Review

Kinetics of transmitter release at the calyx of Held synapse

By Takeshi SAKABA*1,†

(Communicated by Masanori OTSUKA, M.J.A.)

Abstract: Synaptic contacts mediate information transfer between neurons. The calyx of Held, a large synapse in the mammalian auditory brainstem, has been used as a model system for the mechanism of transmitter release from the presynaptic terminal for the last 20 years. By applying simultaneous recordings from pre- and postsynaptic compartments, the calcium-dependence of the kinetics of transmitter release has been quantified. A single pool of readily releasable vesicles cannot explain the time course of release during repetitive activity. Rather, multiple pools of vesicles have to be postulated that are replenished with distinct kinetics after depletion. The physical identity of vesicle replenishment has been unknown. Recently, it has become possible to apply total internal reflection fluorescent microscopy to the calyx terminal. This technique allowed the visualization of the dynamics of individual synaptic vesicles. Rather than recruitment of vesicles to the transmitter release sites, priming of tethered vesicles in the total internal reflection fluorescent field limited the number of readily releasable vesicles during sustained activity.

Keywords: synapse, synaptic transmission, presynaptic, exocytosis, transmitter release

Introduction: The quantal hypothesis of synaptic transmission

When an action potential (AP) arrives at the presynaptic terminal, calcium (Ca) influx through voltage-gated Ca channels triggers synaptic vesicle fusion and the release of transmitters stored in the synaptic vesicles. The released transmitters act on the postsynaptic receptors and elicit synaptic responses. Presynaptic mechanisms have remained obscure because direct presynaptic recordings of the presynaptic terminal have been difficult, because the terminal is usually small, with a diameter of approximately 1 µm. Nevertheless, important observations were done in the 1950–60s by Bernard Katz and colleagues, using model preparations, such as squid giant synapses and frog neuromuscular junctions.1)

At the neuromuscular junction, when a nerve was stimulated in low Ca conditions (when transmitter release probability is low), the postsynaptic responses fluctuated from trial to trial (Fig. 1A).

Amplitude histograms of the postsynaptic responses showed several peaks. From this observation and later observations that a single peak was mediated by the release of thousands of transmitter molecules,2) it has been postulated that transmitter molecules are released in a quantal manner. The amplitude distribution is fitted best by a sum of Gaussian distributions with increased width, and the amplitude of these Gaussians follow a Poisson distribution (Fig. 1A). This observation led Katz and colleagues to propose the following hypothesis (the so called quantal hypothesis, Fig. 1B). There are a fixed number of transmitter release sites (N), where transmitter release takes place with some probability (Pr). Released transmitter from a single release site causes the postsynaptic response with certain amplitudes (q). Thus,

\[ \text{Synaptic responses amplitude} = N \cdot P_r \cdot q \]

Nowadays, it is considered that transmitter molecules are stored in synaptic vesicles, and vesicle fusion with the plasma membrane mediates transmitter release. Vere-Jones (1966)3) and subsequently Zucker (1973)4) suggested that Pr is a product of Pocc and Pves. Pocc is the probability that synaptic vesicles are docked with the transmitter release sites (some sites may be empty) and Pves is the release

*1 Graduate School of Brain Science, Doshisha University, Kyoto, Japan.
† Correspondence should be addressed: T. Sakaba, Graduate School of Brain Science, Doshisha University, Kyo-Tanabe, Kyoto 610-0394, Japan (e-mail: tsakaba@mail.doshisha.ac.jp).

doi: 10.2183/pjab.94.010
©2018 The Japan Academy
probability of docked synaptic vesicles. Separation of $P_{occ}$ and $P_{ves}$ is difficult, and $P_{occ}$ has often been neglected in subsequent studies. However, recently some attempts have been made to dissect these two parameters.5)

Quantum analysis at the calyx of Held synapse

Subsequent to the observations by Katz and colleagues, quantum analysis was applied to mammalian central nervous system (CNS) synapses in the 1990s, when the patch clamp technique was successfully applied to slice preparations.6) At the neuromuscular junction, it has been accepted that postsynaptic responses follow the amounts of transmitter released linearly.7) In mammalian synapses, some deviations from the quantal hypothesis have been noted. For example, postsynaptic receptors are sometimes saturated by the release of single synaptic vesicles and the postsynaptic responses do not respond linearly to the amounts of transmitter released (receptor saturation8)), whereas receptors are not saturated in other central synapses.9) In addition, it has been postulated that the release probability is not uniform across release sites, but instead is it rather heterogeneous.10)–12) On the other hand, in some central synapses, the simplest quantal hypothesis holds true.13),14) At the neuromuscular junction (NMJ) in low Ca solution the distribution of postsynaptic responses usually follows Poisson’s law (because $Pr < 1$) and at central synapses, binomial statistics are required.

The analysis of central synapses often stimulated presynaptic axons and recorded postsynaptic responses. This analysis has the advantage that the presynaptic compartments are intact. However, because presynaptic terminals are not directly controlled because of their small size (1 µm), it is uncertain if the terminals are reliably stimulated. It has been sometimes postulated that fluctuations in postsynaptic responses are caused by the stochastic nature of presynaptic nerve activity in addition to that of transmitter release.15) In other words, AP's

---

**Fig. 1.** Quantal hypothesis of synaptic transmission. A. Left: At neuromuscular junctions, the amplitudes of endplate potentials (EPPs) fluctuate from trial to trial in conditions with a low release probability. Right: Amplitude histogram of EPPs. The distribution is described by convolution of the Poisson distribution (explaining peaks) with Gaussian distribution (explaining dispersion of unit responses). Unit response is defined as 1. Some trials failed to elicit responses. From Del Castillo and Katz (1954).2) B. Scheme of synaptic transmission. Synaptic responses are a product of $N$ (the number of release sites or docking sites), $Pr$ (transmitter release probability per site), and $q$ (quantal size, a unit response). Reproduced from Trigo et al. (2012) Proc. Natl. Acad. Sci. U.S.A. 109, 18138–43.8) C. Top: Transmission image of the calyx of Held synapse. A presynaptic terminal wraps directly onto the soma of the postsynaptic cell. Bottom: Fluorescence image of the calyx of Held terminal.
Kinetic analysis has been carried out to elucidate the properties of ion channels.\textsuperscript{20} In this type of analysis, the kinetics and voltage dependence of activation (open), deactivation (close), and inactivation (close after sustained activation) of sodium and potassium channels are quantitatively dissected using the voltage clamp method. Because channels are voltage dependent, it was essential to clamp the voltage and to measure ionic conductance in voltage-control conditions. Moreover, the input (voltage) – output (currents) relationship was examined by varying the amplitudes of voltage pulses systematically, and the mechanisms of channel activation, deactivation, and inactivation were analyzed by formulating a kinetic model. In principle, similar kinetic analyses can be applied to dissect the mechanism of synaptic transmission, if one can control the input (voltage of the presynaptic terminal) and see the consequence in postsynaptic responses. Simultaneous recordings of pre- and postsynaptic compartments have been applied to squid giant synapses by Augustine, Llinas, and colleagues\textsuperscript{21,22} and subsequently, since the 1990s the recordings have been successfully applied to the calyx of Held,\textsuperscript{17,23} which allows kinetic analysis of transmitter release at a mammalian central synapse.

During an AP, the membrane potential reaches over 0 mV during overshoot, followed by repolarization, and this functionally is well approximated by a short depolarization with a duration of 1 ms or less (Fig. 2A). If the duration of the depolarization is prolonged, then all the releasable vesicles will be used up, and the time course of cumulative release should follow a single exponential. Then, the plateau value should be equal to \( N \), and the time constant will be the inversely related to \( P_{occ} \), in the simplest form (Fig. 2A). \( P_{occ} \) (or cumulative release) = \( 1 - \exp(-d/\tau) \), where \( d \) is the duration of the pulse, and \( \tau \) is the time constant (Fig. 2A). It should be noted that this simplest form assumes that \( P_{occ} = 1 \), and if \( P_{occ} < 1 \), then the plateau value should be \( N P_{occ} \).

Nevertheless, if \( P_{occ} \) is not uniform across release sites, then the deviation from a single exponential should be evident. Simultaneous voltage clamps were performed on the pre- and postsynaptic compartments of the calyx of Held, and the presynaptic terminal was depolarized to 0 mV for 50 ms (Fig. 2B). EPSCs were evoked in response to presynaptic Ca influx. In this particular set of experiments, cyclothiazide and kynurenic acid were present to block postsynaptic AMPA receptor desensitization and saturation, respectively. The deconvolution method, which extracts the transmitter release rates from the EPSC waveforms, was applied to estimate the time course of cumulative release.\textsuperscript{24} It turned out that cumulative release reached a plateau value of 3000 (2000–5000) vesicles, meaning that the calyx terminal has 3000 releasable vesicles (readily releasable pool [RRP] of synaptic vesicles). Moreover, the time course of release was rather biexponential, with time constants of 2–3 ms (fast releasing pool [FRP] of vesicles) and 20–30 ms (slowly releasing pool [SRP] of vesicles) at room temperature.\textsuperscript{25,26} Given the finding by electron microscopy that the calyx of Held has about 500 active zones, it followed that there are 3 tightly docked and 3 loosely docked vesicles within an
active zone, which corresponded to the FRP and SRP of vesicles, respectively.\(^{27,28}\) Therefore, Pr is heterogeneous temporally at the calyx of Held. The number of readily releasable synaptic vesicles (FRP + SRP) is twice or three times as large as N estimated from the variance mean analysis.\(^{18,19}\) This is because only the FRP is mainly responsible for the synchronous release of vesicles in response to an AP, whereas the SRP is not involved in AP-evoked release and may not be “seen” when stimulated at low frequency.\(^{20}\) The SRP may be more relevant for asynchronous release after tetanic stimulation.\(^{30}\) Recent studies suggested that vesicles in the SRP are rapidly converted to FRP vesicles during repetitive stimulation.\(^{5,31}\) In this context, a major role of the SRP is to feed vesicles rapidly to the FRP during trains of APs. In conclusion, the simplest quantal hypothesis seems a good approximation for the transmitter release process in response to a single AP or during low-frequency stimulation, but heterogeneous Pr has to be postulated during repetitive stimulation.

What is the main mechanism of heterogeneous Pr? We can consider two possibilities: one is that some vesicles are located close to Ca channels whereas others are located further away (positional heterogeneity). Because synaptic vesicles are located farther away from Ca channels, the Ca concentration around synaptic vesicles will be lower and accordingly Pr will also be lower. A second possibility is that Ca-sensitivity for transmitter release is heterogeneous (intrinsic heterogeneity) so that some vesicles have lower Ca sensitivity (Fig. 3A). To distinguish these two possibilities, Ca uncaging methods were used\(^{22}\) (see also Woelfel \textit{et al.}, 2007\(^{33}\)). The caged Ca-compound DM-Nitrophen was introduced into the presynaptic terminal via a patch pipette, and flash photolysis was applied, which elevated Ca uniformly (Fig. 3A). If positional heterogeneity is responsible for heterogeneous Pr, the release rate constants will be similar between vesicles in the FRP and SRP, because spatial gradients of Ca will be minimized in Ca-uncaging conditions. If

---

**Fig. 2.** Transmitter release time course. **A.** In response to an AP-like stimulation (a 1–2 ms depolarizing pulse), the cumulative release (the number of vesicles releasing neurotransmitter) can be determined by N Pr. The real AP has a repolarizing phase after depolarization. When the pulse duration is prolonged, all the vesicles release neurotransmitter, and the cumulative release should be N (Pr = 1). The release time course should be fitted by an exponential if Pr is homogeneous. From top, presynaptic potential, cumulative release, and schematic drawing of vesicle release are shown. The schematic drawing illustrates the case for N = 3 and Pr = 0.3 or 1. Adapted from Ref. 69. **B.** The time course of release at the calyx of Held. The presynaptic terminal was depolarized from −80 mV to 0 mV for 50 ms (after a 2 ms pulse to +70 mV to activate Ca currents maximally). The EPSC was observed at the postsynaptic side, and the cumulative release was estimated by the deconvolution method and was fitted by a dual exponential, which suggests that Pr is heterogeneous among synaptic vesicles. From top, presynaptic membrane potential (Vpre), presynaptic Ca currents (ICa), EPSC, and cumulative release, schematic drawing of vesicle release, are shown. FRP and SRP are the fast releasing pool and slowly releasing pool of synaptic vesicles, respectively. N: number of vesicles in each pool. PAP: release probability per an AP. Reproduced from Sakaba and Neher (2001), with permission from Elsevier.\(^{26}\)
intrinsic heterogeneity is responsible, fast and slow components of the release should still exist upon flash photolysis. Figure 3B shows such an experiment. When a 10 ms depolarizing pulse was applied, vesicles in the FRP were depleted. A subsequent 50 ms pulse only released vesicles in the SRP. When flash photolysis was applied instead of the second pulse, only the fast release was observed, suggesting that vesicles in the SRP can be released rapidly when sufficiently high Ca was applied. It was concluded that positional heterogeneity is mainly responsible for heterogeneous Pr. In contrast, Woelfel et al. (2008) suggested that intrinsic Ca sensitivity is responsible for heterogeneous Pr by doing similar experiments with the same preparation. The slow component of release may consist of several populations of vesicles, one located far away from Ca channels (with a Ca sensitivity similar to those of fast vesicles) and another representing intrinsically slow vesicles. Alternatively, some of the slow component reflects the release of newly replenished vesicles, in which case the Ca sensitivity of vesicle replenishment rather than that of vesicle fusion determines the apparent Ca sensitivity. Because vesicle replenishment can be rapid (time constant of 100 ms) at the calyx in conditions of flash photolysis (or during strong stimulus such as depolarizing pulses), it is difficult to distinguish between slow-release of release-ready vesicles at the release site and the release of newly replenished vesicles in the time frame of several hundreds of ms (the experimental condition in Fig. 3). Furthermore, adaptation of the release machinery might happen during prolonged exposure of Ca. This issue needs to be carefully addressed.

Later, it was postulated that even vesicles in the FRP are not homogeneous. Muller et al. (2010) described that a train of 20 Hz presynaptic stimulation gave rise to synaptic depression, but
a subsequent 200 Hz stimulation caused synaptic facilitation. If Pr is high, synaptic responses should exhibit depression because releasable vesicles are depleted during the stimulus train. Conversely, facilitation of synaptic responses is seen at synapses with low Pr. Therefore, the findings of Muller et al. can be best explained by postulating that the conditioning 20 Hz train depletes high Pr vesicles without depleting low Pr vesicles in the FRP. When flash photolysis was applied instead of 20 Hz stimulation, the Ca sensitivity of release was decreased by a factor of 2, indicating that the intrinsic Ca sensitivity of fast releasing vesicles is heterogeneous (the remaining vesicles after 20 Hz stimulation are slower by a factor of 2). Conversely, it may be concluded that during 20 Hz stimulation a population of ‘superprimed’ vesicles has high release probability and are released rapidly. More recently, Taschenberger et al. (2017) suggested that superprimed vesicles contribute to the amplitudes of EPSCs in the initial phase of repetitive stimulation. This contribution differs between synapses. Their results also suggested that post-tetanic potentiation is mediated by an increase in the number of superprimed synaptic vesicles through activation of second messenger systems. In other words, superpriming is responsible for variability among synapses, short-term presynaptic plasticity such as post-tetanic potentiation, and its degree of modulation. Therefore, the consensus is that a subtle difference in intrinsic release kinetics (2-fold) has a large impact on synaptic strength in physiological conditions.

**Synaptic vesicle replenishment to the readily releasable pool of synaptic vesicles**

After depletion of the RRP, new vesicles have to be recruited. In low frequency stimulation, vesicle recruitment does not have to be considered, and the simple model of the quantal hypothesis can account for synaptic responses (Fig. 4C, low-frequency stimulation). Vesicle recruitment is particularly important during high-frequency stimulation (>100 Hz at the calyx of Held), because the rate of recruitment determines synaptic strengths in such conditions (Fig. 4A). In other words, the transmitter release model has to be extended. It has been postulated that physical translocation of synaptic vesicles to the release site is the limiting process of vesicle recruitment to the release site (Fig. 4B, tethering in Fig. 4C). Zenisek et al. (2000) showed that translocation of vesicles to the release sites determined the rate of sustained release after the depletion of release-ready vesicles at retinal bipolar cell synapses. In addition, recent studies by Silver and colleagues have led to the postulate that the rate of diffusion of vesicles determines the rate of vesicle replenishment at cerebellar synapses. Alternatively, the readiness of the release machinery for fusion, or priming of synaptic vesicles, which must follow vesicle tethering, may be a limiting factor (docking/priming in Fig. 4C). In neurons, deficiencies in priming molecules (such as Munc13s, a protein family implicated in vesicle priming/maturation before exocytosis), caused the RRP to be almost lost, whereas the number of morphologically docked vesicles remained intact. In more recent studies, the distance of synaptic vesicles from the plasma membrane was found to be slightly increased (<100 nm) after the deletion of priming molecules, suggesting that priming and docking happen at the same time.

The rate of vesicle recruitment can be determined by a dual pulse experiment at the calyx of Held, in which two pulses are applied with varied intervals, which depleted the RRP. When such an experiment was carried out on the calyx of Held synapse (a 50 ms pulse to 0 mV × 2), recovery of the SRP was rapid with a time constant of 100 ms, whereas that of the FRP had rapid and slow components. Rapid recovery is Ca and calmodulin dependent, and it has been shown that Munc13 has a calmodulin binding site, is responsible for rapid recovery from depletion (see also Fig. 4A). Because Munc13 has been postulated as a priming molecule, the result indicated that priming is a limiting step for vesicle replenishment. However, in addition to priming, the availability of release sites (acceptor complex) becomes rate limiting in certain conditions (Fig. 4C). Consistently, it has also been shown that recovery of the fast component of release (the FRP) was slowed when synaptic vesicle endocytosis was inhibited. Release sites have to be cleared by endocytosis before new vesicles can be docked or become release ready. Such release site clearance has been supported by several other studies. It is also notable that mutation of the calmodulin-binding site of Munc13-1 slows not only recovery from vesicle depletion but also retrieval of the excyotosed synaptic vesicle protein synaptotagmin 2 (a putative Ca sensor for vesicle fusion), which may mean that Munc13-1 may be important for site clearance in addition to its well-known role in priming. We should note that a 50 ms pulse is a strong stimulus, and the replenishment seen by such strong stimulation is different from that seen by APs. Perhaps, the rapid vesicle
recruitment observed by Saviane and Silver (2006). A time constant of 10 ms may be quite similar to the rapid conversion of slow to fast releasing vesicles, as seen by Lee et al. (2013). Nevertheless, the relative importance of the three steps (tethering, priming, and availability of release sites; Fig 4C) can be determined at the calyx of Held by direct observation of synaptic vesicles and release sites (see below).

Fig. 4. The mechanism of synaptic vesicle replenishment. A. During a 100 Hz stimulation of APs, postsynaptic responses are depressed, which are due to the depletion of the RRP (FRP). After depression, synaptic responses recover with a time constant of seconds. Mutation of the calmodulin-binding domain of Munc13-1 slows recovery from synaptic depression. Reproduced from Lipstein et al. (2013), with permission from Elsevier. B. After fusion of release-ready vesicles at the release site, new vesicles are recruited to the release site for the next round of fusion. The rate of recruitment (tethering) may determine the rate of vesicle release during sustained activity. Reproduced from Furukawa et al. (1982). C. Synaptic vesicle cycle and release site cycle. Synaptic vesicles are tethered to the release site before priming, after which vesicles are ready for fusion. Priming involves closer apposition of vesicles to plasma membrane in the order of tens of nm (docking). After fusion, vesicles are retrieved through endocytosis and neurotransmitters have to refill the vesicle before the next round of exocytosis. In addition to synaptic vesicle cycles, the release sites at the plasma membrane have to be ready to accept synaptic vesicles. After fusion, the release sites have to be cleared and reused. For synaptic vesicle replenishment, tethering, priming/docking, and the reuse of release sites may be involved. In low-frequency stimulation, the RRP size (N in the quantal hypothesis) and Pr determine synaptic strengths; therefore, the quantal hypothesis explains synaptic responses well. In high-frequency stimulation, more processes are involved.

Direct observation of synaptic vesicle dynamics using total internal reflection fluorescence microscopy

So far, this review has looked at the kinetics of release at the calyx of Held synapse using electrophysiology. Although electrophysiology has two advantages over other methods: (1) looking at
physiological responses and (2) good temporal resolution, the physical identities of the release components are unclear. To investigate this, imaging methods would be more appropriate. One bottleneck, however, is that single synaptic vesicles are so small (30–50 nm in diameter) that conventional methods are unlikely to capture vesicle dynamics. Nevertheless, single vesicle dynamics could be monitored by labeling synaptic vesicles sparsely using pHluorin-based methods and quantum dots. Alternatively, total internal reflection fluorescence (TIRF) microscopy has been used to look at the dynamics of synaptic vesicles near the plasma membrane, which was pioneered by the Almers lab. In this method, cells adhering to a glass coverslip were used, and an evanescent wave excited the fluorophore just above the glass interface within a depth of 50–100 nm. Although x- and y-axes are diffraction limited and vesicles have to be localized by sparse labeling, the z-axis has good resolution, and it is possible to monitor the dynamics of synaptic vesicle docking and fusion precisely near the plasma membrane. Midorikawa and Sakaba (2015) have used this method to look at the dynamics of synaptic vesicles at the acutely dissociated calyx of Held synaptic terminal.

Synaptic vesicles were visualized using the styryl dye FM1–43, which had been taken up into cells through endocytosis. Stained synaptic vesicle membrane and individual vesicles close to the plasma membrane could be visualized (Fig. 5, top panel in A–C). Three types of events (diffuse, bounce, and tethering events) were observed, when we performed TIRF measurements. In diffuse events, dyes spread at the plasma membrane upon fusion, representing synaptic vesicle fusion (Fig. 5A). Bouncing events represent the approach to and immediate retraction from the plasma membrane. These events are not relevant for vesicle fusion or replenishment (Fig. 5B). Tethering events represent the approach and stabilization of synaptic vesicles at the plasma membrane. Diffuse events happen during depolarizing pulses, whereas tethering events take place after pulses, suggesting that tethering events represent synaptic vesicle replenishment at the release site (Fig. 5C). Midorikawa and Sakaba (2015) also observed that many vesicles were stabilized in the TIRF field and were not responsive to stimulation. Although some
vesicles are not located at release sites or else are “dead” due to dissociation processes, some are likely to be tethered, but unprimed vesicles. Because the processes of tethering, priming, and fusion could be seen directly, it was possible to address the following question: Which of the two processes, tethering or priming, determines the rate of recovery after depletion of the RRP?

In order to address this issue, we performed Ca uncaging experiments (Fig. 6A). In this experiment, Ca was uniformly elevated to a high level (>10 µM) within the terminal by flash photolysis of a caged-Ca compound. In response to flash photolysis, all the readily releasable vesicles were depleted rapidly (<10 ms). Depletion of the RRP was followed by a sustained component of release presumably reflecting the release of newly replenished synaptic vesicles. The release rate during the sustained component was similar to that of rapid vesicle replenishment to the RRP (time constant of 300 ms). If tethering underlines synaptic vesicle replenishment to the RRP, the rate of tethering and the sustained component should be similar. However, in reality, they differed by 5–10-fold. We found that newly tethered vesicles could not participate in release events, and the rate of tethering (time constant of seconds) was slower than that of replenishment (time constant of hundreds of ms, Fig. 6A, bottom). The sustained component was found to be mediated by the fusion of already tethered vesicles. In other words, not all the vesicles in the TIRF field could fuse immediately upon stimulation, and some vesicles were molecularly unprimed for fusion. Figure 6B summarizes the results from TIRF measurements at the calyx of Held. Following depletion of the RRP after high-frequency stimulation (>100 Hz), synaptic vesicles are rapidly replenished to the RRP through the priming of already tethered vesicles. To compensate for the slow rate of tethering, the number of unprimed vesicles at the plasma membrane outnumbers that of the RRP (3 times the RRP).
The mechanism of synaptic vesicle replenishment may be different among synapses. Some direct observations of this process have been made in a limited number of synapses. In hippocampal mossy fiber synapses, vesicle dynamics essentially similar to those at the calyx synapse have been obtained recently using TIRF microscopy, except that the rate of tethering (seconds) was somewhat slower (Fig. 7A). In retinal bipolar cells, the rate of tethering was much faster, and newly tethered vesicles could release neurotransmitter rapidly, suggesting that tethering rather than priming was rate limiting. In hair cells and cerebellar mossy fiber synapses, it seems that synaptic vesicle mobility determined the rate of synaptic vesicle replenishment. A, Top: At the calyx of Held, synaptic vesicles are tethered and primed before entering the SRP. The SRP can be rapidly converted to the FRP. The FRP vesicles are located closer to Ca channels compared with the SRP. Some vesicles in the FRP are converted to superprimed vesicles, which can release transmitter rapidly (in a ms). Release site clearance may affect the rate of conversion from the SRP to the FRP. Middle: In hippocampal mossy fiber synapses, synaptic vesicles are tethered and primed before entering the RRP. The release time course of the RRP is relatively slow, perhaps suggesting that RRP vesicles in mossy fiber synapses are similar to SRP vesicles at the calyx of Held. PKA tightens spatial coupling between Ca channels and synaptic vesicles. Bottom: In cerebellar basket cell synapses, the RRP is relatively uniform and synaptic vesicle replenishment has a time constant of hundreds of ms. However, detailed analysis has not been carried out, and the mechanism of release may be more complex. B, Amplitudes of synaptic responses during repetitive stimulation are plotted against the stimulus number. At the calyx of Held, the responses at high frequency are larger than those expected from a single pool model, because multiple release processes are involved. From lower to higher frequency, several vesicle pools (superprimed, FRP, and SRP) are involved. Note that the SRP vesicles are likely to be converted to the FRP vesicles before release. >100 Hz, synaptic vesicle replenishment mediated by priming and release clearance is involved.

Fig. 7. Synaptic vesicle dynamics at various types of synapses. A. Top: At the calyx of Held, synaptic vesicles are tethered and primed before entering the SRP. The SRP can be rapidly converted to the FRP. The FRP vesicles are located closer to Ca channels compared with the SRP. Some vesicles in the FRP are converted to superprimed vesicles, which can release transmitter rapidly (in a ms). Release site clearance may affect the rate of conversion from the SRP to the FRP. Middle: In hippocampal mossy fiber synapses, synaptic vesicles are tethered and primed before entering the RRP. The release time course of the RRP is relatively slow, perhaps suggesting that RRP vesicles in mossy fiber synapses are similar to SRP vesicles at the calyx of Held. PKA tightens spatial coupling between Ca channels and synaptic vesicles. Bottom: In cerebellar basket cell synapses, the RRP is relatively uniform and synaptic vesicle replenishment has a time constant of hundreds of ms. However, detailed analysis has not been carried out, and the mechanism of release may be more complex. B. Amplitudes of synaptic responses during repetitive stimulation are plotted against the stimulus number. At the calyx of Held, the responses at high frequency are larger than those expected from a single pool model, because multiple release processes are involved. From lower to higher frequency, several vesicle pools (superprimed, FRP, and SRP) are involved. Note that the SRP vesicles are likely to be converted to the FRP vesicles before release. >100 Hz, synaptic vesicle replenishment mediated by priming and release clearance is involved.

The mechanism of synaptic vesicle replenishment may be different among synapses. Some direct observations of this process have been made in a limited number of synapses. In hippocampal mossy fiber synapses, vesicle dynamics essentially similar to those at the calyx synapse have been obtained recently using TIRF microscopy, except that the rate of tethering (seconds) was somewhat slower (Fig. 7A). In retinal bipolar cells, the rate of tethering was much faster, and newly tethered vesicles could release neurotransmitter rapidly, suggesting that tethering rather than priming was rate limiting. In hair cells and cerebellar mossy fiber synapses, it seems that synaptic vesicle mobility determined the rate of synaptic vesicle replenishment. A, Top: At the calyx of Held, synaptic vesicles are tethered and primed before entering the SRP. The SRP can be rapidly converted to the FRP. The FRP vesicles are located closer to Ca channels compared with the SRP. Some vesicles in the FRP are converted to superprimed vesicles, which can release transmitter rapidly (in a ms). Release site clearance may affect the rate of conversion from the SRP to the FRP. Middle: In hippocampal mossy fiber synapses, synaptic vesicles are tethered and primed before entering the RRP. The release time course of the RRP is relatively slow, perhaps suggesting that RRP vesicles in mossy fiber synapses are similar to SRP vesicles at the calyx of Held. PKA tightens spatial coupling between Ca channels and synaptic vesicles. Bottom: In cerebellar basket cell synapses, the RRP is relatively uniform and synaptic vesicle replenishment has a time constant of hundreds of ms. However, detailed analysis has not been carried out, and the mechanism of release may be more complex. B, Amplitudes of synaptic responses during repetitive stimulation are plotted against the stimulus number. At the calyx of Held, the responses at high frequency are larger than those expected from a single pool model, because multiple release processes are involved. From lower to higher frequency, several vesicle pools (superprimed, FRP, and SRP) are involved. Note that the SRP vesicles are likely to be converted to the FRP vesicles before release. >100 Hz, synaptic vesicle replenishment mediated by priming and release clearance is involved.

Comparison with other synapses: can the observations at the calyx of Held be generalized?

In this manuscript, the distinct kinetics of transmitter release at the calyx of Held have been examined, and there are three components of release (at least): superprimed, primed (the FRP), and slowly released components (the SRP) (Fig. 7A). Superprimed vesicles are relevant for the AP-evoked release, and the initial phase of synaptic depression during low-frequency stimulation. The remaining vesicles in the FRP have lower Pr, compared with
superprimed vesicles, and are, therefore, released less in response to a single AP. However, they are more resistant to depletion. The SRP is more relevant for the steady state during high-frequency stimulation through rapid conversion to the FRP. Finally, priming and release site clearance (synaptic vesicle replenishment to the RRP) help sustained release during very-high-frequency stimulation (>100 Hz) and recovery from synaptic depression. If only a single pool of vesicles was involved in transmitter release, the synaptic responses would be proportional to the inverse of frequency (1/f), and transmitter release would stop at high frequency nerve activity. Multiple release processes allow more sustained responses during repetitive stimulation (Fig. 7B). In other words, the simplest quantal hypothesis holds true or is a very good approximation for synaptic responses in response to single APs or to low frequency stimulus trains, but additional processes have to be considered to explain synaptic responses in the high frequency range (Fig. 4C).

Do these multiple components exist at other synapses? Perhaps, the calyx of Held may require multiple release components to adapt itself for high-frequency signal transmission in the auditory pathway.

In hippocampal mossy fiber synapses, which show facilitation during repetitive activity, the rates of synaptic vesicle tethering and priming were essentially similar to those at the calyx of Held. Release probability was lower and was very sensitive to EGTA, which is a Ca buffer with the slow binding rate, suggesting that distances between Ca channels and vesicles were large. Activation of protein kinase A (PKA), an important regulator of LTP, abolished the sensitivity of transmitter release to EGTA, suggesting tighter coupling between Ca channels and synaptic vesicles (Fig. 7A). Perhaps, vesicles in the RRP at mossy fiber synapses are similar to the SRP vesicles at the calyx synapse, and PKA converts these vesicles to the FRP vesicles.

At the inhibitory synapse between the basket cell and Purkinje cells in the cerebellum, the RRP seemed more homogeneous, and heterogeneous Pr was not observed. Similar observations have been made at the inhibitory synapse between a Purkinje cell and the neuron in the deep cerebellar nuclei. Synaptic vesicle replenishment to the RRP had a single time constant of hundreds of ms. Therefore, the release mechanism seemed much simpler at the inhibitory synapse (Fig. 7A). However, other studies were inconsistent with this idea and Pr may be more heterogeneous. Because detailed kinetic analysis has not been carried out at inhibitory synapses because of technical difficulties with direct presynaptic patch clamp, it remains to be seen if the release process is relatively simple or as complex as that at the calyx of Held.

Evidence supporting multiple pools of synaptic vesicles has also been obtained at other synapses. Superprimed vesicles were originally postulated by Schlueter et al. in hippocampal cultures lacking Rab3, as well as by Hanse and Gustafsson et al. (2001) using minimal stimulation. In addition, slowly releasing vesicles have been postulated in hippocampal and cerebellar synapses. On the other hand, anatomically, the active zone structure is different among preparations, and kinetically, the time course of replenishment was considerably different, ranging from tens of ms to seconds. Therefore, the observations at the calyx of Held may not be always applicable to other synapses, or in other words, synapses are adapted to the demands that each synapse may have.

**Outlook**

The calyx of Held is one of the best characterized synaptic preparations in the mammalian CNS, and multiple components of transmitter release have been dissected using electrophysiology. The next step is to examine the molecular mechanisms of transmitter release and vesicle replenishment. Recent developments in slice and dissociated culture, together with imaging and possible molecular perturbations offer unique possibilities for gaining detailed knowledge on the molecular mechanisms of synaptic transmission. Recent developments in optical microscopy have also revealed physical correlates of the transmitter release processes, as postulated by electrophysiology. Vesicle dynamics can be readily observed using TIRF microscopy, and the physical identities of vesicle tethering and priming are beginning to be understood. Compared with synaptic vesicles, the dynamics of the release machinery are more difficult to observe directly, and further technical developments are needed. In other synapses, which are smaller than the calyx of Held, direct terminal recordings using the subcellular patch clamp technique and the development of super-resolution microscopy will allow similar types of biophysical analyses, which may reveal the general mechanisms of transmitter release and how synapses adapt to functional demands in the context of a neural network.
Acknowledgment

I thank Masao Tachibana, Erwin Neher, Tomoyuki Takahashi and Yasuo Ihara for encouragement, and Shin-ya Kawaguchi and Mitsuharu Midorikawa for collaboration. I also thank Erwin Neher for comments on the manuscript. This study was supported by MEXT/JSPS KAKENHI (15H04261, 15K14321, 17H05753), Takeda Science Foundation, and the JSPS core-to-core program A advanced research networks.

References


65) Young, S.M. Jr. and Neher, E. (2009) Synaptotagmin has an essential function in synaptic vesicle positioning for synchronous release in addition to its role as a calcium sensor. Neuron 63, 482–496.


Profile

Takeshi Sakaba was born in Tokyo in 1972. He graduated from the Department of Psychology, the University of Tokyo in 1994. Subsequently, he obtained a Masters degree (1996) and a Doctoral degree (1998) at the Department of Psychology, the University of Tokyo under the supervision of Masao Tachibana. In 1998, he joined the Department of Membrane Biophysics, Max Planck Institute for Biophysical Chemistry in Goettingen, Germany as a post-doc, under the supervision of Erwin Neher. In 2006, he became an independent junior group leader of the Max Planck Society. Since 2011, he has been a full professor at Doshisha University, Graduate School of Brain Science. In 2016, he received the Tsukahara prize from the Brain Science Foundation in Japan. His main interest is the mechanism of transmitter release and synaptic plasticity in the mammalian CNS.