TITLE: SYNAPTIC RIBBON ENABLES TEMPORAL PRECISION OF HAIR CELL AFFERENT SYNAPSE BY INCREASING THE NUMBER OF READILY RELEASABLE VESICLES: A MODELING STUDY

AUTHORS John H. Wittig Jr¹ and Thomas D. Parsons²,³

AFFILIATION ¹Department of Bioengineering, School of Engineering and Applied Science, ²Department of Clinical Studies, New Bolton Center, School of Veterinary Medicine, and ³Department of Otorhinolaryngology, Head and Neck Surgery, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104 USA

RUNNING HEAD SYNAPTIC RIBBON’S ROLE IN PRECISE SPIKE TIMING

CONTACT INFORMATION Dr. Thomas D. Parsons,
University of Pennsylvania
School of Veterinary Medicine
382 West Street Road
Kennett Square, PA 19348.
E-mail: thd@vet.upenn.edu
ABSTRACT

Synaptic ribbons are classically associated with mediating indefatigable neurotransmitter release by sensory neurons that encode persistent stimuli. Yet when hair cells lack anchored ribbons, the temporal precision of vesicle fusion and auditory nerve discharges are degraded. A rarified statistical model predicted increasing precision of first-exocytosis latency with the number of readily releasable vesicles. We developed an experimentally constrained biophysical model to test the hypothesis that ribbons enable temporally precise exocytosis by increasing the readily releasable pool size. Simulations of calcium influx, buffered calcium diffusion, and synaptic vesicle exocytosis were stochastic (Monte Carlo), and yielded spatiotemporal distributions of vesicle fusion consistent with experimental measurements of exocytosis magnitude and first-spike latency of nerve fibers. No single vesicle could drive the auditory nerve with requisite precision, indicating a requirement for multiple readily releasable vesicles. However, plasmalemma docked vesicles alone did not account for the nerve’s precision --- the synaptic ribbon was required to retain a pool of readily releasable vesicles sufficiently large to statistically ensure first-exocytosis latency was both short and reproducible. The model predicted that at least sixteen readily releasable vesicles were necessary to match the nerve’s precision, and provided insight into interspecies differences in synaptic anatomy and physiology. We confirmed that ribbon-associated vesicles were required in disparate calcium buffer conditions, and irrespective of the number of vesicles required to trigger an action potential. We conclude that one of the simplest functions ascribable to the ribbon, the ability to hold docked vesicles at an active zone, accounts for the synapse’s temporal precision.

Word count = 250

KEYWORDS (three to five words or short phrases not in title or running head):

Ribbon synapse
Exocytosis
Synaptic jitter
Monte Carlo
First latency
INTRODUCTION

Synaptic ribbons are found in photoreceptors and bipolar cells in the retina, and mecano-sensory hair cells in the cochlea and vestibular end organs --- all cells that encode sensory stimuli with a graded receptor potential instead of action potentials, or spikes (reviewed in Sterling and Matthews 2005). Ribbon synapses are characterized by an enigmatic electron-dense structure anchored to the presynaptic density and surrounded by a halo of tens to hundreds of synaptic vesicles. While these cells can exocytose a large readily releasable pool of vesicles within tens of milliseconds, they also maintain high rates of sustained exocytosis for seconds (reviewed in Nouvian et al. 2006). Surprisingly, the sustained component of exocytosis was unaffected in goldfish retinal bipolar cells when ribbons were dissociated from active zones during the diurnal cycle, though rapid exocytosis was compromised (Hull et al. 2006). Furthermore, both rapid exocytosis and precise spike timing in the auditory nerve were degraded in a mouse mutant lacking anchored hair cell ribbons (Khimich et al. 2005), leading to the hypothesis that the ribbon stabilizes a large readily releasable pool thereby enabling the temporal precision of auditory coding (Moser et al. 2006).

The hair cell ribbon synapse’s temporal precision is revealed by the spike timing of eighth nerve fibers. Cat auditory nerve fibers encode the onset of sound with a precisely timed first spike; the latency of this first spike decreases with increasing stimulus intensity, and the fibers’ minimum latencies range from 1 to 3 milliseconds (ms) with a standard deviation, or jitter, as low as 0.1 ms (Heil and Irvine 1997). First-spike latency of frog nerve fibers are also precise, with jitter as low as 0.3 ms (estimated from data in Feng 1982; Ronken et al. 1993). These measures are based on spike timing, and are notable compared to the 0.5 to 0.8 ms jitter of postsynaptic currents in cortical synapses (Jonas et al. 1993; Markram et al. 1997), where any temporal variation caused by fluctuations in action potential generation are absent.

The ribbon’s molecular composition is largely unknown (Lenzi and von Gersdorff 2001), and provides limited clues to its role in precise spike timing. Several possibilities have been suggested. It might facilitate precisely timed exocytosis by co-localizing specialized proteins within the active zone (Hibino et al. 2002) or increasing calcium levels by restricting calcium diffusion (Roberts 1994). Perhaps it enables a precisely timed postsynaptic response by facilitating concurrent fusion of multiple vesicles (Glowatzki and Fuchs 2002) to rapidly depolarize the afferent fiber (Parsons and Sterling 2003). Or the ribbon might facilitate rapid signaling despite slow exocytosis kinetics by simply increasing the number of available vesicles (Almers 1994).

We tested the hypothesis that the synaptic ribbon enables precise first-spike latency simply by increasing the size of the readily releasable pool. We derived a statistical model relating the number of readily releasable vesicles to the precision of first latency, assuming a single vesicle could trigger an action potential. As vesicle count increased latency became more precise. To examine the ribbon’s role in this relationship, we created a biophysical model of the hair cell ribbon synapse with synaptic vesicles distributed according to detailed anatomical measurements. The temporal precision of simulated first-exocytosis latency was sufficient to account for observed first-spike latency --- only when ribbon-associated vesicles contributed to the readily releasable pool. The dependence on ribbon-associated vesicles held for a variety of perturbations including stimulus intensity, calcium buffer conditions, and the number of vesicles required to trigger a spike. Thus our results are applicable to hair cell ribbon synapses in multiple species, as well as other synapses that drive precisely timed spikes with the synchronous fusion of multiple synaptic vesicles.
METHODS

Statistical model of first latency

We defined the exocytic latency of a single vesicle as a random variable \( L \) with the probability density function \( f_L \) and cumulative distribution function \( F_L \). Initially, we estimated \( f_L \) by integrating the kinetic equations describing calcium-dependent exocytosis (Beutner et al. 2001) during a step in calcium concentration from 0 to 50 \( \mu M \) (Fig. 1a, blue trace) using a fifth order Runge-Kutta algorithm with a variable time-step (Press et al. 1992). Subsequently, we estimated \( f_L \) by smoothing the simulated latency distribution of a single vesicle with a three point moving average (Fig. 6a, blue trace). Random variables representing the exocytic latency of \( N \) readily releasable vesicles made the set \([L_1, L_2, \ldots, L_N]\), each drawn from the distribution \( f_L \). First-exocytosis latency was the random variable, \( T \), where \( T = \text{minimum} \ [L_1, L_2, \ldots, L_N] \). The cumulative first latency distribution, \( F_T \), as a function of the number of readily releasable vesicles was:

\[
F_T = P(T <= t) = 1 - P(T > t) \\
= 1 - P(L_1 > t \& L_2 > t \& \ldots \& L_N > t) \\
= 1 - P(L_1 > t) * P(L_2 > t) * \ldots * P(L_N > t) \\
= 1 - (1 - F_L)^N
\]

The mean and jitter of \( T \) were equal to its expected value, \( E(T) \), and the square root of its variance (\( \sigma^2 \)). To compare how vesicle populations with different kinetics contributed to first latency, we modified the final equation to include additional latency distributions.

Simulated active-zone exocytosis

We simulated calcium influx, buffered diffusion, and synaptic vesicle exocytosis following Bennett and colleagues’ (2000) Monte Carlo model of frog NMJ. Their model accounted for mobile (diffusible) calcium buffers without resorting to the computationally intense approach of tracking diffusion and interaction of multiple species of molecules (e.g. Shahrezaei and Delaney 2004). This sacrifice in spatiotemporal accuracy for computational efficiency was essential for our analysis of variance between trials. Here we briefly describe the efficient Monte Carlo implementation and then focus on our adaptations to the ribbon synapse in leopard frog (rana pipiens) saccular hair cells, particularly the distribution of synaptic vesicles and gating of calcium channels.

Bennett and colleagues’ (2000) model combined spatially resolved diffusion of individual calcium ions with coarse buffer molecule locations discretized into three-dimensional grids of volume elements. Each time step, ions probabilistically entered through open calcium channels, attempted to bind or unbind buffer molecules in their local volume element, and diffused. Ions bound to mobile buffers diffused according to the buffer’s diffusion coefficient, whereas ions bound to fixed buffers or vesicle calcium sensors did not diffuse. Ions reflected from the edges of the simulated volume but diffused freely through synaptic vesicles and the synaptic ribbon. Background levels of calcium and occupied buffer molecules were removed from the simulation. This simplification had minimal impact on our mobile buffer concentration (3.1% decrement in total concentration) or synaptic vesicle calcium loading (only 0.3% of all sites would be occupied at background calcium levels). The modeling approach and assumptions were validated (Bennett et al. 2000), though a subsequent computationally intensive synapse model has been used to explore how synaptic vesicles themselves can act as calcium diffusion barriers (Shahrezaei and Delaney 2004).

We adapted Bennet and colleague’s model to the anatomy and physiology of the hair cell by modifying parameters to match published data (Table 1). We solved buffered diffusion with a 0.1 \( \mu s \) time step. Every time step, diffusing calcium ions jumped an average of 5.3 nm in each coordinate (compared to 14.5 nm in Bennett et al. 2000; and 1.7 nm in Shahrezaei and Delaney 2004). The frog’s endogenous mobile buffer, calretinin, has at least five binding sites, four of which are high affinity and are likely cooperative. In absence of detailed measurements of calretinin’s cooperativity, we implemented a simplified model with a single high affinity binding site at four times the measured concentration of the protein (see similar implementation in Roberts 1994). We used (40 nm)\(^3\) volume elements for fixed and mobile buffers, which were distributed uniformly throughout the cytosol, and (10 nm)\(^3\) elements for synaptic vesicle calcium sensors. Exocytosis required the cooperative binding of five calcium ions. We did not include calcium extrusion mechanisms due to the brevity of our stimuli (10 ms) and the size of our simulated volume (1.6x1.6x0.8 \( \mu m \)). The active-zone center was in the middle of the 1.6x1.6 \( \mu m \) surface representing the plasmalemma; the synaptic ribbon was centered above this point and 84 calcium channels were distributed up to 160 nm from this point to match freeze-fracture analysis of the pre-synaptic membrane (Roberts 1994; Roberts et al. 1990).
Synaptic vesicles were distributed randomly according to vesicle packing densities measured in isolated frog hair cells after 30 minutes of continuous inhibition with zero external calcium (Lenzi et al. 2002). We converted density per spatial unit (either volume or area) to vesicle count mean and SD, then drew Gaussian random deviates to find the number of vesicles per spatial unit each trial. Docked vesicles within a spatial unit that contained calcium channels (i.e. ≤200 nm from the active-zone center) were preferentially co-localized with those channels based on the possibility of their physical coupling (Hibino et al. 2002), given similar findings at other synapses (Rettig et al. 1996; Sheng et al. 1996). Otherwise, vesicle locations were randomly chosen within each spatial unit assuming uniform density. Locations were discretized into (10 nm)³ volume elements. We restricted overlap of docked vesicles by assuming each occupied a 3x3x3 cube of volume elements, slightly underestimating their 40 nm diameter. Non-docked vesicles required a smaller 2x2x2 cube for successful placement all vesicles. We first placed all docked-not-tethered vesicles, then docked-and-tethered, then tethered, then outlier (see Fig. 2). Vesicle membranes contain multiple copies of calcium-sensing proteins thought to trigger exocytosis (Takamori et al. 2006), and we assumed the copy responsible for exocytosis sampled the volume element at the vesicle’s plasmalemma-facing pole (see alternatives in Shahrezaei and Delaney 2005; Shahrezaei and Delaney 2004). One side of this volume element was the plasmalemma for docked vesicles. Co-localized vesicles contained a calcium channel within that 10x10 nm swath of plasmalemma. Therefore, the center of a co-localized vesicle’s calcium sensor was between 5 and 8 nm from its nearest calcium channel, within the range of values estimated experimentally in rat inner hair cells (~3 and 23 nm, Goutman and Glowatzki 2007).

Docked-and-tethered vesicles were co-localized more frequently than docked-not-tethered in our model because experimental evidence demonstrates that there are more docked-and-tethered vesicles in the region of the active zone where calcium channels are located (Lenzi et al. 2002; Roberts et al. 1990). Our model’s vesicle distributions included 50 +/- 13 docked vesicles within 200 nm of the active-zone center. In this area, docked-and-tethered vesicles outnumbered docked-not-tethered vesicles ~4:1. On average, 21 +/- 7 vesicles in this area were not co-localized to channels despite the relative excess of channels, in part because most channels were within 20 nm of their nearest neighbor, such that if one was co-localized with a vesicle a second vesicle could not physically fit above the neighboring channel. Rather than choosing vesicles to co-localize among the two populations, we attempted to increase the functional efficacy of docked-not-tethered vesicles by preferentially co-localizing them. As a result, 95% of docked-not-tethered vesicles within 200 nm of the active-zone center were co-localized, and the ratio of co-localized docked-and-tethered to docked-not-tethered was ~2:1 in our simulations.

Calcium channel gating kinetics were based on a combination of single-channel (Rodriguez-Contreras and Yamoah 2001) and whole-cell (Armstrong and Roberts 1998) current analysis. Opening and closing rate constants were derived assuming a four state Markov model, with three independent closed states and an open state. We computed a sigmoid conductance--voltage curve from whole-cell currents, and then scaled it by the maximum open probability of a single channel to create an open probability--voltage curve. This and the activation time constant of whole-cell currents were used to find voltage-dependent rate constants (Supplemental Fig. 1). Each channel was initialized in the fully closed state, then state transitions were evaluated probabilistically every 1 μs, subdividing when necessary to insure all transition probabilities remained below 0.1 (Gil et al. 2000). This method yielded realistic whole-cell current kinetics and single-channel open probabilities. In contrast, whole-cell current magnitudes estimated by scaling our single active zone current by 19 (Roberts et al. 1990) was one fourth of the experimentally observed value (~83 pA versus -363 pA in Armstrong and Roberts 1998), suggesting a significant number of calcium channels are extra-synaptic (see Rodriguez-Contreras and Yamoah 2001), an underestimate of the number of channels within each active zone, heterogeneity in calcium channel properties, or a discrepancy in single-channel conductance between preparations.

Three sets of source data were estimated from published graphical data: the location of calcium channels (Roberts 1994), the spatial distribution of synaptic vesicle densities (Lenzi et al. 2002), and the voltage-dependence of calcium current (Armstrong and Roberts 1998). For each of these data sets, we imported digital copies of published images into a graphics program (Photoshop, Adobe Inc.), and then manually converted the graphical information into numeric values.
RESULTS

STATISTICAL MODEL OF MULTIPLE RELEASED VESICLES

As a first step in understanding the ribbon’s role in precise spike timing, we sought to identify a mathematical relationship between the number of readily releasable vesicles and precision of auditory nerve spikes. Hair cell ribbon synapses are characterized by multiple readily releasable synaptic vesicles that exocytose in response to a few milliseconds of depolarization (Goutman and Glowatzki 2007; Keen and Hudspeth 2006; Moser and Beutner 2000), though a single vesicle may be sufficient to trigger an action potential in eighth nerve fibers (Geisler 1997). We hypothesized that the post-synaptic fiber responds to the first of several vesicles to yield a precise first-spike latency.

We estimated the latency of a single synaptic vesicle by integrating a kinetic model of calcium-dependent exocytosis in hair cells (Beutner et al. 2001). We assumed the vesicle was subject to a step increase to 50 μM calcium concentration based on physiological estimates from frog hair cells (Roberts 1993), though estimates from other ribbon synapse range from tens to hundreds of μM (Beutner et al. 2001; Heidelberger 2001). The stimulus caused a rapid onset of release probability that peaked 1.9 ms after the concentration change (Fig. 1a, blue trace). A single vesicle’s mean exocytic latency was 3.05 ms, and its latency jitter (standard deviation) was 1.52 ms.

Inverse Relationship between latency and the number of readily releasable vesicles

A mathematical relationship between exocytic latency of a single vesicle and the first-exocytosis latency of multiple vesicles can be derived from the statistics of random variables (see METHODS). A similar mathematical formalism was used to model how postsynaptic inputs contribute to spike-timing precision in the auditory brainstem (Xu-Friedman and Regehr 2005b). Synaptic vesicle exocytosis is quantal, therefore a vesicle’s exocytic latency on any one trial could be represented by a random variable that sampled the smooth release probability. With two readily releasable vesicles, the first latency distribution had a higher peak and shorter tail than that of a single vesicle, because the probability that both vesicles exocytosed late in the same trial was low (Fig. 1a, red trace). With five vesicles, first latency distribution had an even higher peak and truncated tail, and was therefore even more precise (black trace).

The statistical model predicted a monotonic decrease in latency mean and jitter with increasing number of readily releasable vesicles (Fig. 1b,c). Both first-latency mean and jitter decreased rapidly with the first few vesicles added, and tapered off with increasing vesicle counts. Similar relationships were found when the calcium level sampled by each vesicle was either lower or higher than that estimated for frog (Supplemental Fig. 2). Although these inverse relationships lend mathematical justification to the intuitive hypothesis put forth previously by others (Almers 1994; Moser et al. 2006), several simplifying assumptions required further investigation. For instance, the model assumed a step
increase in calcium concentration that stayed fixed throughout the stimulus. But the time course of calcium concentration should fluctuate in time due to the stochastic nature of calcium influx, and the actual concentration sensed by synaptic vesicles should depend on many factors, including their distance from calcium channels and properties of the endogenous calcium buffer. The model also assumed all vesicles sampled the same release probability, whereas variation in vesicle-channel distances could cause drastic differences in the underlying release probability of each vesicle.

To investigate the validity of these simplifying assumptions, we developed a stochastic (Monte Carlo) computational model of the hair cell ribbon synapse that tracked individual calcium channels, ions, and synaptic vesicles during stimulation. This approach enabled us to distinguish the contribution of distinct vesicle populations at a ribbon synapse. We modeled the ribbon synapse with the most thorough anatomical characterization to date --- the hair cell afferent synapse from the leopard frog (Rana pipiens) sacculus. In addition to the anatomical characterization of this synapse, the temporal precision of its afferent fiber has been extensively

FIG. 2: Spatial distribution of synaptic vesicles, free calcium, and exocytosis.
Top row is a cross section of the active zone (see inset) with synaptic ribbon in gray and skewed representation of the plasmalemma surface below to indicate depth. Middle row is full view of plasmalemma surface. Top and middle rows are from a single trial of a 10 ms depolarizing step to -20 mV, bottom row is the mean +/- SD of 250 trials. (a,b,c) Distribution of synaptic vesicles (circles), color coded into four morphologically distinct populations indicated in c. (d,e) Free calcium concentration averaged in (10 nm)³ volume elements during the last 0.1 ms of a 10 ms depolarization to -20 mV. Calcium channels are open squares. Color coding indicates concentration is equal to or greater than value indicated in legend (d, inset). (f) The number of vesicles that sampled the calcium concentrations indicated in the color-coded legend above, where each vesicle's sampled concentration was averaged across the entire 10 ms depolarization. (g,h) Vesicles that exocytosed by the end of the stimulation are filled with yellow, averages in (i).
documented, and provided a means of validating both our statistical and computational models.

**PROPERTIES OF SIMULATED ACTIVE-ZONE EXOCYTOSIS**

Three-dimensional reconstructions of this ribbon synapse revealed the distribution of synaptic vesicles in the active zone (Lenzi et al. 2002). The ribbon is spherical, ~400 nm across, and is surrounded by a halo of ~400 synaptic vesicles either touching or tethered to it by filaments. Additionally, a significant number of vesicles in the active zone are not associated with the ribbon and are either floating in the cytosol or docked to the plasmalemma. The location-dependent density of four morphologically distinct populations of synaptic vesicles were reported (see Table 2): docked vesicles not associated with the ribbon (docked-not-tethered), docked vesicles associated with the ribbon (docked-and-tethered), non-docked vesicles associated with the ribbon (tethered), and non-docked vesicles floating freely in the cytosol (outliers). Our model active zone matched the reported spatial distribution of vesicles from each population (Fig. 2a–c).

**Spatiotemporal modeling of intracellular calcium dynamics**

To drive exocytosis, we simulated calcium influx through 84 L-type calcium channels whose number, locations and physiology were inferred from experimental observations (see METHODS and Table 1). The channels were distributed in the plasmalemma directly below the synaptic ribbon, within 160 nm of the active-zone center (Fig. 2b, squares are calcium channels). Most channels were less than 20 nm from their nearest neighbor. Synaptic vesicles docked near the active-zone center were preferentially co-localized to calcium channels such that their calcium sensors were within 5 to 8 nm of a channel (Fig. 2b, circles surrounding squares) based on the possibility of their physical coupling (Hibino et al. 2002) and physiological estimates of their spacing in hair cells (~3 and 23 nm, Goutman and Glowatzki 2007). Maximal calcium influx was achieved with a voltage step to ~20 mV, which drove ~13,000 ions/ms into the active zone, distributed among the flickering channels.

As calcium ions entered the active zone they diffused freely and were bound by a combination of fixed and mobile (diffusible) calcium buffers. The experimentally identified endogenous mobile buffer included in our model, calretinin (Edmonds et al. 2000), should be kinetically equivalent to a millimolar of the fast exogenous buffer BAPTA according to in vitro buffer titration experiments (Roberts 1993). Previous modeling studies predicted that such a buffer limits free calcium ions to within 50 nm of open calcium channels, with a theoretical space constant of ~20 nm (Roberts 1994). Our probabilistic model of calcium influx and diffusion agreed with these predictions; free calcium levels were greater than 50 μM only within 20 nm of the open channels (Fig. 2d and e, red and pink squares are 10x10 nm). A few calcium ions diffused deeper into the cytosol, leading to isolated volume elements with calcium concentrations up to 10 μM as far as 400 nm from the calcium channel array (Fig. 2d and e, blue squares). Calcium buffers sequestered more than 99% of the calcium ions after they entered the cell, such that at the end of a 10 ms stimulus only 110 ions remained free in the cytosol.

We quantified the calcium concentrations sampled by morphologically distinct vesicle populations during a 10 ms step to ~20 mV (Fig. 2f). On average, each population contained at least 3 vesicles per trial that were exposed to calcium concentrations greater than 1 μM, with the docked-and-tethered population having the most and the outlier population having the fewest. Only vesicles co-localized with calcium channels sampled concentrations in excess of 25 μM, as expected from the location of calcium peaks in Fig. 2e. Only 10% of docked-not-tethered vesicles are found within 200 nm of the active-zone center (Lenzi et al. 2002) and could thus co-localize with calcium channels (see METHODS), therefore most of these vesicles were too far from the source of influx to experience calcium levels as moderate as 0.25 μM. Overall, the distribution of calcium sampled by each population quantifies the extent that docked vesicles (docked-not-tethered and docked-and-tethered populations) sampled higher calcium concentrations than non-docked vesicles (tethered and outlier populations), despite the fact that a majority of the vesicles in the active zone were non-docked (Fig. 2c).

**Vesicles co-localized with calcium channels dominate exocytosis**

Our model is the first to combine vesicle location with the spatiotemporal distribution of calcium to predict the spatial distribution of exocytosis at the hair cell active zone (see Supplemental Movies 1 & 2). Previous models of other active zones have combined these computations (e.g. Bennett et al. 2000; Gil et al. 2000; Shahrezaei and Delaney 2004), though they lacked the unique three-dimensional distribution of vesicles found at ribbon synapses. Previous hair cell models either computed spatially detailed calcium dynamics without exocytosis (e.g.
averaged calcium levels throughout the active zone (e.g. Sumner et al. 2002), computed exocytosis without considering calcium dependencies (Schnee et al. 2005), or embedded this step within a phenomenological model of signal transduction (e.g. Zhang et al. 2001).

We probabilistically simulated calcium-dependent exocytosis according to a kinetic model derived from mouse inner hair cell capacitance measurements (Beutner et al. 2001). Molecular mechanisms of regulated secretion are conserved across species (Bennett and Scheller 1993), therefore we assumed the rate constants derived from mouse cells in vitro at room temperature approximated those expected in vivo in the poikilothermic frog. Capacitance data suggests that vesicles not morphologically docked to the membrane can rapidly fuse when subjected to elevated calcium levels (Beutner et al. 2001). We therefore assumed all synaptic vesicles in our model were capable of calcium-dependent exocytosis irrespective of their location. In our simulations, an average of 30 synaptic vesicles exocytosed by the end of the stimulus (Fig. 2g and h, yellow filled vesicles reveal a representative trial). This amount of exocytosis was within the range observed experimentally with capacitance measurements of hair cells from the frog saccus after normalizing for differences in whole-cell calcium current (Edmonds et al. 2004; Rutherford and Roberts 2006).

Each morphologically distinct population contributed a different number of vesicles to the overall response of the active zone (Fig. 2i). On average, most exocytosed vesicles originated from either the docked-not-tethered or docked-and-tethered populations (9.4 and 19.3 vesicles/trial, respectively). The tethered and outlier populations contributed little to the active zone’s response (1.2 and 0.1 vesicles/trial). The minimal contribution of vesicles not docked at the active zone suggests that multi-quantal release in the frog synapse (Keen and Hudspeth 2006) is supported by parallel release of multiple docked vesicles rather than the compound or cumulative fusion of vesicles tethered along the ribbon (Edmonds et al. 2004; Parsons and Sterling 2003). The distribution of exocytosed vesicles among each population resembled the distribution of vesicles that sampled elevated calcium concentrations in Fig. 2f. Ninety-one percent of all exocytosed vesicles were co-localized with calcium channels, and of the vesicles co-localized with calcium channels, 98% exocytosed (Table 2). Based on this reliability of fusion in the presence of the endogenous buffer, we defined the co-localized vesicles as readily releasable.

**Temporal precision of synaptic vesicle exocytosis**

Our simulations predicted the time course of synaptic vesicle exocytosis at the hair cell ribbon synapse with sub-millisecond resolution (Fig. 3a). Each trial yielded a slightly different temporal pattern and number of exocytosed vesicles. After the stimulus onset there was a brief delay before any vesicle exocytosed, followed by an onslaught of vesicle fusions that subsided within 5 ms. Sometimes a vesicle fused before the stimulus started due to random channel openings at the -80 mV holding potential. The stochastic nature of synaptic vesicle exocytosis was captured by our probabilistic modeling technique, and the exocytosis times from any one trial were qualitatively similar to the limited examples of postsynaptic current onsets measured in frog hair cell afferents (Keen and Hudspeth 2006).

We computed the average exocytosis rate by combining the results from 250 trials of depolarization (Fig. 3b). It peaked at ~9 vesicles/ms.
within 3 ms of stimulus onset and rapidly declined during the stimulus. This decline, or adaptation, was due to the depletion of readily releasable vesicles. The half width of the peak at half height was less than 2 ms. Thus, the time course closely resembled measurements from afferent fibers from immature rat inner hair cells (Goutman and Glowatzki 2007), which had a peak rate of ~1 EPSC/ms. Furthermore, the time course was qualitatively similar to the initial time course of exocytosis rate measured with postsynaptic recordings during step depolarization of retinal ribbon synapses (Borges et al. 1995) and non-ribbon synapses with multiple contributing active zones (Sakaba and Neher 2001; Stevens and Tsujimoto 1995). However, the peak in exocytosis rate was an order of magnitude less precise than the first-spike latency of frog nerve fibers.

**First-exocytosis latency accounts for observed first-spike latency of frog afferent fibers**

The latency histogram of first-exocytosed vesicles from 250 trials was precisely timed, with a mean latency of 0.88 ms and jitter (standard deviation) of 0.28 ms (Fig. 3c). Reported first-spike latency of acoustically driven leopard frog nerve fibers can have mean values as low as 1 ms (Feng 1982). Although jitter was not directly reported, our lowest estimate from published data of several individual fibers was 0.3 ms (Feng 1982; Ronken et al. 1993). Thus our simulations predicted that the temporal precision of first-exocytosis latency was sufficient to account for the observed first-spike latency of frog nerve fibers.

The hair cell is a non-spiking neuron; its synapses are driven by a graded receptor potential proportional to the intensity of an acoustic stimulus (Crawford and Fettiplace 1980). We simulated the synapse’s response to voltage steps ranging from -65 to +25 mV that elicited calcium currents between 10 and 100% of the maximum. First-exocytosis latency was sufficiently precise to account for frog’s minimum first-spike latency with steps between -35 and +5 mV, when calcium current was as low as 80% of the maximum (Supplemental Fig. 3). Given the robustness of our results, we limited our remaining analyses to exocytosis elicited during maximal calcium influx.

**The anatomical basis of precise first latency**

We attempted to identify which vesicle(s) were required for precise first-exocytosis latency. We first compared the nerve’s spike latency to exocytosis latency of a single vesicle, then increased the readily releasable pool to include docked vesicles not associated with the ribbon (docked-not-tethered), and finally docked vesicles that were ribbon associated (docked-and-tethered). We then used our simulation results to update our statistical model relating the number of readily releasable vesicles to the precision of first-exocytosis latency. With the updated model, we predicted the minimum number of readily releasable vesicles required to account for the temporal precision of frog nerve fibers.

**Co-localized vesicle tracks calcium channel gating**

![Co-localized vesicle tracks calcium channel gating](image)
Perhaps the temporal precision of first-exocytosis latency was conferred by a single vesicle that exocytosed rapidly. To test this hypothesis, we evaluated the latency of a single vesicle in the middle of the calcium channel array that was co-localized within 8 nm of one channel and less than 20 nm from a second channel. Previous theoretical studies predicted that calcium concentration increases rapidly and to high levels near open calcium channels (e.g., Roberts 1994). Therefore a vesicle located within a few nanometers of one or more open calcium channels should rapidly undergo calcium-dependent exocytosis and likely contributes to the temporal precision of the synapse (Brandt et al. 2005), though the advantages of co-localization may be diminished by the calcium channel’s tendency to flicker open and closed.

During a representative single trial, the closest channel opened soon after the stimulus onset, then went through a series of open and closed states (Fig. 4a, red trace). Across 250 trials, the average single channel current increased to its steady-state value in less than 0.5 ms, remained relatively constant throughout the duration of the stimulus, and exhibited a brief tail current upon repolarization (Fig. 4a, black trace). The sub-unity maximum open probability of hair cell calcium channels (Rodriguez-Contreras and Yamoah 2001) caused frequent channel closures and resulted in an average steady-state current significantly less than that through an open channel (-0.05 versus -0.13 pA).

The single trial of channel gating in Fig. 4a yielded a similar (albeit noisier) time series of local calcium concentration, with peaks in the calcium concentration coinciding with when the channel was open and dips coinciding with when the channel was closed (compare red traces in Fig. 4a and b). When the channel was open, calcium concentration in a (10 nm)$^3$ volume of cytoplasm containing the channel was 120 μM, equivalent to a single free ion in the volume 7.2% of the time. When it was closed, concentration dropped to 22 μM, equivalent to an ion in the volume 1.3% of the time. The concentration was non-zero when the channel was closed due to diffusion of calcium from neighboring open channels and release of calcium from buffer molecules. The average calcium concentration mirrored the kinetics of the average single channel current and had a steady state value of 55 μM (Fig. 4b, black trace), a concentration sufficient to elicit rapid exocytosis in mouse inner hair cells (Beutner et al. 2001).

**Single vesicle is insufficient for temporal precision**

In most trials, the vesicle exocytosed during the stimulus, though on one trial it fused five milliseconds before the stimulus started due to spontaneous channel openings at the holding potential (data not shown). We evaluated the binding and unbinding of calcium ions that lead to vesicle exocytosis in a single trial (Fig. 4c, * marks vesicle fusion). The vesicle bound ions rapidly when the channel was open, but less frequently when the channel was closed, and sometimes released ions. During the initial channel opening the vesicle bound three calcium ions, one of which was released once
the channel closed. After a series of binding and unbinding ions, the vesicle eventually regained a fifth ion and fused 0.14 ms later. The single-vesicle latency distribution had a mean and jitter of 2.89 and 1.70 ms (Fig. 4d, gray bars), and was thus substantially slower and less precise than the first-spike latency of eighth nerve fibers. Thus, the evolution of calcium binding during channel openings and unbinding during channel closures significantly prolonged the latency of exocytosis compared to the theoretical minimum. This highlights the detrimental affect of low maximum open probability on exocytic latency. These findings portend the need for more than one readily releasable vesicle to precisely encode acoustic transients.

The transition from five bound ions to exocytosis (Fig. 4c) was governed by the rate constant of the calcium independent step in hair cell exocytosis. In theory, a vesicle subject to an instantaneous increase to infinitely high calcium levels would only be limited by this rate constant and exocytose with a latency mean and jitter of 0.7 ms, which is still less precise than the first-spike latency of eighth nerve fibers. Therefore a single vesicle was incapable of accounting for the nerve’s temporal precision.

Ribbon-associated vesicles are required for temporal precision

We tested the hypothesis that synaptic ribbons confer precise timing by increasing the number of readily releasable vesicles. We simulated the absence of an anchored synaptic ribbon by evaluating exocytosis of only the docked-not-tethered population, such that all vesicles normally associated with the ribbon were absent, including both the docked-and-tethered and tethered populations. On average, 91 docked vesicles remained in our simulations, nine of which were co-localized with calcium channels (Table 2). Anatomically, this amounted to a 30% reduction in the number of docked vesicles in our simulations, which was comparable to the 33% reduction in docked vesicle count measured at hair cell afferent synapses lacking ribbons (Khimich et al. 2005). We did not add vesicles to the cytosol normally occupied by the ribbon because our simulations of the intact synapse indicated that few non-docked vesicles exocytosed (Fig. 2i), therefore we likely underestimated the number of cytosolic vesicles in the ribbon-free active zone.

In the absence of a synaptic ribbon, exocytosis rate resembled a scaled version of the response of the intact active zone (Fig. 5a, compare green histogram and black trace). Both peaked within 3 ms of the stimulus onset, but magnitude decreased by ~66% without the ribbon-associated vesicles, as might be predicted from the relative contribution of the docked-and-tethered population to exocytosis in Fig. 2c. This decrease in response magnitude was comparable to the observed 75% reduction in readily releasable pool size in mice lacking anchored ribbons (Khimich et al. 2005). This decrease also matched a ~66% reduction in the fast component of exocytosis.
at active zones temporarily lacking anchored ribbons in goldfish bipolar cells, which we estimated from published data (Hull et al. 2006) assuming that exocytosis at intact synapses was unchanged (i.e. Khimich et al. 2005) and extra-synaptic exocytosis was negligible (Zenisek et al. 2003). Our ability to model the absence of a synaptic ribbon was essential to our analysis of its function, and the congruence between both anatomical and physiological measures of these specialized synapses when they lack ribbons validated our approach.

First-exocytosis latency was slower and less precise than predicted for the intact active zone, with a mean and jitter of 1.36 and 0.69 ms (Fig. 5b). Although the first latency histogram had a peak near 1 ms, several trials yielded first-exocytosis latencies between 2 and 5 ms, thereby increasing latency mean and jitter. We concluded that a readily releasable pool consisting solely of docked-not-tethered vesicles was insufficient for achieving the observed spike-timing precision of frog nerve fibers.

The synapse’s temporal precision was regained when our analysis also included docked vesicles associated with the ribbon (Fig. 5c,d). Exocytosis rate of all docked vesicles (docked-not-tethered and docked-and-tethered) was almost identical to that of the intact active zone (Fig. 5c, compare solid histogram and black trace). The histograms differed slightly only after the stimulus had been on for more than 3 ms, owing to the limited contribution of tethered and outlier populations. First-exocytosis latency with all docked vesicles matched that of the intact active zone, with a mean and jitter of 0.88 and 0.28 ms (compare Figs. 5d and 3c). The addition of ribbon-associated vesicles reduced first-exocytosis latency by increasing the likelihood of an early first fusion and eliminating late first fusions. We concluded that docked vesicles associated with the synaptic ribbon must contribute to the readily releasable vesicle pool in order to account for first-spike latency precision of frog nerve fibers.

**Comparison of statistical model and biophysical simulations**

Our simulations of active-zone exocytosis predicted a reduction in first-exocytosis latency mean and jitter as the size of the readily releasable vesicle pool increased. This result was in accordance with our preliminary statistical model that assumed all vesicles were subject to a step increase of 50 μM calcium concentration (Fig. 1). Using the results of our exocytosis simulations, we updated our statistical model to more accurately reflect the stochasticity of the underlying biophysical processes. The statistical model requires a single-vesicle latency distribution and the number of vesicles sampling that distribution; we replaced the single-vesicle latency distribution estimated for a constant step to 50 μM with the distribution resulting from biophysical simulations of a single co-localized vesicle in the frog active zone (Fig. 6a, black trace from Fig. 4d).

As with the preliminary statistical model, when two or five readily releasable vesicles sampled the latency distribution of a co-localized vesicle, the resultant first-latency distributions had increasingly higher peaks and shorter tails --- latency mean and jitter decreased (Fig. 6a, red and blue traces). The inverse relationships between latency and the number of readily releasable vesicles (Fig. 6b and c, gray lines) were very similar to those predicted with a calcium step (i.e. Fig. 1b,c) because the average calcium concentration sampled by the co-localized vesicle was 55 μM. We estimated the number of vesicles required to match frog nerve fiber data with the updated statistical model (red dashed lines). At least 15 and 16 vesicles were required to match the precision of first-spike latency mean and jitter, respectively.

We compared the statistical model’s predictions to exocytosis simulations. We first considered the
simulated latency of a single available vesicle, just *docked-not-tethered* vesicles, and all docked vesicles, again defining the vesicles co-localized with calcium channels as readily releasable (Fig. 6b and c, *large squares*). The number of readily releasable vesicles was always one during the single vesicle simulations but was variable in the *docked-not-tethered* and all docked simulations. The simulation results followed the trends of the statistical model’s predictions. However, the latency jitter of the *docked-not-tethered* population was markedly higher than predicted (Fig. 6c, compare green square and gray line).

To understand this apparent discrepancy, we refined our analysis by subdividing the original categories of vesicle populations into groups of trials with the same number of readily releasable vesicles (Fig. 6b and c, *small squares*). For example, we identified 7 of 250 trials with exactly 16 *docked-not-tethered* vesicles, and computed the latency of those 7 trials to be 1.02 +/- 0.35 ms. In these limited trials where there were sufficient numbers of *docked-not-tethered* vesicles in the active zone, latency and jitter nearly matched the performance of the nerve. The redefined groups of trials exquisitely tracked the statistical model’s predictions, and support our hypothesis that first-exocytosis latency can be reasonably inferred from the latency distribution of a single readily releasable vesicle and the number of readily releasable vesicles. This highlights the point that ribbon-associated vesicles did not constitute a physiologically unique pool of release-ready synaptic vesicles; ribbon-associated vesicles increased the size of a physiologically homogenous pool so it was consistently large enough for precision.

**Ribbon statistically ensures precise timing**

Simulations with and without ribbon-associated docked vesicles straddled the theoretical minimum number of vesicles required to match the nerve’s latency precision (Fig. 6b,c), substantiating the ribbon’s role in increasing the number of readily releasable vesicles to drive precisely timed auditory nerve spikes. Variance in the number of docked vesicles not associated with the ribbon (*docked-not-tethered*) made this population susceptible to elevated latency (mean and jitter) from trials with few readily releasable vesicles. Adding ribbon-associated docked vesicles shifted the hair cell synapse away from the steep part of the theoretical curves, such that variability in the number of all docked vesicles had little effect on the synapse’s latency. Therefore, ribbon-associated vesicles enhanced the precision of first-exocytosis latency not just by increasing the mean number of releasable vesicles, but also by reducing sensitivity to variance in readily releasable vesicle count.

**Ribbon-associated vesicles are essential for temporal precision even if more than one vesicle is required to trigger an action potential**

Physiological measurements support our assumption that a single vesicle was sufficient to elicit an action potential in the afferent (Glowatzki and Fuchs 2002; Siegel 1992). However, multi-quantal release is an emerging characteristic of the hair cell synapse (Glowatzki and Fuchs 2002; Keen and Hudspeth 2006) that might enable precise spike-time timing by rapidly depolarizing the post-synaptic fiber with neurotransmitter from multiple vesicles (Parsons and Sterling 2003). Accordingly, we examined temporal precision when more than one vesicle was required to elicit a spike. An analogous relationship between first-spike latency and the number of converging synaptic inputs, rather than the number of converging synaptic vesicles, was examined in neurons that respond to the temporal coincidence of two or more inputs (Xu-Friedman and Regehr 2005a). These coincidence detecting neurons behave like ideal integrators on short time scales, responding to the precise timing of the Nth input when exactly N inputs were required to elicit a spike.
We reevaluated our active-zone simulations assuming two or more vesicles were required to elicit a spike. The temporal precision of the second or fifth exocytosed vesicle also improved with increasing readily releasable pool size. Increasing the number of vesicles required to trigger a spike increased latency mean (Fig. 7a). In contrast, latency jitter was largely independent of coincident vesicle fusion requirements, especially when the readily releasable pool was greater than ~15 vesicles (Fig. 7b). Thus a large readily releasable pool was necessary for precise spike timing even when the afferent was a coincidence detector. We concluded that ribbon-associated vesicles were essential for exceeding first-spike latency precision of frog nerve fibers, irrespective of the number of vesicles required to trigger a spike.

**Ribbon-associated vesicles enhance temporal precision even with a slow calcium buffer**

A salient difference between hair cells for different species is the endogenous calcium buffer. In frog, turtle, and chicken hair cells, millimolar concentrations of the fast-binding buffer, BAPTA, were required to match the endogenous buffer’s affect on calcium dependent physiological processes (e.g. Roberts 1993). In contrast, the endogenous calcium buffer in mouse hair cells was matched with ~1 mM of the slow-binding buffer EGTA (Moser and Beutner 2000). Replacing the endogenous frog buffer with EGTA drastically increased the distribution of free calcium in the active zone, such that vesicles up to 200 nm from the calcium channels exocytosed (Supplemental Fig. 4, Supplemental Movies 3 and 4). Vesicles close to calcium channels exocytosed more rapidly due to higher average calcium levels whereas the more distant vesicles had slower kinetics (Supplemental Fig. 5). The contents of these more distant vesicles could in theory reach the synaptic cleft via calcium-dependent mechanisms of multi-quantal exocytosis such as cumulative or compound fusion (Neef et al. 2007; Parsons and Sterling 2003).

The time course of synaptic vesicle exocytosis in our simulations was dramatically different when EGTA replaced the frog’s endogenous buffer (Fig. 8a). The peak exocytosis rate with EGTA was more than two-fold greater and transpired almost a millisecond before the peak with calretinin. The high exocytosis rate was maintained throughout the stimulus and exhibited little adaptation; the average exocytosis rate during the last two milliseconds of the stimulus was 70% of the peak rate, compared to 5% with calretinin. After the stimulus ceased, the synapse exocytosed 20 vesicles per trial with EGTA, compared to 0.3 vesicles with calretinin. First-exocytosis latency mean and jitter was 0.6 and 0.15 ms, 30% faster and more precise than with calretinin (Fig. 8b), albeit at the expense of continued exocytosis after the stimulus terminated.

As with calretinin, we evaluated the anatomical basis of temporally precise exocytosis by examining the contribution of different vesicle populations. With EGTA, docked-not-tethered vesicles alone resulted in the exocytosis of 27 docked-not-tethered vesicles (Fig. 8c) and a first-exocytosis latency mean and jitter of 1.11 and 1.14 ms (Fig. 8d). The EGTA buffered synapse recovered its temporal precision when the docked-and-tethered population was
included (Fig. 8e,f). The exocytosis rate of all docked vesicles had a prominent peak between one and two milliseconds then adapted to a low level. The difference between exocytosis rate of all docked vesicles and the intact active zone demonstrated that tethered and outlier populations exocytosed late, thereby limiting adaptation. First-exocytosis latency of all docked vesicles matched the intact synapse, with a mean and jitter of 0.60 and 0.16 ms (compare Figs. 8b and 8f). We concluded that ribbon-associated vesicles also enhanced the precision of first-exocytosis latency with EGTA.
DISCUSSION

The spike-timing jitter of leopard frog eighth nerve fibers can be as low as 0.3 ms (estimated from data in Feng 1982; Ronken et al. 1993). Surprisingly, the theoretical lower bound of exocytic jitter for an individual vesicle is 0.7 ms. Jitter increases to 1.7 ms in simulations accounting for calcium channel gating and buffered diffusion (Fig. 4). Thus no single vesicle is capable of driving the nerve with requisite precision. The hair cell ribbon synapse compensates for the imprecise timing of individual vesicle fusions by encoding stimulus onset with the precisely timed fusion of the first of many readily releasable vesicles (Fig. 3). The synaptic ribbon is required to retain a pool of readily releasable vesicles sufficiently large to statistically ensure first-exocytosis latency is both short and reproducible (Fig. 6).

Validating the computational model

We carefully selected an experimental species and stimulus paradigm to minimize complications when juxtaposing simulation results derived from in vitro data to auditory nerve responses recorded in vivo. In particular, we chose the frog not only because its ribbon synapse is well characterized, but because analyzing data collected at room temperature from a poikilothermic species obviated the need for temperature compensation between preparations. Furthermore, we limited our analysis to minimum first-spike latencies because the high-intensity acoustic transients that drive such responses could be reasonably simulated with a brief square-pulse voltage step. To allay concern about the physiological relevance of our test voltages, we confirmed that the model could match the nerve’s minimum latency jitter for a range of simulated receptor potentials (Supplemental Fig. 3). Our selection of species and stimulus minimized the model’s free parameters, though our conclusions are drawn from computational results, and as such can only be as strong as the assumptions in the model.

One key assumption was that removing a synaptic ribbon could simply be modeled by removing all ribbon-associated vesicles from our analysis. Possibly, the ribbon acts as a calcium diffusion barrier that increases local calcium concentration near the channels (Roberts 1994). In our simulations, calcium ions rarely diffused more than 20 nm from open channels due to the frog’s endogenous calcium buffer (Fig. 2g,h), therefore the ribbon’s affect on local calcium levels should be negligible. Furthermore, simply removing ribbon-associated vesicles from our analysis was sufficient to account for experimentally observed reduction in both the number of docked vesicles (Khimich et al. 2005) and size of the readily releasable pool (Hull et al. 2006; Khimich et al. 2005). This validated our computational approach to removing the synaptic ribbon thereby enabling us to make conclusions about its role in hair cells.

Another key assumption was the co-localization of synaptic vesicles and calcium channels, as we defined these vesicles as readily releasable. There is no direct measure of the distance between vesicles and channels; we placed co-localized vesicles’ calcium sensors 5 to 8 nm from a channel, which was between two experimental estimates from rat inner hair cells (~3 and 23 nm, Goutman and Glowatzki 2007). Variation in this distance can significantly affect release probability due to the short space constant of calcium decay in frog hair cells (~20 nm, Roberts 1993). For example, in additional simulations where docked vesicles were not explicitly co-localized with channels, such that on average 28 vesicles were between 5 to 30 nm of their nearest channel, exocytosis magnitude decreased by more than 40% and first-exocytosis latency mean and jitter more than doubled (data not shown). Furthermore, when vesicles act as calcium diffusion barriers, release probability depends on the specific location of the sensor on the vesicle’s surface (Shahrezaei and Delaney 2004). Without more precise experimental measurements of the spacing between vesicles and calcium channels or the location of calcium sensors on a vesicle, our estimate of these parameters was necessarily approximate. The overall agreement between our standard simulations and multiple experimental results argues that our model was sufficiently accurate to support the hypothesis that ribbon-associated vesicles are essential for precise spike timing.

A ribbon synapse lacking a synaptic ribbon is not equivalent to a conventional synapse

It is important to distinguish the difference between a hair cell active zone lacking a synaptic ribbon and a conventional (non-ribbon) synapse. Some conventional active zones have many more docked vesicles than are found at an intact ribbon synapse (e.g. Gustafsson et al. 2002). Increasing vesicle count alone is insufficient to cause a reduction in first-exocytosis latency; our statistical model relates the number of physiologically active (readily releasable) vesicles to temporal precision. A ubiquitous characteristic of ribbon synapses is the exorbitant speed and quantity of synaptic vesicle exocytosis during brief stimuli (Edmonds et al. 2004; Moser and Beutner 2000; von Gersdorff et al. 1996) relative to most conventional synapses. These
physiological measures indicate that ribbon synapses have higher numbers of readily releasable vesicles than most conventional synapses, and perhaps are specialized to exocytose a greater proportion of anatomically docked vesicles (Moser et al. 2006).

**First- versus coincident-exocytosis detection**

Although a single vesicle likely triggers a spike in the postsynaptic afferent fiber (Geisler 1997), unitary postsynaptic current amplitudes measured at hair cell afferents are variable (Glowatzki and Fuchs 2002; Keen and Hudspeth 2006) and some may trigger a spike while others do not. Specialized neurons in the auditory system detect the coincidence of multiple subthreshold postsynaptic currents to precisely encode acoustic transients (e.g. Oertel et al. 2000). In our simulations, latency jitter can be equally precise if more than one vesicle is required to trigger a spike (Fig. 7b), suggesting the afferent fiber could detect the coincident exocytosis of synaptic vesicles to establish temporal precision. Other features of the afferent fiber’s response may provide clues to whether it detects first- or coincident-exocytosis events. Based on analogous studies of an auditory brainstem neuron (Rothman and Young 1996), a first-exocytosis detector should have a shorter mean latency, a higher spontaneous rate, and higher inter-spike interval variance than a coincident-exocytosis detector.

**Slow buffer improves temporal precision of first-exocytosis latency**

Replacing the endogenous calcium buffer in our simulations with a slow-binding surrogate causes a dramatic increase in the magnitude and temporal precision of active-zone exocytosis. With calretinin, vesicles co-localized to calcium channels dominate the exocytic response to a 10 ms stimulus. With EGTA, in addition to co-localized vesicles, many other vesicles exocytose. One possible explanation for increased temporal precision with EGTA is a functional increase in number of readily releasable vesicles. However, there are distinct kinetic components comprising active-zone exocytosis in the presence of EGTA (i.e. Fig. 8c). Docked vesicles within 200 nm of the active-zone center constitute the fastest kinetic component, while vesicles further out sample lower concentrations of calcium and have slower kinetics. Slightly modifying our statistical model of first-exocytosis latency to analyze the impact of vesicle populations residing different distances from the active-zone center, we found that those with the fastest kinetics almost exclusively determine the first-exocytosis latency distribution (data not shown). Therefore, the improvement of first-exocytosis latency with EGTA was due primarily to faster exocytosis kinetics of co-localized vesicles that sampled higher calcium concentrations (Supplemental Fig. 4).

**Rapid replenishment of the readily releasable pool for precise timing**
The predicted relationship between readily releasable pool size and temporal precision has several important implications for auditory nerve function. Release ready vesicles in hair cells are hypothesized to be transiently depleted during stimulation that results in neural adaptation (Furukawa et al. 1982; Goutman and Glowatzki 2007; Moser and Beutner 2000; Spassova et al. 2004). If temporal precision of the hair cell synapse depends on the size of the readily releasable pool, then pool depletion should degrade temporal precision (Fig. 9a). However, chick auditory nerve fibers maintain temporal precision during adaptation (Avissar et al. 2007). If the readily releasable pool is depleted sufficiently during stimulation to cause neural adaptation, but is replenished rapidly enough to maintain the pool size above a critical threshold, then temporal precision of ongoing tonal stimuli could be maintained (Fig. 9b). Our predictions of an asymptotic relationship between temporal precision and readily releasable pool size (Figs. 1c and 6c) suggests that reducing the pool size from its maximum value could reduce firing rate without sacrificing timing. Perhaps the rapid refilling of the hair cell’s readily releasable pool compared to other ribbon synapses (Edmonds et al. 2004; Moser and Beutner 2000; Spassova et al. 2004) is essential not only to support bouts of sustained transmitter release (Griesinger et al. 2005), but also to maintain its temporal fidelity.

We have argued that the ribbon is important for maintaining a large readily releasable pool that insures a rapid and reliable release event at sound onset (Figs. 5 and 8). Auditory nerve fibers from mice lacking anchored synaptic ribbons exhibit increased first-latency jitter, which is consistent with our predictions. However, the fibers phase-lock to transposed tones with the same temporal precision as fibers from mice with anchored ribbons (N. Strenzke, B. Buran, D. Khimich, M.C. Liberman, T. Moser, abstract). We propose that during stimulation, rapid recruitment of synaptic vesicles to the active zone could sufficiently increase the readily releasable pool size to encode prolonged acoustic stimuli with temporal precision (Fig. 9c). This proposition is bolstered by observations that readily releasable pool refilling is calcium dependent (Moser and Beutner 2000; Spassova et al. 2004) and that the sustained component of release, often interpreted as the mobilization of synaptic vesicles to the readily releasable pool, is normal in the mice lacking anchored ribbons (Khimich et al. 2005).

Thus we hypothesize that one important function of synaptic ribbons in hair cells is to catch and hold vesicles that are mobilized to the active zone during refilling. The ribbon could function as a depot to retain a sufficiently large number of mobilized vesicles to statistically ensure that first spike latency is both short and reproducible at times when refilling is not active --- such as at stimulus onset.

**Spike-rate adaptation in the frog auditory nerve**

One prominent feature of our simulations is the decrease in exocytosis rate during stimulation due to the exhaustion of the readily releasable pool (Figs. 3 and 5). Such exhaustion has been proposed to underlie rate adaptation of eighth nerve fibers (Furukawa et al. 1982; Goutman and Glowatzki 2007; Moser and Beutner 2000; Spassova et al. 2004). Thus it is interesting to consider how the extent and time course of exocytosis-rate adaptation in our model compares to spike-rate adaptation of frog nerve fibers. Unlike mammalian auditory nerve fibers, which often share a stereotyped adaptation profile, frog fibers vary widely in their degree and time course of adaptation (Lewis 1986; Megela and Capranica 1981; Ronken 1990) depending on which end organ the fiber innervates (reviewed in Smotherman and Narins 2000).

In contrast to our simulations of hair cell exocytosis, the fibers innervating saccular hair cells often exhibit no adaptation at all (Lewis 1986; Smotherman and Narins 2000). This apparent discrepancy might be explained by a post-synaptic mechanism that blunts the onset response of the fibers, such as neurotransmitter receptor saturation or refractoriness of the spike generator. Instead we favor the idea that rapid replenishment of the readily releasable pool in vivo counteracts exocytosis-rate adaptation (i.e. Fig. 9b) and would thus limit neural adaptation. Furthermore, our simulations with a slow-binding calcium buffer exhibited significantly less adaptation than those with the endogenous buffer (Fig. 8a). Thus, the differences in neural adaptation observed across the end organs and between fibers in frog auditory nerve might result from different endogenous buffering properties or different rates of vesicle replenishment in the hair cells that innervate them.

**Strategies for reducing synaptic jitter**

Our analysis of exocytic latency facilitates interpretation of how other synapses might be configured to achieve precise spike timing. Essentially there are two ways to improve timing: increasing the number of readily releasable vesicles (Fig. 6), or decreasing latency of each individual vesicle via faster exocytosis kinetics (Fig. 4 versus...
Supplemental Fig. 4). The former could be achieved by increasing the size of the synaptic ribbon, or perhaps by combining readily releasable vesicles from separate active zones. Afferent fibers in amphibians and avians receive inputs from several synaptic ribbons (Fischer 1992; Simmons et al. 1992). Perhaps the barn owl, whose auditory nerve fibers are among the most precisely timed (Koppl et al. 1993), detects first-exocytosis events across many active zones in its basilar papilla’s fovea (Fischer 1994). Another example where multiple active zones might combine their outputs to increase the number of readily releasable vesicles is in the calyx of Held, where a single presynaptic terminal contains up to 600 active zones (Satzler et al. 2002) and drives postsynaptic currents with a jitter of 0.02 ms (Borst et al. 1995).

The other means of improving timing is to increase the rate of exocytosis for each readily releasable vesicle. Hair cells may increase exocytosis kinetics relative to conventional synapses with specialized proteins, such as oterferlin, the putative hair cell specific calcium sensor (Roux et al. 2006). Another means of increasing exocytosis rate is to increase local calcium concentration by reducing the binding kinetics of the endogenous calcium buffer (Fig. 8). Calcium levels can also be increased by increasing the conductance of each calcium channel, increasing the number of channels, increasing the open probability of each channel, decreasing the distance between channels and vesicles, or creating local regions of high calcium with diffusion barriers such as the ribbon (Roberts 1994) or vesicles themselves (Shahrezaei and Delaney 2004). Furthermore, hair cell exocytosis kinetics increase with temperature (Nouvian 2007). Thus, the possibility of slow calcium buffers (Beutner et al. 2001) in combination with elevated body temperature is consistent with the remarkably precise first-spike latencies recorded from mammalian afferents (Heil and Irvine 1997).
This work was supported by the National Institute on Deafness and Other Communication Disorders Awards T32 DC 005363 (fellowship to J.H.W.), R01 DC 003783 (T.D.P.), and the Pennsylvania Lions Hearing Research Foundation (T.D.P.).
REFERENCES


FIGURE LEGENDS

FIG. 1: Statistical model predicts inverse relationship between readily releasable pool size and first-exocytosis latency.
(a) Latency histogram for a single synaptic vesicle subjected to a step increase in calcium concentration to 50 μM (blue line). Statistical estimate of first-exocytosis latency distributions if two (red) or five (black) vesicles sampled the single-vesicle distribution. (b) First-exocytosis latency mean and (c) jitter values calculated across a range of readily releasable vesicles (black points).

FIG. 2: Spatial distribution of synaptic vesicles, free calcium, and exocytosis.
Top row is a cross section of the active zone (see inset) with synaptic ribbon in gray and skewed representation of the plasmalemma surface below to indicate depth. Middle row is full view of plasmalemma surface. Top and middle rows are from a single trial of a 10 ms depolarizing step to -20 mV, bottom row is the mean +/- SD of 250 trials. (a,b,c) Distribution of synaptic vesicles (circles), color coded into four morphologically distinct populations indicated in c. (d,e) Free calcium concentration averaged in (10 nm)^3 volume elements during the last 0.1 ms of a 10 ms depolarization to -20 mV. Calcium channels are open squares. Color coding indicates concentration is equal to or greater than value indicated in legend (d, inset). (f) The number of vesicles that sampled the calcium concentrations indicated in the color-coded legend above, where each vesicle’s sampled concentration was averaged across the entire 10 ms depolarization. (g,h) Vesicles that exocytosed by the end of the stimulation are filled with yellow, averages in (i).

FIG. 3: Active-zone exocytosis is temporally precise.
(top) Stimulus bar indicates step to -20 mV from -80 mV. (a) Exocytosis raster from 25 trials. Each mark represents the exocytosis of a single synaptic vesicle. (b) Average exocytosis rate from 250 trials. (c) Histogram of first-exocytosis latency after the stimulus onset. Arrows indicate latency mean (blue) and jitter (red). Histograms binned at 0.2 ms.

FIG. 4: Individual vesicle exocytosis is not temporally precise.
(top) Stimulus bar indicates step to -20 mV from -80 mV. Panels (a,b,d) compare the results from a single trial (red lines) to the average of 250 trials (black lines in a,b; gray bars in d). (a) Influx through a single channel. (b)
Calcium concentration within a (10 nm)$^3$ volume element sampled by a co-localized vesicle, averaged over 10 μs. Mean +/- SD calcium concentration of this single trial during channel open or closed times during the 10 ms stimulation, and the average concentration during stimulation of 250 trials. (c) State transitions as vesicle bound five calcium ions (states 1 through 5) and exocytoosed (state 6, *). Initially the vesicle has no bound ions (state 0). (d) Distribution of exocytosis times, binned at 0.2 ms.

FIG. 5: Ribbon-associated vesicles are required for temporal precision.
Comparison of exocytosis rates (top) or first-exocytosis latency histograms (bottom) for different populations of vesicles: (a,b) docked-not-tethered population, (c,d) all docked vesicles (docked-not-tethered and docked-and-tethered populations). In a and c, results from the intact active zone (all vesicles) are depicted by the solid line.

FIG. 6: Exocytosis simulations validate the statistical model.
(a) Latency histogram for a single co-localized vesicle (from Fig. 4d) smoothed with a three-point triangle filter (blue line). Statistical estimate of first-exocytosis latency distributions if two (red) or five (black) vesicles sampled the single-vesicle distribution. (b,c) First latency mean and jitter. Gray lines are statistical model predictions updated with biophysical model results. Large squares are stimulation results from Figs. 4 and 5, where the number of readily releasable vesicles equals the number of vesicles co-localized with channels. Small squares are simulation results divided into groups with the same number of readily releasable vesicles. Only groups with five or more trials were included (mean 13 trials/group). Horizontal dashed lines are minimum values observed in frog nerve fibers. Vertical dashed lines are the corresponding number of predicted vesicles.

FIG. 7: Ribbon-associated vesicles are required for temporal precision during coincidence detection.
Comparison of the latency mean (a) and jitter (b) of first exocytoosed vesicle, second vesicle, and fifth vesicle. Each point estimated from small groups of simulations with the same number of readily releasable vesicles as in Fig. 6. Dashed lines are minimum reported frog values.

FIG. 8: Ribbon-associated vesicles are required for temporal precision with a slow calcium buffer.
(a,b) Comparison of exocytosis rate and first-exocytosis latency with the standard buffer, calretinin, and the slow buffer EGTA. Simulations otherwise identical. (c--f) Replicates Fig. 5, but with EGTA.
FIG. 9: Readily releasable pool depletion and temporal precision of the auditory nerve.

(a) Readily releasable pool depletion degrades temporal precision during prolonged stimuli. Cumulative synaptic vesicle exocytosis (green line) increases during stimulation as readily releasable vesicles (black and red line) exocytose without replenishment. At stimulus onset the large readily releasable pool (black line) exceeds the minimum size required for temporal precision (gray dashed line), as defined in Figure 6c. During prolonged stimuli the readily releasable pool is depleted below this critical size (arrow), such that it is no longer large enough to support precise timing (red line). (b) Rapid replenishment of the readily releasable pool curtails depletion so temporal precision is maintained. (c) In the absence of a synaptic ribbon, the readily releasable pool is initially small and temporal precision is poor (red line). Stimulus-induced recruitment of vesicles to the active zone increases the size of the readily releasable pool so it is sufficiently large (arrow) to enable temporally precise responses to prolonged stimuli (black line). All traces are based on a three state kinetic model describing the number of readily releasable vesicles, \( V_{rr} \), where \( V_{rr} \) is irreversibly replenished and depleted with rate constants \( k_r \) and \( k_d \), respectively: 
\[
\frac{1}{(k_r + k_d)} \frac{d V_{rr}}{dt} + V_{rr} = V_{rr-max} \frac{k_r}{(k_r+k_d)}.
\]

Supplemental FIG. 1: Derivation of calcium current voltage-dependent rate constants.

Calcium channel gating kinetics were based on a combination of single-channel (Rodriguez-Contreras and Yamoah 2001) and whole-cell (Armstrong and Roberts 1998) current analysis. Opening and closing rate constants were derived assuming a four state Markov model, with three independent closed states and an open state. (a) We computed a sigmoid conductance--voltage curve from whole-cell currents, and then scaled it by the maximum single-channel open probability of 0.4 to create the open probability--voltage curve. (b) Time constant fit to experimental data using a third-order exponential (Armstrong and Roberts 1998). (c & d) Algebraic manipulation of the open probability and activation kinetic values were used to find voltage-dependent opening and closing rate constants for the current, where the opening rate equals the open probability divided by the time constant, and the closing rate equals one minus the open probability divided by the time constant (Hudspeth and Lewis 1988). Model fits (solid lines) are compared to data points (triangles) extrapolated from experimental data. Opening rate = 
\[
a \times \frac{(V_m - V_o)}{(e^{V_o - V_m} - 1)},
\]
where \( a = 220 \text{ mV}^{-1} \text{s}^{-1} \), \( V_o = -40 \text{ mV} \), \( V_a = 11.5 \text{ mV} \), \( a = 200 \text{ mV} \). Closing rate = 
\[
B_o + \frac{b}{(1 + e^{(V_m - V_o)/V_b})},
\]
where \( B_o = 1900 \text{ s}^{-1} \), \( b = 3200 \text{ mV} \text{ s}^{-1} \), \( V_o = 2 \text{ mV} \), \( V_b = 9 \text{ mV} \).
Supplemental FIG. 2: Inverse relationship between readily releasable pool size and first-exocytosis latency is robust to different calcium concentrations. The statistical model predicted a monotonic decrease in latency mean and jitter with increasing number of readily releasable vesicles when the calcium sensor was subjected to 50 μM calcium (Fig. 1). We extended this finding by examining how the latency and its dependency on vesicle number were impacted by different concentrations of calcium at the release site. (a) Latency histogram for a single synaptic vesicle subjected to a variety of step increases in calcium concentration including: 25μM (light blue line), 50 μM (blue line), 100 μM (light red line), and 200 μM (red line). Increasing calcium concentration decreased latency mean and jitter of a single vesicle, however, even with a concentration of 200 μM these values were not sufficient to account for the observed precision of auditory nerve first-spike latencies (as low as 1.0 +/- 0.3 ms in Feng 1982; Ronken et al. 1993). (b & c) First-exocytosis latency mean and jitter values calculated across a range of readily releasable vesicle counts and calcium concentrations. For all concentrations, first-latency mean and jitter decreased rapidly with the first few vesicles added, and then tapered off with increasing vesicle counts. Interestingly, at 25 μM even a large number of readily releasable vesicles were insufficient to account for the nerve’s response, suggesting that release sites in the frog hair cell active zone are subject to higher concentrations. Such elevated concentrations are consistent with the previous findings in frog (Roberts 1993), and emphasize the necessity of calcium channel-synaptic vesicle co-localization (nanodomains) for hair cell synaptic function (Brandt et al. 2005; Roberts 1994).

Supplemental FIG. 3: Voltage-dependence of calcium influx, exocytosis magnitude and temporal precision. The hair cell is a non-spiking neuron; its synapses are driven by a graded receptor potential proportional to the intensity of an acoustic stimulus (Crawford and Fettiplace 1980). We simulated the synapse’s response to 10 ms voltage steps ranging from -65 to +25 mV and documented calcium current amplitude (a), exocytosis magnitude (b), and first-exocytosis latency mean (c) and jitter (d). Calcium current amplitude was averaged during the entire 10 ms stimulus and exocytosis magnitude expressed as numbers of vesicles. Red dashed lines in panels c and d are minimum latency values from frog nerve fibers. First-exocytosis latency was sufficiently precise to account for frog’s minimum first-spike latency with steps between -35 and +5 mV, when calcium current was as low as 80% of the maximum. Results from 100 trials, error bars are SD.

Supplemental FIG. 4: Spatial distribution of synaptic vesicles, free calcium, and exocytosis in the presence of a slow calcium buffer.
A salient difference between hair cells for different species is the endogenous calcium buffer. In frog, turtle, and chicken hair cells, millimolar concentrations of the fast-binding buffer, BAPTA, were required to match the endogenous buffer’s affect on calcium dependent physiological processes (e.g. Roberts 1993). In contrast, the endogenous calcium buffer in mouse hair cells was matched with ~1 mM of the slow-binding buffer EGTA (Moser and Beutner 2000). Replacing the endogenous frog buffer (equivalent of 1.2 mM calretinin with 4 independent binding sites per molecule) with EGTA drastically increased the distribution of free calcium in the active zone and increased exocytosis. Top row is a cross section of the active zone (see inset) with synaptic ribbon in gray and skewed representation of the plasmalemma surface below to indicate depth. Middle row is full view of plasmalemma surface. Top and middle rows are from a single trial, bottom row is the mean +/- SD of 250 trials.

(a,b,c) Distribution of synaptic vesicles (circles), color coded into four anatomically distinct populations indicated in c. (d,e) Free calcium concentration averaged in (10 nm)³ volume elements during the a 10 ms depolarization to -20 mV. Calcium channels are open squares. Color coding indicates concentration is equal to or greater than value indicated in legend (d, inset). Open squares are channels. (f) The number of vesicles sampling each color-coded calcium concentration during the 10 ms step to -20 mV. (g,h) Vesicles that exocytose by the end of the stimulation are filled with yellow, averages in (i). In contrast to calretinin (Fig. 2, Supplemental Movies 1 & 2), which limited exocytosis to calcium channels sites, EGTA allowed vesicles up to 200 nm from the calcium channels to exocytose (also see Supplemental Movies 3 and 4). The contents of these more distant vesicles could in theory reach the synaptic cleft via calcium-dependent mechanisms of multi-quantal exocytosis such as cumulative or compound fusion (Neef et al. 2007; Parsons and Sterling 2003)

**Supplemental FIG. 5: Individual synaptic vesicle exocytosis is not temporally precise in the presence of a slow calcium buffer.**

We also examined the impact of calcium buffer on the exocytosis kinetics of a single vesicle by replacing the endogenous buffer calretinin with the slow buffer EGTA. (top) Stimulus bar indicates step to -20 mV from -80 mV. Panels (a,b,d) compare the results from a single trial (red lines) to the average of 250 trials (black lines in a,b; gray bars in d). (a) Influx through a single channel. (b) Calcium concentration within a (10 nm)³ volume element sampled by a co-localized vesicle, averaged over 10 μs. Mean +/- SD calcium concentration from this single trial during channel open or closed times, and the mean concentration averaged across the full 10 ms of each of the 250 trials. Notice a momentary increase in local calcium levels 1 ms before the voltage stimulus due to diffusion of
calcium ions from other nearby channels that randomly flickered open. (c) State transitions as vesicle bound five calcium ions (states 1 through 5) and exocytosed (state 6, *). In this example, the vesicle bound at least three ions during the 10 ms preceding the voltage step due to random channel opening and diffusion from nearby channels. (d) Distribution of exocytosis times, binned at 0.2 ms. Single-vesicle exocytosis is almost two-times faster and more precise with EGTA than in the presence of the slow endogenous buffer (cf. Fig. 4) due to the higher average calcium levels achieved close to calcium channels. However, even in the presence of a slow buffer, the single-vesicle latency is not sufficiently precise to account for the temporal precision of auditory nerve first-spike latencies.

**Supplemental Movie 1: Spatiotemporal dynamics of buffered diffusion and exocytosis with calretinin.**

A cross section of the active zone with synaptic ribbon in gray, calcium channels as open squares, and synaptic vesicles color coded by their morphological population (*docked-and-tethered* in black, *docked-not-tethered* in green, *tethered* in blue, and *outliers* in gray). Each frame of the movie shows the spatial distribution of free calcium ions, averaged in (10 nm)$^3$ volume elements for 0.1 ms (>100 μM in dark red and >10 μM in dark blue; see legend on right for more color-coding of concentrations). A vesicle is filled with yellow when it exocytosed anytime up to the displayed time. The stimulus is a 10 ms voltage step to -20 mV, starting at Time = 0 ms, from a holding potential of -80 mV. The mobile calcium buffer is 1.2 calretinin which is modeled as 4.8 mM of independent, high-affinity calcium binding sites.

**Supplemental Movie 2: Spatiotemporal dynamics of buffered diffusion and exocytosis at the plasmalemma surface with calretinin.**

The surface of the active zone with outline of the synaptic ribbon’s shadown in gray, calcium channels as open squares, and synaptic vesicles color coded by their morphological population (*docked-and-tethered* in black, *docked-not-tethered* in green). Each frame of the movie shows the spatial distribution of free calcium ions, averaged in (10 nm)$^3$ volume elements for 0.1 ms (>100 μM in dark red and >10 μM in dark blue; see legend in Fig. 1d or Supplemental Fig. 4d for more color-coding of concentrations). A vesicle is filled with yellow when it exocytosed anytime up to the displayed time. The stimulus is a 10 ms voltage step to -20 mV, starting at Time = 0 ms, from a holding potential of -80 mV. The mobile calcium buffer is 1.2 calretinin which is modeled as 4.8 mM of independent, high-affinity calcium binding sites. With calretinin (Fig. 2, Supplemental movie 1 & 2), exocytosis is tightly linked to locations of calcium channels. All exocytosed vesicles lay atop a channel, and only two vesicles
above channels did not exocytose. These simulations graphically reinforce the importance of calcium channel-
synaptic vesicle co-localization at this synapse, which might be achieved through physical interactions identified at
other synapses (Rettig et al. 1996; Sheng et al. 1996).

**Supplemental Movie 3: Spatiotemporal dynamics of buffered diffusion and exocytosis with EGTA.**

A cross section of the active zone with synaptic ribbon in gray, calcium channels as open squares, and synaptic
vesicles color coded by their morphological population (*docked-and-tethered* in black, *docked-not-tethered* in green,
tethered in blue, and *outliers* in gray). Each frame of the movie shows the spatial distribution of free calcium ions,
averaged in (10 nm)$^3$ volume elements for 0.1 ms (>100 μM in dark red and >10 μM in dark blue; see legend on
right for more color-coding of concentrations). A vesicle is filled with yellow when it exocytosed anytime up to the
displayed time. The stimulus is a 10 ms voltage step to -20 mV, starting at Time = 0 ms, from a holding potential of
-80 mV. The mobile calcium buffer is EGTA, and this slow-binding buffer allows calcium ions to diffuse several
hundred nm into the cytosol before they are captured, resulting in high calcium levels far from the calcium channels
and ultimately in the exocytosis of distant synaptic vesicles. The contents of these more distant vesicles could in
theory reach the synaptic cleft via calcium-dependent mechanisms of multi-quantal exocytosis such as cumulative or
compound fusion (Neef et al. 2007; Parsons and Sterling 2003). After the stimulus ceases, the slow-binding kinetics
of EGTA are incapable of rapidly removing excess ions, and the cytosol’s concentration remains elevated for several
ms.

**Supplemental Movie 4: Spatiotemporal dynamics of buffered diffusion and exocytosis at the plasmalemma
surface with EGTA.**

The surface of the active zone with outline of the synaptic ribbon’s shadow in gray, calcium channels as open
squares, and synaptic vesicles color coded by their morphological population (*docked-and-tethered* in black, *docked-
not-tethered* in green). Each frame of the movie shows the spatial distribution of free calcium ions, averaged in (10
nm)$^3$ volume elements for 0.1 ms (>100 μM in dark red and >10 μM in dark blue; see legend in Fig. 1d or
Supplemental Fig. 4d for more color-coding of concentrations). A vesicle is filled with yellow when it exocytosed
anytime up to the displayed time. The stimulus is a 10 ms voltage step to -20 mV, starting at Time = 0 ms, from a
holding potential of -80 mV. The mobile calcium buffer is EGTA. The extensive and far-reaching exocytosis with
EGTA as the endogenous buffer (Supplemental Fig. 5, Supplemental Movies 3 & 4) dramatically contrasts to the
presence of calretinin (Fig. 2, Supplemental Movie 2), where exocytosis is limited to calcium channel sites as expected with a nanodomain organization of the active zone. These findings highlight the critical role that the endogenous buffer plays in creating nanodomains.
TABLE LEGENDS

Table 1: Active-zone simulation parameters.
List of values and citations employed to parameterize our stochastic model.

* Estimated from calmodulin, a comparably sized calcium buffer.

** Estimated in (Edmonds et al. 2000) from cardiac cell calretinin.

*** Estimated from (Lenzi et al. 2002).

Table 2: Distribution of synaptic vesicles and exocytosis by population.
Experimental vesicle counts from (Lenzi et al. 2002), where they used the nomenclature: Outlying, docked; Synaptic body associated, docked; Synaptic body associated, not docked; Outlying, not docked. We modified their anatomical nomenclature in an effort to better reflect each group’s functional significance in our active-zone simulations, specifically contrasting the two groups that dominate exocytosis: docked-not-tethered versus docked-and-tethered vesicles. Co-localized vesicles were 5-8 nm from the nearest calcium channel (see METHODS). Simulated exocytosis was driven by 10 ms of maximal calcium influx. Percent exocytosis computed on trial-by-trial basis then averaged across trials. All values are mean +/- SD of 250 trials, except experimental vesicle count variations, which are SEM.
(a) Statistical First Latency

- 1 vesicle, $3.0 \pm 1.5$ ms
- 2 vesicles, $2.2 \pm 1.0$ ms
- 5 vesicles, $1.4 \pm 0.6$ ms

(b) Latency Mean (ms)

(c) Latency Jitter (ms)

Number of Readily Releasable Vesicles
(a)

(b) Exocytosis per Trial
30.5 +/- 4.5 vesicles

(c) First Latency
mean 0.88 ms
jitter 0.28 ms

Time (ms)
DOCKED-not-TETHERED

Exocytosis per Trial
9 +/- 6 vesicles

First Latency
mean 1.36 ms
jitter 0.69 ms

ALL DOCKED VESICLES

29 +/- 4 vesicles

mean 0.88 ms
jitter 0.28 ms
(a) Statistical First Latency

- 1 vesicle, 2.9 +/- 1.7 ms
- 2 vesicles, 2.2 +/- 1.0 ms
- 5 vesicles, 1.5 +/- 0.6 ms

(b) Latency Mean (ms)

(c) Latency Jitter (ms)
a) Ribbon, Without Refilling
   \textit{temporal precision degraded}

b) Ribbon, With Refilling
   \textit{temporal precision maintained}

c) No Ribbon, With Refilling
   \textit{temporal precision rescued}
Table 1: Active-zone simulation parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Calcium Current</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Channel count</td>
<td>84 channels</td>
<td>(Roberts et al. 1990)</td>
</tr>
<tr>
<td>Channel locations</td>
<td>Fig. 2e</td>
<td>Replicate location of presumed calcium channels in (Roberts 1994)</td>
</tr>
<tr>
<td>Channel conductance</td>
<td>2.1 pS</td>
<td>(Rodriguez-Contreras et al. 2002)</td>
</tr>
<tr>
<td>Maximum Open Probability</td>
<td>0.4</td>
<td>(Rodriguez-Contreras and Yamoah 2001)</td>
</tr>
<tr>
<td>Reversal Potential</td>
<td>41.7 mV</td>
<td>Fit to (Armstrong and Roberts 1998)</td>
</tr>
<tr>
<td>Kinetic Model States</td>
<td>3 close, 1 open</td>
<td>(Hudspeth and Lewis 1988)</td>
</tr>
<tr>
<td>Kinetic Model Rate Constants</td>
<td>Supplemental Fig. 1</td>
<td>Fit to (Armstrong and Roberts 1998; Rodriguez-Contreras and Yamoah 2001)</td>
</tr>
<tr>
<td><strong>Calcium</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diffusion Coefficient</td>
<td>223 µm²s⁻¹</td>
<td>(Allbritton et al. 1992)</td>
</tr>
<tr>
<td>Resting Concentration</td>
<td>48 nM</td>
<td>(Lumpkin and Hudspeth 1998)</td>
</tr>
<tr>
<td><strong>Fixed Buffer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{on}$</td>
<td>1357 µM⁻¹s⁻¹</td>
<td>(Lumpkin and Hudspeth 1998)</td>
</tr>
<tr>
<td>$K_D$</td>
<td>200 nM</td>
<td>(Lumpkin and Hudspeth 1998)</td>
</tr>
<tr>
<td>Concentration</td>
<td>0.610 mM</td>
<td>(Lumpkin and Hudspeth 1998)</td>
</tr>
<tr>
<td><strong>Calretinin</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Diffusion Coefficient</strong></th>
<th>32 µM²s⁻¹</th>
<th>(Falke et al. 1994)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 µM⁻¹s⁻¹</td>
<td>(Dong et al. 1996)**</td>
</tr>
<tr>
<td></td>
<td>1.5 µM (high affinity) 0.5 mM (low affinity)</td>
<td>(Schwaller et al. 1997)</td>
</tr>
<tr>
<td><strong>Binding sites</strong></td>
<td>4 (high affinity) 1 (low affinity)</td>
<td>(Schwaller et al. 1997)</td>
</tr>
<tr>
<td><strong>Protein Concentration</strong></td>
<td>1.2 mM</td>
<td>(Edmonds et al. 2000)</td>
</tr>
<tr>
<td><strong>EGTA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Diffusion Coefficient</strong></td>
<td>113 µM²s⁻¹</td>
<td>(Tsien 1980)</td>
</tr>
<tr>
<td></td>
<td>2.5 µM⁻¹s⁻¹</td>
<td>(Naraghi and Neher 1997)</td>
</tr>
<tr>
<td></td>
<td>180 nM</td>
<td>(Naraghi and Neher 1997)</td>
</tr>
<tr>
<td><strong>Concentration</strong></td>
<td>1 mM</td>
<td>(Moser and Beutner 2000)</td>
</tr>
<tr>
<td><strong>Vesicle Distribution</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Location-dependent packing densities</strong></td>
<td>***</td>
<td>(Lenzi et al. 2002), inhibited values</td>
</tr>
<tr>
<td><strong>Ribon diameter</strong></td>
<td>395 nm</td>
<td>(Lenzi et al. 2002)</td>
</tr>
<tr>
<td><strong>Ribon anchor</strong></td>
<td>50 nm</td>
<td>***</td>
</tr>
<tr>
<td><strong>Vesicle diameter</strong></td>
<td>40 nm</td>
<td>***</td>
</tr>
<tr>
<td><strong>Tether length</strong></td>
<td>20 nm</td>
<td>***</td>
</tr>
<tr>
<td><strong>Exocytosis Kinetics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Kinetic Model States</strong></td>
<td>0–5 ions bound, fused</td>
<td>(Beutner et al. 2001)</td>
</tr>
<tr>
<td><strong>Binding rate</strong></td>
<td>27.6 µM⁻¹s⁻¹</td>
<td>(Beutner et al. 2001)</td>
</tr>
<tr>
<td><strong>Unbinding rate</strong></td>
<td>2150 s⁻¹</td>
<td>(Beutner et al. 2001)</td>
</tr>
<tr>
<td><strong>Cooperativity constant</strong></td>
<td>0.4</td>
<td>(Beutner et al. 2001)</td>
</tr>
<tr>
<td><strong>Calcium-independent fusion</strong></td>
<td>1695 s⁻¹</td>
<td>(Beutner et al. 2001)</td>
</tr>
</tbody>
</table>
rate
Table 2: Distribution of synaptic vesicles and exocytosis by population.

<table>
<thead>
<tr>
<th></th>
<th>Docked-not-Tethered</th>
<th>Docked-and-Tethered</th>
<th>Tethered</th>
<th>Outlier</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experimental Vesicle Counts</strong></td>
<td>81 +/- 37</td>
<td>43 +/- 8</td>
<td>357 +/- 42</td>
<td>1997 +/- 320</td>
</tr>
<tr>
<td><strong>Simulated Vesicle Counts</strong></td>
<td>91 +/- 31</td>
<td>39 +/- 11</td>
<td>361 +/- 38</td>
<td>1914 +/- 615</td>
</tr>
<tr>
<td><strong>Co-localized with channels</strong></td>
<td>9 +/- 6</td>
<td>19 +/- 7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Simulated Exocytosis:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(vesicles)</td>
<td>9.3 +/- 6.4</td>
<td>19.8 +/- 6.9</td>
<td>1.3 +/- 1.4</td>
<td>0.1 +/- 0.4</td>
</tr>
<tr>
<td>(% of population)</td>
<td>10 +/- 7</td>
<td>52 +/- 18</td>
<td>0.39 +/- 0.46</td>
<td>0.01 +/- 0.02</td>
</tr>
<tr>
<td>(% of co-localized)</td>
<td>98 +/- 5</td>
<td>98 +/- 3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>