool of releasable vesicles (Ryan et al., 1993). These experiments provide a dynamic view of the life cycle of synaptic vesicles.
8. RELEASE OF NEUROTRANSMITTERS

1. The presence of Ca$^{2+}$ channels in presynaptic terminals is shown by the ability to stimulate local action potentials that trigger release in a high-[Ca$^{2+}$] medium when Na$^+$ action potentials are blocked with tetrodotoxin and K$^+$ channels are blocked with tetraethylammonium.

2. Action potentials trigger no phasic release of transmitter when Ca$^{2+}$ influx is blocked, even when presynaptic Ca$^{2+}$ is tonically elevated by photolysis of photosensitive Ca$^{2+}$ chelators; however, the elevated presynaptic [Ca$^{2+}$] accelerates the frequency of mEPSPs (Mulkey and Zucker, 1991).

3. Vesicles Are Released by Calcium Microdomains

Single action potentials generate a Ca$^{2+}$ rise of about 10 nM, which lasts a few seconds (Charlton et al., 1982; Zucker et al., 1991). This increment in [Ca$^{2+}$] is a small fraction of the typical resting [Ca$^{2+}$] of 100 nM. How can such a tiny change in [Ca$^{2+}$] trigger a massive synchronous release of quanta, and why is secretion so brief compared with the duration of the [Ca$^{2+}$] change? As mentioned earlier, postsynaptic responses begin only 0.5 ms after an action potential invades nerve terminals. This synaptic delay includes the time taken for Ca$^{2+}$ channels to begin to open after the peak of the action potential (300 µs) (Llinás et al., 1981), leaving only about 200 µs after that for transmitter secretion and the start of a postsynaptic response. At this time, Ca$^{2+}$ has barely begun to diffuse away from Ca$^{2+}$ channel mouths. In an aqueous solution, [Ca$^{2+}$] would be confined mainly to within 1 µm channel mouths estimated roughly from the solution of the diffusion equation for a brief influx of $M$ moles of Ca$^{2+}$,

$$[\text{Ca}^{2+}] = \frac{M}{8(\pi Dt)^{3/2}} e^{-r^2/4Dt}$$

where $t$ is time after the influx, $r$ is distance from the channel mouth, and $D$ is the diffusion constant for Ca$^{2+}$, $\sim 6 \times 10^{-6}$ cm$^2$ s$^{-1}$. In the cytoplasm, Ca$^{2+}$ diffusion is retarded by intracellular organelles and the presence of millimolar concentrations of fast-acting protein-associated Ca$^{2+}$ binding sites with an average dissociation constant of a few micromolar. Together, these effects restrict Ca$^{2+}$ microdomains to about 50 nm around channel mouths.

Furthermore, the 200 µs preceding the postsynaptic response must include not only the time required for Ca$^{2+}$ to reach its target but also the time required for Ca$^{2+}$ to bind and initiate exocytosis and for transmitter to diffuse across the synaptic cleft, bind to receptors, and begin to open channels. Thus, the presynaptic Ca$^{2+}$ targets must be located within a few tens of nanometers of Ca$^{2+}$ channel mouths. Neuromuscular junctions that are fast-frozen during the act of secretion show vesicle fusion images in freeze-fracture planes of the presynaptic membrane about 50 nm from intramembranous particles thought to be Ca$^{2+}$ channels (see Fig. 8.1C). Solution of the diffusion equation for a steady-point source of Ca$^{2+}$ influx in the presence of a nearly immobile fast-binding Ca$^{2+}$ buffer reveals that approximately 100 µs after a Ca$^{2+}$ channel opens, [Ca$^{2+}$] increases to more than 10 mM at 50 nm from its source and to more than 100 µM at a distance of 10 nm (Llinás et al., 1981).

This calculation considers only what happens in the neighborhood of a single open Ca$^{2+}$ channel. However, when individual Ca$^{2+}$ channels are labeled with biotinylated α-conotoxin tagged with colloidal gold particles, more than 100 channels per active zone are seen in terminals of chick parasympathetic ganglia (Haydon et al., 1994). Any vesicle docked at such an active zone is likely to be surrounded by as many as 10 Ca$^{2+}$ channels within a 50 nm distance. Even though not all these channels will open during each action potential, more than one channel is likely to open, so a vesicle will be influenced by Ca$^{2+}$ entering through several nearby channels. At the squid giant synapse, more than 50 channels open in each ~0.6
µm² active zone, whereas 10 channels open within the more compact active zones of frog saccular hair cells (Roberts, 1994; Yamada and Zucker, 1992). The Ca²⁺ microdomains of these channels overlap at single vesicles, and they cooperate in triggering secretion of a vesicle. Calculations of diffusion of Ca²⁺ ions from arrays of Ca²⁺ channels in the presence of a saturable buffer indicate that the [Ca²⁺] at sites where neurotransmitter release is triggered may reach 10–200 µM or higher (Fig. 8.3).

Three indications that [Ca²⁺] in fact reaches very high levels in active zones during action potentials are:

1. Highly localized [Ca²⁺] levels greater than 100 µM have been measured in presynaptic submembrane regions of squid giant synapses likely to be within active zones by using the low-affinity Ca²⁺-sensitive photoprotein n-aequorin-J (Llinás et al., 1992).

FIGURE 8.3 Microdomains with high Ca²⁺ concentrations form in the cytosol near open Ca²⁺ channels and trigger the exocytosis of synaptic vesicles. (A) In this adaptation of a model of Ca²⁺ dynamics in the terminal, a set of Ca²⁺ channels is spaced along the x-axis, as if in a cross section of a terminal. The channels have opened and, while they are open, the cytosolic Ca²⁺ concentration (y-axis) is spatially inhomogeneous. Near the mouth of the channel, the influx of Ca²⁺ drives the local concentration to as high as 800 µM, but within just 50 nm of the channel, the concentration drops off to about 25 µM. The channels are irregularly spaced but are often sufficiently close to one another that their clouds of Ca²⁺ can overlap and sum. (B) In the active zone (gray), an action potential has opened a fraction of the Ca²⁺ channels and microdomains of high cytosolic Ca²⁺ (pink) arise around these open channels as Ca²⁺ flows into the cell. In the rest of the cytoplasm, the Ca²⁺ concentration is at resting levels (0.10 µM), but within these microdomains, and particularly near the channel mouth, Ca²⁺ concentrations are much higher, as in (A). Synaptic vesicles docked and primed at the active zone may come under the influence of one or more of these microdomains and thereby be triggered to fuse with the membrane. (C) A few milliseconds after the action potential, the channels have closed and the microdomains have dispersed. The overall Ca²⁺ concentration in the terminal is now slightly higher (0.11 µM) than before the action potential. If no other action potentials occur, the cell will pump the extra Ca²⁺ out across the plasma membrane and restore the initial condition after several hundred milliseconds. (A) is adapted from Roberts, W.M. (1994): Localization of calcium signals by a mobile calcium buffer in frog saccular hair cells. J. Neurosci. 14, 3246-3262. The numerical details are for hair cell synapses only.
2. Estimates of \([\text{Ca}^{2+}]\) based on the activity of \(\text{Ca}^{2+}\)-activated \(K^+\) channels in active zones of mechanosensory hair cells are similar (Roberts, 1990).

3. At most rapidly transmitting synapses, transmitter release is blocked only by presynaptic injection of at least millimolar concentrations of fast high-affinity \(\text{Ca}^{2+}\) chelators, indicating that release is triggered locally by high concentrations of \(\text{Ca}^{2+}\) (Adler et al., 1991).

Vesicle Exocytosis Is Normally Triggered by Overlapping \(\text{Ca}^{2+}\) Channel Microdomains of High \(\text{Ca}^{2+}\)

Although release of a quantum of transmitter subsequent to the opening of a single presynaptic \(\text{Ca}^{2+}\) channel has been observed (Stanley, 1993), exocytosis may normally be due to \(\text{Ca}^{2+}\) entering through clusters of \(\text{Ca}^{2+}\) channels in active zones and contributing to local high \(\text{Ca}^{2+}\) at docked vesicles:

1. When transmitter release is increased under voltage clamp with pulses of increasing amplitude, a third-order power law relationship exists between presynaptic \(\text{Ca}^{2+}\) current and postsynaptic response (Augustine and Charlton, 1986). If each vesicle were released by \(\text{Ca}^{2+}\) entering through a single \(\text{Ca}^{2+}\) channel, then increasing depolarizations would recruit additional channel openings and proportionally more vesicle releases (Simon and Llinás, 1985). However, if \(\text{Ca}^{2+}\) channel microdomains from neighboring clustered \(\text{Ca}^{2+}\) channels overlap at docked vesicles, the \(\text{[Ca}^{2+}\) at each vesicle will rise with increasing depolarization as more channels are recruited, and some cooperativity of \(\text{Ca}^{2+}\) action in triggering secretion will be expressed (Zucker and Fogelson, 1986).

2. In some neurons, more than one \(\text{Ca}^{2+}\) channel type contributes to secretion (Dunlap et al., 1995; Wheeler et al., 1994). When contributions of each channel type are isolated pharmacologically, their combined effects add nonlinearly, much as would be predicted by a fourth-order cooperativity, indicating that the \(\text{Ca}^{2+}\) microdomains of different channels overlap and summate at the \(\text{Ca}^{2+}\) sensors at vesicles docked within individual active zones.

3. At brain stem calyx of Held synapses (Borst and Sakmann, 1996), presynaptic injection of the slow-acting \(\text{Ca}^{2+}\) buffer ethylene glycol bis (b-aminoethyl ether)-\(N,N,N',N'\)-tetraacetic acid (EGTA) reduces transmission, indicating that the target for \(\text{Ca}^{2+}\) action is not particularly close to any one \(\text{Ca}^{2+}\) channel, but rather is affected by \(\text{Ca}^{2+}\) ions entering through many channels (Meinrenken et al., 2002).

4. When transmitter release is increased by prolonging presynaptic depolarizations (e.g., by broadening action potentials with \(K^+\) channel blockers), more channels are not likely to be opened simultaneously. Rather, as channels that open early in the action potential close, others open; so the pattern of presynaptic \(\text{Ca}^{2+}\) microdomains is not so much intensified as prolonged, leading to a more nearly linear relationship between increases in \(\text{Ca}^{2+}\) influx and transmitter release (Zucker et al., 1991).

5. Large depolarizations admit little \(\text{Ca}^{2+}\) as they approach the \(\text{Ca}^{2+}\) equilibrium potential; they are therefore accompanied by a reduced \(\text{Ca}^{2+}\) current and reduced transmitter release during a pulse. However, large depolarizations can release more transmitter than can small depolarizations evoking a given macroscopic \(\text{Ca}^{2+}\) current (Augustine et al., 1985; Llinás et al., 1981). This apparent voltage dependence of transmitter release may be due to the different spatial profiles of \(\text{[Ca}^{2+}\) in the active zone, with greater overlap of \(\text{[Ca}^{2+}\) from the larger number of more closely apposed open \(\text{Ca}^{2+}\) channels during large depolarizations (Zucker and Fogelson, 1986).

The Exocytosis Trigger Must Have Fast, Low-Affinity, Cooperative \(\text{Ca}^{2+}\) Binding

The brevity of the synaptic delay implies not only that \(\text{Ca}^{2+}\) acts near \(\text{Ca}^{2+}\) channels to evoke exocytosis but also that \(\text{Ca}^{2+}\) must bind to its receptor extremely rapidly. This is confirmed by the finding that presynaptic injection of relatively slow \(\text{Ca}^{2+}\) buffers such as EGTA have almost no effect on transmitter release to single action potentials. Only millimolar concentrations of fast \(\text{Ca}^{2+}\) buffers such as 1,2-bis(2-amino-phenoxy)ethane-\(N,N,N',N'\)-tetraacetic acid (BAPTA), with on-rates of about \(5 \times 10^8 \text{M}^{-1} \text{s}^{-1}\), can capture \(\text{Ca}^{2+}\) ions before they bind to the secretory trigger (Adler et al., 1991), indicating that the on-rate of \(\text{Ca}^{2+}\) binding to this trigger is similarly fast. At a rate of \(5 \times 10^8 \text{M}^{-1} \text{s}^{-1}\), 10–100 \(\mu\text{M} [\text{Ca}^{2+}]\) reach equilibrium with its target in about 50–500 \(\mu\text{s}\).

From the dependence of transmitter release on external \(\text{[Ca}^{2+}\), it is known that at least four \(\text{Ca}^{2+}\)
ions cooperate in the release of a vesicle. The off-rate of Ca\(^{2+}\) dissociation from these sites also must be fast, at least $10^3$ s\(^{-1}\), to account for the rapid termination of transmitter release ($0.25$ ms time constant) after Ca\(^{2+}\) channels close and Ca\(^{2+}\) microdomains collapse. The high temperature sensitivity of the time course of transmitter release ($Q_{10} \approx 3$) indicates that exocytosis is rate limited by a step with a high-energy barrier (Yamada and Zucker, 1992). This step is likely to be the process of exocytosis itself. If Ca\(^{2+}\) binding is not rate limiting, its dissociation rate must be substantially faster than $10^3$ s\(^{-1}\). This means that the affinity of the secretory trigger for Ca\(^{2+}\) is low, with a dissociation constant ($K_D$) above $10$ $\mu$M.

The Ca\(^{2+}\)-binding trigger is not saturated under normal conditions, because increasing [Ca\(^{2+}\)] in the bath increases release. Furthermore, because of the speed with which Ca\(^{2+}\) binds to its sites, this reaction nearly equilibrates during the typical $0.5$–$1.0$ ms that [Ca\(^{2+}\)] remains high before Ca\(^{2+}\) channels close at the end of an action potential. If [Ca\(^{2+}\)] reaches 10–100 $\mu$M or more in equilibrium with unsaturated release sites, the affinity of at least some of those sites binding Ca\(^{2+}\) must be $\sim 10$–$100$ $\mu$M or lower.

These predictions are consistent with experiments in which neurosecretion is triggered by photolysis of caged Ca\(^{2+}\) chelators such as DM-nitrophen. Partial photolysis of partially Ca\(^{2+}\)-loaded DM-nitrophen (Landó and Zucker, 1994) or addition of other Ca\(^{2+}\) buffers (Bollmann and Sakmann, 2005) generates a [Ca\(^{2+}\)] “spike” of a duration similar to the lifetime of Ca\(^{2+}\) microdomains around Ca\(^{2+}\) channels opened by an action potential. This spike results in a postsynaptic response that closely resembles the normal EPSC at crayfish neuromuscular junctions, confirming that no presynaptic depolarization is necessary to obtain high levels of phasic transmitter release. Secretion depended on the fourth power of peak [Ca\(^{2+}\)], and about 10–25 $\mu$M Ca\(^{2+}\) activated release at the same rate as an action potential.

In similar experiments on retinal bipolar neurons from fish (Heidelberger et al., 1994), fully loaded DM-nitrophen was photolysed to produce a stepped increase in [Ca\(^{2+}\)] while secretion was monitored as an increase in membrane capacitance, a measure of cell membrane area increased by fusion of vesicles. The Ca\(^{2+}\) concentration had to be raised by more than 20 $\mu$M before a fast phase of secretion developed. The sharp Ca\(^{2+}\) dependence of release and short synaptic delays were fitted by a model with a high degree of positive Ca\(^{2+}\) cooperativity, in which four successive Ca\(^{2+}\) ions bind with affinities increasing (or $K_D$ decreasing) from 140 to 9 $\mu$M, followed by a Ca\(^{2+}\)-independent rate-limiting step. In contrast, at the calyx of Held synapse, a [Ca\(^{2+}\)] level of about 10 $\mu$M was sufficient to activate release at a rate similar to that in an EPSP (Schneggenburger and Neher, 2000), possibly without positive cooperativity (Bollmann et al., 2000). These experiments differentiate Ca\(^{2+}\) receptors triggering release at various synapses.

### Calcium Ions Must Mobilize Vesicles to Docking Sites at Slowly Transmitting Synapses

Most peptidergic synapses and some synapses releasing biogenic amines display kinetics remarkably different from those of fast synapses. In these slower synapses, single action potentials often have no discernible postsynaptic effect. During repetitive stimulation, postsynaptic responses rise slowly, often with a delay of seconds from the beginning of stimulation, and persist just as long after stimulation ceases. Such slow responses are due to many factors: the postsynaptic receptors may have intrinsically sluggish second messengers or G proteins (Chapters 11 and 16); the postsynaptic receptors are often distant from release sites, so extracellular diffusion takes significant time; and release starts after the beginning of stimulation and continues after stimulation stops. Given these limitations, it is not surprising that single quanta are never discernible, either as spontaneous PSPs or as components of evoked responses.

The ultrastructural anatomy of presynaptic terminals of slowly transmitting synapses also is different from that of fast synapses. Transmitter is stored in large, dense core vesicles scattered randomly throughout the cytoplasm; vesicles do not tend to cluster at active zones or line up at the membrane, docked and ready for release (De Camilli and John, 1990; Leenders et al., 1999). Nevertheless, there is no doubt that transmitter is released from vesicles, because it is both stored and often synthesized in them, and they can be seen to undergo exocytosis during high-frequency stimulation causing high rates of release (Verhage et al., 1994).

### A High-Affinity Calcium Binding Step Controls Secretion of Slow Transmitters

Calcium ions are required for excitation–secretion coupling in slow synapses, but the dependence of release on [Ca\(^{2+}\)] is linear, in contrast with fast synapses (Sakaguchi et al., 1991). Furthermore, because few vesicles are pre docked at active zones, most of those released by repetitive activity are not exposed to the local high [Ca\(^{2+}\)] near Ca\(^{2+}\) channels. Thus, an important event triggered by Ca\(^{2+}\) influx in action potentials is likely to be the translocation of dense core vesicles to
plasma membrane release sites, followed by exocytosis. This process has a very different dependence on \([\text{Ca}^{2+}]\) than does the release of docked vesicles. Measurements of \([\text{Ca}^{2+}]\) during stimulation indicate that release correlates well with \([\text{Ca}^{2+}]\) levels in the low micromolar range above a minimum, or threshold, level of a few hundred nanomolar (Lindau et al., 1992; Peng and Zucker, 1993). Such a high affinity docking process may also be the rate-limiting step for secretion from some tonic ribbon synapses, such as photoreceptors (Thoreson et al., 2004).

A striking difference between the release of fast transmitters, such as \(\gamma\)-aminobutyric acid (GABA) and glutamate, and peptide transmitters, such as cholecystokinin, was found in studies of synaptosomes, isolated nerve terminals prepared from homogenized brain tissue by differential centrifugation (Verhage et al., 1991). When terminals were depolarized to admit \(\text{Ca}^{2+}\) through \(\text{Ca}^{2+}\) channels, the amino acid transmitters GABA and glutamate were released at much lower levels of bulk cytoplasmic \([\text{Ca}^{2+}]\) than when \(\text{Ca}^{2+}\) was admitted more uniformly and gradually across the membrane by use of the \(\text{Ca}^{2+}\)-transporting ionophore ionomycin. Peptides were released at the same low levels of \([\text{Ca}^{2+}]\) no matter which method was used to elevate \([\text{Ca}^{2+}]\). Thus, only amino acids were sensitive to the difference in \([\text{Ca}^{2+}]\) gradients imposed by the two methods and were preferentially released by local high submembrane \([\text{Ca}^{2+}]\) caused by depolarization. Apparently, peptides are released by a high-affinity rate-limiting step not especially sensitive to submembrane \([\text{Ca}^{2+}]\) levels.

**Slow and Fast Transmitters May Be Co-Released from the Same Neuron Terminal**

Some neurons have both small synaptic vesicles containing acetylcholine or glutamate and large, dense core vesicles containing neuropeptides (Lundberg and Hökfelt, 1986). Often, the two transmitters act on different targets. Single action potentials release only the fast transmitter, so different patterns of activity can have very different relative effects on the targets. For example, postganglionic parasympathetic nerves to the salivary gland release acetylcholine, which stimulates salivation, and vasoactive intestinal peptide, which stimulates vasodilation. Many examples of the corelease of multiple transmitters have been described.

**Summary**

\(\text{Ca}^{2+}\) acts as an intracellular messenger tying the electrical signal of presynaptic depolarization to the act of neurosecretion. At fast synapses, \(\text{Ca}^{2+}\) enters through clusters of channels near docked synaptic vesicles in active zones. It acts at extremely short distances (tens of nanometers) in remarkably little time (200 µs) and at very high local concentrations (10–100 µM), in calcium microdomains, by binding cooperatively to a low-affinity receptor with fast kinetics to trigger exocytosis. Some transmitters, such as peptides and some biogenic amines, are stored in larger, dense core vesicles not docked at the plasma membrane in active zones. Release of these transmitters, as well as their diffusion to postsynaptic targets and their postsynaptic actions, is much slower than that of transmitters such as acetylcholine and amino acids at fast synapses. Release of slow transmitters depends linearly on \([\text{Ca}^{2+}]\) and may be governed by a \(\text{Ca}^{2+}\)-sensitive rate-limiting step different from that triggering exocytosis of docked vesicles at fast synapses.

**THE MOLECULAR MECHANISMS OF THE NERVE TERMINAL**

To release neurotransmitter in response to an action potential, a synaptic vesicle must fuse with the plasma membrane with great rapidity and fidelity and thus the synapse requires an effective and well-regulated molecular machine. The mechanisms of the terminal must also include the means to load the vesicle with transmitter, to dock the vesicle near the membrane so that it can fuse with a short latency, to define a release site on the plasma membrane, and to restrict fusion to the active zone rather than other points on the surface of the terminal or axon. Additionally, a reserve of synaptic vesicles must be held near the active zone and those vesicles must be recruited to the plasma membrane as needed. The number of vesicles that are ready and waiting to fuse must be strictly determined, and the protein and lipid components of the vesicle must be recycled to form a new vesicle after fusion has occurred. For each of these processes, a molecular understanding remains incomplete, but rapid scientific progress in this field has made considerable headway.

**A Cycle of Membrane Trafficking**

Active neurons need to secrete transmitter in a constant, ongoing fashion. A bouton in the CNS, for example, may contain a store of 200 vesicles, but if it releases even one of these with each action
potential and if the cell is firing at an unexceptional rate such as 5 Hz, the store of vesicles would be consumed within less than a minute. This calculation exemplifies the need for an efficient mechanism to recycle and reload vesicles within the terminal. Transport of newly synthesized vesicles from the cell body would be far too slow to support such a demand, and the axonal traffic needed to supply an entire arbor of nerve terminals would be staggering. As a consequence, only peptide neurotransmitters are supplied in this fashion, because they must be synthesized in the endoplasmic reticulum and then sent down the axon. Not surprisingly, therefore, vesicles with peptide transmitters are released at very low rates. For most neurotransmitters, however, the exocytosis of a synaptic vesicle is rapidly followed by its endocytosis and within approximately 30 s the vesicle is again available for release (Ryan and Smith, 1995). This pathway is sometimes referred to as the exo–endo-cytic cycle (Fig. 8.2). Moreover, each step in this cycle represents a potential control point for modulating the efficacy of the synapse. Modulation of the strength or fidelity of synaptic signaling, commonly known as synaptic plasticity, plays an important role both in the development of synaptic connections and in the functioning of the mature nervous system (see also Chapter 19). Indeed, regulation by Ca\(^{2+}\) and other second messenger systems is known to affect the docking, fusing, and recycling of vesicles at some synapses and is also likely to regulate the balance between reserve stores and those vesicles actively engaged in the exo–endo-cytic cycle (Beutner et al., 2001; Dinkelacker et al., 2000; Neher and Zucker, 1993; Smith et al., 1998; Wierda et al., 2007). Understanding the mechanisms of this modulation is an important goal and will certainly require a detailed understanding of the fundamental machinery itself. The question has been approached through the combined use of biochemical, genetic, and biophysical techniques.

**Transmitter Release Is Rapid**

As discussed earlier in this chapter, the delay between the arrival of an action potential at a terminal and the secretion of the transmitter can be less than 200 \(\mu\)s. This places some severe constraints on the fusion mechanism. Vesicles must already be present at the release sites, as there is no time to mobilize them from a distance. A catalytic cascade during fusion, such as that involved in phototransduction or in excitation–contraction coupling in smooth muscle, would also be far too slow for excitation–secretion coupling at the nerve terminal. Indeed, even a single bimolecular catalytic step might be too slow for such short latencies. Models therefore favor the idea that a fusion-ready complex of the vesicle and plasma membrane is preassembled at release sites and that Ca\(^{2+}\) binding need only trigger a simple conformation change in this complex to open a pathway for the transmitter to exit the vesicle. Because the volume of the synaptic vesicle is small, the diffusion of transmitter from the vesicle proceeds almost instantaneously as soon as a pore has opened between the vesicle lumen and the extracellular space. This structure is referred to as the fusion pore, but its biochemical nature is unknown. Thus, the time-critical step comes between the influx of Ca\(^{2+}\) and the formation of the fusion pore. The complete merging of the vesicle and plasma membrane, if it occurs at all, can occur on a slower time course. The movement of the vesicle to the release site and any biochemical events that need to occur to reach the fusion-ready state can also be slower. These largely theoretical steps are sometimes referred to as “docking” and “priming.” Docking and priming a vesicle cannot be too slow, however; a CNS synapse is estimated to have 2 to 20 vesicles in this fusion-ready state (Harris and Sultan, 1995; Stevens and Tsujimoto, 1995), and, therefore, if a synapse is to respond faithfully to a sustained train of action potentials, it must be able to replace the fusion-ready vesicles with a time course of seconds. The rate at which this occurs may determine some of the dynamic properties of the synapse.

The short latency of transmission would seem to preclude the involvement of ATP hydrolysis at the fusion step. This is born out by numerous physiological studies in which exocytosis persists after Mg\(^{2+}\)-ATP has been dialyzed from the cell and in which all the Mg\(^{2+}\) has been chelated (which would render any residual ATP inert to most enzymes) (Ahnert-Hilger et al., 1985; Hay and Martin, 1992; Holz et al., 1989; Parsons et al., 1995). Thus, if energy is needed to fuse the membranes, it should be stored in the fusion-ready state of the vesicle–membrane complex and released on addition of Ca\(^{2+}\).

**Transmitter Release Is a Cell Biological Question**

Exocytosis and endocytosis are not unique to neurons; these processes go on in every eukaryotic cell. Moreover, exocytosis itself is only one representative of a general class of membrane-trafficking steps in which one membrane-bound compartment must fuse with another. Other examples would include the fusion of recycling vesicles with endosomes, transport from the endoplasmic reticulum (ER) to the Golgi, or transport
from endosomal compartments to lysosomes. In each case, the same biophysical problem must be overcome, and the mechanisms for all these membrane fusion steps appear to have much in common.

For a vesicle to fuse with the plasma membrane, or any other target, there is a large energy barrier to surmount. To bring the lipid bilayers within a few nanometers of one another so that they can fuse, the hydration shell around the polar lipid head groups must be disrupted. Simply to split open each membrane so that the bilayer of one could be connected to the bilayer of the other would require exposing the hydrophobic core of each membrane to the aqueous milieu of the cytoplasm, and this barrier is sufficiently great that it does not occur under normal conditions. Thus a specialized mechanism is required to bring the membranes close together and then drive fusion. At present, it appears that all the membrane-trafficking steps within the cell use a similar set of proteins to accomplish this task (Fig. 8.4). As is discussed later, homologues of proteins found at the synapse and known to be essential for exocytosis have also been shown to function in ER-to-Golgi transport in mammalian cells, in endosomal fusion, and in exocytosis in yeast. Indeed, it appears that representatives of this core set of proteins are present.
on every trafficking vesicle or target membrane and required for every fusion step (Chen and Scheller, 2001). This discovery, which grew from the conjunction of independent studies of different model systems (Bennett and Scheller, 1993; Fischer von Mollard et al., 1991), has led to the exciting hypothesis that all intracellular membrane fusions will be united by a single and universal mechanism. To the extent that this proves true, it will be a great boon to cell biology: experiments in one model system, for example, the highly developed genetic analysis of membrane fusion in yeast, can be absorbed into neuroscience. Similarly, the abundance of synaptic vesicles for biochemical analysis and the unparalleled precision of electrophysiological assays of single vesicle fusions can deepen the understanding of other cellular events. Will these disparate membrane fusion events be truly identical in their mechanisms? The jury is still out. Most likely, however, different membrane fusions will present variations on a common theme; the fundamental processes will be adapted to the specific requirements of each physiological step. Perhaps an instructive analogy may be drawn with muscle contraction (Leavis and Gergely, 1984): actin and myosin offer a fundamentally conserved mechanism for producing force and movement, and every muscle contains isoforms of actin, heavy and light myosin chains, troponins, and tropomyosins. But skeletal muscle and smooth muscle differ enormously in the details of how Ca\(^{2+}\) interacts with these proteins and regulates their activity. Regulation via troponins or via myosin light chain kinase, differences in the intrinsic ATPase rates of the myosins, and other critical features have adapted these related sets of proteins to their particular purpose in the individual type of muscle. Similarly, the need for extremely tight control of fusion in the nerve terminal—for rapid rates of fusion with very short latencies to occur in brief bursts at precise points on the plasma membrane—may cause some profound differences in how the synaptic isoforms of these proteins function compared with the isoforms involved in general cellular traffic.

**Identifying the Synaptic Proteins: Vesicle Purification**

The most important starting point in the elucidation of the machinery of the nerve terminal was the observation that synaptic vesicles can be purified. Synaptic vesicles are abundant in nervous tissue (Fig. 8.5) and, owing to their unique physical properties (uniform small diameter and low buoyant density), can be purified to homogeneity by simple subcellular fractionation techniques (Carlson et al., 1978; Nagy et al., 1976). As a result, at a biochemical level, synaptic vesicles are among the most thoroughly characterized organelles. One of the first sources for the purification of synaptic vesicles was the electric organ of marine elasmobranchs. This structure, a specialized adaptation of the neuromuscular junction, is highly enriched in synaptic vesicles. It has also proven simple to isolate synaptic vesicles from mammalian brain. The protein compositions of synaptic vesicles from these sources are remarkably similar, demonstrating the evolutionary conservation of synaptic vesicle function. This similarity also points to the fact that many of the proteins present on the synaptic vesicle membrane perform general functions that are not restricted to a single class of transmitter.

Our knowledge of many of the important proteins in vesicle fusion commenced with the purification and characterization of synaptic vesicles. These vesicles contained a discrete set of abundant proteins, and individual proteins could be isolated and subsequently cloned. This analysis has recently been elevated to a more comprehensive level by the application of mass spectrometry to purified vesicles (Takamori et al., 2006). The major constituents of the synaptic vesicle are shown in Fig. 8.4. For some of these proteins there are well-established functions, but for others the functions remain uncertain (Table 8.1). The proton transporter, for example, is an ATPase that acidifies the lumen of the vesicle. The resulting proton gradient provides the energy by which transmitter is moved into the vesicle. This task is carried out by another vesicular protein, a vesicular transporter that allows protons to move down their electrochemical gradient (leaving the vesicle) in exchange for moving the transmitter into the lumen. Several vesicular transporters, all structurally related, are known, and by their substrate selectivity they help to specialize a terminal for release of the appropriate transmitter (see later). Two other large proteins with multiple transmembrane domains are abundant in vesicles: SV2 (synaptic vesicle protein 2) and synaptophysin. The functional significance of these proteins remains elusive despite extensive biochemical and genetic characterization (Bajjalieh et al., 1992; Brose and Rosenmund, 1999; Buckley et al., 1997; Crowder et al., 1999; Custer et al., 2006; Feany et al., 1992; McMahon et al., 1996; Südhof et al., 1987). Other vesicular proteins are discussed later in this chapter.

Transmitter release depends on more than just vesicular proteins; proteins of the plasma membrane and cytoplasm are also important. In many cases, the identification of these additional components or an appreciation of their importance to the synapse...
FIGURE 8.5  Structure of a synapse between a parallel fiber and a Purkinje cell spine (S) in the cerebellum. The sample was rapidly frozen and then freeze-fractured, shallow-etched, and rotary-shadowed to reveal the details of the synaptic architecture. Inset shows synaptic vesicles (V), and arrows indicate actin attachments. From Landis et al. (1988), Neuron 1: 201–209. Copyright © by Cell Press.
derived from investigations of vesicular proteins. Two examples serve to illustrate this point. The first example begins with synaptotagmin, an integral membrane protein of the synaptic vesicle that was purified from synaptic vesicles and cloned (Bixby and Reichardt, 1985; Perin et al., 1990). The portion of synaptotagmin that extends into the cytoplasm (the majority of the protein; see Fig. 8.4) was subsequently used for affinity column chromatography. In this manner, a protein called syntaxin was identified as a synaptotagmin-binding protein (Bennett et al., 1992). This protein resides in the plasma membrane and is now appreciated as one of the critical players in vesicle fusion (see later). More recently, biochemical experiments have identified a new protein, called complexin, that binds to a syntaxin-containing complex. Complexin, as discussed later, appears to have a crucial function in regulating the fusion of vesicles (Giraudo et al., 2006; McMahan et al., 1995). Thus, subsequent to the isolation of an abundant vesicular protein, biochemical assays have led to a fuller picture. A second and similar example began with the realization that an abundant small GTP-binding protein, rab3, was present on the vesicle surface (Südhof et al., 1987). This protein, discussed further later, may regulate vesicle availability or docking to release sites. Subsequently, rabphilin was identified on the basis of its affinity for the GTP-bound form of rab3 (Shirataki et al., 1993). Rabphilin lacks a transmembrane domain but is recruited to the surface of the synaptic vesicle by binding to rab3. Rabphilin may have a role in modulating transmission, particularly in mossy fiber terminals of the hippocampus (Lonart and Südhof, 1998). Subsequently, many additional rab3-binding proteins have been identified, including RIM (rab3-interacting molecule), a component of the active zone (Dresbach et al., 2001; Wang et al., 1997). Rab3 can exist in either GTP- or GDP-bound states, and additional factors that regulate these states were also identified: a GDP dissociation inhibitor (GDI), GDP/GTP exchange protein (GEP), and GTPase-activating protein (GAP) are found in the synaptic cytosol (Südhof, 1997). Thus, from the identification of a synaptic vesicle protein, an array of additional factors have come to light, all of which are likely to figure in the exocytotic cycle.

**Genetics Identifies Synaptic Proteins**

Genetic screens have provided an independent method for identifying the machinery of transmitter release. One of the most fertile screens was carried out not in the nervous system per se, but rather in yeast. Because membrane trafficking in yeast is closely parallel to vesicle fusion at the terminal, mutations that alter the secretion of enzymes from yeast can be a springboard for the identification of synaptic proteins. In the early 1980s a series of such screens were carried out (Novick et al., 1980, 1981) and a collection of more than 50 mutants was obtained. In many of these mutants, post-Golgi vesicles accumulated in the cytoplasm and thus the mutation appeared to block a late stage of transport, such as the targeting or fusion of these vesicles at the plasma membrane. Screens for suppressors and enhancers of these secretion mutations uncovered further components (Aalto et al., 1993). Subsequently, excellent in vitro assays have been established in which to study the fusion of vesicles derived from yeast with their target organelles (Conradt et al., 1994; Mayer et al., 1996; Wickner and Haas, 2000). Among the secretion mutants and their interacting genes were homologues of some of the proteins discussed previously: sec4 encodes a small GTP-binding protein like rab3, and Sso1 and Sso2 encode plasma membrane proteins that are homologues of syntaxin (Aalto et al., 1993; Salminen and Novick, 1987). The sec1 gene encodes a soluble protein with a very high affinity for Sso1, and the mammalian homologue of this protein, n-sec1, is tightly bound to syntaxin in nerve terminals and has an essential function in transmission (Aalto et al., 1991; Pevsner et al., 1994). Yeast is
TABLE 8.2 Additional Proteins Implicated in Transmitter Release

<table>
<thead>
<tr>
<th>Protein</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntaxin</td>
<td>SNARE protein present on plasma membrane (and on synaptic vesicles to a lesser extent); forms core complex with SNAP-25 and VAMP/synaptobrevin; essential for late step in fusion</td>
</tr>
<tr>
<td>SNAP-25</td>
<td>SNARE protein present on plasma membrane (and on synaptic vesicles to a lesser extent); forms core complex with syntaxin and VAMP/synaptobrevin; essential for late step in fusion</td>
</tr>
<tr>
<td>Nsec-1/munc-18</td>
<td>Syntaxin-binding protein required for all membrane traffic to the cell surface; likely bound to syntaxin before and after formation of SNARE complex</td>
</tr>
<tr>
<td>Complexin</td>
<td>SNARE-binding protein; inhibits vesicle fusion(?)</td>
</tr>
<tr>
<td>Tomosyn</td>
<td>SNARE-binding protein; inhibits vesicle fusion</td>
</tr>
<tr>
<td>Snapin</td>
<td>Binds SNAP-25; associated with synaptic vesicles; unknown function</td>
</tr>
<tr>
<td>NSF</td>
<td>ATPase that can disassemble the SNARE complex; likely to disrupt complexes after exocytosis</td>
</tr>
<tr>
<td>α-SNAP</td>
<td>Cofactor for NSF in SNARE complex disassembly</td>
</tr>
<tr>
<td>unc-13/munc-13</td>
<td>Active zone protein; vesicle priming for release; modulation of transmission</td>
</tr>
<tr>
<td>Rabphilin</td>
<td>C2 domain protein; Ca(^{2+})-binding protein; binds rab3 and associates with synaptic vesicle; modulation of transmission</td>
</tr>
<tr>
<td>DOC2</td>
<td>C2 domain protein; Ca(^{2+})-binding protein; binds munc-18; unknown function</td>
</tr>
<tr>
<td>RIM1 and related proteins</td>
<td>Active zone proteins; bind rab3; modulation of transmission</td>
</tr>
<tr>
<td>Piccolo</td>
<td>Likely scaffolding protein to tether vesicles near active zone</td>
</tr>
<tr>
<td>Bassoon</td>
<td>Likely scaffolding protein to tether vesicles near active zone</td>
</tr>
<tr>
<td>Exocyst (sec6/8 complex)</td>
<td>Marks plasma membrane sites of vesicle fusion in yeast; synaptic role uncertain</td>
</tr>
</tbody>
</table>

The Mechanism of Membrane Fusion: SNAREs and the Core Complex

Three synaptic proteins, VAMP (vesicle-associated membrane protein)/synaptobrevin, syntaxin, and SNAP-25 (synaptosome-associated protein of 25 kDa), are capable of forming an exceptionally tight complex that is generally referred to as either the core complex or SNARE complex (Söllner et al., 1993; Sutton et al., 1998). The interaction of these three is essential for synaptic transmission and likely to lie very close to or indeed at the final fusion step of exocytosis (Fig. 8.6 and Fig. 8.7). What are these proteins? VAMP (also called synaptobrevin) was among the first synaptic vesicle proteins to be cloned (Trimble et al., 1988). It is anchored to the synaptic vesicle by a single transmembrane domain and has a cytoplasmic domain that contributes a coiled-coil strand to the core complex. Syntaxin has a very similar structure but is located primarily in the plasma membrane (though some is present on vesicles as well) (Bennett et al., 1992). SNAP-25 is also a protein primarily of the plasma membrane but, unlike the others, lacks a transmembrane domain and is instead anchored in its central region by acylations (Chapman et al., 1994). SNAP-25 contributes two strands to the SNARE complex. Thus the interactions of these proteins can be envisioned as bringing closely together the two membranes that are meant to fuse. The proteins of this complex are archetypes of a class of membrane-trafficking proteins collectively called the SNAREs. The vesicle-associated proteins, such as VAMP, are referred to as v-SNAREs, and those of the target membrane, such as SNAP-25 and syntaxin, are referred to as t-SNAREs.

A SNARE complex is found at each membrane-trafficking step within a eukaryotic cell (Fig. 8.7).
FIGURE 8.6  SNARE proteins and the action of the clostridial neurotoxins. The SNARE complex shown on the left brings the vesicle and plasma membranes into close proximity and likely represents one of the last steps in vesicle fusion. Vesicular VAMP, also called *synaptobrevin*, binds with the syntaxin and SNAP-25 that are anchored to the plasma membrane. Transmitter release can be blocked by tetanus toxin and the botulinum toxins—proteases that cleave specific SNARE proteins.

FIGURE 8.7  Neurotransmitter release probably shares a core mechanism with all membrane fusion events within eukaryotic cells. The fusion of synaptic vesicles (A) is driven by a particular complex of four coiled-coil domains contributed by three different proteins. Exocytosis in yeast (B), the fusion of late endosomes in mammalian cells (C), and the fusion of vacuolar vesicles in yeast (D) exemplify the closely related four-stranded coiled-coil complexes that are required to drive fusion in other membrane-trafficking steps.
Through a combination of SNARE proteins on the opposing membranes, a four-stranded coiled coil is formed. Alongside the example of the synapse in Fig. 8.7A, three analogous cases are shown from exocytosis in yeast (Fig. 8.7B), the fusion of late endosomes with one another (Fig. 8.7C) (Antonin et al., 2000), and the fusion of vesicles that form the yeast vacuole (Fig. 8.7D) (Sato et al., 2000; Wickner and Haas, 2000). Abundant genetic and biochemical data argue for an essential role for SNAREs, and the combination of yeast genetics, in vitro assays, synaptic biochemistry, and synaptic physiology has had a synergistic effect in advancing the field. It was the discovery that secretory mutations in yeast were homologues to the synaptic SNAREs that provided some of the first functional data on these proteins (Bennett and Scheller, 1993), while data on synapses first placed the proteins on vesicles and the plasma membrane (Bennett et al., 1992; Chapman et al., 1994; Trimble et al., 1988). Exploring the mechanism of the SNAREs similarly brought together neurons and yeast. The crystal structure, for example, was solved first for the synaptic proteins (Sutton et al., 1998). For crystallization, however, the transmembrane domains and, of course, the membranes themselves were absent. Thus it is not possible to conclude from the structure alone that the SNAREs formed a bridge between the membranes rather than complexes within the same membrane. In an in vitro assay with purified yeast vesicles, however, it was possible to establish this point (Nichols et al., 1997).

Some of the strongest evidence for an essential role for SNAREs at the synapse has come from the study of a potent set of eight neurotoxins produced by clostridial bacteria. These toxins (tetanus toxin and the family of related botulinum toxins) have long been known to block the release of neurotransmitter from the terminal. The discovery that they do so by proteolytically cleaving individual members of the SNARE complex (Link et al., 1992; Schiavo et al., 1992) provided neurobiologists with a set of tools with which to probe SNARE function (Fig. 8.6). Each of the toxins comprises a heavy chain and a light chain that are linked by disulfide bonds (Pellizzari et al., 1999). The heavy chain binds the toxin to surface receptors on neurons and thereby enables the toxin to be endocytosed. Once inside the cell, the disulfide bond is reduced and the free light chain enters the cytoplasm of the cell. This light chain is the active portion of the toxin and is a member of the Zn\(^{2+}\)-dependent family of proteases. The catalytic nature of the toxin accounts for its astonishing potency; a few tetanus toxin light chains, for example, at a synapse can suffice to proteolyze all the VAMP/synaptobrevin and thereby shut down transmitter release. The toxins are highly specific, recognizing unique sequences within an individual SNARE protein as summarized in Fig. 8.6. VAMP/synaptobrevin can be cleaved not only by tetanus toxin but by the botulinum toxin types B, D, F, and G, and each toxin cleaves at a different peptide bond within the structure (Schiavo et al., 1992; Yamasaki et al., 1994). SNAP-25 is cleaved by botulinum toxins A and E, but again at different sites (Blasi et al., 1993a; Schiavo et al., 1993). Botulinum toxin C1 cleaves both syntaxin (Blasi et al., 1993b) and, less efficiently, SNAP-25 (Foran et al., 1996).

What precisely is the function of the SNARE proteins in promoting transmitter release? How does the assembly of this complex relate to membrane fusion? Studies with botulinum toxins, yeast mutants, mutants of \textit{Drosophila} and \textit{C. elegans}, and permeabilized mammalian cells all place the SNAREs late in the process. Synapses that lack an individual SNARE, for example, synapses whose VAMP or syntaxin have been mutated or cleaved by toxin, have the expected population of synaptic vesicles, and these vesicles accumulate at active zones in the expected manner. There is, however, disagreement as to whether these vesicles are “docked” in a tight association with the plasma membrane in the absence of the SNARE complex (Hammarlund et al., 2007; Hunt et al., 1994). Despite the presence of vesicles, these synapses are incapable of secreting transmitter. Thus the SNAREs appear to be essential for a step concurrent either with or after the docking of the vesicle at the release site, but before the fusion pore opens and transmitter can diffuse into the cleft. The details of how the complex functions, however, are less certain. One attractive model (Chen and Scheller, 2001) is that the energy released by the formation of this very high affinity complex is used to drive together the two membranes. A loose complex of the SNARES would form and then “zipper up” and pull the membranes together. This may correspond to the actual fusion of the membranes or, alternatively, to a priming step that requires a subsequent rearrangement of the lipids to open the pore that will connect the vesicle lumen to the extracellular space. Evidence from the fusion of yeast vacuolar vesicles (Peters and Mayer, 1998; Peters et al., 2001) implicates a distinct downstream step, regulated by Ca\(^{2+}\)/calmodulin and involving subunits of the proton pump, but whether or not this is true of neurons as well remains unknown. Vesicles consisting of only lipid bilayers and SNAREs can fuse in vitro, suggesting that no other proteins will be essential for the fusion step (Nickel et al., 1999).

One additional function may reside with the SNARE proteins: the identification of an appropriate target membrane (McNew et al., 2000; Rothman and
Warren, 1994). Within a cell, there are myriad membrane compartments with which a transport vesicle might fuse: How then is specificity achieved? The great diversity of SNAREs (Fig. 8.7) may account for some of this specificity because not all combinations of v- and t-SNAREs form functional complexes. This potential mechanism, however, is likely to be only a part of the story. Particularly in the nerve terminal, it appears that synaptic vesicles can find the active zone even in the absence of the relevant SNAREs (Broadie et al., 1995; Hunt et al., 1994). In addition, the t-SNAREs syntaxin and SNAP-25 can be present along the entire axon and thus are inadequate in explaining the selective release of transmitter at synapses and active zones (Garcia et al., 1995; Sesack and Snyder, 1995).

NSF: An ATPase for Membrane Trafficking

At some point as vesicles move through their exoendocytic cycle, energy must be added to the system. N-Ethylmaleimide-sensitive factor (NSF), an ATPase involved in membrane trafficking, is one likely source. NSF was first identified as a required cytosolic factor in an in vitro trafficking assay (Block et al., 1988). The importance of NSF was confirmed when it was found to correspond to the yeast sec18 gene, an essential gene for secretion (Eakle et al., 1988). NSF hexamers bind a cofactor called a-SNAP (soluble NSF attachment protein) or sec17, and this complex, in turn, can bind to the SNARE complex. When Mg-ATP is hydrolyzed, the SNARE complex is disrupted into its component proteins (Söllner et al., 1993). Originally it was speculated that this disassembly of the complex might correspond to fusion—that the action of NSF might catalytically wrench the SNAREs so that the membranes were brought together. As discussed previously, however, a late role for ATP is unlikely in transmitter release, and in vitro SNAREs can stimulate fusion without NSF present. More recent models put NSF action well before or after the fusion step (Banerjee et al., 1996; Mayer et al., 1996). If SNARE complexes form between VAMP, syntaxin, and SNAP-25 all in the same membrane, these futile complexes can be split apart by NSF so that productive complexes bridging the membrane compartments can be formed. After fusion, the tight SNARE complex needs to be disrupted so that the VAMP can be recycled to synaptic vesicles while the other SNAREs remain on the plasma membrane. If, indeed, the energy of forming a tight SNARE complex is part of the energy that drives fusion, NSF, by restoring the SNAREs to their dissociated, high-energy state, will be an important part of the energetics of membrane fusion.

Docking and Priming the Vesicles for Fusion

Although SNAREs are thought to act late in the fusion reaction and fusion itself is an extremely rapid state, many preparatory and regulatory steps may precede the action of the SNAREs. These steps must tether the vesicle at an appropriate release site in the active zone and hold the vesicle in a fusion-ready state. These mechanisms remain among the most obscure at the synapse, but the list of proteins that may participate is growing. One example is the protein n-sec1 (also called munc18), the neuronal homologue of the product of the yeast sec1 gene. Mutations of this protein prevent trafficking in both yeast and higher organisms. This protein binds to syntaxin with a very high affinity and, when so bound, may hold syntaxin in a closed state that prevents syntaxin from binding to SNAP-25 or VAMP (Garcia et al., 1994; Pevsner et al., 1994). However, n-sec1 also can bind to syntaxin and the assembled SNARE complex (Dulubova et al., 2007). It appears likely that n-sec1 serves two functions: it may promote membrane fusion by priming syntaxin and interacting with the SNARE complex during fusion, but n-sec1 may also be a negative regulator, keeping syntaxin closed and inert until an appropriate vesicle or signal displaces n-sec1 and allows a SNARE complex to form. Another synaptic protein, tomosyn, also serves as a negative regulator of SNARE complex formation. Tomosyn, a soluble protein, can bind to complexes of syntaxin and SNAP-25 and prevent the incorporation of VAMP/synaptobrevin into the complex and thereby inhibit release (McEwen et al., 2006).

One of the most interesting of the proteins that are thought to regulate docking and priming via the SNARE complex is unc-13, a protein first discovered in C. elegans by virtue of its neurological phenotype when mutated (Maruyama and Brenner, 1991). Unc-13 and its mammalian homologs (Munc-13s) appear necessary in the transformation of syntaxin from a closed state to an open state that permits SNARE complex assembly (Hammarlund et al., 2007; Richmond et al., 2001). Unc-13 contains a binding site for diacylglycerol and for Ca$^{2+}$/calmodulin and has a critical role in the modulation of transmitter release by those second messengers (Junge et al., 2004; Rosenmund et al., 2002; Wierda et al., 2007). Thus unc-13 is not only an important piece of the release machinery but also a control point for synaptic plasticity.

Rab3 is another protein for which a priming or regulatory role is often invoked at the synapse. This small GTP-binding protein, mentioned previously, is the homologue of the yeast sec4 gene product. In yeast, and at other membrane-trafficking steps within mammalian cells, rab proteins have essential roles
(Gorvel et al., 1991; Salminen and Novick, 1987). They appear to help a vesicle to recognize its appropriate target and begin the process of SNARE complex formation. At the synapse, however, the significance of rab3 is still uncertain. Though it is clearly associated with synaptic vesicles, genetic disruption of all four isoforms of rab3 in mice causes only subtle alterations in synaptic properties (Schluter et al., 2004). Whether this is due to additional, redundant rab proteins or whether the rab family has been relegated to a more minor role at the synapse remains to be determined.

Because synaptic vesicles dock and fuse specifically at the active zone, this region of the nerve terminal membrane must have unique properties that promote docking and priming. The special nature of this domain is easily discernible in electron micrographs: a “fuzz” of electron-dense material can be observed opposite the postsynaptic density. In some synapses, such as photoreceptors, hair cells, and many insect synapses, the structures are more elaborate and include ribbons, dense bodies, and T-bars that extend into the cytoplasm and appear to have a special relationship with the nearby pool of vesicles. Recent advances in electron microscopy have allowed a more detailed look at the association of vesicles and plasma membrane at the active zone (Harlow et al., 2001). At the neuromuscular junction of the frog (Fig. 8.1), it has been revealed that the electron-dense “fuzz” adjacent to the presynaptic plasma membrane is actually a highly ordered structure—a lattice of proteins that connect the vesicles to a cytoskeleton, to one another, and to the plasma membrane (Fig. 8.8). The molecules that correspond to these structures are not yet known. A few proteins, however, are known to be concentrated in the active zone or in the cloud of vesicles near the active zone. These proteins, piccolo, bassoon, RIM1, Cast/ELKS, and unc-13, may be a part of the machinery that defines the active zone as the appropriate target for synaptic vesicle fusion (Garner et al., 2000; Ohtsuka et al., 2002).
In addition to the specializations of the active zone, additional machinery must be present to preserve a dense cluster of synaptic vesicles extending approximately 200 nm back from the active zone (Fig. 8.1). The vesicles in this domain are not likely to be releasable within microseconds of the arrival of an action potential but instead are likely to represent a reserve pool from which vesicles can be mobilized to release sites on the plasma membrane. The equilibrium between this pool and the vesicles actually at the membrane may be an important determinant of the number of vesicles released per impulse but remains poorly understood. One protein that is likely to play a role in the maintenance of the reserve pool is synapsin, a family of peripheral membrane proteins on synaptic vesicles (De Camilli et al., 1983; Sudhof et al., 1989). Synapsins can also bind actin filaments and thus may provide a linker that tethers the vesicles in the cluster to the synaptic cytoskeleton (Fig. 8.5). Disruption of this link can cause the vesicle cluster to be diminished (Pieribone et al., 1995) and reduce the number of vesicles in the releasable pool of the terminal (Ryan et al., 1996). Synapsin has attracted considerable interest because it is the substrate for phosphorylation by both cAMP and Ca\(^{2+}\)-dependent protein kinases (Greengard et al., 1993). These phosphorylations may influence the availability of reserve vesicles for recruitment to release sites.

### Coupling the Action Potential to Vesicle Fusion

The most striking difference between synaptic transmission and traffic between other cellular compartments is the rapid triggering of fusion by action potentials. As already discussed, the opening of Ca\(^{2+}\) channels and the focal rise of intracellular Ca\(^{2+}\) activate the fusion machinery. How does the terminal sense the rise in Ca\(^{2+}\)? What is the Ca\(^{2+}\) trigger and how does it open the fusion pore? Is it a single Ca\(^{2+}\)-binding protein or do several components respond to the altered Ca\(^{2+}\) concentration? Does Ca\(^{2+}\) remove a brake that normally prevents a docked, primed vesicle from fusing, or does Ca\(^{2+}\) induce a conformational change that is actively required to promote fusion? Whence does the steep, exponential relationship of release to intracellular Ca\(^{2+}\) arise? These questions are an active area of investigation and debate.

Some clues may come from other systems: Ca\(^{2+}\)-dependent membrane fusion is not unique to the synapse. Ca\(^{2+}\) can trigger both exocrine and endocrine secretion. Furthermore, Ca\(^{2+}\) released from intracellular stores now appears to be essential in trafficking steps in yeast that previously had been viewed as constitutive and unregulated. In yeast vacuolar fusion, for example, calmodulin senses a local increase in Ca\(^{2+}\) and triggers fusion (Peters and Mayer, 1998). Calmodulin, however, has not been the leading candidate for the synaptic trigger. The affinity of calmodulin for Ca\(^{2+}\) has generally been taken to be too high to explain the relatively high levels of Ca\(^{2+}\) (at least 10 \(\mu\)M) that are needed to evoke transmitter release. However, the apparent affinity of calmodulin for Ca\(^{2+}\) is very dependent on the proteins to which calmodulin binds, and the four Ca\(^{2+}\)-binding sites of free calmodulin probably have an average affinity of about 11 \(\mu\)M (Jurado et al., 1999). Evidence for involvement of calmodulin in synaptic transmission continues to arise (Arredondo et al., 1998; Chamberlain et al., 1995; Chen and Scheller, 2001; Junge et al., 2004; Quetglas et al., 2000), suggesting that yeast vacuolar traffic and synaptic transmission may not be as divergent as one might think. At present, however, a modulatory role for calmodulin is favored over a requirement in the final triggering step at the synapse.

The leading candidate for synaptic Ca\(^{2+}\) sensor is synaptotagmin, an integral membrane protein of the synaptic vesicle (Fig. 8.4) (Brose et al., 1992; Geppert et al., 1994; Li et al., 1995; Perrin et al., 1991). Synaptotagmin has a large cytoplasmic portion that comprises two Ca\(^{2+}\)-binding C2 domains, called C2A and C2B. These domains can also interact with the SNARE complex proteins and with phospholipids in a Ca\(^{2+}\)-dependent manner. Consistent with the hypothesis that synaptotagmin is the trigger for vesicle fusion, mutations that remove the most abundant isofrom of synaptotagmin, synaptotagmin 1, profoundly reduce synaptic transmission in flies, worms, and mice (Broadie et al., 1994; Di Antonio et al., 1993; Di Antonio and Schwarz, 1994; Geppert et al., 1994; Littleton et al., 1993; Nonet et al., 1993) while having little effect on or enhancing the rate of spontaneous release of transmitter. Mutations of the Ca\(^{2+}\)-binding residues of the C2B domain have a very potent effect in disrupting transmission (Mackler et al., 2002), more so than changes to the C2A domain (Fernandez-Chacon et al., 2002; Robinson et al., 2002), suggesting a particularly critical role for Ca\(^{2+}\)-binding at that site. However, the loss of synaptotagmin 1 does not completely abolish Ca\(^{2+}\)-evoked synaptic transmission (Broadie et al., 1994; Geppert et al., 1994; Nonet et al., 1993), indicating that either an unidentified additional synaptotagmin isoform drives this residual vesicle fusion or that an additional Ca\(^{2+}\) sensor is also involved. Synaptotagmin is likely to be involved in endocytosis and potentially in vesicle docking as well (Haucke and De Camilli, 1999; Jorgensen et al., 1995; Reist et al., 1998; Zhang et al., 1994), which has complicated the analysis. These processes, as mentioned previously, are also likely to be regulated by Ca\(^{2+}\), and the Ca\(^{2+}\)-binding sites on synaptotagmin may be relevant for this regulation as well.
Fresh insight into how synaptotagmin promotes vesicle fusion has come from studies of complexin, a small cytoplasmic protein that can bind to SNARE complexes. Complexin appears to prevent membrane fusion when it is bound to the complex. In the presence of Ca\(^{2+}\), however, and only in the presence of Ca\(^{2+}\), synaptotagmin will displace complexin from the SNAREs and promote fusion. Thus the triggering event may involve the synaptotagmin-dependent removal of a brake (complexin) and subsequent interactions of synaptotagmin with both the SNAREs and membrane lipids that lead to the fusion of the membranes (Giraudo et al., 2006; McMahon et al., 1995).

**Packaging Transmitter into the Vesicle**

A central requirement of quantal synaptic transmission is the synchronous release of thousands of molecules of transmitter from the presynaptic nerve terminal. This requirement is partly met by the capacity of synaptic vesicles to accumulate and store high concentrations of transmitter. In cholinergic neurons, the concentration of acetylcholine within the synaptic vesicle can reach 0.6 M, more than 1000-fold greater than that in the cytoplasm (Carlson et al., 1978; Nagy et al., 1976). Two synaptic vesicle proteins mediate the uptake of transmitter: the vacuolar proton pump and a family of transmitter transporters. The vacuolar proton pump is a multisubunit ATPase that catalyzes the translocation of protons from the cytoplasm into the lumen of a variety of intracellular organelles, including synaptic vesicles (Nelson, 1992). The resulting transmembrane electrochemical proton gradient is used as the energy source for the active uptake of transmitter-by-transmitter transporters. Transmitter uptake has been characterized in isolated synaptic vesicle preparations in which at least four types of distinct transporters have been identified: one for acetylcholine, another for biogenic amines (catecholamines and serotonin), a third for the excitatory amino acid glutamate, and the fourth for inhibitory amino acids (GABA and glycine) (Edwards, 1992). At least one gene for each of these classes of transporter has now been cloned (Bellocchio et al., 2000; Reimer et al., 1998; Takamori et al., 2000). As expected, these distinct transporters are differentially expressed by neurons. The type of transporter expressed in a cell dictates the type of transmitter stored in the synaptic vesicles of a particular neuron, and when investigators drive the expression of a glutamate transporter, for example, in a GABA-releasing neuron, they can trick the cell into now releasing glutamate.

The vesicular transporters are integral membrane proteins with 12 membrane-spanning domains that display sequence similarity with bacterial drug resistance transporters. The synaptic vesicle transporters are clearly distinct from the plasma membrane transporter that remove transmitter from the synaptic cleft and thereby contribute to the termination of synaptic signaling (see Fig. 8.2 and Chapter 9). The distinguishing characteristics include their transport topology, energy source, pharmacology, and structure (Reimer et al., 1998).

In contrast to small chemical transmitters, proteinaceous signaling molecules, including neuropeptides and hormones, are typically stored in granules that are larger and have a higher electron density than synaptic vesicles. The contents of these granules are not recycled at the release site; as a result, their replenishment requires new protein synthesis followed by packaging into secretory vesicles in the cell body. Because of the slow kinetics of their release, the slow responsiveness of their postsynaptic receptors, and their inability to be locally recycled, proteinaceous signaling molecules typically mediate regulatory functions.

**Endocytosis Recovers Synaptic Vesicle Components**

After exocytosis, the components of the synaptic vesicle membrane must be recovered from the presynaptic plasma membrane, as discussed previously. Vesicle recycling is accomplished by either of two mechanisms. The first is simply a reversal of the fusion process. In this case, a fusion pore opens to allow transmitter release and then rapidly closes to re-form a vesicle. Often nicknamed “kiss and run,” it has the theoretical advantage that it would allow all the vesicular components to remain together on a single vesicle that would be immediately available for reloading with transmitter. The mechanism is potentially quick and energetically efficient. This mechanism is employed in some systems (Monck and Fernandez, 1992), but its relevance for the synapse is questionable and perhaps quite limited (Dickman et al., 2005; Granseth et al., 2006; Sankaranarayanan and Ryan, 2001).

The predominant pathway for synaptic vesicle recycling is more likely to be endocytosis. Endocytosis of synaptic vesicle components, like receptor-mediated endocytosis in other cell types, is mediated by vesicles coated with the protein clathrin (Granseth et al., 2006; Maycox et al., 1992). Accessory proteins select the cargo incorporated into these vesicles as they assemble. One class of accessory proteins known as AP–2 displays a high affinity for synaptotagmin (Zhang et al., 1994), which may be important, therefore, in recruiting the clathrin coat to the vesicle. The final pinching off of the clathrin-coated vesicle requires the protein dynamin, which can form a ringlike collar around the neck of an
endocytosing vesicle (De Camilli et al., 1995; Warnock and Schmid, 1996). A crucial role for dynamin in synaptic vesicle recycling is most clearly demonstrated in a temperature-sensitive Drosophila mutant known as shibire. The shibire gene encodes the Drosophila homologue of dynamin (Chen et al., 1991; Van der Bliek and Meyerowitz, 1991). At the nonpermissive temperature, the shibire mutant flies rapidly become paralyzed owing to nearly complete depletion of synaptic vesicles from their nerve terminals. Dynamin is a GTPase whose activity is modulated by calcium-regulated phosphorylation and dephosphorylation (Robinson et al., 1994). Thus, the endocytic recycling of synaptic vesicles provides another site at which calcium regulates the synaptic vesicle exo–endoctytic cycle. When the components of the synaptic vesicle membrane are recovered in clathrin-coated vesicles, recycling is completed by vesicle uncoating and, perhaps, passage through an endosomal compartment in the nerve terminal (see Fig. 8.2). Other proteins, such as synaptojanin and endophilin, accelerate the process of retrieving vesicles by this classical pathway (Dickman et al., 2005).

Summary

The life of the synaptic vesicle involves much more than just the Ca$^{2+}$-dependent fusion of a vesicle with the plasma membrane. It is a cyclical progression that must include endocytosis, transmitter loading, docking, and priming steps as well. In many regards, this cell biological process shares mechanistic similarities with membrane trafficking in other parts of the cell and with simpler organisms such as yeast. The interaction of the vesicular and plasma membrane proteins of the SNARE complex—VAMP/synaptobrevin, syntaxin, and SNAP–25—is an essential, late step in fusion. Many other proteins have been identified that are likely to precede the action of the SNAREs, regulate the SNAREs, and recycle them. Components of the synaptic vesicle membrane, the presynaptic plasma membrane, and the cytoplasm all contribute to the regulation of synaptic vesicle function. Together, these proteins build on the fundamental core apparatus to create an astonishingly accurate, fast, and reliable means of delivering transmitter to the synaptic cleft.

**QUANTAL ANALYSIS**

A quantitative description of the signal passing across a synapse is of utmost importance in understanding the function of the nervous system. This signal is the final output of all the integrative processes taking place in the presynaptic cell, and a complete statistical description should be able to capture the flow of information between neurons, as well as between neurons and effector cells. At many chemical synapses, a quantitative description is also a source of unique insight into the biophysics of transmission. The postsynaptic signal often fluctuates from trial to trial in a quantal manner; that is, it adopts preferred levels, which arise from the summation of various numbers of discrete events, thought to result from the release of individual vesicles of neurotransmitter (Katz, 1969, Martin, 1977). Examination of the trial-to-trial amplitude fluctuation of the synaptic signal allows the size of the quantum, as well as the average number of quanta released for a given presynaptic action potential, to be estimated. This approach is also a source of insight into the probabilistic processes underlying transmitter release from the presynaptic terminal and into the mechanisms by which transmission can be modified by physiological, pharmacological, and pathological phenomena.

**Transmission at the Frog Neuromuscular Junction Is Quantized**

The quantal nature of transmission was first demonstrated in the early 1950s by Bernard Katz and his colleagues, who studied the frog motor end plate. By recording from a muscle fiber immediately under a branch of the motor axon, they measured the postsynaptic potential both at rest and in response to stimulation of the axon. Spontaneous signals (mEPPs) were observed to occur at random intervals, measuring between 0.1 and 2 mV in amplitude (Fatt and Katz, 1952). Stimulating the presynaptic axon produced a postsynaptic signal up to 100 times larger. At any one site, the spontaneous mEPPs were of roughly the same amplitude, with a coefficient of variation of 30%, compatible with the intermittent release of multimolecular packets of transmitter from the presynaptic terminal. When the presynaptic axon was stimulated under conditions designed to depress transmitter release to very low levels, the evoked EPP fluctuated from trial to trial between preferred amplitudes, which coincided with integral multiples of the mEPP amplitude (Boyd and Martin, 1956; del Castillo and Katz, 1954; Liley, 1956) (see Fig. 8.9). This implied that the mEPP was a quantal building block, variable numbers of which were released to make up the evoked signal. The relative numbers of trials resulting in 0, 1, 2, . . . , quanta were well described by a Poisson distribution, a statistical distribution that arises in many instances where a random process operates (Box 8.4), implying that the process governing quantal release may also depend on a simple underlying mechanism.
On the basis of this evidence, Katz and colleagues proposed the following model of transmission, which has gained wide acceptance and is referred to as the standard Katz model.

- Arrival of an action potential at the presynaptic terminal briefly raises the probability of release of quanta of transmitter.
- Several quanta are available to be released, and every quantum gives roughly the same electrical signal in the postsynaptic cell. This is the quantal amplitude, $Q$, which sums linearly with all other quanta released.
- The average number of quanta released, $m$, is given by the product of $n$, the number of available quanta, and $p$, the average release probability: $m = np$. The relative probability of observing $0, 1, 2, \ldots, n$ quanta released is then given by a binomial distribution, with parameters $n$ and $p$ (see Box 8.4).

- Under conditions of depressed transmission, $p$ is low, and the system approximates a Poisson process. This is the limiting case of the binomial distribution where $p$ tends to 0 and $n$ tends to $\infty$ and is determined by the unique parameter $m$ (see Box 8.4).

The standard Katz model has been supported by similar experiments in other preparations, which have shown that evoked EPSPs, IPSPs, EPSCs, or IPSCs tend to cluster near preferred values corresponding to integral multiples of a unit. In many cases, this quantal amplitude corresponds closely to the amplitude of spontaneous miniature postsynaptic signals (potentials or currents) occurring in the absence of presynaptic action potentials (Fig. 8.10) (Paulsen and Heggelund, 1994).

A major impetus for accurate measurement of quantal parameters is that it may help determine the locus of modulatory effects on synaptic transmission.

---

**FIGURE 8.9** Quantal transmission at the neuromuscular junction. (A) Intracellular recordings from a rat muscle fiber in response to repeated presynaptic stimulation of the motor axon. Extracellular $[Ca^{2+}]$ and $[Mg^{2+}]$ were kept low and high, respectively, to depress transmission to a very low level. The size of the postsynaptic response, seen after the stimulus (not shown) that occurs at the dashed vertical line, fluctuated from trial to trial, with some trials giving failures of transmission. Spontaneous mEPs, occurring in the background, had approximately the same amplitude as the smallest evoked EPPs, implying that they arose from the release of single quanta of acetylcholine. (B) The peak amplitudes of 200 evoked EPPs from a similar experiment, plotted as an amplitude histogram. Eighteen trials resulted in failures of transmission (indicated by the bar at 0 mV), and the rest gave EPPs whose amplitude tended to cluster at integral multiples of 0.4 mV. This coincides with the mean amplitude of the spontaneous mEPs, whose amplitude distribution is shown in the inset together with a Gaussian fit. The shading through the EPP histogram is a fit obtained by assuming a Poisson model of quantal release (see Box 8.4). The parameters describing this model were the average number of quanta released, $m$, obtained by dividing the mean EPP amplitude by the mean mEP amplitude, and the quantal amplitude, $Q$, and its variance, $Var(Q)$, obtained from the mEP amplitude distribution. There is a good agreement with the observed amplitude distribution. The Poisson model predicts 19 failures. Roman numerals indicate the number of quanta corresponding to each component in the distribution. (A) Adapted from Liley, A.W. (1956). The quantal components of the mammalian end-plate potential. J. Physiol. (B) from Boyd and Martin (1956).
According to the binomial description, \( n \) quanta can be released in response to a presynaptic action potential, each of which has a probability, \( p \), of being discharged. For convenience, let us define \( q \) as the probability that a given quantum is not discharged in a given trial: \( q = 1 - p \). In any given trial, the number of quanta observed is between 0 and \( n \). Imagine that only three quanta are available (\( n = 3 \)), each of which has a 40% chance of being discharged in response to the action potential (\( p = 0.4, q = 0.6 \)). The average number of quanta released is \( m = np = 1.2 \).

The probability that no quanta are released is \( q^3 = 0.216 \).

The probability that only one quantum is released is the sum of the probabilities that only the first is released, only the second is released, and only the third is released: \( 3pq^2 = 0.432 \).

The probability that two of three are released is, conversely, \( 3p^2q = 0.288 \).

Finally, the probability that all three are released is \( p^3 = 0.064 \).

The relative probabilities of observing 0, 1, ..., \( n \) quanta are the coefficients of the expansion of \( G = (p + qz)^n = 0.216 + 0.432z + 0.288z^2 + 0.064z^3 \). This is known as the generating function for the binomial distribution, and \( z \) is simply a "dummy variable": that is, it serves no function except to allow the binomial expansion. These coefficients can be obtained from the binomial distribution, which gives the probability of observing 0, 1, ..., \( n \) quanta released for any \( n \) and \( p \). Writing \( P_x \) for the probability of observing \( x \) quanta, we obtain

\[
P_x = \frac{n!}{(n-x)!x!} p^x q^{n-x}.
\]

**Poisson Model**

As \( n \) becomes very large and \( p \) very small, the probability of observing 0, 1, ..., quanta is equally well described by a Poisson distribution. This is a limiting case of the binomial distribution when \( n \) tends to \( \infty \) and \( p \) tends to 0, and, instead of two parameters (\( n \) and \( p \)), it is described by the sole parameter \( m \). Again, \( m \) is the average value for \( P_x \). The relative probability of observing 0, 1, ..., quanta is now given by

\[
P_x = \frac{m^x e^{-m}}{x!}
\]

For the same average quantal content \( m \) as in the binomial model, the relative probabilities of observing 0, 1, ..., quanta are now approximately

\[
P_0 \approx 0.30, \quad P_1 \approx 0.36, \quad P_2 \approx 0.21, \quad P_3 \approx 0.08, \quad P_4 \approx 0.02, \quad \ldots
\]

**Compound or Nonuniform Binomial Model**

In the binomial model, what happens if different release sites have different, but still independent, probabilities? Let us take the following example: \( p_1 = 0.1, p_2 = 0.2, p_3 = 0.9 \) (again defining \( q_1 = 1 - p_1, q_2 = 1 - p_2, q_3 = 1 - p_3 \)). The average quantal content is again \( p_1 + p_2 + p_3 = 1.2 \).

\[
P_0 = q_1q_2q_3 = 0.07, \quad P_1 = p_1q_2q_3 + q_1p_2q_3 + q_1q_2p_3 = 0.67, \quad P_2 = q_1p_2p_3 + p_1q_2p_3 + p_1p_2q_3 = 0.23, \quad P_3 = p_1p_2p_3 = 0.01.
\]

Generally, \( P_x \) are again obtained from the coefficients of the polynomial expansion of the generating function:

\[
G = (p_1 + q_1z)(p_2 + q_2z)\ldots(p_n + q_nz) = \Pi_k(p_k + q_kz), \quad k = 1, \ldots, n.
\]
transmission. An increase or decrease in the probability of presynaptic transmitter release should be detected as a change in quantal content, \( m \), whereas an alteration in the postsynaptic density or efficacy of receptors should be detected as a change in quantal amplitude, \( Q \).

**Biophysical Phenomena Underlie the Quantal Parameters**

Considerable effort has been directed at establishing the physical correlates of the quantal parameters, \( Q \), \( n \), and \( p \).

**Quantal Parameter \( Q \)**

A mEPP was thought to be too large to be accounted for by the release of an individual molecule of acetylcholine, because low concentrations of exogenous acetylcholine generated responses much smaller than an mEPP (Fatt and Katz, 1952). Vesicles were subsequently observed in electron micrographs of presynaptic terminals (see Chapter 1). Because these vesicles could act as packaging devices for transmitter, it was proposed that the quantal amplitude, \( Q \), represents the discharge of one vesicle into the synaptic cleft. A large body of evidence (Box 8.1) has since led to almost universal agreement that \( Q \) is indeed the postsynaptic response to exocytosis of a single vesicle of transmitter.

Figure 8.11 (Bartol et al., 1991) shows a computer simulation of the diffusion of molecules of acetylcholine and binding to receptors at the neuromuscular junction: acetylcholine spreads out in a disk in the synaptic cleft and binds most of the underlying postsynaptic receptors, although many more spare receptors beyond the edge of the disk remain unoccupied by transmitter (Matthews-Bellinger and Salpeter, 1973).

**Quantal Parameter \( n \)**

The physical correlate of \( n \) has been more elusive. In the original description of quantal transmission, \( n \) represented the number of releasable quanta. As the ultrastructure of the presynaptic terminal was elucidated, however, vesicles were found to be clustered near specializations, or active zones, in the presynaptic membrane. Each active zone is composed of a dense bar on the cytoplasmic face of the terminal membrane, bordered by rows of intramembranous particles, thought to be voltage-sensitive calcium channels (Heuser and Reese, 1973). As stated earlier, entry of calcium ions through these channels raises the intracellular calcium concentration in a small volume immediately adjacent to the channels, triggering the exocytosis. Can \( n \) then be the number of active zones, if they are release sites? At the frog neuromuscular junction, simultaneously evoked quanta seem to summate linearly; however, when acetylcholine is added to the bath the relationship between depolarization and acetylcholine concentration is nonlinear (Hartzell et al., 1975). The discrepancy between the linear summation of quanta and the nonlinear dose dependency of the response to acetylcholine can be explained by proposing that the multiple quanta making up an EPP activate separate populations of
receptors. This could occur if each active zone normally releases only one quantum and is consistent with the view that $n$ is indeed the number of active zones.

**Quantal Parameter $p$**

Parameter $p$ is the probability of exocytosis in response to a presynaptic action potential. Because this interval is of finite duration, it is more strictly a time integral of the probability during this transient event (Katz and Miledi, 1965). Moreover, because discharge of a quantum may leave a release site empty, $p$ should be treated as a product of two probabilities: (1) that a release site is occupied by a quantum ($p_1$) and (2) that a presynaptic action potential evokes release ($p_2$) (Zucker, 1973).

With improved ultrastructural resolution of presynaptic terminals, it has become apparent that some
vesicles are especially intimately related to the active
zone. The number of such “docked” vesicles may rep-
resent a readily releasable pool (Schikorski and Ste-
vens, 1997), which corresponds to the product of $n$
and $p_1$.

The Standard Katz Model Does Not Always Apply

Before accepting the standard Katz model in all its
details, we more closely examine some of its implica-
tions to see how they tally with our knowledge of the
underlying molecular mechanisms.

Quantal Uniformity

At the vertebrate neuromuscular junction, the
amount of acetylcholine released into the synaptic
cleft determines the quantal size (Fletcher and For-
rester, 1975; Kuffler and Yoshikami, 1975; Whittaker, 1988). Thus, for the quantal amplitude to
be constant at different release sites, a uniform popu-
lation of vesicles must be available to be released, and
similar number of receptors with identical properties
must exist opposite each release site. Electron micro-
scopic images of vesicles in the presynaptic terminal
indicate that their diameters are indeed remarkably
uniform, although whether the neurotransmitter
content of the vesicles is unvarying is not known. Simi-
larly, although postsynaptic receptors are clustered
opposite the active zone, their density and properties
may not be uniform between different sites.

Uniform and Independent Release Probabilities

The release sites must be identical, with a uniform
probability of exocytosis. If this condition were not
satisfied, evoked signals would still cluster at integral
multiples of the quantal amplitude, but the relative
proportion of trials resulting in $0, 1, \ldots, n$ quanta would
no longer be described by a simple binomial (or Pois-
on) distribution. As a limiting case, if $p$ at some sites
is effectively $0$, then the meaning of $n$ is questionable.

Rapid and Synchronous Transmitter Release

All-or-none exocytosis is clearly necessary for quan-
tization and supported by freeze-fracture images of
terminals taken during intense evoked release (Heuser
et al., 1979). However, all-or-none exocytosis may not
be the only mode of transmitter release, because secre-
tory vesicles in mast cells can release some of their con-
ents through a fusion pore that opens reversibly
without necessarily leading to full exocytosis (Alvarez
de Toledo et al., 1993). Whether this mode of release
also occurs in synapses remains to be determined.

Quantization in the size of the evoked response
can also be concealed by asynchrony of transmitter
release from individual sites. (Isaacson and Walmsley,
1995).

Ion Channel Noise

Stochastic properties of postsynaptic ligand-gated
ion channels must not add excessive variability to
the size of the postsynaptic signal. Again, if this con-
dition were not satisfied, it would be difficult to iden-
tify the quantal amplitude, and clustering of
amplitudes at integral multiples of the quantum
would be concealed. If we assume that individual
ionophores act independently of one another, the var-
iance of the quantal current arising from their sto-
chasitic opening is described by the binomial formula

$$Var = i^2 k p_o (1 - p_o),$$

where $i$, $k$, and $p_o$ refer to the single-channel current,
the number of channels, and their probability of open-
ing in response to transmitter release, respectively.
Because the average quantal current amplitude is $i k p_o$,
the coefficient of variation of the quantal amplitude is

$$C_v = \sqrt{(1 - p_o) / k p_o}.$$
release probability, the trial-to-trial amplitude fluctuation of the postsynaptic signal is unlikely to be described by a binomial or Poisson model. Indeed, the evidence that these simple probabilistic models are correct is far from compelling. On the other hand, the fact that evoked synaptic signals are often found clustered at integral multiples of an underlying unit strongly argues that vesicle filling and the postsynaptic phenomena determining the quantal amplitude are sufficiently uniform to ensure that a more general quantal description of transmission applies.

Central Nervous System Synapses Behave Differently from the Frog End Plate

A number of differences have emerged between quantal transmission at the neuromuscular junction and in central synapses in vertebrates.

One-Quantum Release

A correlation of histological and electrophysiological evidence obtained in the same preparation has led to the proposal that many individual terminals in the CNS have only one release site, which releases at most one vesicle at a time (Gulyas et al., 1993; Korn et al., 1981; Silver et al., 2003; Walmsley, 1991). There are some notable exceptions to this rule. Calyceal synapses in the mammalian brain stem auditory pathway, for instance, have multiple active zones, and glutamate released from one release site can interact with transmitter released from neighboring sites (Trussell et al., 1993). Other “giant” synapses in the CNS include those formed by climbing fibers on Purkinje cells, and by mossy fibers on pyramidal neurons in the CA3 region of the hippocampus. Evidence for multivesicular release has, however, also been obtained at small glutamatergic synapses, where the postsynaptic signal amplitude can, under certain conditions, vary with the release probability even when failures of release are excluded (Christie and Jahr, 2006; Oertner et al., 2002). Near-synchronous release of two vesicles of GABA has also been reported at small inhibitory synapses in the cerebellar cortex (Auger et al., 1998).

Nonuniform Release Probabilities

In the mammalian spinal cord, release probabilities may vary between individual sites supplied by an individual muscle afferent. Postsynaptic signals have been shown to fluctuate in a manner that cannot be described by a binomial model, unless the individual release probabilities are allowed to vary (Jack et al., 1981; Redman, 1990; Walmsley, 1991). Because the release sites are often segregated in different terminals, they may be subject to differing amounts of tonic presynaptic inhibition mediated by axo-axonic synapses.

Relatively Few Receptors Are Available to Detect Presynaptic Transmitter Release

A major difference between vertebrate CNS synapses and the neuromuscular junction is that the quantal amplitude is often determined not only by the vesicle contents but also by the number of available receptors (Edwards et al., 1990; Jonas et al., 1993; Korn and Faber, 1987, 1991; Redman, 1990; Walmsley, 1991). At some excitatory synapses, the glutamate content of a quantum appears to be sufficient to bind a large proportion of the available postsynaptic receptors (Clements et al., 1992; Tang et al., 1994; Tong and Jahr, 1994). Fewer than 100 receptors open, compared with 1000–2000 at the neuromuscular junction (Edwards et al., 1990; Jonas et al., 1993).

Different Receptors May Sample Different Quantal Contents

In contrast to the neuromuscular junction, CNS synapses frequently have several pharmacologically distinct postsynaptic receptors. Although little is known of quantal signaling via metabotropic receptors, glutamatergic synapses contain different combinations of AMPA, kainate, and N-methyl-D-aspartate (NMDA) receptors, all of which can open in response to glutamate release (see Chapter 11). The quantal content sampled by NMDA receptors is often larger than that sampled by AMPA receptors (Isaac et al., 1995; Kullmann, 1994; Liao et al., 1995). Because NMDA receptors are unable to open at hyperpolarized membrane potentials, such synapses are functionally silent in the absence of postsynaptic depolarization. This discrepancy in signaling by AMPA and NMDA receptors is partly explained by differential expression of receptors at synapses (Nusser et al., 1998), but, in addition, differences in the affinity of AMPA and NMDA receptors for glutamate may also play a role (Kullmann, 2003).

Spontaneous Miniature Postsynaptic Signals

In the central nervous system, spontaneous mEPSCs and mIPSCs vary widely in amplitude (Edwards et al., 1990; Jonas et al., 1993; Manabe et al., 1992) implying a quantal coefficient of variation considerably greater than that at the neuromuscular junction: between 40 and 80% instead of 30%. At first sight, this would preclude unambiguous peaks in histograms of evoked signals. However, quantal variability must be divided
into variability from trial to trial at an individual release site (intrasite) and variability among sites (intersite). Spontaneous mEPSCs and mIPSCs arise from a large number of different sites, so their amplitude range includes both sources of quantal variability. If intrasite quantal variability were very large, then we would not expect to be able to detect quantal clustering in the amplitudes of evoked synaptic signals. The fact that such clustering is sometimes seen (see Fig. 8.10) implies that intrasite variability can be modest, and the wide range of amplitudes of miniature events principally indicates a large intersite variability. Thus, the mean amplitude of spontaneous miniature events cannot be used as a guide to the quantal amplitude underlying an evoked synaptic signal.

Quantal Parameters Can Be Estimated from Evoked and Spontaneous Signals

The goal is to establish, with a reasonable degree of precision, the quantal amplitude, \( Q \), the average quantal content, \( m \), and, if appropriate, the number of release sites, \( n \), and the average release probability, \( p \). The realization that release sites in the CNS may not always be uniform (Redman, 1990; Walmsley, 1991) makes a complete statistical description of transmission much more difficult to achieve. The parameters that need estimation must then include the release probability and the quantal amplitude and variability at each site. In principle, it should be possible to estimate the quantal parameters from the probability density of a statistic measured from the evoked postsynaptic signal. Most workers have measured the peak amplitude of the postsynaptic voltage or current on a large number of trials and displayed the results in the form of a histogram. Considerable information can also be obtained from the amplitude distribution of spontaneous miniature events. Two major obstacles, sampling artifact and noise, immediately arise.

Because the data sample is finite, the true probability density of a desired statistic is not known; only an approximation can be obtained from the recordings. This problem can be mitigated by obtaining a larger sample, but the sample is generally limited by nonstationarity in the recording and time constraints imposed by the experiment. Noise also conceals the true amplitude of spontaneous or evoked signals. If the noise amplitude is comparable to the quantal amplitude, then not only can spontaneous miniature events be missed, but features of the probability density of evoked signals also can be concealed. Relying entirely on visual inspection to determine whether the peaks and troughs in an amplitude histogram are “genuine” (i.e., that they do not arise from sampling artifact and noise) is misleading. A number of different computational approaches has been developed to overcome this obstacle. These approaches differ in the degree to which they rely on assumptions about the underlying probabilistic process. Clearly, if the assumptions are incorrect, then nothing has been achieved, because the parameters will have been estimated incorrectly.

Spontaneous Miniature Signals

If \( Q \) can be obtained from the amplitude distribution of mEPSCs or mIPSCs, then the average quantal content, \( m \), can be obtained by dividing the average evoked signal amplitude by \( Q \). This can be done only when the amplitude distribution of spontaneous signals is narrow, and when there is no a priori reason that the quanta underlying the evoked signals should be different.

Spontaneous miniature signals can also be used to detect changes in quantal parameters caused by a conditioning treatment affecting a large number of synapses; if the average amplitude of mEPSCs or mIPSCs becomes larger after an experimental perturbation, the implication is that a widespread increase in quantal amplitude has been distributed among the synapses that give rise to the miniature currents. Changes in the frequency of spontaneous miniature events, on the other hand, usually imply an alteration in the average release probability (Van der Kloot, 1991; Van der Kloot and Molgò, 1994). An important difficulty with analysis of spontaneous miniature signals in CNS neurons is that their amplitude distribution is generally skewed, with a long tail toward larger values. At the other end of the distribution, small events often fall at the threshold for detection. To compare mEPSCs or mIPSCs obtained before and after a manipulation, cumulative distributions are generally easier to interpret than raw histograms (Van der Kloot, 1991). A genuine widespread change in amplitude is then seen as a shift in the position of the cumulative distribution, whereas a change in frequency should have no effect on the position of the line, other than that which can be accounted for by sampling artifact (Fig. 8.12). The Kolmogorov–Smirnov test can then be applied to test the hypothesis that any difference between the two curves arose by chance.

Multimodal amplitude distributions have occasionally been described for spontaneous miniature signals (Edwards et al., 1990; Jonas et al., 1993). When the modes are at equal intervals on the amplitude axis, multiquantal release, possibly arising from regenerative processes in the presynaptic terminal, is
implied, although sampling error as a source of spurious peaks must be ruled out.

The principal shortcoming of miniature spontaneous signal analysis is that, generally, the synapses giving rise to the spontaneous signals cannot be identified unambiguously, although localized application of hypertonic sucrose or barium can selectively increase the frequency of discharge in relatively restricted areas (Bekkers and Stevens, 1989; Fatt and Katz, 1952).

Quantal Amplitude Estimation

When Q cannot be estimated from the amplitude distribution of spontaneous mEPSCs or mIPSCs, it can often be obtained from the positions of the peaks in histograms of evoked signals. However, plotting data in the form of histograms can result in misleading estimates of Q in the presence of noise and finite sampling; spurious peaks and troughs can emerge, depending on the position and width of the bins. An approach for attacking this problem is to convolve the data with a kernel, generally a Gaussian function (Silverman, 1986). Convolution describes the mathematical equivalent of “smearing” one function across another. Gaussian kernel convolution has the effect of producing a smooth function with inflections that are no longer susceptible to binning artifact. The main pitfall of this approach is that the optimal width of the smoothing kernel cannot be known a priori, and, if it is too narrow, artifactual inflections, arising from noise and finite sampling, will still appear in the convolved function.

Whatever method is used to display the data, the question remains whether all the inflections in the resulting function could have arisen by chance because of finite sampling. A testable null hypothesis is that the underlying function is in fact continuous; that is, transmission is not quantal. This hypothesis can be modeled by choosing a unimodal distribution with the same overall shape as that of the data distribution, but without any peaks or troughs. If this hypothesis can be rejected at a given degree of confidence, the implication is that the clustering in the data sample does indeed indicate a genuine underlying quantized process. An easily implemented general method is to draw random samples repeatedly from the smooth function, with a sample size equal to that of the data. If the peaks and troughs in these random samples are never or only rarely as prominent as those in the data sample, then the null hypothesis can be rejected. This is an example of a Monte Carlo test (Horn, 1987).

Poisson Model

If the Poisson model holds, m can be estimated by counting the quanta released on each of a large number of trials (Isaacson and Walmsley, 1995). Alternatively, since the variance of a Poisson distribution is equal to its mean, m can be estimated from the trial-to-trial variability in the number of quanta released. However, it is rarely possible to count quanta unambiguously, principally because recordings are affected by noise and because the quantal amplitude has some intrinsic variability, and so indirect methods frequently have to be used to estimate m. The first of
these is to count the proportion of trials that result in a failure of transmission \((N_0/N\) of \(N\) trials). The first term of the Poisson expansion (Box 8.4) gives the probability of observing zero quanta released and is equal to \(e^{-m}\). Therefore, \(m\) is given by taking the natural logarithm of the inverse of this ratio: 

\[
m = \log_e (N/N_0).
\]

The second indirect approach to estimating \(m\) is to measure the coefficient of variation of the postsynaptic response, \(CV_R\). \(CV_R\) must be corrected for two other sources of variability in the postsynaptic response: (1) variability in quantal size, expressed as the quantal coefficient of variation, \(CV_Q\); and (2) background noise variance, \(Var_t\), which can be measured separately by collecting data in the absence of evoked activity. If we assume that the variances arising from the Poisson process, quantal variability, and noise add linearly, \(m\) is then given by the formula

\[
m = \frac{1 + CV_Q^2}{CV_R^2 - Var_t/R^2},
\]

where \(R\) is the average evoked response amplitude (McLachlan, 1978). The variance method of estimation is often used incorrectly when the necessary corrections for quantal size and noise fluctuations are ignored.

Agreement among the estimates of \(m\) obtained with all these methods constitutes circumstantial evidence in favor of the model.

**Binomial Models**

The binomial model has more parameters than the Poisson \((n\) and \(p\) replace \(m\)). If the number of trials resulting in \(0, 1, \ldots, n\) trials is known unambiguously, then \(n\) and \(p\) can be estimated as follows. From the binomial theorem, the variance of the number of quanta, \(Var_m\), is equal to \(np(1 - p)\). It follows that \(p = 1 - Var_m/m\) and \(n = m/p\). If the number of quanta released cannot be estimated unambiguously, then \(p\) can be obtained from the proportion of failures of transmission. Because \(N_0 = N(1 - p)^n\), it follows that \(p = 1 - (N_0/N)^{1/n}\). The usefulness of this method is limited by the requirements that there be an appreciable proportion of failures and that \(n\) be known. Alternatively, the variance method may be used, again with an appropriate correction for the quantal variability and background noise (McLachlan, 1978):

\[
p = 1 + CV_Q^2 - \frac{(R \cdot CV_R^2 - Var_t/R)}{Q}.
\]

This method requires that estimates be made of both the average quantal size, \(Q\) and its coefficient of variation, \(CV_Q\), severely limiting its usefulness.

If we relax the assumption of uniform release probabilities while continuing to assume that different sites are independent of one another, then a nonuniform or compound binomial model must be applied (Jack et al., 1981). In this case, the desired parameters include the individual release probabilities: \(p_1, p_2, \ldots, p_n\). If the sample were perfect, they could be obtained by treating the observed proportions of trials resulting in \(0, 1, \ldots, n\) quanta as the polynomial expansion of \(\Pi_k (p_k + q_kz)\) (see Box 8.4). Solving the polynomial would then yield \(p_1, p_2, \ldots, p_n\) (Jack et al., 1981). In practice, however, the sample is incomplete; that is, some rare events may never have been observed, and others may be spuriously overrepresented. Root-finding algorithms generally yield complex roots in this situation. An alternative approach is to use a numerical optimization, that is, to find the release probabilities that give the best agreement with the data, taking into account the fact that the data sample is incomplete (Kullmann, 1989).

**Noise Deconvolution**

In many cases, the proportion of trials resulting in \(0, 1, \ldots, n\) trials cannot be determined unambiguously because of excessive noise, and the assumption that Poisson or simple binomial statistics apply is untenable. How then can one resolve the underlying quantal process at the synapses under investigation? The method of noise deconvolution (Edwards et al., 1976) again relies on the assumption that noise adds linearly to the synaptic signal, which means that the sampled probability density function is a convolution of the underlying quantal density function with the noise density function. Because the noise can be measured independently, by recording the background signal in the absence of evoked synaptic activity, it should be possible to undo the convolution to reconstruct the probability density function that describes the underlying signal.

This operation is not trivial, because the evoked signal and noise samples are finite: their true probability density functions are not known, and only an approximation can be obtained from the measured signals. The underlying noise-free probability density function must therefore be estimated by applying an optimization method. The underlying function is generally assumed to comprise a number of discrete components representing different numbers of quanta released. The task is then formally equivalent to solving a mixture problem, in which the data are sampled from a mixture of overlapping distributions, or components, each having a membership (probability), mean amplitude, and variance that need to be estimated. Optimization algorithms work as follows: (1) the data distribution is compared with an initial solution reconvolved with the noise function; (2) the
solution is then adjusted to improve the goodness of fit; and (3) the cycle is repeated until no further improvement is detectable. The best results are obtained by maximizing likelihood, and a robust and versatile algorithm to use for this purpose is known as the expectation–maximization algorithm (Kullmann, 1989; Stricker and Redman, 1994). A number of constraints can be imposed on the solution to accommodate physiological assumptions. As a rule, as more constraints are imposed, the quantal parameters are more accurately estimated, but only as long as the underlying assumptions are justified.

A major obstacle is that the number of components in the solution, which cannot generally be known a priori, is a critical parameter. As the number of parameters available to fit the data is increased, the maximum likelihood value increases, because finer details of the data distribution, many of which are due to sampling error and noise, can be accounted for. An alternative approach, which avoids the problem of overfitting, is to treat the underlying probability density function not as a mixture of discrete components but as a continuous function. The solution is biased toward the flattest, most featureless function that is just compatible with the data. This method, known as maximum entropy noise deconvolution, can give an estimate of quantal amplitude if there are periodic inflections in the solution (Kullmann and Nicoll, 1992).

Figure 8.13 shows the results of applying several deconvolution methods to an amplitude histogram.

**Model Discrimination**

An important goal of parameter estimation is to choose between different models of transmission. The simplest approach is to ask if a given model is able to fit the data, by applying a conventional goodness-of-fit test, such as the $\chi^2$ (chi squared) test. If the fit is unsatisfactory, the model can tentatively be rejected with the corresponding degree of confidence. However, the model’s being in good agreement with the data does not necessarily mean that the assumptions underlying the solution are correct, because many alternative models also may give adequate fits.

**Confidence Intervals**

Confidence intervals must be estimated for quantal parameters, as for any statistic. Such an estimation can be difficult for any but the simplest model because the parameter space has many dimensions, and even the number of dimensions is often unknown. Resampling methods that rely on repeating the optimization on a large number of random samples drawn from the original data set (Efron and Tibshirani, 1993) must be used with caution, and it is important in all cases to be aware of the limitations and biases of optimization algorithms by testing them extensively with Monte Carlo simulations.

**Quantal Analysis Can Shed Light on the Mechanisms of Modulation of Synaptic Transmission**

As mentioned previously, a comparison of quantal parameters before and after a treatment that alters synaptic strength can potentially indicate how this alteration is expressed. Ideally, the mean release probability, number of release sites, and quantal amplitude could be estimated at different time points to determine how they change. In practice, it is often difficult to establish all these parameters unambiguously. This difficulty does not preclude separating changes in one parameter from changes in another parameter, because certain statistics that reflect the trial-to-trial variability of transmission change in characteristic ways. This approach is known as variance analysis.

**Coefficient of Variation**

With the use of the coefficient of variation of the evoked signal, the binomial and Poisson models allow inferences to be made about the site of modulation of transmission without the need to estimate the quantal parameters (McLachlan, 1978). To correct for the background noise variance, $\text{Var}_B$, the coefficient of variation of the underlying signal, $CV_S$, is given by

$$CV_S = \sqrt{\frac{\text{Var}_R - \text{Var}_I}{R}}.$$

$CV_S$ is determined by probabilistic quantal release as well as by the quantal variability and is a useful statistic because it is dimensionless. It can be used to distinguish between changes in quantal amplitude and changes in quantal content. Briefly, if $CV_S$ changes with a conditioning treatment that alters the average amplitude of the postsynaptic signal, the implication is a change in quantal content. If, conversely, $CV_S$ is unaffected, the implication is that the conditioning treatment altered quantal amplitude.

If a Poisson model is assumed, further information can be obtained by plotting the ratio of $1/CV_S^2$ before and after a manipulation against the corresponding ratio of mean amplitude $\bar{R}$ (Fig. 8.14) (Manabe et al., 1993). Because the variance of a Poisson distribution is equal to its mean, a change in quantal content, $m$, should cause an excursion along the line of identity. A change in quantal amplitude, $Q$, on the other hand, should have no effect on $1/CV_S^2$, so the data points...
FIGURE 8.13 Noise deconvolution. (A) Amplitude histogram for 400 EPSCs recorded in a CA1 cell in a hippocampal slice in response to repeated stimulation of afferent fibers. EPSCs appear to cluster at integral multiples of approximately 3.6 pA. The continuous line in (B) is the maximum entropy noise deconvolution solution. This function, convolved with the noise, just fits the data at the 5% level of confidence (i.e., a curve any smoother and more featureless would have to be rejected at the 5% level). The periodic inflections seen in this function imply that clustering results not simply from noise, sampling, and binning artifact but also from an underlying quantal process. (C) The result of maximum likelihood deconvolution, with nine underlying components, each with the same variance as the background noise but with no constraint on their amplitudes and probabilities. The continuous line represents the solution reconvolved with the noise, showing a very good fit to the data. The underlying components (dashed lines) are plotted in (D), together with their sum (continuous line). Although the agreement of this solution with the data is excellent, some of its features may arise from sampling artifact and noise. (E, F) A quantal model has been fitted to the data; that is, the components have been constrained to occur at equal intervals, with the first component at 0, and with a quantal coefficient of variation ($CV_Q$) of 0.15 (E) or 0.3 (F). The maximum likelihood solution in (F), but not in (E), can be rejected at the 5% level, implying that $CV_Q < 0.3$. 

8. RELEASE OF NEUROTRANSMITTERS
should fall on the horizontal line. If the points fall between the line of identity and the horizontal line, we can conclude that both quantal content and quantal amplitude changed. If a simple binomial model is assumed, the ratio should fall on the 45° line (b) implies an increase in $n$. If a simple binomial model holds, the ratio should fall on the 45° line for an increase in $n$ and above it (c) for an increase in $p$. Ratios falling below the 45° line and above the horizontal (d) imply an increase in both quantal content and quantal amplitude. Conversely, modulations that decrease synaptic strength should cause the ratios to fall to the left of the point (1, 1) and below it. These rules apply only if a Poisson or simple binomial model holds. (B) Experimental results obtained by recording from CA1 hippocampal pyramidal cells with various manipulations. Each point shows the mean of several cells (with standard errors). Increasing the driving force for the synaptic current by changing the postsynaptic membrane potential (MP) produces no change in $1/CV_s^2$ as expected from a purely postsynaptic modification. Extracellular theophylline or Cs+, an increase in the extracellular $[Ca^{2+}]/[Mg^{2+}]$ ratio, and facilitation by a conditioning prepulse (PPF) cause the ratios to fall in region (c) of (A), as expected from an increase in binomial $p$. Increasing the stimulus strength also causes the points to fall in region (c), although a simple binomial model predicts that an increase in $n$ should cause the ratio to fall on the 45° line. Long-term potentiation of transmission causes the points to fall in region (d), implying an increase in both quantal content and quantal amplitude. Conversely, baclofen decreases transmitter release, and the glutamate receptor antagonist CNQX decreases quantal amplitude. (B) Reproduced with permission from Macmillan Publishers Ltd. Manabe, T., Renner, P., and Nicoll, R. (1992). Postsynaptic contribution to long-term potentiation revealed by the analysis of miniature synaptic currents. *Nature (London)* 355, 50-55.

**FIGURE 8.14** Coefficient of variation method for determining the site of modulation of synaptic transmission. (A) The possible excursions in the ratio of $1/CV_s^2$ plotted against the ratio of the mean response amplitude $R$ for a manipulation that increases synaptic strength. Subscripts "c" and "t" refer to the control and test conditions, respectively. A horizontal excursion (a) implies an increase in quantal amplitude, $Q$, for instance, through a postsynaptic increase in the number of receptors. If a Poisson model applies, an excursion along the 45° line (b) implies an increase in $m$. If a simple binomial model holds, the ratio should fall on the 45° line for an increase in $n$ and above it (c) for an increase in $p$. Ratios falling below the 45° line and above the horizontal (d) imply an increase in both quantal content and quantal amplitude. Conversely, modulations that decrease synaptic strength should cause the ratios to fall to the left of the point (1, 1) and below it. These rules apply only if a Poisson or simple binomial model holds. (B) Experimental results obtained by recording from CA1 hippocampal pyramidal cells with various manipulations. Each point shows the mean of several cells (with standard errors). Increasing the driving force for the synaptic current by changing the postsynaptic membrane potential (MP) produces no change in $1/CV_s^2$ as expected from a purely postsynaptic modification. Extracellular theophylline or Cs+, an increase in the extracellular $[Ca^{2+}]/[Mg^{2+}]$ ratio, and facilitation by a conditioning prepulse (PPF) cause the ratios to fall in region (c) of (A), as expected from an increase in binomial $p$. Increasing the stimulus strength also causes the points to fall in region (c), although a simple binomial model predicts that an increase in $n$ should cause the ratio to fall on the 45° line. Long-term potentiation of transmission causes the points to fall in region (d), implying an increase in both quantal content and quantal amplitude. Conversely, baclofen decreases transmitter release, and the glutamate receptor antagonist CNQX decreases quantal amplitude. (B) Reproduced with permission from Macmillan Publishers Ltd. Manabe, T., Renner, P., and Nicoll, R. (1992). Postsynaptic contribution to long-term potentiation revealed by the analysis of miniature synaptic currents. *Nature (London)* 355, 50-55.
compared with the effect of the new manipulation under investigation. If the observed ratio of \(1/CV_s^2\) plotted against the ratio \(R\) runs along one of the trajectories established for manipulations of either \(n, p,\) or \(Q,\) then it can be inferred that an alteration in the corresponding parameter has occurred. Although this is a potentially powerful approach to verify the validity of variance analysis, it is still potentially flawed if the synaptic plasticity under investigation is not reproduced by any of the experimental alterations of \(n, p,\) or \(Q.\) A notable example is the controversy over the site of expression of long-term potentiation (LTP): although variance analysis consistently shows an increase in quantal content, implying a presynaptic alteration in \(n\) and/or \(p,\) an alternative explanation that has received substantial support from alternative methods is that postsynaptic receptor clusters are uncovered at previously silent sites (see Chapter 19). That is, a postsynaptic modification, with \(Q\) at some sites switching to nonzero values, mimics a presynaptic alteration.

### Manipulating the Release Probability Experimentally Can Yield an Estimate of Quantal Parameters

The previous section described how experimental manipulation of the quantal parameters can be used to test the validity of the variance method. Although this method can shed light on how quantal parameters change with an alteration in synaptic strength, it is not designed to determine the absolute values of these parameters. Paradoxically, manipulating the release probability can actually allow an estimate to be obtained of the quantal parameters under conditions where they cannot be ascertained from the other methods previously described (Clements and Silver, 2000). This approach is analogous to nonstationary variance analysis, a powerful method used to estimate single-channel conductance from membrane currents. For a simple binomial model, the trial-to-trial variance of the postsynaptic response is given by: \(Var_R = npQ^2(1 + CV_Q^2) - np^2Q^2.\) (The term \((1 + CV_Q^2)\) takes into account the contribution of quantal variability within release sites, assuming that the quantal coefficient of variation \(CV_Q\) is uniform across different sites.) This relationship can be rewritten as a function of the mean postsynaptic response amplitude \(R: Var_R = AR - BR^2.\) The quantal parameter estimates are then given by \(Q = A/(1 + CV_Q^2), p = B - R(1 + CV_Q^2)/A,\) and \(n = 1/B.\) In practice the method works as follows. The transmitter release probability is manipulated by varying the extracellular \([Ca^{2+}]/[Mg^{2+}]\) ratio or by applying drugs that act presynaptically, to construct a variance–mean plot (Fig. 8.15). If the binomial model holds, this has a parabolic shape. Fitting the formula \(Var_R = AR - BR^2\) then yields estimates for \(Q, n,\) and \(p.\) If the simple binomial assumption is relaxed, the parabolic shape is often preserved or can become skewed. Under these conditions, the estimates for \(Q\) and \(p\) can still be informative, although they should be more correctly interpreted as weighted means of the underlying parameters, because release sites with low probabilities and/or low quantal amplitudes contribute less to the postsynaptic response variability. This method can be useful even if it is not possible to increase the mean release probability sufficiently to obtain a clear parabolic variance–mean relationship: the initial slope can be fitted with \(Var_R = AR,\) yielding an estimate of \(Q.\) In this situation, however, it is not possible to estimate \(n\) and \(p.\)
Estimation of Transmitter Release Probability Does Not Rely Exclusively on Quantal Analysis of Evoked Postsynaptic Signals

With the development of fluorescent dyes that label presynaptic vesicles as they are recycled (Box 8.3), it has become possible to estimate the probability with which vesicles are released following presynaptic action potentials. This approach has given an elegant confirmation of the quantal hypothesis (Ryan et al., 1997).

An alternative method for estimating the release probability at glutamatergic synapses makes use of the pharmacological properties of N-methyl-D-aspartate (NMDA) receptors (Rosenmund et al., 1993). After application of MK-801, an irreversible open channel blocker, the size of the population synaptic signal mediated by NMDA receptors gradually decays with repeated stimulation of presynaptic fibers. Because the rate at which the postsynaptic signal decays is related to the probability that postsynaptic receptors are activated by released glutamate, this method gives an indirect estimate of the probability of transmitter release. Interestingly, the time course of the decay cannot be described by a single exponential, implying that the release probability must vary considerably across the population of synapses contributing to the signal.

Another indirect method for monitoring changes in transmitter release probability is to examine the short-term facilitation or depression when two or more presynaptic stimuli are delivered in rapid succession, and then to repeat this measurement after a manipulation that alters the strength of the synapse. In general, either increasing or reducing the probability of transmitter release has a relatively greater effect on the size of the response to the first pulse than on that to the second. Therefore, a change in the ratio of the responses to the two pulses implies that at least part of the effect of the manipulation is mediated presynaptically.

Summary

Quantal analysis has greatly improved our understanding of biophysical and pharmacological mechanisms of transmission. Numerical methods can be used to estimate the probability of transmitter release and the size of the postsynaptic effect of an individual quantum of neurotransmitter. Although these methods must be applied with caution, they yield a unique insight into the mechanisms of synaptic plasticity, both at the neuromuscular junction and in the central nervous system.

SHORT-TERM SYNAPTIC PLASTICITY

Chemical synapses are not static transmitters of information. Their effectiveness waxes and wanes, depending on frequency of stimulation and history of prior activity (Zucker and Regehr, 2002). At most synapses, repetitive high-frequency stimulation (called a tetanus) is initially dominated by a growth in successive PSP amplitudes, called synaptic facilitation. This process builds to a steady state within about 1 s and decays equally rapidly when stimulation stops. Decay is measured by single test stimuli given at various intervals after a conditioning train. Facilitation can often be divided into two exponential phases, called its first and second components, and may reach appreciable levels (e.g., a doubling of PSP size) after a single action potential. At most synapses, a slower phase of increase in efficacy, which has a characteristic time constant of several seconds and is called augmentation, succeeds facilitation. Finally, with prolonged stimulation, some synapses display a third phase of growth in PSP amplitude that lasts minutes and is called potentiation.

Often, a phase of decreasing transmission, called synaptic depression, is superimposed on these processes. Synaptic depression leads to a dip in transmission during repetitive stimulation, which often tends to overlap and obscure the augmentation and potentiation phases. When stimulation ceases, recovery from the various processes occurs in the same order as their development during the tetanus, with facilitation decaying first, then depression and augmentation, and finally potentiation (Fig. 8.16). Thus, potentiation is often visible in isolation only long after a tetanus and is thus called post-tetanic potentiation (PTP). At some synapses, even longer-lasting effects (persisting for hours), named long-term potentiation (LTP) (Baxter et al., 1985; Minota et al., 1991), have been observed. This long-term potentiation should not be confused with the form of synaptic plasticity bearing the same name and prominent at mammalian cortical synapses (Chapter 19). In almost all synapses in which a quantal analysis has been done, all these forms of synaptic plasticity (except some forms of cortical LTP) are due to changes in the number of quanta released by action potentials. When binomial parameters were estimated, correlated changes in \( p \) and \( n \) were usually observed. This result is expected if release sites with nonuniform \( p \) become more effective while “silent” sites are recruited during enhanced transmission, and vice versa during depression.
Depression May Arise from Depletion of Readily Releasable Transmitter or from Autoinhibition

In contrast with the growth phases in synaptic plasticity, the rate at which depression develops usually depends on stimulation frequency, whereas recovery from depression proceeds with a single time constant of seconds to minutes in different preparations. At many synapses, depression is relieved when transmission is reduced by lowering \([\text{Ca}^{2+}]\) or raising \([\text{Mg}^{2+}]\) in the medium. These characteristics are consistent with depression being due to the depletion of a readily releasable store of docked or nearly docked vesicles and recovery being due to their replenishment from a supply store (Fig. 8.17) (Zucker, 1989). In fact, styryl dyes appear to label distinct reserve and readily releasable pools of vesicles at \textit{Drosophila} neuromuscular junctions (Kuromi and Kidokoro, 1998). The parameters of such a depletion model can be estimated from the rate of recovery from depression, which gives the rate of refilling the releasable store, and the fractional drop in PSPs given at short intervals, which gives the fraction of the releasable store liberated by each action potential.

Recent data suggest that the classic depletion model is too simple. As the readily releasable store is depleted, both the probability of its releasing a vesicle (Dobrunz and Stevens, 1997; Wu and Borst, 1999) and its maximum capacity (Dobrunz and Stevens, 1997) appear to drop. More realistic models take account of the finite capacity of the readily releasable vesicle pool (Stevens and Wesseling, 1999; Wu and Betz, 1998). At the frog neuromuscular junction depression also involves a reduction in the rates of endocytotic vesicle recovery and of their transport into the readily releasable store (Wu and Betz, 1998). At other synapses a \textit{Ca}^{2+}/calmodulin-dependent target (Sakaba and Neher, 2001) speeds mobilization of vesicles into the readily releasable pool to partially compensate for its depletion (Dittman and Regehr, 1998; Stevens and Wesseling, 1998; Wang and Kaczmarek, 1998). This target may be \textit{Ca}^{2+}/calmodulin-dependent phosphorylation of synapsin to mobilize vesicles by releasing them from tethering to actin filaments (Sun et al., 2006).

It is now clear that vesicle depletion cannot account for all forms of synaptic depression. At some synapses, depression is due to an inhibitory action of released transmitter on presynaptic receptors called \textit{autoreceptors}. For example, in rat hippocampal cortex, depression of GABA responses is blocked by antagonists of presynaptic GABA\textsubscript{B} receptors (Davies et al., 1990). At other synapses made by dorsal root ganglion neurons and by some brainstem neurons, depression appears to be due to inactivation of \textit{Ca}\textsuperscript{2+} channels during repetitive activity (Jia and Nelson, 1986; Xu and Wu, 2005). This mechanism, however, has been specifically rejected at other synapses (Charlton et al., 1982). Moreover, in other brain-stem neurons and at some central synapses in the sea slug \textit{Aplysia}, depression is due in part to desensitization of postsynaptic cholinergic and AMPA-type glutamatergic receptors (Antzoulatos et al., 2003; Otis et al., 1996; Wachtel and Kandel, 1971). Depression can also be caused by a reduction in sensitivity of NMDA-type
glutamatergic receptors arising from postsynaptic Ca\(^{2+}\) accumulation (Mennerick and Zorumski, 1996).

**Facilitation, Augmentation, and Potentiation Are Due to Effects of Residual Ca\(^{2+}\)**

With few exceptions (Wojtowicz and Atwood, 1988), all the phases of increased short-term plasticity are Ca\(^{2+}\)-dependent in the sense that little or no facilitation, augmentation, or potentiation is generated by stimulation in Ca\(^{2+}\)-free medium. Originally, these phases of increased transmission were thought to be due to the effect of residual Ca\(^{2+}\) remaining in active zones after presynaptic activity and summating with Ca\(^{2+}\) influx during subsequent action potentials to generate slightly higher peaks of [Ca\(^{2+}\)] (Katz and Miledi, 1968; Zucker, 1989). Owing to the highly nonlinear dependence of transmitter release on [Ca\(^{2+}\)], a small residual [Ca\(^{2+}\)] could activate a substantial increase in phasic transmitter release during an action potential while simultaneously increasing MPSP frequency. Temporal correlations between increased spike-evoked transmission after single action potentials or tetani and increases in mPSP frequency supported this idea. But the tetanic accumulation of facilitation, augmentation, and potentiation did not accord quantitatively with predictions of a model of the accumulation of Ca\(^{2+}\) acting at one site (Magleby and Zengel, 1982). Simulations of expected levels of peak and residual [Ca\(^{2+}\)] levels also were unable to account for the full magnitude of facilitation (Yamada and Zucker, 1992). Presynaptic [Ca\(^{2+}\)] measurements using fura-2 confirmed the persistence of residual [Ca\(^{2+}\)] after repetitive stimulation during facilitation, augmentation, PTP, and LTP but showed that it was too weak to explain augmentation or potentiation by simply summating with peak [Ca\(^{2+}\)] from Ca\(^{2+}\) channels at release sites (Atluri and Regehr, 1996; Delaney et al., 1991; Delaney and Tank, 1994). These findings led to the proposal that, in addition to summating with peak [Ca\(^{2+}\)] transients at release sites, Ca\(^{2+}\) acts to increase transmission at one or more targets distinct from the sites triggering exocytosis.

Two possibilities exist. Residual-free Ca\(^{2+}\) could act in equilibrium with such sites to increase transmission (Matveev et al., 2002), or Ca\(^{2+}\) could bind to these targets and activate processes that increase release after residual Ca\(^{2+}\) has dissipated (Yamada and Zucker, 1992). The latter idea is suggested by experiments in which facilitation, augmentation, or potentiation persists when residual Ca\(^{2+}\) should be absorbed by presynaptic introduction of exogenous chelators; however, results of this sort of experiment have not been consistent (Zucker, 1994). The former idea is supported by experiments in which photolabile BAPTA derivatives were injected presynaptically at crayfish neuromuscular junctions (Kamiya and Zucker, 1994). Post-tetanic photolysis to produce enough chelator to suddenly remove residual Ca\(^{2+}\) after conditioning stimulation sharply reduced facilitation within a few milliseconds, and it reduced augmentation and potentiation within about 1 s. These results suggest that Ca\(^{2+}\) can “prime” subsequent phasic release by action potentials by acting at two additional targets distinct from the exocytosis trigger: a fast site responsible for facilitation and a slow one for augmentation and potentiation.

Several additional indications of separate sites of Ca\(^{2+}\) action in synaptic plasticity are as follows:

1. Facilitation grows while secretion decays after one action potential at very low temperature at frog neuromuscular junctions (Van der Kloot, 1994).
2. Sr\(^{2+}\) and Ba\(^{2+}\) selectively enhance facilitation and augmentation, respectively, of both evoked transmitter release and mEPP frequency at frog neuromuscular junctions (Zengel and Magleby, 1980, 1981).
4. Other transformed *Drosophila* carrying an inhibitor of Ca\(^{2+}\)-calmodulin-dependent protein kinase II have impaired facilitation, augmentation, and potentiation (Wang et al., 1994).
5. Mice lacking synapsin I show a specific defect in facilitation but not potentiation in hippocampal pyramidal cells (Rosahl et al., 1993).
6. Facilitation is approximately linearly related to residual [Ca\(^{2+}\)] while secretion involves a high degree of Ca\(^{2+}\) cooperativity (Korogod et al., 2005; Vyshedsiky and Lin, 1997; Wright et al., 1996).
7. Some synapses have an extremely low basal probability of vesicular release to an action potential, which grows dramatically during repetitive activity. At such strongly facilitating synapses, vesicles that are releasable in a train (the readily releasable pool) are not all immediately releasable but first must undergo a final Ca\(^{2+}\)-dependent “priming” step (Millar et al., 2005).
8. Calmodulin-dependent binding of Ca\(^{2+}\) to munc13 isoforms that interact with RIM1 mediates at least the augmentation phase of post-tetanic enhancement of transmission (Junge et al., 2004, Rosenmund et al., 2002).
9. One mechanism by which Ca\(^{2+}\) can facilitate release is by binding to highly saturable endogenous buffers like calbindin-D28k, reducing its capability of rapidly capturing Ca\(^{2+}\) entering in successive action potentials (Matveev et al., 2004). This is the dominant mechanism for facilitation at some (Blatow et al., 2003), but not other (Rozov et al., 2001), vertebrate central synapses.

The duration of augmentation is set by the time course for removal of residual Ca\(^{2+}\) from boutons after a moderate tetanus. Potentiation lasts longer after a strong tetanus because residual Ca\(^{2+}\) is present longer, owing to overloading of the processes responsible for removing excess Ca\(^{2+}\) from neurons. These processes include Ca\(^{2+}\) extrusion pumps, such as the plasma membrane ATPase and Na\(^{+}\)-Ca\(^{2+}\) exchange, and Ca\(^{2+}\) uptake into organelles such as endoplasmic reticulum and mitochondria (David et al., 1998; Fossier et al., 2000; Tang and Zucker, 1997). The accumulation of Ca\(^{2+}\) in mitochondria during a tetanus and its gradual post-tetanic release into cytoplasm produce a prolonged elevation in residual [Ca\(^{2+}\)], and lead to PTP (Tang and Zucker, 1997). In addition, Na\(^{+}\) accumulation during tetanic activity prolongs residual Ca\(^{2+}\) by reducing its extrusion by Na\(^{+}\)-Ca\(^{2+}\) exchange (Zhong et al., 2001). PTP can also arise from a modulation of Ca\(^{2+}\) channels (Habets and Borst, 2006).

Potentiation has also been attributed to Ca\(^{2+}\)-calmodulin-dependent kinase II phosphorylation of synapsin I (Greengard et al., 1993). However, calmodulin inhibitors and kinase inhibitors and mutants failed to affect transmission, facilitation, augmentation, or potentiation at crayfish neuromuscular junctions and mammalian cortical synapses (Zucker, 1994). But short-term synaptic enhancement in Drosophila (see earlier) and LTP in bullfrog sympathetic ganglia may depend on this enzymatic pathway (Baxter et al., 1985; Minota et al., 1991).

Summary

Short-term synaptic plasticity allows synaptic strength to be modulated as a function of prior activity. Synapses may show a decline in transmission (depression) or an increase in synaptic efficacy, with time constants ranging from seconds (facilitation and augmentation) to minutes (potentiation or PTP) to hours (LTP); many synapses show a mixture of several of these phases. Depression may be due to depletion of a readily releasable supply of vesicles or to the inhibitory action of transmitter or enzymatically produced transmitter products on presynaptic autoreceptors. Depression makes synapses selectively responsive to brief stimuli or to changes in level of activity. Frequency-dependent increases in synaptic efficacy are due to the effects of residual presynaptic Ca\(^{2+}\) acting to modulate the release process. At least part of Ca\(^{2+}\) action is mediated by separate targets for facilitation and for augmentation and potentiation. PTP is prolonged after a long tetanus because residual Ca\(^{2+}\) remains in synaptic terminals for minutes after such stimulation. These frequency-dependent increases in synaptic efficacy allow synapses to distinguish significant signals from noise and respond to selected patterns of activity.

References


REFERENCES


