Chapter 2
Multiple Modes of Fusion and Retrieval at the Calyx of Held Synapse

Liming He, Benjamin D. McNeil, and Ling-Gang Wu

Abstract
Neurotransmitter in vesicles is released through a fusion pore when vesicles fuse with the plasma membrane. Subsequent retrieval of the fused vesicle membrane is the key step in recycling exocytosed vesicles. Recent application of advanced electrophysiologic techniques to a large nerve terminal, the calyx of Held, has led to direct recordings of endocytosis, individual vesicle fusion and retrieval, and the kinetics of the fusion pore opening process and the fission pore closure process. These studies have revealed three kinetically different forms of endocytosis—rapid, slow, and bulk endocytosis—and two forms of fusion—full collapse and kiss-and-run. These research advancements are reviewed in this chapter.

Keywords
Vesicle fusion, vesicle endocytosis, exocytosis, full collapse fusion, kiss-and-run, capacitance recording, calyx of Held, synaptic transmission, quantal response.

Neurons release neurotransmitter through synaptic vesicle exocytosis, a specialized form of vesicle trafficking whereby synaptic vesicles fuse with the presynaptic...
plasma membrane at the active zone and release their contents into the synaptic cleft. Following exocytosis, vesicles are retrieved from the plasma membrane in a process called endocytosis and refilled with neurotransmitter, forming new vesicles that can be used for further release (1). Thus, coordination of the exocytosis of neurotransmitter and the endocytosis of vesicular components sustains the membrane trafficking of synaptic vesicles.

The modes of exocytosis and endocytosis depend on the behavior of the fusion pore, a molecular structure that forms to connect vesicle membrane and presynaptic plasma membrane during synaptic vesicle fusion. The initial fusion pore may expand rapidly, allowing the vesicles to fully collapse into the plasma membrane (2). This mode is called “full collapse” or “full fusion.” Alternatively, the initial fusion pore opens for a short time without dilating and then closes again, allowing the vesicle to retain its integrity when it discharges its contents. This process is called “kiss-and-run” fusion (3,4). Full fusion leads to a rapid release of all neurotransmitter in a bolus, while kiss-and-run may regulate the rate of neurotransmitter discharge by opening the fusion pore to different degrees (5). The rate of neurotransmitter release may affect the postsynaptic response, and possibly contribute to synaptic plasticity (6).

The modes of fusion may also determine a vesicle’s fate after neurotransmitter release. It is thought that full collapse fusion is followed by compensatory endocytosis, a clathrin-mediated process in which vesicles are reformed from the plasma membrane and severed by the guanosine triphosphatase (GTPase) dynamin (1). The endocytic rate is much slower than kiss-and-run endocytosis, which simply involves closing of the fusion pore. It is plausible that the differing rates of endocytosis may affect synaptic function, as endocytosis is critical for replenishment of various synaptic vesicle pools, which influences the ability of the nerve terminal to maintain transmitter release during repetitive firing (7). Regulation of vesicle availability through the rate of endocytosis, and thus the rate of vesicle cycling, could contribute to the generation of some forms of synaptic plasticity (7). Owing to these potential important roles, the kinetics of endocytosis and its regulation have been intensively studied in the past decade.

Both fusion and retrieval can be monitored in live synapses with imaging and electrophysiologic techniques (8). Compared to imaging techniques, electrophysiologic techniques, including whole-cell and cell-attached capacitance recording techniques, provide faster time resolution and allow for the measurement of the fusion pore opening process and the fission pore closure kinetics. In the past few years, these advanced electrophysiologic techniques have been applied to a large nerve terminal, the calyx of Held, to study the rates and modes of endocytosis. These studies have provided direct recordings of fast endocytosis and individual vesicle fusion and retrieval. Further, the kinetics of the fusion pore opening process and the fission pore closure process have been measured. These results have not been reported at any other synapse, which is at least partly due to the small size of most other synapses that has precluded the application of the electrophysiologic techniques. This chapter discusses what we have learned from electrophysiologic studies of fusion and retrieval at the calyx of Held. We believe that the calyx of Held synapse is an excellent model for the study of vesicle fusion and retrieval, and that
the results obtained at this large synapse can provide useful lessons for further studies of fusion and retrieval at most small conventional synapses.

The Calyx of Held Synapse and the Whole-Cell Capacitance Measurement

The calyx of Held is a glutamatergic nerve terminal that forms part of the relay pathway involved in sound localization in the auditory brainstem (Fig. 2.1a) (9). The calyx of Held arises from globular bushy cells in the anterior ventral cochlear nucleus (aVCN), which project onto principal neurons of the contralateral medial nucleus of the trapezoid body (MNTB; Fig. 2.1b) (9). Electron microscopy (EM) reconstruction of the calyx of Held has shown the presence of approximately 300 to 700 individual active zones, each with about two morphologically docked vesicles (10,11). The large number of active zones helps ensure rapid signaling, as a presynaptic action potential (AP) releases hundreds of vesicles, generating a large

![Figure 2.1](image-url)

**Fig. 2.1** Anatomy of the calyx of Held synapse. (a) A diagram of the adult calyx of Held, a glutamatergic synapse in the mammalian auditory brainstem. Note the large-caliber axon (4 to 12 µm in diameter) that gives rise to the calyx. The postsynaptic cell has relatively short dendrites and an axon with a collateral branch. MNTB, medial nucleus of the trapezoid body. (b) A diagram of the circuitry of the superior olivary complex (SOC), which is involved in computing sound localization from the auditory inputs from both ears. Auditory information from the cochlea is transmitted to the ipsilateral anterior ventral cochlear nucleus (aVCN) through excitatory synapses onto spherical bushy cells (SBCs) (black arrow) and globular bushy cells (GBCs) (lighter neuron beside “GBC” label). The GBCs synapse onto the contralateral medial nucleus of the trapezoid body (MNTB), which makes an inhibitory synapse onto the lateral superior olive (LSO). The LSO also receives excitatory input from the ipsilateral SBCs. It is at the level of the LSO that discharges evoked by interaural intensity differences are first represented as differences in the timing of excitatory and inhibitory inputs. So, the LSO is thought to function as a coincidence detector of binaural signals, whereas the main role of the MNTB is simply to act as a fast, sign-inverting relay station. The principal cells of the MNTB, however, also receive strong inhibitory input from unknown sources, and are organized in a tonotopic map of characteristic sound frequencies. The MNTB is therefore not a functionally homogeneous nucleus, and its output might be modulated by other brainstem nuclei. MSO, medial superior olive. (Adapted from Von Gersdorff and Borst [9], with permission.)
excitatory postsynaptic current (EPSC) that rapidly depolarizes the postsynaptic neuron to the threshold for generating action potentials (9).

Except for the large size of the terminal and the large quantal synaptic output, the calyx of Held synapse acts in a similar way to a conventional fast synapse in the central nervous system (CNS). For example, the terminal releases glutamate in response to presynaptic action potential stimulation. The terminal contains many spherical, clear-core vesicles with a diameter of about 50 nm (10,12). The delay between the peak of a presynaptic action potential and the onset of the EPSC is less than 1 ms (13). Multiple types of voltage-dependent Ca\(^{2+}\) channels control transmitter release at single release sites in the calyx of Held synapse (14,15), as in many other CNS synapses (16). The synapse exhibits short-term synaptic depression and facilitation (9,17). Since the MNTB synapse functionally resembles other fast central synapses in many aspects, the experimental results obtained from this synapse not only may be applied to calyceal synapses, but also may have a more general significance.

Presynaptic patch-clamp measurements of exocytosis and endocytosis rely on the increase in surface membrane area when vesicles fuse with the plasma membrane during exocytosis, and the decrease in area when membrane is retrieved by endocytosis. The changes in surface area can be monitored electrically as changes in membrane capacitance (\(C_m\)) (18). Whole-cell capacitance measurement is best achieved in round cells that are electrically equivalent to a membrane capacitance in parallel with a membrane resistance (Fig. 2.2a) (18). Recent studies demonstrated that this technique can also be applied to the calyx of Held (19–21). Although the calyx is connected with an axon, simulation suggests that the axon does not significantly affect the measurement of the capacitance at the calyx when Lindau-Neher’s technique is used (20); thus, changes in membrane size at the calyx can be accurately measured. Experimental results have confirmed this simulation result. First, when both the EPSC and the presynaptic capacitance were simultaneously recorded at the same synapse, the EPSC amplitude or the charge increased as the capacitance jump (\(\Delta C_m\)) increased (Fig. 2.2b). The relation between EPSC amplitude and presynaptic capacitance jump could be fit by a linear regression line with a slope of about 148 picoamperes per femtofarad (pA/fF) (Fig. 2.2c) (19). Furthermore, by averaging about 2.7 million spontaneous miniature EPSCs (mEPSCs) and the corresponding presynaptic capacitance traces from 459 individual synapses, we found that the presynaptic membrane capacitance jumped by about 65 atofarads (aF) within 1 ms before the onset of the mean mEPSC (Fig. 2.2d) (22,23). As the specific membrane capacitance is 9 fF \(\mu\text{m}^{-2}\) (24), 65 aF corresponds to a vesicle with a diameter of 48 nm, which is similar to the estimate from EM (10). We concluded that the capacitance jump accurately reflects vesicle fusion.

Whole-cell recordings of fusion have advantages and drawbacks, compared to the more common measurement of postsynaptic currents induced by transmitter binding. The drawbacks are that the signal-to-noise ratio is not as high as postsynaptic current recordings, and recording release during a stimulus is not possible. However, one significant advantage is that capacitance measurements provide a better estimate of total release, since they are independent of the functional state of the postsynaptic receptors, while postsynaptic currents are complicated by the effects of receptor saturation, desensitization, and inactivation that skew the relationship between release
and postsynaptic response. In addition to its utility in the study of vesicle fusion, time-resolved capacitance measurement at the calyx provides a powerful technique to study synaptic vesicle endocytosis at a central synapse.

**Fig. 2.2** Capacitance jumps reflect exocytosis at the calyx of Held. (a) An equivalent circuit of a cell in the whole-cell recording configuration. $C_m$ is the membrane capacitance, $R_A$ is the access resistance, and $R_n$ is the membrane resistance. (b) Sample recordings of the presynaptic $\text{Ca}^{2+}$ currents and excitatory postsynaptic currents (EPSCs) induced by 1-ms steps to $+10 \text{mV}$, $0 \text{mV}$, and $-6 \text{mV}$ at the calyx. Presynaptic $\text{Ca}^{2+}$ currents and EPSCs are plotted at the same time scale, whereas capacitance changes are plotted at a different time scale. BL, baseline capacitance; a, b, c, data points. (c) The linear relation between the EPSC amplitude (EPSC Amp, left y axis) and $\Delta C_m$ evoked by a series of 1-ms step depolarizations from $-80 \text{mV}$ to a voltage ranging from $-10$ to $+20 \text{mV}$ at a synapse. Right y axis is approximate vesicle number, calculated from the $\Delta C_m$, assuming an average vesicle diameter of 50 nm. (B,C: Adapted from Sun and Wu [19], with permission.) (d) The mean mEPSC and $C_m$ averaged from 2.66 million fusion events obtained from 459 paired recordings. Traces are shown without filtering. (Adapted from Wu et al [23].)

**A Linear Relation Between the Time Constant of Endocytosis and the Amount of Exocytosis**

Synaptic vesicle endocytosis can be detected as the decay of the capacitance jump to baseline levels after a stimulus. At the calyx of Held this decay is exponential, and is described by the time constant ($\tau$) in an exponential equation. The $\tau$ is the
time at which ~63% of the jump has decayed. Recent studies at the calyx of Held showed that the time course of endocytosis, as measured by whole-cell capacitance recordings, depended on the stimulation intensity. The capacitance change after a single action potential-equivalent (AP-e) stimulus was measured by two labs, using slightly different stimuli. One lab used a 1-ms step depolarization from −80 mV to +7 mV, which induces the release of the same number of vesicles as an action potential (22). After this stimulus, the capacitance jump was about 20 fF, and decayed with a \( \tau \) of about 2.2 seconds (Fig. 2.3a) (21). Earlier measurements indicated that the decay was much faster, at \( \tau = 115 \text{ ms} \) (22), but it was later shown that they were contaminated by the presence of an artifactual capacitance jump with a decay time in the hundreds of milliseconds, which persists after abolishing exocytosis with botulinum neurotoxin E (25) or C (21). Another lab approximated an AP-e with a 4-ms step depolarization from −80 mV to +80 mV, followed by a 1-ms step to +40 mV (25). This was found to evoke approximately three times the amount of release as an action potential (21), and induced a capacitance jump of about 60 fF. After this stimulus, the decay was fit with a \( \tau \) of about 10.4 seconds when the first 500 ms after stimulation was ignored (25). The difference in the \( \tau \) reported by these two labs is at least partly due to the difference in the stimulation intensity and thus the capacitance jump, because when the first lab changed the stimulus protocol and increased the capacitance jump from about 20 fF to about 60 fF, the \( \tau \) increased from about 2.2 to 4.2 seconds (21).

The \( \tau \) after short AP-e trains increased in linear proportion to the net capacitance jump at the end of the stimulus (Fig. 2.3b,c) (21,22). This relationship cannot be attributed to an increase in stimulus frequency alone, as similar capacitance jumps elicited by AP-e trains at frequencies from 20 to 333 Hz had similar decay times (21,22). As further evidence, the trend was also seen when longer AP-e trains or continuous step depolarizations from −80 mV to 0 to 10 mV with durations from 2 to 20 ms were used to evoke exocytosis (Fig. 2.3c) (21,22,25). The lengthening of the time constant also cannot be explained by the increase in global calcium influx due to increased stimulation, because ethylene glycol tetraacetic acid (EGTA), a slow-binding calcium chelator that eliminates the residual calcium transient, did not reduce the \( \tau \) (22). Moreover, clamping presynaptic calcium levels at −1 \( \mu \)M, which exceeds the peak global calcium concentration after 10 AP-e at 333 Hz, did not lengthen the \( \tau \) (22). The duration of elevated local calcium domains can also be ruled out, as the \( \tau \) was similar after 10- and 30-ms step depolarizations from −80 mV to +10 mV, which elicited similar capacitance jumps and similar time courses of endocytosis (22). Thus, it appears that the net accumulation of fused vesicle membrane at the end of a stimulus is itself the cause of the increase in the \( \tau \).

This linear relation is also seen at other synapses. Styryl dye uptake studies at the frog neuromuscular junction show that the endocytic \( \tau \) increases as the duration of a 30-Hz stimulus train of APs increases (26). Similarly, at cultured hippocampal neuron boutons, studies using the genetically encoded exo-/endocytic marker synaptopHluorin demonstrated that the \( \tau \) increased linearly with the number of APs in a 10-Hz train, from <10 seconds after 20 APs to ~90 seconds after 600 APs (27). There is some question about what happens at very mild stimulation, as another
Fig. 2.3 Slow and fast endocytosis at the calyx of Held. (a) Sampled whole-cell capacitance (Cm) response to an action potential-equivalent (AP-e) (1-ms step to +7 mV). The capacitance decay, starting from 200 ms after stimulation, was fit with an exponential function with a time constant of 2.9 seconds. (Adapted from Wu et al [21].) (b) Sampled Cm response to 10 AP-e at 333 Hz. Note that the retrieval time is significantly longer than after 1 AP-e. (Adapted from Sun et al [22].) (c) The relation between the time constant and the amplitude of endocytosis. A plot of endocytosis time constant versus its amplitude after various stimuli. Endocytosis after step depolarizations are indicated by circles, and endocytosis after AP-e is indicated by triangles. Average amount of endocytosis and time constant are as follows: 1-ms step to +7 mV: 20 fF, 2.2 seconds; 1-ms step to +15–30 mV: 57 fF, 4.2 seconds; 20–ms step to +10 mV: 473 fF, 12.0 seconds; ten 20-ms steps to +10 mV at 10 Hz: 880 fF (slow component), 18.4 seconds; ten 20-ms steps to +10 mV at 1 Hz: 1390 fF (slow component), 23.0 seconds; 50 AP-e at 100 Hz: 710 fF, 16.4 seconds; 200 AP-e at 100 Hz: 1025 fF (slow component), 19.6 seconds; 50 AP-e at 100 Hz repeated five times at 1 Hz: 1060 fF (slow component), 18.8 seconds; 50 AP-e at 30 Hz repeated five times at 0.4 Hz: 599 fF (slow component), 18.0 seconds. The data were fit with a linear regression line with a slope of 1.4 seconds/100 fF. (d) Sampled Cm response to 200 AP-e at 100 Hz (bars). The capacitance decay was fit with a biexponential function [gray, $\tau_1 = 1.8$ seconds (377 fF), $\tau_2 = 24.0$ seconds (1020 fF)]. (e) Sampled Cm response to 200 AP-e at 100 Hz with 10 mM ethyleneglycoltetraacetic acid (EGTA) in the pipette. The capacitance decay was fit with a monoeponential function (gray, $\tau = 24.7$ seconds). (c–e: Adapted from Wu et al [21], with permission.)
study with a similar marker, called sypHy, found that this linearity did not hold for trains of up to 40 APs at a 20-Hz frequency (28). However, the authors did see a lengthening of the $\tau$ with stronger stimulation.

A hypothesis to explain this correlation between $\tau$ and net exocytosis has been described (27). In this model, the endocytic apparatus has a fixed rate and limited capacity; when the amount of fusion exceeds the capacity for retrieval, membrane accumulates at the plasma membrane and is cleared at a constant rate. Thus, addition of more fused vesicles will lengthen the $\tau$ in a linear fashion. This model predicts the observed linear relationship between the $\tau$ and the net exocytosis after mild-to-moderate stimulation at the calyx. However, it also predicts a linear decay, while at the calyx (and in cultured hippocampal neurons), membrane retrieval is best fit with an exponential function. Thus, this model may be similar, but not identical, to the physiologic mechanism in the calyx of Held at mild-to-moderate stimulation.

Intense Stimulation Activates Rapid Endocytosis by Increasing the Calcium Influx

Under stronger stimulus conditions at the calyx, the linear relationship between the $\tau$ and net exocytosis no longer applies (21). When a 20-ms pulse stimulation is repeated 10 times at 10 Hz, a fast component with a $\tau$ of 1 to 2 seconds is evident immediately after the end of stimulation. After several seconds it gives way to slow endocytosis, and the total capacitance decline is best fit with a double exponential. The fast component can also be evoked by a train of 200 AP-e at 100 Hz (Fig. 2.3d), suggesting that it is physiologically important.

Slowing the frequency of the 20-ms pulses to 1 Hz allows capacitance measurements to be made between pulses, enabling an examination of how the fast component develops. Under this condition, the initial rate of endocytosis increased from about 28 fF/s after the first pulse to a plateau of about 208 fF/s after the sixth pulse, which corresponds to about six vesicles per second per active zone (21). It is estimated that this high retrieval rate is mostly (two thirds) due to rapid endocytosis, indicating that the fast component of endocytosis becomes dominant during stimulation.

The trigger for fast endocytosis is calcium. EGTA (10 mM) added to the presynaptic pipette blocked the fast component induction during a train of 200 AP-es at 100 Hz (Fig. 2.3e). Likewise, lowering the depolarization voltage during a train of ten 20-ms pulses at 1 Hz from 90 mV (−80 mV to +10 mV) to 75 mV (−80 mV to −5 mV) reduced the evoked calcium current and eliminated the fast component. Membrane accumulation is not the trigger for fast endocytosis, because 20 pulses at the reduced 75-mV jump at 1 Hz failed to elicit a fast component, though this stimulus caused a net accumulation (≈1.5 picofarad [pF]) similar to the control condition (using 90-mV depolarization steps) in which fast endocytosis was observed (21).
The role of calcium in endocytosis has been investigated at several other synapses. At the frog neuromuscular junction, only a slow form of endocytosis is detected, and its rate is not sensitive to raised intracellular calcium levels (26). In cultured hippocampal neurons, experiments using synaptophosphoryl or synaptophysin have not detected a fast kinetic component of endocytosis, either (27–29). The rate of the single component is reduced when extracellular calcium is reduced, but raising Ca\(^{2+}\) levels above physiologic concentrations does not accelerate the endocytic process (29). At goldfish bipolar cells, endocytosis is fast following a brief depolarization (30,31), and can be slowed by adding EGTA or 1,2-bis(o-aminophenoxy)ethane-N,N,N',N'-tetraacetic acid (BAPTA), another calcium chelator, to the patch pipette. However, during intensive stimulation, fast endocytosis appears to be slowed down or even blocked (30,32), which may be due to a buildup in intracellular calcium (30). The reasons for this apparent discrepancy have not been discovered, but it has been proposed that intensive stimulation causes nonsynchronous release to occur, far away from central active zones, and that these vesicles can only be retrieved through a slower pathway (32). A flash photolysis study in mouse cochlear inner hair cells has shown that a fast form of endocytosis is activated at internal calcium concentrations above 15\(\mu\)M, and that above this level the percentage of total retrieval conducted by the fast component increases with calcium levels, though the \(\tau\) stays the same (33). Thus, it appears that calcium regulation of endocytosis depends on the type of neuron.

**Bulk Endocytosis and the Measurement of the Fission Pore Formation and Closure**

An early, pioneering study of endocytosis using EM noted the appearance of large endosome-like structures in the nerve terminal after strong stimulation, from which small vesicles bud off (2). Further studies led to the widely held hypothesis that these endosome-like structures are generated slowly (~1 min) from the plasma membrane, a process called bulk endocytosis (34–39). However, the kinetic evidence indicating the instant of bulk membrane fission was missing at synapses. This section discusses a recent study at the calyx of Held that provides this missing piece of evidence (40).

Bulk membrane uptake could be detected after stimulation by step depolarization and AP-e trains, and events were characterized by downward capacitance shifts (DCSs) with a 10% to 90% decay time between approximately 30 and 500 ms (Fig. 2.4a). The sizes of the DCSs ranged from the detection limit of ~20 fF up to 500 fF, with an average size of about 131 fF. These values were much larger than the membrane capacitance (~65 aF) of a single vesicle. Their size distribution was eccentric and peaked at the detection limit of ~20 fF, suggesting that there were possibly many more events that were too small to be detected (40).

The occurrence of DCSs increased with stimulation. The DCSs were detected during baseline recordings at a frequency of ~0.003 Hz; after ten 20-ms pulses (from −80 mV to +10 mV) at 10 Hz, the frequency increased to about 0.021 Hz in
the first 10 seconds after stimulation (Fig. 2.4b). The frequency decayed to baseline levels within 80 seconds and had a half decay time of <20 seconds. Vesicle fusion was required for this increase, as the increase was not seen after stimulation when exocytosis was blocked by botulinum neurotoxin C. The total amount of retrieval conducted by bulk endocytosis was about ~10% of the net exocytosis. This is likely an underestimate, since, as noted above, events below 20 fF in size could not be detected (40).

The diameter of the fission pore and the rate of closure were determined using the measured pore conductance, and with the assumption that the pore is cylindrical. The initial diameter ranged from approximately 3 to 19 nm, and decreased to undetectable levels within 500 ms (Fig. 2.4c). The slope of the 20% to 80% decrease in diameter was about 39 nm/s (Fig. 2.4c). The slope for each event was not correlated with the DCS size, suggesting that the fission step is separate from the fission pore formation step. Consistent with this suggestion, bulk membrane fission can occur as early as a few seconds after stimulation, and the rate of fission pore closure is much lower than the rate needed to form the fission pore in this time. Thus, bulk endocytosis is composed of two kinetically different steps: a membrane invagination step that forms the fission pore, and the closure of the pore that completes the fission process (40).

It should be noted that the frequency of DCSs peaked in less than 10 seconds after stimulation. Some bulk endocytosis events occurred at only a few seconds

---

**Fig. 2.4** Bulk endocytosis at the calyx of Held. (a) Three kinetic forms of endocytosis (arrows)—rapid, slow, and bulk endocytosis—were observed by whole-cell capacitance (Cm) recordings at the calyx of Held. The stimulus (vertical arrow) was 10 depolarizing pulses of 20 ms from −80 to +10 mV at 10 Hz. (b) The whole-cell membrane capacitance (Cm) response to 10 depolarizing pulses at 10 Hz (vertical arrow). The arrows indicate bulk endocytosis that occurred within a few seconds after the stimulus. (c) The whole-cell membrane capacitance (Cm), the fission pore conductance \( (G_p) \), and the fission pore diameter \( (D_p) \) during bulk endocytosis. (a–c: Adapted from Wu and Wu [40].)
after the stimulus (Fig. 2.4b). Such a rapid time course is in sharp contrast to the
current view that endosome-like structures are generated on a time scale of minutes
(2,38,41,42). This discrepancy is likely due to methodologic differences. Electron
microscopy has a low time resolution, whereas the capacitance measurement tech-
nique used in the present work provides a time resolution of milliseconds. Electron
microscopy measures the lifetime of endosome-like structures, whereas the capaci-
tance measurement indicates the time course of generating endosome-like struc-
tures from the plasma membrane. We suggest modifying the current view to a rapid
generation of endosome-like structures, followed by the slow budding off of small
vesicles from these structures.

Resolving Full Collapse Fusion and Kiss-and-Run
with Cell-Attached Capacitance Recordings

We have discovered at least three kinetic forms of endocytosis at the whole-cell
configuration at the calyx of Held. These forms are rapid, slow, and bulk endocy-
tosis (Fig. 2.4a). Rapid endocytosis implies a kiss-and-run form of fusion,
whereas slow endocytosis is consistent with full collapse fusion (43). However,
other interpretations are possible. For example, imaging studies at goldfish retinal
bipolar synapses raise the possibility that rapid endocytosis could be a result of
full collapse fusion followed by rapid endocytosis (44). The key difference
between these two modes of fusion is that kiss-and-run opens a fusion pore and
closes the pore rapidly, whereas full collapse fusion fully expands the fusion
pore. The only unambiguous way to distinguish these is to record fusion pore
kinetics at synapses, which is technically challenging and rarely performed; cor-
respondingly, whether kiss-and-run exists at synapses is currently under intense
debate (45). Fusion pore kinetics in nonsynaptic preparations is most commonly
measured using the cell-attached technique. When using this technique, the secre-
tory vesicle membrane is modeled as a capacitor $C_v$, with a conductance $G_p$ that
is determined by the fusion pore. Both elements form a $G_p C_v$ series circuit that is
added to the cell membrane capacitance $C_m$ upon fusion (Fig. 2.5a). This section
describes progress in adapting this technique to synaptic preparations, and a
recent study using cell-attached recordings, which has resolved the fusion pore
conductance at the calyx (46).

The cell-attached capacitance measurement technique provides a high enough
resolution to detect single vesicle fusion events, which appear as unitary capaci-
tance steps directly proportional to the vesicle size. Full vesicle incorporation into
the plasma membrane produces an upward capacitance step, whereas kiss-and-run
fusion produces an up-step followed within a few seconds by a down-step, called a
capacitance flicker (5). In some of these fusion events, fusion pore conductance can
be measured, allowing for an estimate of the fusion pore size (5).

Measurements of capacitance steps during large dense core vesicle fusion in
endocrine and immune cells have provided a detailed picture of exocytosis of this
Fig. 2.5 Kiss-and-run and full collapse fusion recorded in the cell-attached mode at the calyx of Held. (a) Equivalent circuit of a patch in the cell attached mode, when a vesicle is fused with the plasma membrane and has opened a fusion pore. Double line indicates patch pipette. $C_m$, $G_m$, $C_v$, and $G_p$ are cell membrane capacitance, cell membrane conductance, vesicle capacitance, and fusion pore conductance, respectively. (b) The procedure to perform cell-attached recordings at the release face of a calyx. A calyx associated with a postsynaptic neuron was first identified (a). The postsynaptic neuron was sucked and pulled away by a large pipette (b). After the postsynaptic neuron was removed (c), another pipette was positioned at the release face of the calyx membrane (d) for cell-attached recordings. Individual vesicle fusion events, like those shown in c and d, were recorded. (c) Im (imaginary component of the admittance, reflecting capacitance), Re (real component of the admittance, reflecting conductance) traces from cell-attached recordings during application of 25 mM KCl. The inset shows discrete capacitance up-steps. (d) Im, Re, and $G_p$ during a capacitance flicker. A nonflicker up-step (right) occurring 10 seconds later was not accompanied by detectable Re changes, indicating proper phase adjustment. (e) Im, Re, and $G_p$ during a full collapse fusion. Note that $G_p$ was detected in this fusion event. (f) Top: Illustration of a full-collapse event. Middle:
vesicle class in nonneuronal cells (5). Extending this approach to synapses has been frustrated by two problems: First, the smaller size of vesicles makes resolving individual vesicle fusion more difficult. Second, the postsynaptic neuron apposing the presynaptic release site might make the release site inaccessible to the patch pipette. These two hurdles can be overcome. First, in pituitary nerve terminals (which do not form synapses), fusion of individual microvesicles similar in size to synaptic vesicles were resolved, demonstrating that these small events can in fact be detected (47,48). About 5% of fusion events were capacitance flickers with a fusion pore conductance of ~19 pico siemens (pS), indicating the existence of kiss-and-run fusion (47). Second, direct patch-clamp recordings of nerve terminals that are not associated with their postsynaptic counterpart have been made at a chick calyx-type synapse (49), synaptosomes (50), and the rat calyx of Held (46).

At the calyx of Held synapse, the release site can be exposed by pulling out the postsynaptic neuron using a large pipette (Fig. 2.5b). At the exposed release sites, cell-attached recordings reveal individual capacitance up-steps reflecting single vesicle fusion during high potassium application (Fig. 2.5c) (46). About 20% of fusion events were capacitance flickers. The capacitance flicker duration ranged from 10 ms to 2 seconds with a mean of ~300 ms (Fig. 2.5d). For most capacitance flickers, the fusion pore conductance was larger than 288 pS. The exact size could not be detected owing to the resolution limit. In a small fraction of capacitance flickers, however, a fusion pore conductance ranging from 15 to 288 pS with a mean of ~66 pS was observed, which might correspond to a fusion pore with a mean diameter of ~1.1 nm (Fig. 2.5d). These results suggest that a minor fraction of fusion events are kiss-and-run during high potassium application (46).

Most capacitance up-steps are not followed in a brief time by an equal size down-step, and thus reflect full-collapse fusion (46). Their initial fusion pores were often too large or too fast to resolve (Fig. 2.5d). However, in a small fraction of up-steps, an initial fusion pore conductance of ~250 pS was resolved, which was followed in approximately 10 to 300 ms by a rapid pore expansion (Fig. 2.5e) (46). These results provide the first kinetic evidence revealing the instant of full collapse fusion at synapses.

Fig. 2.5 (continued) The simulated mEPSC caused by a $G_p \geq 288$ pS, as observed in 97% of fusion events. This trace was the average of the simulated mEPSC resulting from initial $G_p$ values of $\infty$ and 288 pS. The scale bars also apply to g. Bottom: Three experimentally observed individual mEPSCs at the calyx of Held with a rapid rise time. The scale bars also apply to panel g. (g) Top: Illustration of a kiss-and-run event. Middle: The simulated mEPSC caused by a $G_p$ of 66 pS, as observed in 3% of fusion events. Bottom: Three experimentally observed individual mEPSCs displaying a 10% to 90% rise time slower than 0.8 ms and an amplitude smaller than 25 pA. These represent about 1.1% of all observed events in ref. 46. (Adapted from He et al [46]).
The Impact of Kiss-and-Run Fusion on the Quantal Response

Synapses that are able to use kiss-and-run fusion may have two advantages, compared to those that use only full collapse. First, it allows for rapid and economical vesicle recycling, perhaps preventing some of the rundown in release during strong stimulation. Second, its narrow fusion pore could limit the rate of transmitter discharge out of the vesicle, resulting in a slower and smaller quantal response (Fig. 2.5g) compared to full collapse fusion (Fig. 2.5f). Control over the amplitude and the kinetics of the quantal response by regulation of the two fusion modes may contribute to synaptic plasticity (6). However, whether the fusion pore size is small enough to slow down transmitter diffusion was largely unclear. We have recently attempted to address this issue at the calyx-type synapse.

During capacitance flickers at the calyx-type synapse, the fusion pore conductance of most events was more than 288 pS, whereas the conductance in a minor fraction of events was on average ∼66 pS. Knowing the fusion pore conductance \( G_p \), the fusion pore diameter \( D_p \) could be estimated by this equation (51,52):

\[
D_p = (4G_p \rho \lambda / \pi)^{0.5},
\]

where \( \rho \) is the saline resistivity (100 Ω cm); and \( \lambda \), the pore length, is taken as the length of a gap junction channel (15 nm). According to this equation, kiss-and-run fusion with a \( G_p \) of 66 pS corresponds to a \( D_p \) of 1.1 nm, and kiss-and-run fusion with a \( G_p \) more than 288 pS corresponds to a \( D_p \) of more than 2.3 nm. The values of the fusion pore diameter can be used to estimate the time constant (\( \tau_{glu,v} \)) of transmitter diffusion from the vesicle to the synaptic cleft by the equation

\[
\tau_{glu,v} = (\pi D_v^3 / 6) / (\rho K_d G_p),
\]

where \( D_v \) is the vesicle diameter and \( K_d \), the diffusion constant, is 3.3 × 10^{-6} cm^2 s^{-1} for glutamate in the synaptic cleft. Based on this equation, \( \tau_{glu,v} \) is 2.3 ms for a \( G_p \) of 66 pS, and 0.54 ms for a \( G_p \) of 288 pS. These calculations suggest that kiss-and-run with a small fusion pore can slow down the diffusion of transmitter out of the vesicle. This suggestion was further confirmed by Monte Carlo simulations of quantal events with MCell 2.50, a program that models the three-dimensional random walk diffusion and reaction kinetics in complex spatial environments reflecting realistic cellular ultrastructure (53). The simulation shows that kiss-and-run fusion with a \( D_p \) of 1.1 nm would cause an mEPSC with a much slower rise and decay (Fig. 2.5g), and a smaller amplitude as compared to full collapse fusion with a \( D_p \) that is too large or too fast to resolve (Fig. 2.5f). Since kiss-and-run with a \( D_p \) of about 1.1 nm was detected in only ∼3% of the total fusion events, small mEPSCs with a slow rise and decay must represent a very minor fraction of mEPSCs. Consistent with this prediction, less than ∼1% of the measured mEPSCs were small and slow in both rise and decay (46). These results suggest that kiss-and-run with a small fusion pore may induce small and slow mEPSCs.

It should be pointed out that a small and slow mEPSC is not necessarily the result of kiss-and-run fusion with a small fusion pore. Many other mechanisms may also determine the amplitudes and the kinetics of mEPSCs. These mechanisms include variation in the vesicle size (54–56), the vesicular transmitter content (57–60), the distance between release sites and glutamate receptor clusters (61), differences in receptor subunit compositions (62), and release from boutons other than
the calyx that form synapses at the principal cell in the medial nucleus of the trapezoid body (63).

**Conclusion**

Both whole-cell and cell-attached capacitance measurements have been successfully applied to the calyx of Held synapse in the past few years. Whole-cell recordings reveal three kinetically different forms of endocytosis: rapid, slow, and bulk endocytosis. Intense stimulation triggers rapid endocytosis by increasing the calcium influx, which may speed up vesicle recycling to catch up with the rapid rate of exocytosis. Bulk endocytosis was shown to occur faster than previously estimated, and to carry $\sim 10\%$ of the total endocytic load. Cell-attached recordings show two modes of fusion: kiss-and-run and full collapse fusion. Kiss-and-run fusion is followed by rapid endocytosis, whereas full collapse fusion is not. Kiss-and-run with a small fusion pore is likely to produce a small and slow quantal response. Switch between kiss-and-run and full collapse may thus be a mechanism by which synaptic plasticity can be achieved. Further experiments are needed to determine whether rapid and slow endocytosis observed at the whole-cell level reflect kiss-and-run and full collapse fusion, respectively.

**References**
