CHAPTER 15

Release of Neurotransmitters

Robert S. Zucker, Dimitri M. Kullmann and Pascal S. Kaeser

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 9). Electrical connections via gap junctions are 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. However, a presynaptic spike will be filtered by the membrane time constant of the postsynaptic membrane, and thus attenuated in the postsynaptic cell, where it is not likely to generate a change in membrane potential of more than a few millivolts.

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 7, 10, 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 (see Chapters 18-20).

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 accelerates tremendously, but very briefly, 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. 15.1A). 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 Ca$^{2+}$ channels that are strategically colocalized with the synaptic vesicles in the active zone (Robitaille et al., 1990). The local intense rise in Ca$^{2+}$ concentration triggers the fusion of docked vesicles with the plasma membrane (called exocytosis; Figs. 15.1B 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 15.1.
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 Kloot, 1991).
6. Quantal size is independent of membrane potential or cytoplasmic acetylcholine concentration altered osmotically (Van der Kloot, 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-amino-pyridine 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. 15.1C).
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 shiibe (Van der Kloot 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 15.3 and section titled “Quantal Analysis”) (Murthy and Stevens, 1998; Ryan et al., 1997).

At neuromuscular junctions, transmitter from one vesicle diffuses across the synaptic cleft in 2 ms and reaches a concentration of about 1 mM at the postsynaptic receptors (Matthews-Bellinger and Salpeter, 1978). 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 10 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 (see 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 (Clements et al., 1992; Tang et al., 1994; Tong and Jahr, 1994) in the cleft to about 1 mM and activates about 30 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 (Clements 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. 15.2.

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 (McGuinness 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 Ca^{2+} channels (Haydon et al., 1994; Robitaille et al., 1990) (mainly N- and P/Q-type Ca^{2+} channels, depending on the synapse) (Dunlap et al., 1995; Wheeler et al., 1994) (see Chapters 11 and 12). Depolarization of the plasma membrane by an invading action potential opens these voltage-dependent Ca^{2+} channels to admit Ca^{2+} ions in the neighborhood of docked vesicles. The local high concentration of Ca^{2+} resulting from the opening of multiple Ca^{2+} channels triggers exocytosis. After exocytosis, some vesicles may rapidly reclose, but most fuse fully with the plasma membrane (Granseth 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 is also regulated by presynaptic [Ca^{2+}] (Hosoi et al., 2009; Leitz and Kavalali, 2011; Schweizer and Ryan, 2006; Yao et al., 2009). 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 7 and 8). 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 are 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 transmitter precursors) in vesicles within about 15 s (Edwards, 1992; Hori and Takahashi, 2012). 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 7). 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. 15.2 comes from ultrastructural and pharmacological experiments. Some of this evidence is outlined in Box 15.2 and Box 15.3.
FIGURE 15.2 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 antiporter, (6) synapsin I-dependent anchoring of vesicles to actin filaments near active zones, (7) releasable vesicles docked and primed in active zones near Ca²⁺ channels, (8) depolarization of nerve terminal and presynaptic bouton by action potential, (9) opening of Ca²⁺ channels and formation of regions of local high [Ca²⁺] ("Ca²⁺ microdomains") in active zones, (10) triggering of exocytosis of docked vesicles comprising quantal units of XMTR released by overlapping Ca²⁺ microdomains, (11) exposure of vesicle proteins and transporters to the synaptic cleft after vesicle fusion, (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.
15. RELEASE OF NEUROTRANSMITTERS

448

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 (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-
BOX 15.3

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.

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 (Rahamimoff et al., 1980; Steinbach and Stevens, 1976). Stimulation in a [Ca\(^{2+}\)]-free medium reduces intracellular [Ca\(^{2+}\)] and MEPSP frequency.
4. 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.
5. Divalent cations that permeate Ca\(^{2+}\) channels, such as Ba\(^{2+}\) and Sr\(^{2+}\), support transmitter release, although only weakly. Cations that block Ca\(^{2+}\) channels, such as Co\(^{2+}\) and Mn\(^{2+}\), block transmission (Augustine et al., 1987); Mg\(^{2+}\) reduces transmission, perhaps by screening fixed surface charge and effectively hyperpolarizing the nerve (Muller and Finkelstein, 1974).
6. Transmission depends nonlinearly on [Ca\(^{2+}\)] in the bath, varying with the fourth power of [Ca\(^{2+}\)], whereas Ca\(^{2+}\) influx remains a linear function of [Ca\(^{2+}\)], indicating a high degree of Ca\(^{2+}\) cooperativity in triggering exocytosis (Llinás et al., 1981).
7. At giant synapses in the stellate ganglion of squid, voltage-clamp recording of the presynaptic Ca\(^{2+}\) current reveals a close correspondence between Ca\(^{2+}\) influx and transmitter release, including an association between the off-EPSP and a delay in Ca\(^{2+}\) current until the end of large pulses (called a tail current, see Chapter 14) (Llinás et al., 1981).
Vesicles are Released by Calcium Microdomains

Single action potentials generate a Ca$^{2+}$ rise as little as 10 nM, which lasts up to 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 of 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$^{2+}$ 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. 15.1C). 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 μM 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, sometimes referred to as a Ca$^{2+}$ channel nanodomain. However, when individual Ca$^{2+}$ channels are labeled with biotinylated ω-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). A similar density of Ca$^{2+}$ channels is seen in active zones of vertebrate central synapses (Holderith et al., 2012; Sheng et al., 2012). Any vesicle docked at such an active zone is likely to be surrounded by several Ca$^{2+}$ channels within a 50–100 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 μm$^2$ active zone, whereas 10 channels open within the more compact active zones of frog saccular hair cells (Roberts, 1994; Yamada and Zucker, 1992). These clusters of overlapping Ca$^{2+}$ nanodomains are often called Ca$^{2+}$ microdomains, and they cooperate in triggering secretion of a vesicle. Calculations of diffusion of Ca$^{2+}$ ions from arrays of Ca$^{2+}$ channels in the presence of a saturable buffer indicate that the [Ca$^{2+}$] at sites where neurotransmitter release is triggered may reach 10–200 μM or higher (Fig. 15.3).

Three indications that [Ca$^{2+}$] locally reaches very high levels in active zones during action potentials are:

1. Highly localized [Ca$^{2+}$] 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 Ca$^{2+}$-sensitive photoprotein n-aequorin-J (Llinás et al., 1992).

2. Estimates of [Ca$^{2+}$] based on the activity of Ca$^{2+}$-activated K$^+$ channels in active zones of mechanosensory hair cells are similar (Roberts et al., 1990).

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

Vesicle Exocytosis is Normally Triggered by Overlapping Ca$^{2+}$ Channel Microdomains of High [Ca$^{2+}$]

Although release of a quantum of transmitter by the sharp local rise of [Ca$^{2+}$] in a calcium nanodomain
Microdomains with high Ca\(^{2+}\) concentrations form in the cytosol near open Ca\(^{2+}\) channels and trigger the exocytosis of synaptic vesicles. (A) In this adaptation of a model of Ca\(^{2+}\) dynamics in the terminal, a set of Ca\(^{2+}\) 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 Ca\(^{2+}\) concentration (y axis) is spatially inhomogeneous. Near the mouth of the channel, the influx of Ca\(^{2+}\) drives the local concentration to as high as 800 \(\mu\text{M}\), but within just 50 nm of the channel the concentration drops off to about 25 \(\mu\text{M}\). The channels are irregularly spaced but are often sufficiently close to one another that their clouds of Ca\(^{2+}\) can overlap and sum. (B) In the active zone (gray), an action potential has opened a fraction of the Ca\(^{2+}\) channels and microdomains of high cytosolic Ca\(^{2+}\) (pink) arise around these open channels as Ca\(^{2+}\) flows into the cell. In the rest of the cytoplasm, the Ca\(^{2+}\) concentration is at resting levels (0.10 \(\mu\text{M}\)), but within these microdomains, and particularly near the channel mouth, Ca\(^{2+}\) 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 Ca\(^{2+}\) concentration in the terminal is now slightly higher (0.11 \(\mu\text{M}\)) than before the action potential. If no other action potentials occur, the cell will pump the extra Ca\(^{2+}\) out across the plasma membrane and restore the initial condition after several hundred milliseconds. (A) Adapted from Roberts, W.M. (1994). The numerical details are for hair cell synapses only.
subsequent to the opening of a single presynaptic Ca\(^{2+}\) channel has been observed (Stanley, 1993), in most synapses exocytosis appears to normally be due to Ca\(^{2+}\) entering through clusters of Ca\(^{2+}\) channels in active zones and contributing to local high [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 Ca\(^{2+}\) current and postsynaptic response (Augustine and Charlton, 1986). If each vesicle were released by Ca\(^{2+}\) entering through a single Ca\(^{2+}\) channel, then increasing depolarizations should recruit additional channel openings and proportionally more vesicle releases (Simon and Llinás, 1985). However, if Ca\(^{2+}\) channel microdomains from neighboring clustered Ca\(^{2+}\) channels overlap at docked vesicles, the [Ca\(^{2+}\)] at each vesicle will rise with increasing depolarization as more channels are recruited, and some cooperativity of Ca\(^{2+}\) action in triggering secretion will be expressed (Zucker and Fogelson, 1986).

2. In some neurons, more than one 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 non-linearly, much as would be predicted by a fourth-order cooperativity, indicating that the Ca\(^{2+}\) microdomains of different channels overlap and summate at the 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 Ca\(^{2+}\) buffer ethylene glycol bis (b-aminoethyl ether)-N,N',N',N'-tetraaetic acid (EGTA) reduces transmission, indicating that the target for Ca\(^{2+}\) is low, with a dissociation constant (K\(_D\)) above 10\(^{-3}\) M. These theoretical predictions accord well with the measured binding characteristics of the C2 domains of synaptotagmin (Millet et al., 2002), the putative Ca\(^{2+}\) target in triggering neurotranssecretion (see below).

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 Ca\(^{2+}\) microdomains is not so much intensified as prolonged, leading to a more nearly linear relationship between increases in Ca\(^{2+}\) influx and transmitter release (Zucker et al., 1991).

5. Large depolarizations admit little Ca\(^{2+}\) as they approach the Ca\(^{2+}\) equilibrium potential; they are therefore accompanied by a reduced 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 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 [Ca\(^{2+}\)] in the active zone, with greater overlap of [Ca\(^{2+}\)] from the larger number of more closely apposed open Ca\(^{2+}\) channels during large depolarizations (Zucker and Fogelson, 1986).

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

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

From the dependence of transmitter release on external [Ca\(^{2+}\)], it is known that at least four Ca\(^{2+}\) ions cooperate in the release of a vesicle. The off-rate of Ca\(^{2+}\) dissociation from these sites must be 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. These theoretical predictions accord well with the measured binding characteristics of the C2 domains of synaptotagmin (Millet et al., 2002), the putative Ca\(^{2+}\) target in triggering neurotranssecretion (see below).

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 \(\sim 10-100 \, \mu\text{M}\), 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 or more in equilibrium with unsaturated release sites, 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 \(\mu\text{M}\) 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 photolyzed 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 \(\mu\text{M}\) 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 [Ca\(^{2+}\)] decreasing) from 140 to 9 \(\mu\text{M}\), followed by a Ca\(^{2+}\)-independent rate-limiting step. In contrast, at the calyx of Held synapse, a [Ca\(^{2+}\)] level of about 10 \(\mu\text{M}\) 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 10 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 to line up at the membrane, docked and ready for release (De Camilli and Jahn, 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 prepacked 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 [Ca\(^{2+}\)] than does the release of docked vesicles. Measurements of [Ca\(^{2+}\)] during stimulation indicate that release correlates well with [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 (Thorson 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 Ca\(^{2+}\) through Ca\(^{2+}\) channels, the amino acid transmitters GABA and glutamate were released at much lower levels of bulk
cytoplasmic $[Ca^{2+}]$ than when $Ca^{2+}$ was admitted more uniformly and gradually across the membrane by use of the $Ca^{2+}$-transporting ionophore ionomycin. Peptides were released at the same low levels of $[Ca^{2+}]$ no matter which method was used to elevate $[Ca^{2+}]$. Thus, only amino acids were sensitive to the difference in $[Ca^{2+}]$ gradients imposed by the two methods and were preferentially released by local high submembrane $[Ca^{2+}]$ caused by depolarization. Apparently, peptides are released by a high-affinity rate-limiting step not especially sensitive to submembrane $[Ca^{2+}]$ 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 co-release of multiple transmitters have been described.

### Summary

$Ca^{2+}$ acts as an intracellular messenger tying the electrical signal of presynaptic depolarization to the act of neurosecretion. At fast synapses, $Ca^{2+}$ 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 $[Ca^{2+}]$ and may be governed by a $Ca^{2+}$-sensitive rate-limiting step different from that triggering exocytosis of docked vesicles at fast synapses.

### THE MOLECULAR MECHANISMS OF NEUROTRANSMITTER RELEASE

When an action potential reaches the nerve terminal, synaptic vesicles fuse with the plasma membrane with great rapidity, and thus the synapse requires an effective and well-regulated molecular machine. Processes in the terminal must prepare and maintain vesicles for fast, $Ca^{2+}$-triggering of fusion. The presynaptic compartment is equipped with the mechanisms to load vesicles with transmitters, to dock vesicles near the membrane so that the latency to fusion is short, and to rapidly fuse vesicles with the plasma membrane at the active zone upon $Ca^{2+}$-entry. Additionally, a reserve of synaptic vesicles is held in the nerve terminal, and recruited to the plasma membrane during enhanced activity. After fusion, recycling of the protein and lipid components generates a new vesicle for a next round of transmitter loading and exocytosis. For each of these processes, a molecular understanding remains incomplete, but rapid scientific progress in this field has made considerable headway.

### Neurotransmitter Release Requires a Rapid Cycle of Membrane Trafficking

Active neurons sustain the release of transmitter in an ongoing fashion. A bouton in the CNS, for example, may contain a store of 200 vesicles. If a cell fires action potentials at 5 Hz and releases one vesicle with each action potential, the store of vesicles is depleted within less than a minute. Transport of newly synthesized vesicles from the cell body would be too slow to support such a demand. Thus, the nerve terminal has mechanisms to recycle and reload vesicles. Exocytosis of a synaptic vesicle is rapidly followed by its endocytosis and within 15 s the vesicle is again available for release (Balaji and Ryan, 2007; Dittman and Ryan, 2009; Granath et al., 2006; Ryan and Smith, 1995). This pathway is referred to as the exo–endocytic cycle (Fig. 15.2), and this cycle contains multiple control points for modulating the efficacy of the synapse (Castillo et al., 2012; Neher and Sakaba, 2008; Regehr et al., 2009). Modulation of synaptic transmission, commonly known as synaptic plasticity, plays an important role in the development of synaptic connections and in the functioning of the mature nervous system (see also Chapter 18). Understanding the mechanisms of presynaptic modulation is an important goal and requires a detailed understanding of the fundamental machinery itself.

To fuse a synaptic vesicle with the plasma membrane, the lipid bilayers must be brought within a few nanometers of one another. Thus, the hydration shell
around the polar lipid head groups has to be disrupted, which requires overcoming a large energy barrier. A specialized mechanism has evolved to bring the membranes close together to drive fusion. This mechanism is not unique to neurons; membrane fusion occurs in every eukaryotic cell. Other examples include the fusion of recycling vesicles with endosomes, transport from the endoplasmic reticulum (ER) to the Golgi apparatus, or transport from endosomal compartments to lysosomes. At present, it appears that all membrane trafficking steps depend on a core set of proteins (Jahn et al., 2003). 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 intracellular membrane fusion uses a universal mechanism. To the extent that this is true, it is a great boon to cell biology: findings in different systems—for example, the highly developed genetic analysis of membrane fusion in yeast—are relevant to neuroscience. Similarly, the abundance of synaptic vesicles for biochemical analysis and the unparalleled precision of electrophysiological assays to probe vesicle fusion can deepen the understanding of other cellular events.

Although it is clear that different membrane fusion events present variations on a common theme, important details may differ. The nerve terminal is in need of specialized machinery that enables extremely rapid fusion with short latencies. The delay between the arrival of an action potential at a terminal and the secretion of the transmitter is often less than 500 μs (Meinrenken et al., 2003; Sabatini and Regehr, 1996). This places constraints on the neuronal fusion mechanism. As there is no time to mobilize vesicles from a distance, they must be present in a ready-to-fuse state at the release sites. Thus, a fusion-ready complex of the vesicle and plasma membrane is preassembled at release sites. Ca²⁺ binding to this preassembled complex triggers a conformational change to open a pathway for the transmitter to exit the vesicle (Jahn and Fasshauer, 2012; Südhof and Rothman, 2009). Because the volume of the synaptic vesicle is small, the diffusion of transmitter from the vesicle proceeds almost instantaneously as soon as a fusion pore has opened between the vesicle lumen and the extracellular space. The time-critical step comes between the action potential-induced influx of Ca²⁺ and the formation of the fusion pore. The complete merging of the vesicle and plasma membrane can occur on a slower time course. The movement of the vesicle to the release site and biochemical events that need to occur to reach the fusion-ready state of a vesicle can also be slower. A typical CNS synapse only has a few vesicles in this fusion-ready state (Harris and Sultan, 1995; Stevens and Tsujimoto, 1995). Therefore, to respond faithfully to a sustained train of action potentials, docking and priming of synaptic vesicles, the processes that render vesicles fusion-ready, must occur within a few seconds.

Fusion of two lipid bilayers requires energy. The latency of transmission is too short for ATP hydrolysis during fusion, and endocrine exocytosis persists after ATP has been removed (Ahnert-Hilger et al., 1989; Hay and Martin, 1992; Holz et al., 1989; Parsons et al., 1995). Thus, the energy needed to fuse the membranes is stored in the fusion-ready state of the vesicle-membrane complex (Chen and Scheller, 2001; Südhof and Rothman, 2009). This energy is released upon Ca²⁺-influx, providing the force to fuse membranes.

Identifying Presynaptic Proteins: Vesicle Purification and Genetics

An important step in understanding how neurotransmitters are released is the elucidation of the protein constituents of the release machinery in a presynaptic nerve terminal. Two key approaches identified a number of molecules that organize release: a first step toward a molecular understanding of the release machinery was the observation that synaptic vesicles can be purified (Nagy et al., 1976). In parallel, and as described later, genetic screens have provided an independent method for identifying the machinery of transmitter release.

Synaptic vesicles are abundant in nervous tissue (Fig. 15.1). Their unique physical properties (uniform small diameter and low buoyant density) permit purification to homogeneity by 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 (Takamori et al., 2006). One of the first sources for the purification of synaptic vesicles was the electric organ of the Torpedo, a marine elasmobranch. The electric organ is a specialized adaptation of the neuromuscular junction and contains a large amount of protein in synaptic vesicles. Synaptic vesicles can also be isolated from mammalian brain. The protein compositions of synaptic vesicles from these different sources are remarkably similar, demonstrating the evolutionary conservation of synaptic vesicle function. This similarity indicates that many proteins present on the synaptic vesicle membrane perform general functions that are not restricted to a single class of transmitter or specific synapses.

Synaptic vesicles contain a discrete set of abundant proteins, and from the purified vesicles individual proteins could be isolated (Fig. 15.4). Application of mass
spectrometry to a purified synaptic vesicle fraction has led to a comprehensive, quantitative analysis of the synaptic vesicle proteome and the lipid composition of the vesicular membrane (Takamori et al., 2006). The major constituents of a synaptic vesicle are shown in the stoichiometric illustration of a typical synaptic vesicle in Fig. 15.4B. For some proteins, functions are well established, but for others they remain uncertain (Table 15.1). The proton transporter, for example, is an ATPase that acidifies the vesicle by pumping H⁺ protons into the vesicular lumen (Fig. 15.4A). A vesicular neurotransmitter transporter then exchanges protons flowing down their electrochemical gradient with a transmitter to fill the vesicle.

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. Other vesicular proteins are discussed later in this chapter.

Transmitter release depends on many proteins in the nerve terminal in addition to the vesicular proteins; proteins near the plasma membrane and in the cytoplasm are also important. In many cases, the identification of these additional components derived from investigations of vesicular proteins. Two examples serve to illustrate this point. The first example begins with synaptotagmin, a membrane protein that was purified from synaptic vesicles (Matthew et al., 1981; Perin et al., 1990). Immunopurification of synaptotagmin from brain was used to identify its interaction partners. In this manner, a protein called syntaxin was identified as a synaptotagmin-binding protein (Bennett et al., 1992).
Syntaxin resides in the plasma membrane and is now appreciated as one of the critical players in vesicle fusion (see below). As a second example the vesicle protein rab3 (Fischer von Mollard et al., 1990) was found by protein interaction screens to bind to a number of rab3 interacting proteins. Among these, RIM (rab3-interacting molecule) proteins were identified as binding rab3 in a yeast-two-hybrid screen (Wang et al., 1997). Rab3 is a small G-protein that exists in either GTP- or GDP-bound forms. Studies have demonstrated that RIMs are now recognized as a central component of a specialized protein network that forms sites for synaptic vesicle exocytosis in presynaptic nerve terminals. These examples demonstrate that from the identification of synaptic vesicle proteins, an array of additional components of the presynaptic release machinery have been identified, many of which figure in the exo–endocytic cycle.

Genetic screens have provided an independent method for identifying the machinery of transmitter release. One of the most fertile screens for proteins involved in membrane trafficking was performed in yeast. Given that membrane trafficking in yeast operates on principles similar to synaptic 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 was carried out (Novick et al., 1981; Novick et al., 1980) 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. Genetic screens have provided an independent method for identifying the machinery of transmitter release. One of the most fertile screens for proteins involved in membrane trafficking was performed in yeast. Given that membrane trafficking in yeast operates on principles similar to synaptic 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 was carried out (Novick et al., 1981; Novick et al., 1980) 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. Genetic screens for suppressors and enhancers of these secretion mutations uncovered additional components. Subsequently, in vitro assays have been established to study the fusion of vesicles derived from yeast with their target organelles (Conradt et al., 1994; Mayer et al., 1996). Among the secretion mutants and their interacting genes were homologs of some of the proteins discussed above: sec4 encodes a small GTP-binding protein like rab3, and sso1 and sso2 encode plasma membrane proteins that are homologs of syntaxin (Aalto et al., 1993; Salminen and Novick, 1987). Thus, genetic screening for trafficking mutants in yeast isolated a number of proteins that are critical for presynaptic function. This occurred in parallel with the description of the synaptic vesicle proteome, and the two approaches complemented each other in identifying molecular components of the nerve terminal.

The nematode worm Caenorhabditis elegans is another organism in which genetic screens uncovered important proteins for the synapse. An essential screen in C. elegans was carried out in 1974 (Brenner, 1974). The worms were exposed to the potent mutagen ethyl methansulphonate (EMS), and behavioral phenotypes were described upon mutagenesis. This screen yielded important proteins for presynaptic functions. For example, the unc–13 mutation of C. elegans identified a new component of the nerve terminal. Unc-13, and its mammalian homolog Munc13, are important for vesicle priming and for the modulation of synaptic vesicle exocytosis (Augustin et al., 1999; Maruyama and Brenner, 1991; Richmond et al., 1999; Rosenmund et al., 2002). Another critical protein for vesicle fusion, Unc-18, was identified in this screen (Hosono et al., 1992). Parallel experiments in yeast, C. elegans, and vertebrates identified unc-18 gene family proteins (named sec1 in yeast and Munc-18 and nSec-1 in mammals) as essential components of the secretory pathway with a binding activity to syntaxin and its homologs (Aalto et al., 1991; Hata et al., 1993; Pevsner et al., 1994; Verhage et al., 2000). These proteins are generally called SM-proteins (Sec1/Munc18-like proteins).

SNARE Complexes Provide the Force for Membrane Fusion

Three synaptic proteins, synaptobrevin/VAMP, 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. Together with Munc18 and the synaptic vesicle protein synaptotagmin they constitute the core protein machinery for fusion of synaptic vesicles with the plasma membrane (Fig. 15.5) (Söllner et al., 1993; Südhof and Rothman, 2009; Sutton et al., 1998). The neuronal SNARE complex is essential for synaptic transmission. What are these SNARE proteins, and how do they work?
SNAP-25 SNARE protein present on plasma membrane; forms core complex with synaptobrevin/VAMP; essential for fusion

Syntaxin SNARE protein present on plasma membrane; forms core complex with syntaxin and synaptobrevin/VAMP; essential for fusion

TABLE 15.2 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; forms core complex with SNAP-25 and synaptobrevin/VAMP; essential for fusion</td>
</tr>
<tr>
<td>SNAP-25</td>
<td>SNARE protein present on plasma membrane; forms core complex with syntaxin and synaptobrevin/VAMP; essential for fusion</td>
</tr>
<tr>
<td>Munc18/ nSec-1</td>
<td>Syntaxin-binding protein required for all membrane traffic to the cell surface; binds to closed syntaxin and to assembled SNARE complexes, required for fusion</td>
</tr>
<tr>
<td>Complexin</td>
<td>Binds to SNARE complexes; inhibits and activates vesicle fusion</td>
</tr>
<tr>
<td>Tomosyn</td>
<td>Binds to SNARE complexes; inhibits vesicle fusion</td>
</tr>
<tr>
<td>NSF</td>
<td>ATPase that disassembles 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>DOC2 &amp; Rabphilin</td>
<td>Ca(^{2+})-binding proteins; promote fusion</td>
</tr>
<tr>
<td>Unc-13/ Munc13</td>
<td>Active zone protein; vesicle priming for release; modulation of transmission</td>
</tr>
<tr>
<td>RIM</td>
<td>Active zone protein; vesicle docking and priming, Ca(^{2+})-channel tethering; modulation of transmission</td>
</tr>
<tr>
<td>ELKS/CAST</td>
<td>Active zone protein; scaffolding and modulation of transmission (?)</td>
</tr>
<tr>
<td>α-liprin</td>
<td>Active zone protein; scaffolding, active zone formation and modulation of transmission (?)</td>
</tr>
<tr>
<td>RIM-binding protein</td>
<td>Active zone protein, Ca(^{2+})-channel tethering (?)</td>
</tr>
<tr>
<td>Neurexin</td>
<td>Transmembrane protein, binds to postsynaptic neuroligin, formation and maintenance of synapses</td>
</tr>
<tr>
<td>Piccolo &amp; Bassoon</td>
<td>Scaffolding proteins in the nerve terminal, possibly tethering vesicles near active zone</td>
</tr>
</tbody>
</table>

Synaptobrevin, also called VAMP (vesicle associated membrane protein) was among the first identified synaptic vesicle proteins (Fig. 15.4) (Südhof et al., 1989; Trimble et al., 1988). It is anchored to the synaptic vesicle by a single transmembrane domain. Its cytoplasmic tail contains an α-helical coiled-coil domain called SNARE-motif, and synaptobrevin contributes this motif to the SNARE complex (Figs. 15.5 and 15.6). Syntaxin has a similar structure, but is located primarily in the plasma membrane (Bennett et al., 1992; Inoue et al., 1992). Syntaxin has a longer cytoplasmic region than synaptobrevin/VAMP. It contains a SNARE-motif similar to synaptobrevin, but it also includes a more N-terminal globular H\(_{abc}\) domain that consists of three additional α-helical regions. The syntaxin H\(_{abc}\) motif folds back onto the SNARE motif, forming a “closed” syntaxin conformation (Dulubova et al., 1999). For SNARE complex formation, syntaxin has to undergo a conformational change; the syntaxin SNARE motif becomes available for SNARE complex assembly when the H\(_{abc}\) domain unbinds from the SNARE motif, a state that is referred to as the syntaxin “open” conformation (Dulubova et al., 1999; Fernandez et al., 1998; Misura et al., 2000). As discussed later, Munc18 binds to closed syntaxin and assembled SNARE complexes, and has critical roles in the conformational switch of syntaxin. SNAP-25, the third presynaptic SNARE protein is also bound to the presynaptic plasma membrane. Unlike syntaxin and synaptobrevin, however, it lacks a transmembrane domain, and is membrane-anchored in its central region by palmitoylation (Chapman et al., 1994; Hess et al., 1992; Oyler et al., 1989). SNAP-25 contributes two α-helical strands to the SNARE complex (Fig. 15.6). The interactions of the SNARE motifs of these three proteins bring the vesicular membranes and the presynaptic plasma membrane close together, as components of this protein complex reside on both opposing membranes (Figs. 15.5 and 15.6). The vesicle-associated SNARE synaptobrevin is also referred to as a v-SNARE, and the target membrane SNAREs, such as SNAP–25 and syntaxin, are called t-SNAREs.

A SNARE complex is found at each membrane trafficking step of eukaryotic cells. Three analogous cases are: exocytosis in yeast, fusion of late endosomes with one another, and fusion of vesicles that form the yeast vacuole. All these examples depend on a four-helix bundle SNARE complex to which proteins on opposing membranes contribute their SNARE motifs (Antonin et al., 2000; Sato et al., 2000; Wickner and Haas, 2000). Abundant genetic and biochemical data argue for an essential role for SNAREs in fusion, and the combination of yeast genetics, in vitro assays,
The molecular mechanisms of neurotransmitter release

FIGURE 15.6 The mechanism of membrane fusion. (A) Fusion of synaptic vesicles is driven by a trans-SNARE complex of four coiled-coil domains contributed by three different proteins located on the vesicular and the target membrane. Zippering of this complex from the N- to the C-terminus forces the two membranes together for fusion. Once fused, the SNARE complex is in a cis configuration. Neurotransmitter release shares this core mechanism with membrane fusion events within eukaryotic cells. (B) Structural ribbon rendition of a tightly assembled SNARE complex. The four-helix bundle consists of one SNARE motif from synaptobrevin and syntaxin, while SNAP-25 contributes two SNARE motifs. Adapted from Rizo and Südhof (2012); structure originally published in Sutton et al. (1998).

synaptic biochemistry, and synaptic physiology has had a synergistic effect in advancing a model for fusion that is generally termed the SNARE hypothesis.

How do SNARE proteins mediate fusion? Early evidence for a key role for SNAREs in synaptic vesicle exocytosis has come from studies with bacterial neurotoxins, and much of our current understanding of how SNAREs contribute to neurotransmitter release is derived from genetic and biochemical experiments. Clostridial 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. These experiments revealed that SNAREs are required for fusion; proteolytic cleavage of any of the SNARE proteins entirely shuts down transmitter release (Blasi et al., 1993a; Blasi et al., 1993b; Schiavo et al., 1993). These toxins (tetanus toxin and various botulinum toxins) are highly specific, recognizing unique sequences within individual SNARE proteins. The catalytic nature of the toxin accounts for its astonishing potency; a few tetanus toxin light chains in a nerve terminal, for example, suffice to proteolyze all synaptobrevin/VAMP. The notion that synaptic vesicle exocytosis is highly dependent on SNAREs is also supported by a multitude of more recent genetic experiments. Mice, flies, or worms in which the SNARE proteins syntaxin, synaptobrevin or SNAP-25 are genetically removed have a dramatic reduction in synaptic vesicle exocytosis (Bronk et al., 2007; Gerber et al., 2008; Nonet et al., 1998; Saifee et al., 1998; Schoch et al., 2001; Schulze et al., 1995; Washbourne et al., 2002; Zhou et al., 2013). However,
in synapses that lack individual SNAREs, vesicles still accumulate near active zones (Broadie et al., 1995; de Wit et al., 2006; Hunt et al., 1994). Despite the presence of vesicles, these synapses are incapable of secreting transmitter. These experiments reveal that SNAREs are necessary for a late step in release before the fusion pore opens and transmitter can diffuse into the cleft. But are they sufficient? This question was answered in a series of in vitro liposome fusion experiments. Although the SNAREs alone cannot explain the fast kinetics of synaptic vesicle fusion, liposomes consisting of only lipid bilayers and SNAREs can fuse slowly in vitro. These experiments suggest that SNARE proteins are sufficient to drive fusion (Nickel et al., 1999).

Although details of the molecular operation of SNAREs are still under investigation, a widely accepted hypothesis of SNARE function in fusion has emerged. The plasma membrane SNAREs syntaxin and SNAP-25 start assembling into a bundle of four α-helices with the vesicular synaptobrevin at the N-terminal ends of their SNARE motifs (Fig. 15.6). SNARE complex assembly then progresses from the N- toward the C-termini of the coiled SNARE motifs. At this stage, the complex is referred to a trans-SNARE complex, as its contributing SNARE motifs reside on opposing membranes. As zippering of the complex proceeds, a tight four-helix bundle forms providing energy for bringing the vesicular and target membranes close enough for fusion (Chen and Scheller, 2001; Cohen and Meliyan, 2004; Hanson et al., 1997; Li et al., 2007; Nichols et al., 1997; Südhof and Rothman, 2009; Sutton et al., 1998). Once fused, the core complex resides on the target membrane, and is referred to as a cis-SNARE complex. As described below, there are specific, ATP-dependent processes that disassemble the SNAREs, and synaptobrevin/VAMP is recycled through endocytosis for additional rounds of neurotransmitter release. The energy for fusion is provided by the disassembly of SNAREs, and is stored in the fusion-ready conformation of the SNAREs before SNARE zippering proceeds.

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? A great diversity of SNAREs in a eukaryotic cell may account for some of this specificity because not all combinations of v- and t-SNAREs form functional complexes. This mechanism, however, is 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; de Wit et al., 2006; Hunt et al., 1994). In addition, the t-SNARE 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). As discussed in the following, the active zone protein machinery of a nerve terminal forms release sites, and likely contributes a substrate for target membrane identification of a vesicle.

**Ca²⁺ Triggering of Neurotransmitter Release by Synaptotagmin**

The most striking difference between synaptic vesicle exocytosis and traffic between other cellular compartments is the rapid triggering of fusion by action potentials. As already discussed, the opening of Ca²⁺ channels and the focal rise of intracellular Ca²⁺ activates the fusion machinery. How does the terminal sense the rise in Ca²⁺? How does the Ca²⁺ open the fusion pore?

The main Ca²⁺ sensor for synaptic vesicle exocytosis is synaptotagmin, an integral membrane protein of the synaptic vesicle (Fig. 15.4) (Brose et al., 1992; Geppert et al., 1994; Perin et al., 1991). Synaptotagmin has a large cytoplasmic portion that comprises modules that bind to Ca²⁺. These conserved domains are called C₂ domains and are found in other Ca²⁺ binding proteins, for example the Ca²⁺-dependent forms of protein kinase C. The more central and C-terminal domains are termed C₂A and C₂B, respectively. The synaptotagmin C₂ domains interact with SNARE proteins and with phospholipids in a Ca²⁺-dependent manner. Consistent with the notion that synaptotagmins are triggers for vesicle fusion, mutations that remove the most abundant isoform of synaptotagmin, synaptotagmin 1, profoundly reduce synchronous synaptic transmission in mice, flies and worms (Broadie et al., 1994; DiAntonio et al., 1993; DiAntonio and Schwarz, 1994; Geppert et al., 1994; Littleton et al., 1993; Nonet et al., 1993). A key experiment to establish the function of synaptotagmin as a Ca²⁺ sensor was to introduce subtle mutations in the C₂ domains of synaptotagmin that shifted the Ca²⁺-binding affinity in vitro. In animals bearing these mutations, presynaptic release probability was changed, indicating that synaptotagmin acts as a Ca²⁺ regulator of synaptic vesicle exocytosis (Fernandez-Chacon et al., 2001; Pang et al., 2006). Each C₂ domain of synaptotagmin contributes to Ca²⁺-triggering. Mutating the Ca²⁺-binding residues of the C₂B domain largely blocks synchronous neurotransmitter release (Mackler et al., 2002), whereas abolishing Ca²⁺ binding of the C₂A domain decreases exocytosis less dramatically, but it changes the Ca²⁺-cooperativity of release and provides a charge switch in its binding
THE MOLECULAR MECHANISMS OF NEUROTRANSMITTER RELEASE

FIGURE 15.7 The exo-endocytic SNARE/SM protein cycle. SNARE and SM proteins change their association with each other and with the membrane during the synaptic vesicle cycle. The cycle illustrates the various transient stages of monomeric SNAREs and SNARE complexes bound to Munc18 and synaptotagmin. Munc18 plays a critical role (see text) in guiding syntaxin from its closed conformation (top) to its open conformation associated with SNARE complexes. Ca$^{2+}$ binding to synaptotagmin is the main trigger for synchronous transmitter release upon opening of voltage-gated Ca$^{2+}$ channels. NSF and α-SNAP are required for ATP-dependent disassembly of SNARE complexes to bring the vesicles and the release site back to a state that allows for a new round of fusion.

pocket that is critical to trigger fusion (Robinson et al., 2002; Shin et al., 2009; Striegel et al., 2012). Together, these observations suggest that Ca$^{2+}$-binding to both synaptotagmin C$_2$ domains contributes to Ca$^{2+}$-triggering of neurotransmitter release (Fig. 15.7).

An intense effort has been made to determine how synaptotagmin triggers fusion, and a working model has been developed. It is based on structural and biochemical studies of synaptotagmin, combined with functional analyses of synaptic transmission in genetically modified mice. The synaptotagmin C$_2$-domains are Janus-shaped protein domains with a top surface that binds Ca$^{2+}$ and membrane phospholipids, and a charged bottom surface that also binds to membrane phospholipids (Arac et al., 2006; Fernandez et al., 2001; Xue et al., 2008). The current model suggests that these two C$_2$ domain surfaces bind to the opposing vesicular and presynaptic membranes when Ca$^{2+}$ enters the nerve terminal. Ca$^{2+}$-binding to the C$_2$ domains increases their membrane penetration, bringing the two membranes close enough to induce SNARE-zippering to generate the force responsible for membrane fusion (Pang and Südhof, 2010; Rizo and Rosenmund, 2008).

The loss of the fast Ca$^{2+}$ sensor synaptotagmin, however, does not completely abolish Ca$^{2+}$-evoked synaptic transmission, indicating that additional Ca$^{2+}$ sensors are involved in synaptic vesicle exocytosis (Broadie et al., 1994; DiAntonio and Schwarz, 1994; Geppert et al., 1994; Sun et al., 2007). These additional Ca$^{2+}$ sensors for synaptic vesicle exocytosis could either be synaptotagmin isoforms, or other proteins with Ca$^{2+}$-binding modules. Intense efforts are being made to identify these additional Ca$^{2+}$ sensors, and to determine their roles in action potential evoked synchronous and asynchronous release, and in
spontaneous miniature neurotransmitter release (Groffen et al., 2010; Wen et al., 2010; Yao et al., 2011).

Synaptotagmin may also have functions that are different from Ca\(^{2+}\)-triggering. At some synapses it is involved in endocytosis and in vesicle docking (de Wit et al., 2009; Haucke and De Camilli, 1999; Jorgensen et al., 1995; Poskanzer et al., 2003; Reist et al., 1998; Wang et al., 2011). Because these additional functions also affect neurotransmitter release, and because they may be Ca\(^{2+}\)-dependent as well, they have complicated the analyses of synaptotagmin as a Ca\(^{2+}\) sensor.

**Molecular Control of SNARE Complex Assembly**

SNARE complex assembly forces the two membranes to fuse in close proximity of each other. In order to control when and where synaptic vesicle release occurs, it is critical to restrict and regulate SNARE complex assembly and disassembly. Mechanisms have been identified by which SNARE complexes are organized, and several classes of proteins have emerged as controllers of SNARE complex assembly (Jahn and Fasshauer, 2012; Südhof and Rothman, 2009). Two fundamental processes have been described: Munc18 proteins operate as clamps that control the assembly of four helix bundles, whereas complexin contains an α-helical domain that incorporates into SNARE complexes.

Munc18, also called nSec-1, is a member of the SM-protein family (Hata et al., 1993; Pevsner et al., 1994). It folds into an arch-shaped clamp (Misura et al., 2000), which apparently evolved to bind to four-helix bundles, a common structure in SNARE complexes (Südhof and Rothman, 2009). Munc18 was initially found to bind directly to syntaxin in its non-fusogenic closed conformation (Dulubova et al., 1999; Hata et al., 1993; Misura et al., 2000), which consists of a four-helix bundle that includes its SNARE motif and its α-helical Habc domain (Fig. 15.6). In this structure Munc18 appears to inhibit SNARE complex assembly, and thus fusion. However, the absence of Munc18 induces a complete loss of synaptic transmission, indicating a positive role in fusion (Verhage et al., 2000). Recent investigations have helped to solve this riddle by uncovering additional modes of interaction between Munc18 and SNAREs. Besides binding the α-helical regions of closed syntaxin, the outer surface of Munc18 also binds to an N-terminal region of syntaxin, and the clamp itself also binds to the four-helix bundle of the assembled SNARE complex (Dulubova et al., 2007; Dulubova et al., 2002; Khvotchev et al., 2007; Shen et al., 2007; Yamaguchi et al., 2002). It is thought that the interaction of Munc18 with the syntaxin N-terminal peptide allows the clamp to unbind from the closed conformation, but to remain bound to syntaxin. As the syntaxin SNARE motif binds with the other SNAREs, Munc18 binds to the four helix-bundle of the assembling trans-SNARE complex. These sequential steps are apparently necessary for SNARE complex assembly, and to explain the requirement for Munc18 in fusion (Jahn and Fasshauer, 2012; Südhof and Rothman, 2009).

Complexin was identified in biochemical experiments as another protein that binds to SNARE complexes (Chen et al., 2002; McMahon et al., 1995), and it has crucial functions in regulating vesicle fusion. Genetic loss-of-function experiments in flies and mice indicate that complexin performs a dual function in release: It clamps SNARE complexes in an arrested state to inhibit spontaneous fusion, and it facilitates SNARE activity during action potential-induced release (Giraudo et al., 2006; Huntwork and Littleton, 2007; Tang et al., 2006; Xue et al., 2007). It is believed that complexin acts either by stabilizing partially zipped SNARE complexes, or by competing directly with synaptobrevin in SNARE complex assembly (Jahn and Fasshauer, 2012; Südhof and Rothman, 2009). In the presence of Ca\(^{2+}\), synaptotagmin triggers fusion, possibly by displacing complexin from SNARE complexes (Krishnakumar et al., 2011; Kummel et al., 2011; Li et al., 2011; Yang et al., 2010).

A third synaptic protein, tomosyn, may also regulate SNARE complexes through a SNARE-like α-helix. Tomosyn can bind to complexes of syntaxin and SNAP-25 and prevent the incorporation of synaptobrevin into the complex, thereby inhibiting SNARE complex assembly, and hence transmitter release (Gracheva et al., 2006; McEwen et al., 2006; Pobbati et al., 2004).

Although many molecular details remain under investigation, SM proteins such as Munc18 are evidently designed to bind to four-helix bundles to guide their assembly, and cycle with SNAREs during the exo-endocytic vesicle cycle (Fig. 15.7). Additional proteins with α-helical domains, such as complexin and tomosyn, also bind to SNARE motif bundles, possibly in a competitive manner with SNAREs and synaptotagmin, to control their assembly and zipping.

**Energy-Dependent Disassembly of SNAREs After Fusion**

Synaptic vesicle fusion requires energy. Thus, as vesicles move through their exo—endocytic cycle, energy must be added to the system. N-ethylmaleimide-sensitive factor (NSF), an ATPase involved in membrane trafficking, is a likely source. NSF was first identified as a required cytosolic factor...
in an \textit{in vitro} trafficking assay (Block et al., 1988). The importance of NSF was confirmed when it was found to correspond to yeast Sec18, the product of a gene essential for secretion (Eakle et al., 1988). NSF hexamers bind a cofactor called o-SNAP (soluble NSF attachment protein, not to be confused with the unrelated SNARE SNAP-25) or Sec17 and this complex, in turn, can bind to the SNARE complex. When ATP is hydrolyzed, the SNARE complex is disassembled into its component proteins (Sollner et al., 1993). NSF acts after the membrane fusion step (Banerjee et al., 1996; Mayer et al., 1996). The tight \textit{cis}-SNARE complex needs to be dismantled so that synaptobrevin/VAMP can be recycled to synaptic vesicles while the other SNAREs remain on the plasma membrane (Fig. 15.7). If the energy of forming a tight SNARE complex is part of the energy that drives fusion, restoring SNAREs to their dissociated state is an important part of the energetics of membrane fusion. Much of the energy required for fusion is added by the ATPase activity of NSF: as it disassembles SNARE complexes, it recharges the individual SNAREs by elevating them into an energetic state that allows for reassembling a SNARE complex capable of a new round of fusion.

**An Active Zone for Docking and Priming Vesicles for Fusion**

SNAREs and synaptotagmin act late in the fusion reaction, and fusion itself is an extremely rapid process. To enable fast action potential-induced Ca$^{2+}$-triggering of fusion, many preparatory and regulatory steps must precede the action of SNAREs and Ca$^{2+}$ sensors. These steps occur at a dense protein network called the active zone of a synapse. The active zone is attached to the presynaptic plasma membrane and forms highly organized sites for the release of neurotransmitters (Fig. 15.8, Couteaux and Pecot-Dechavassine, 1970; Södhof, 2012). Detailed tomographic electron microscopy allows a more precise look at the association of vesicles and plasma membrane at the active zone. At frog neuromuscular junctions an electron-dense “fuzz” adjacent to the presynaptic plasma membrane is a highly ordered structure—a lattice of proteins that connect the vesicles to a cytoskeleton, to one another, and to the plasma membrane (Harlow et al., 2001) (Fig. 15.8). In typical synapses of the vertebrate brain, active zones form a hexagonal grid with depressions for docking and fusion of synaptic vesicles (Fernandez-Busnadiego et al., 2010; Limbach et al., 2011; Pfenninger et al., 1972). In other synapses, such as photoreceptors, hair cells, and many insect synapses, the structures are more elaborate and include ribbons, dense bodies, or T-bars that extend into the cytoplasm and appear to have a special relationship with nearby vesicles (Zhai and Bellen, 2004).

Several proteins are concentrated at the active zones of central synapses. Among them are RIM, Munc13, ELKS/CAST (for glutamine[E], leucine[L], lysine[K], and serine [S]-rich protein, or cytomatrix at the active zone-associated structural protein), liprin-α, RIM binding protein (RIM-BP), and piccolo/bassoon (Table 15.2 and Fig. 15.8). They are part of the machinery that defines the active zone as the target for synaptic vesicle fusion. They fulfill at least three critical functions. During docking, they physically tether the vesicle close to the site of release. Vesicles then undergo a molecular maturation called priming, during which they are elevated into a fusion-ready state. For Ca$^{2+}$ triggering of fusion, docking and priming are mandatory steps that prepare synaptic vesicles for ultrafast fusion when an action potential arrives. Active zones are also important in concentrating Ca$^{2+}$ influx into Ca$^{2+}$ microdomains, and they do so by physically tethering presynaptic Ca$^{2+}$ channels to release sites at the active zone. The underlying molecular mechanisms are only partially understood, and are currently under intense investigation.

A shared property of active zone proteins is that they are large, often around 200 kDa, and they typically consist of protein scaffolding modules that are connected by linker sequences (Fig. 15.8). Active zone proteins are also involved in modulating presynaptic plasticity. Thus, the processes at the active zone that prepare vesicles for fusion are subject to regulation, and are important players in mediating plasticity of synaptic vesicle exocytosis (Gundelfinger and Fejtova, 2012).

One of the most interesting proteins at the active zone is Unc-13 and its mammalian homolog Munc13, a protein first discovered in C. elegans by virtue of its neurological phenotype when mutated (Brose et al., 1995; Maruyama and Brenner, 1991). Munc13 participates in multiple important protein complexes at the active zone. With its N-terminal C$_2$-domain, Munc13 contributes to a tripartite complex with the active zone protein RIM and the synaptic vesicle protein Rab3 (Betz et al., 2001; Dulubova et al., 2005; Lu et al., 2006). With a central region termed the MUN domain, it interacts with SNARE complexes (Basu et al., 2005; Stevens et al., 2005). These interactions orchestrate synaptic vesicle priming, for which Munc13 is required (Andrews-Zwilling et al., 2006; Augustin et al., 1999; Deng et al., 2011; Richmond et al., 1999). Munc13 also contains a binding site for diacylglycerol and for Ca$^{2+}$/calmodulin. Via those second messengers, Munc13 plays a critical role in the modulation of transmitter release during short-term plasticity (Junge et al., 2004; Rhee et al., 2002; Rosenmund et al., 2002; Wierda et al., 2007) (see Chapter 18).
**FIGURE 15.8** Models of the presynaptic active zone. (A–C) Fine structure of the active zone at a neuromuscular junction. (A) Active zone (yellow) as revealed by electron tomography, viewed from the side of the plasma membrane with adjacent synaptic vesicles (blue). Active zone protein material extends from the vesicles and connects to a central beam in an organized fashion. (B) Transverse view of active zone material and docked synaptic vesicles adjacent to the plasma membrane (orange). (C) Schematic rendering of an active zone based on the tomographic analysis. An ordered structure aligns the vesicles and connects them to the plasma membrane and to one another. (D) Protein interactions in the active zone of a central nervous synapse. Several families of large multidomain proteins form a dense network that connects to synaptic vesicles and tethers Ca\(^{2+}\) channels close by. (A–C) Adapted with permission from Macmillan Publishers Ltd. (Harlow et al., 2001). (D) Adapted from Kaefer et al. (2011).
RIMs are another family of active zone proteins with important functions in transmitter release. They were identified via their interaction with the synaptic vesicle protein Rab3, and initial studies showed that RIMs are important for controlling release probability and short-term plasticity (Schoch et al., 2002; Wang et al., 1995). Intriguingly, RIMs are also required for the expression of presynaptic long-term plasticity (Castillo et al., 2002; Kaeser et al., 2008). RIMs promote these functions by contributing to docking, priming, and Ca\(^{2+}\) channel tethering (Han et al., 2011; Koushika et al., 2001). RIMs are involved in docking and priming of synaptic vesicles via their N-terminal interactions with Rab3 and Munc13, and they tether Ca\(^{2+}\)-channels to the active zone by directly binding to their C terminus with the RIM PDZ domain (Deng et al., 2011; Gracheva et al., 2008; Kaeser et al., 2011).

In addition to the specializations of the active zone, additional machinery preserves a dense cluster of synaptic vesicles extending into the presynaptic cytoplasm, and there are mechanisms by which active zones interact with the presynaptic plasma membrane. Synapsin and piccolo/bassoon help to tether vesicles to active zones (Hallermann et al., 2010; Leal-Ortiz et al., 2008; Mukherjee et al., 2010; Ryan and Smith, 1995; Sikou et al., 2007). It is unclear how the active zone protein complex itself is tethered to the presynaptic plasma membrane, but potential mechanism could involve transsynaptic interaction molecules such as neurexins and certain types of phosphotyrosine phosphatases (Südhof, 2012).

Closing the Synaptic Vesicle Cycle: Endocytosis and Refilling of Vesicles

After exocytosis, the components of the synaptic vesicle membrane are recovered from the presynaptic plasma membrane, and synaptic vesicles are refilled for further rounds of exocytosis. Several possible mechanisms of vesicle recycling in a presynaptic nerve terminal are under investigation (Dittman and Ryan, 2009).

A prominent pathway for synaptic vesicle recycling is endocytosis of synaptic vesicles after fully collapsed fusion. Endocytosis of synaptic vesicle components, like receptor-mediated endocytosis in other cell types, is mediated by vesicles coated with the protein clathrin (Granath 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 synaptopagmin (Zhang et al., 1994), which may be important in recruiting the clathrin coat to the vesicle. The pinching off of the clathrin-coated vesicle involves the protein dynamin, a GTPase which forms a ring-like collar around the neck of an endocytosing vesicle (De Camilli et al., 1995; Warrick and Schmid, 1996). A crucial role for dynamin in synaptic vesicle recycling was demonstrated in a temperature-sensitive Drosophila mutant known as shibire. The shibire gene encodes the Drosophila homolog of dynamin. At the nonpermissive temperature, shibire mutants rapidly become paralyzed owing to nearly complete depletion of synaptic vesicles from their nerve terminals (Chen et al., 1991; Koenig and Ikeda, 1989; van der Bie and Meyerowitz, 1991). However, in mice lacking the most prominent dynamins (dynamin 1 and 3 in vertebrate brain), endocytosis was supported quite well during limited activity, but was severely impaired during strong, exogenous stimulation. This suggests that alternative mechanisms support endocytosis in vertebrates, either via other dynamins or by dynamin-independent mechanisms (Ferguson et al., 2007; Raimondi et al., 2011). Due to its Ca\(^{2+}\)-dependence, endocytotic recycling of synaptic vesicles provides another site at which Ca\(^{2+}\) regulates the vesicle 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. 15.2). There are many important additional players in the molecular pathways of endocytosis, and key questions remain to be resolved. For example, it is not known how a synaptic vesicle maintains its lipid and protein composition as it undergoes an exo—endoctytic cycle. Clearly, there must be processes that sort the constituents as vesicles are endocytosed. In addition, endocytosis of synaptic vesicles requires strong bending of the membrane patch that is retrieved, implying that specific proteins control membrane curvature during endocytosis. Proteins such as synaptojanin, amphiphysin and endophilin may support endocytosis of synaptic vesicles by controlling the composition and curvature of the retrieved membrane patch (Dickman et al., 2005; Dittman and Ryan, 2009; Saheki and De Camilli, 2012; Verstreken et al., 2003).

An alternative mechanism for vesicle retrieval is a quick reversal of fusion after fusion pore opening, before the vesicular membrane collapses into the plasma membrane. In this case, a fusion pore opens to allow transmitter release and then rapidly closes to re-form a vesicle. This form of vesicle retrieval is termed “kiss-and-run,” and it has the advantage that all the vesicular components remain together on a single vesicle, immediately available for reloading with transmitter (Alabi and Tsien, 2013; Ceccarelli et al., 1973). The mechanism is potentially quick and energetically efficient, and is employed in nonneuronal exocytotic pathways (Monck and Fernandez, 1992). Its prevalence and relevance at synapses are under intense investigation (Alabi and

II. PHYSIOLOGY OF ION CHANNELS, EXCITABLE MEMBRANES AND SYNAPTIC TRANSMISSION
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 vacular proton pump and a family of vesicular transporter proteins (Fig. 15.4).

The vacular proton pump is a multisubunit ATPase that transports protons from the cytoplasm into the lumen of synaptic vesicles (Nelson, 1992). This results in an electrochemical gradient with an acidic and positively charged vesicular lumen relative to the cytosol. This electrochemical gradient is used as the energy source for active transport of transmitter into the vesicle (Chaudhry et al., 2008).

The vesicular transporters are integral membrane proteins with 12 membrane-spanning domains. The vesicular transporters are distinct from the plasma membrane transporters that remove transmitter from the synaptic cleft and thereby contribute to the termination of synaptic signaling (see Fig. 15.2 and Chapter 7). The distinguishing characteristics include their transport topology, energy source, pharmacology, and structure (Reimer et al., 1998). At least four types of distinct vesicular 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). Genes for each of these classes of transmitter transporters have been cloned (Bellocchio et al., 2000; Blakely and Edwards, 2012; Erickson et al., 1992; Liu et al., 1992; McIntire et al., 1997; Takamori et al., 2000). The type of transporter expressed in a cell dictates the type of transmitter stored in the synaptic vesicles of a particular neuron. For different transmitters, the driving force for loading the vesicles varies according to their charge (Chaudhry et al., 2008). Glutamate is the most common excitatory neurotransmitter in the vertebrate brain. It is negatively charged, and thus is driven predominantly into the vesicle by the vesicular membrane potential. Acetylcholine is the predominant transmitter at the vertebrate neuromuscular junction. Because it is positively charged, like the vesicle lumen, its vesicular loading is driven by the pH gradient. Recent studies suggest that some neurons express multiple transporters, and may co-release classical transmitters (Hnasko et al., 2010; Hnasko and Edwards, 2012; Jonas et al., 1998; Stuber et al., 2010).

In contrast to small chemical transmitters, peptide-signaling molecules, including neuropeptides and hormones, are typically stored in vesicular granules that are larger and have a higher electron density than synaptic vesicles. Thus, they appear as large dense-core vesicles in electron microscopic images. The contents of these granules are not recycled at the release sites; their replenishment requires new protein synthesis in the cell body followed by packaging and transport to sites of release. Because of the slow kinetics of their release, the slow responsiveness of their postsynaptic receptors, and their inability to be locally recycled, these signaling molecules typically mediate regulatory functions.

**Summary**

A critical step in synaptic transmission is fast Ca^{2+} triggering of the fusion of a synaptic vesicle with the plasma membrane. Zippering of the vesicular and plasma membrane proteins of the SNARE complex—synaptobrevin/VAMP, syntaxin, and SNAP-25—is a key step that provides the force for fusion. 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 exocytotic pathway in neurons, however, is an exquisitely fast, highly regulated pathway that must include docking and priming steps before fusion, and is followed by rapid endocytosis and transmitter reloading to recover the vesicles for a next round of fusion. Additional key molecules to mediate these functions are localized in the vesicular membrane, are attached to a highly specialized membrane area called the active zone, or are soluble proteins in the presynaptic cytoplasm. Examples for such proteins are synaptotagmins as Ca^{2+} sensors, complexin, and Munc18 as organizers of SNARE complex assembly, Munc13 and RIM as docking and priming molecules, and dynamin and endophilin as mediators of endocytosis. Together, these proteins build a fundamental core apparatus to create an astonishingly accurate, fast, and reliable means of delivering transmitter to the synaptic cleft in response to an action potential. The synaptic vesicle cycle is also loaded with mechanisms for vesicle membrane and protein recovery. Some of the proteins at the active zone appear to be involved in regulation of release, but we are only beginning to understand the underlying molecular mechanisms.

**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, the detailed trial-to-trial fluctuations in postsynaptic response to presynaptic action potentials also provide a unique insight into the biophysics of transmission. The postsynaptic signal 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, resulting from the release of individual vesicles of neurotransmitter. Examination of the trial-to-trial amplitude fluctuation of the synaptic signal in many cases 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 also reveals the probabilistic processes underlying transmitter release from the presynaptic terminal and the mechanisms by which transmission can be modified by physiological, pharmacological, and pathological interventions.

Transmission at the Frog Neuromuscular Junction is Quantized

Bernard Katz and his colleagues first demonstrated the quantal nature of transmission at the frog neuromuscular junction in the early 1950s. By recording from a muscle fiber immediately under a branch of the motor axon (the “end-plate”), they measured the postsynaptic membrane potential both at rest and in response to stimulation of the axon. Spontaneous miniature end-plate potentials (mEPPs) were observed to occur at random intervals, measuring between 0.1 and 2 mV in amplitude (Fatt and Katz, 1952; Katz, 1969). Stimulating the presynaptic axon produced a postsynaptic signal approximately 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 EPSP 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. 15.9). 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, which is a limiting case of the binomial distribution (Box 15.4), implying that the process governing quantal release may also depend on a similar underlying mechanism.

On the basis of this evidence, Katz and colleagues proposed the following model of transmission, which has gained wide acceptance and is often 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 elicits roughly the same electrical response 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, ... , n quanta released is then given by a binomial distribution, with parameters n and p (see Box 15.4).
- Under conditions of depressed transmission, p is low, and the system approximates a Poisson process. This is the limiting case of the binomial distribution with small p and large n, and is determined by the unique parameter m (see Box 15.4).

The standard Katz model has been supported by similar experiments in other preparations, showing that evoked EPSPs, IPSPs, EPSCs, or IPSCs cluster near integral multiples of a unit. In many cases, this quantal amplitude corresponds to the amplitude of spontaneous miniature postsynaptic signals (potentials or currents) occurring in the absence of presynaptic action potentials (Fig. 15.10) (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. 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 is 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 a mEPP (Fatt and Katz, 1952). Vesicles were subsequently observed in electron
**FIGURE 15.9** 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 \([\text{Ca}^{2+}]\) and \([\text{Mg}^{2+}]\) 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 mEPPs, 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 mEPPs, 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 15.4). The parameters describing this model were the average number of quanta released, \(m\), obtained by dividing the mean EPP amplitude by the mean mEPP amplitude, and the quantal amplitude, \(Q\), and its variance, \(\text{Var}_Q\), obtained from the mEPP 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. (B) from Boyd and Martin (1956). 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 15.1) has since led to almost universal agreement that \(Q\) is the postsynaptic response to exocytosis of a single vesicle of transmitter. Figure 15.11 (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 receptors beyond the edge of the disk remain unoccupied by transmitter (Matthews-Bellinger and Salpeter, 1978).

**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. At neuromuscular junctions, 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 include voltage-sensitive calcium channels (Heuser and Reese, 1973). As described above, entry of calcium ions through these channels raises the intracellular calcium concentration in a small volume immediately adjacent to the channels, triggering 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., 1977). The discrepancy between the
**QUANTAL ANALYSIS**

**BOX 15.4**

**BINOMIAL AND POISSON MODELS**

**Binomial Model**

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, we define *q* as the probability that a given quantum is not discharged in a 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 (a failure of transmission) 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: 3*pq² = 0.432.

The probability that two of three are released, conversely, 3*p²q = 0.288.

Finally, the probability that all three are released is *p*³ = 0.064.

More generally, the relative probabilities *P*ₙ of observing *x* = 0, 1, ..., *n* quanta are given by the binomial distribution:

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

*P*ₓ are also the coefficients of the polynomial expansion of (*p* + *qz*)ⁿ. For the example above, this expansion is 0.216 + 0.432z + 0.288z² + 0.064z³. *G* = (*p* + *qz*)ⁿ 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 coefficients to be calculated.

**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 ∞ 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 considered above (*m* = 1.2), the relative probabilities of observing 0, 1, ... quanta are now approximately

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

**Compound or Nonuniform Binomial Model**

Returning to the binomial model, what happens if different release sites have different, but still independent, probabilities of releasing a quantum? Let us take the following example: *p₁* = 0.1, *p₂* = 0.2, *p₃* = 0.9 (again defining *q₁* = 1 − *p₁*, *q₂* = 1 − *p₂*, *q₃* = 1 − *p₃*). The average quantal content is *P₁* + *P₂* + *P₃* = 1.2.

\[ P_0 = q_1 q_2 q_3 = 0.07, \]
\[ P_1 = p_1 q_2 q_3 + q_1 p_2 q_3 + q_1 q_2 p_3 = 0.67, \]
\[ P_2 = q_1 p_2 p_3 + p_1 q_2 p_3 + p_1 p_2 q_3 = 0.23, \]
\[ P_3 = p_1 p_2 p_3 = 0.01. \]

Generally, the values of *P*ₓ (*x* = 0, 1, ..., *n*) can be obtained from the coefficients of the polynomial expansion of the generating function

\[ G = (p_1 + q_1 z)(p_2 + q_2 z)...(p_n + q_n z) = \prod_k (p_k + q_k z), k = 1, ..., n. \]

**Quantal Parameter p**

Parameter *p* is the probability of exocytosis at an active zone 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. Moreover, because discharge of a quantum may leave a release site empty, *p* should be treated

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 distinct populations of receptors. This could occur if each active zone normally releases only one quantum, and is consistent with the view that *n* is a measure of the number of active zones.
as a product of two probabilities: (1) that a release site is occupied by a quantum \( p_{occ} \), and (2) that a presynaptic action potential evokes release \( p_{ves} \) (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 represent a readily releasable pool (Schikorski and Stevens, 1997), which corresponds to the product of \( n \) and \( p_{occ} \).

The Standard Katz Model Does Not Always Apply

Before accepting the standard Katz model in all its details, we must examine some of its implications more closely 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 Forrester, 1975; Kuffler and Yoshikami, 1975). Thus, for the quantal amplitude to be constant at different release sites, a uniform population of vesicles must be available to be released from each site, as well as receptors with identical properties opposite each release site. Electron microscopic 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. Similarly, although postsynaptic receptors are clustered opposite the active zone, their density and properties may not be uniform among 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 might not be described by a simple binomial (or Poisson) 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 quantization and is 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 secretory vesicles in mast cells can release some of their...
Monte Carlo simulation of quantal release of acetylcholine at the neuromuscular junction. The postsynaptic receptors are represented as a sheet of spheres—white if unliganded, gray if singly liganded, and black if doubly liganded. Presynaptic structures and the ends of the junctional folds are not shown, and the effect of acetylcholinesterase is not modeled. (A) Thirty microseconds after synchronous release of 9500 molecules of acetylcholine (small gray spheres) from a point source opposite the central fold. (B) Postsynaptic response, showing an effectively saturated area at the center, opposite the release site, surrounded by singly bound and unbound receptors. From Bartol et al., 1991.

Contents 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 a subject of intense investigation and debate. 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 condition were not satisfied, it would be difficult to identify the quantal amplitude, and clustering of amplitudes at integral multiples of the quantum would be concealed. If we assume that individual channels act independently of one another, the variance of the quantal current arising from their stochastic opening is described by the binomial formula

\[ \text{Var} = \mu^2 p_0 (1 - p_0), \]

where \( \mu \), \( k \), and \( p_0 \) refer to the single channel current, the number of channels, and their probability of opening in response to transmitter release, respectively. Because the average quantal current amplitude is \( k p_0 \), the coefficient of variation of the quantal amplitude is

\[ \sqrt{\frac{1}{p_0} - 1}. \]

A low quantal variability, which is required for quantal behavior to be detected, therefore implies either a large number of channels, \( k \), or a high probability of opening, \( p_0 \), in response to transmitter release.

Postsynaptic Summation and Distortion of Signals

Postsynaptic currents or potentials arising from different release sites must sum linearly. If this is not satisfied, clustering of evoked postsynaptic signals may not occur at integral multiples of a quantal amplitude. If the postsynaptic membrane becomes appreciably depolarized as a result of the activation of many receptors, quanta may no longer summate linearly, either because the driving force for ion fluxes decreases or because voltage-gated channels open to cause regenerative currents to flow.

Stationarity

The state of the synapse must be relatively stable with time. A drift in the release probability with time could preclude a binomial or Poisson model, and changes in the quantal amplitude could prevent clear clustering in the distribution of evoked signals.

Thus, many of the requirements for the standard Katz model cannot realistically be expected to hold in all cases. In the presence of nonuniformity of 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 release at most one vesicle at a time (Gulyás 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 brainstem 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). At high action potential frequencies, compound vesicle fusion can occur (He et al., 2009) generating giant mEPSCs or evoked EPSCs composed partly of these compound quantal events (see Chapter 18). 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 multisynaptic release has also been obtained at small glutamatergic synapses, where the postsynaptic signal amplitude, and estimates of local glutamate concentration derived from competition with rapidly equilibrating low-affinity antagonists, can, under certain conditions, vary with the release probability (Christie and Jahr, 2006; Huang et al., 2010; 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) and at ribbon synapses of sensory cells (Grant et al., 2010; Graydon et al., 2011; Singer et al., 2004).

Nonuniform Release Probabilities

In the mammalian spinal cord, release probabilities may vary among sites supplied by an individual muscle afferent. Postsynaptic signals have been shown to fluctuate in a manner that is best described by allowing the individual release probabilities in a binomial model to vary (Jack et al., 1981; Silver, 2003). One way in which this can arise is that release sites are often segregated in different terminals that may be subject to
differing amounts of ionic presynaptic inhibition (Walmsley et al., 1987).

**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; Edwards et al., 1976; Faber and Korn, 1982; Jonas et al., 1993). At some excitatory synapses, the glutamate content of a quantum appears to be sufficient to bind a large proportion of the available contents but also by the number of available receptors (Korn, 1982; Jonas et al.; Edwards et al., 1990; Jonas et al., 1993). 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 10). The quantal content sampled by NMDA receptors is often larger than that sampled by AMPA receptors (Kullmann, 1994). Because NMDA receptors are unable to open at hyperpolarized membrane potentials, such synapses are functionally silent in the absence of postsynaptic depolarization (Kullmann, 2003). This discrepancy in signaling 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.

**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 separated into variability from trial to trial at an individual release site (intra-site) and variability among sites (inter-site). 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. 15.10) implies that intra-site variability can be modest, and the wide range of amplitudes of miniature events principally indicates a large inter-site variability. Thus, the mean amplitude of spontaneous miniature events cannot be used as a guide to the quantal amplitude underlying an evoked synaptic signal.

Recently a large body of diverse evidence has accumulated suggesting that spontaneous miniature synaptic events may arise from a pool of vesicles distinct from the vesicles fusing during evoked release (Kavalali et al., 2011). This hypothesis remains controversial, but to the extent that spontaneous and evoked quanta have different amplitudes or kinetics, the use of mPSPs in quantal analysis to derive the statistical properties of evoked release is problematic.

**Quantal Parameters Can Be Estimated from Evoked and Spontaneous Signals**

A frequent goal is to establish, with a reasonable degree of precision, the quantal amplitude, \( Q \), the average quantal content, \( n \), 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 (Jack et al., 1981) makes a complete statistical description of transmission much more difficult to achieve. The parameters that need estimation must then also include the release probability and 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 have 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 Molgó, 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. 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 other than can be accounted for by sampling artifact (Fig. 15.12).

The Kolmogorov—Smirnov test can 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, 1990; Fatt and Katz, 1952). Alternatively, if strontium is substituted for calcium in the extracellular solution, evoked neurotransmission can be desynchronized (McLachlan, 1977) and individual quanta detected (Goda and Stevens, 1994).
Quantal Amplitude Estimation

When $Q$ cannot be estimated from the amplitude distribution of spontaneous mEPSCs or mIPSCs, it can sometimes 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. 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, both these methods require the number of quanta released on any given trial to be estimated reliably. Recordings are often contaminated by noise and because the quantal amplitude has some intrinsic variability, indirect methods often have to be used to estimate $m$. The first of these is to count the proportion of trials resulting in a failure of transmission ($N_0$) of $N$ trials. The first term of the Poisson expansion (Box 15.4) gives the probability of observing zero quanta released and is equal to $e^{-m}$. Therefore, $m = \log(N/N_0)$. 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, $Var$, 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

$$m = \frac{1 + CV_Q^2}{CV_R^2 - Var/\bar{R}}$$

where $\bar{R}$ 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, ..., $n$ quanta released 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 $p = 1 - Var_m/n$ 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 $p = 1 - \sqrt[n]{N_0/N}$. 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):

$$p = 1 + CV_Q^2 - \bar{R} \cdot CV_R^2 - Var/\bar{R}$$

This method requires that both the average quantal size $Q$ and its coefficient of variation $CV_Q$ are known, 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: $p_1, p_2, ..., p_n$. If the sample were perfect, they could be obtained by treating the observed proportions of trials resulting in 0, 1, ..., $n$ quanta as the polynomial expansion of $\prod_k (p_k + q_k)$ (see Box 15.4). Solving the polynomial would then yield $p_1, p_2, ..., p_n$ (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 finite sampling into account (Kullmann, 1989).
Noise Deconvolution

In many cases, the proportion of trials resulting in 0, 1, ..., n trials cannot be determined unambiguously because of excessive noise or quantal variance, 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) 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 noise deconvolution, can give an estimate of quantal amplitude if there are periodic inflections in the solution (Kullmann and Nicoll, 1992).

Figure 15.13 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 \( x^2 \) (Chi squared) test. If the fit is unsatisfactory, the model can tentatively be rejected with the corresponding degree of confidence. However, a 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
FIGURE 15.13 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 artifacts but 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 re-convolved 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 $CV_Q < 0.3$. Data from Kullmann (1993).
difficult to establish all of 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, \( Var_I \) the coefficient of variation of the underlying signal, \( CV_S \), is given by

\[
CV_S = \frac{\sqrt{Var_R - Var_I}}{R}.
\]

\( CV_S \) 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 (Edwards et al., 1989). 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 the quantal amplitude.

If a Poisson model is assumed, further information can be obtained by plotting the ratio of \( 1/CV^2 \) before and after a manipulation against the corresponding ratio of mean amplitude \( R \) (Fig. 15.14) (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 \( 1/CV^2 \), so the data points 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, then the results are slightly different, because the variance \( np(1-p) \) is less than the mean \( np \). A plot of the ratio of \( 1/CV^2 \) against the ratio of \( R \) in this case falls on the line of identity for manipulations that increase \( n \) and above it for manipulations that increase \( p \).

This method, although easy to use, depends heavily on the assumption that a simple binomial or Poisson model applies. As soon as this assumption is relaxed, a wide range of explanations can be put forward for virtually any outcome (Faber and Korn, 1991). The method is also sensitive to changes in the quantal coefficient of variation, so it must be assumed that changes in \( Q \) are accompanied by proportional changes in \( Var_Q \). It is, however, often possible to test whether these assumptions are correct, by deliberately applying manipulations that are known to alter either \( n \), or \( p \), or \( Q \). For instance, \( n \) can sometimes be altered by varying the number of presynaptic axons stimulated, while \( p \) can be manipulated in relative isolation by altering the extracellular \([Ca^{2+}]/[Mg^{2+}]\) ratio or by applying drugs known to act presynaptically. And \( Q \) can be scaled by applying low concentrations of postsynaptic receptor blockers or by manipulating the driving force for the synaptic current. This set of experiments establishes characteristic trajectories for a plot of ratios of \( 1/CV^2 \) against ratios of \( R \), which may or may not coincide with those expected of binomial or Poisson models. These trajectories can then be compared with the effect of the new manipulation under investigation. If the observed ratio of \( 1/CV^2 \) plotted against the ratio of \( R \) runs along one of the trajectories established for manipulations of \( n \), \( p \) 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 NMDA-receptor mediated long-term potentiation at glutamatergic synapses in the mammalian brain: although variance analysis consistently shows an increase in quantal content, implying a presynaptic alteration in \( n \) and/or \( p \) (Bekkers and Stevens, 1990; Malinow and Tsien, 1990), an alternative explanation that has received substantial support from alternative methods is that postsynaptic receptor clusters are uncovered at previously silent sites (Kullmann, 1994, 2003). That is, a postsynaptic modification, with \( Q \) at some sites switching to nonzero values, mimics a presynaptic alteration.

In practice, a presynaptic alteration in transmitter release can only be safely inferred from variance analysis if this is concordant with the results of other methods, such as showing that the rate of failures of transmission has changed. Another important method to differentiate a prep- from a postsynaptic change in transmission 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 (see Chapter 18). Therefore, a change in the ratio of the responses to the two pulses implies that at least part of...
FIGURE 15.14 Coefficient of variation method for determining the site of modulation of synaptic transmission. (A) The possible excursions in the ratio of $1/CV^2$ plotted against the ratio of the mean response amplitude $\bar{R}$ 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 modulations. 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 $\bar{R}$, as expected from a purely postsynaptic modification. Extracellular theophylline or Cs $^+$, an increase in the $1/CV^2$ 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 Manabe et al. (1993).

the effect of the manipulation is mediated presynaptically (Manabe et al., 1993).

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 described above (Clements and Silver, 2000). 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 parabolic function of the mean postsynaptic response amplitude $\bar{R}$: $Var_R = A\bar{R} - BR^2$. The quantal parameter estimates are then given by $Q = A/(1 + CV_Q^2), p = BR(1 + CV_Q^2)/A$, 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. 15.15). This will have a parabolic shape if the binomial model holds and enough data can be obtained under each condition (an important limitation of the method). CVQ can be indirectly estimated from the amplitude distribution of evoked signals when p is so low that multi-quantal release rarely occurs, assuming that these can be differentiated from failures. Alternatively, the distribution of spontaneous action-potential independent signals yields an upper estimate of CVQ. A and B are then obtained by fitting the equation $\text{Var}_R = A\bar{R} - B\bar{R}^2$, and CVQ inserted to yield estimates for Q, n, and p. If the simple binomial assumption does not apply, the variance-mean plot can still have a parabolic shape. 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: fitting the initial slope of the variance-mean plot with $\text{Var}_R = A\bar{R}$ yields 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

An alternative method for estimating the release probability at glutamatergic synapses makes use of the pharmacological properties of 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. With the development of fluorescent dyes that label presynaptic vesicles as they are recycled (Box 15.3), 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). With appropriate calibration, fluorescence methods can yield estimates of total recycling pool of vesicles, the size of the readily releasable pool, and the probability of release of individual vesicles within this pool (Ermolyuk et al., 2012).

Fluorescence methods have also been applied to resolve postsynaptic Ca$^{2+}$ transients at the Drosophila larval neuromuscular junction (Peled and Isacoff, 2011) at the level of distinct release sites supplied by a presynaptic axon. This method demonstrated that release probabilities at different sites are nonuniform.

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.

References


REFERENCES

Harris, K.M., Sultan, P., 1995. Variation in the number, location and size of synaptic vesicles provides an anatomical basis for the non-uniform probability of release at hippocampal CA1 synapses. Neuropharmacology. 34, 1387–1395.

II. PHYSIOLOGY OF ION CHANNELS, EXCITABLE MEMBRANES AND SYNAPTIC TRANSMISSION


II. PHYSIOLOGY OF ION CHANNELS, EXCITABLE MEMBRANES AND SYNAPTIC TRANSMISSION

Rhee, J.S., Betz, A., Pyott, S., et al., 2002. 3-phorbol ester and diacylglycerol-induced augmentation of transmitter release is mediated by Munc13s and not by PKCs. Cell. 108, 121–133.


II. PHYSIOLOGY OF ION CHANNELS, EXCITABLE MEMBRANES AND SYNAPTIC TRANSMISSION


