

# A Subtraction Mechanism of Temporal Coding in Cerebellar Cortex

Brian E. Kalmbach,<sup>1</sup> Horatiu Voicu,<sup>3</sup> Tatsuya Ohyama,<sup>1</sup> and Michael D. Mauk<sup>1,2</sup>

<sup>1</sup>Center for Learning and Memory and <sup>2</sup>Section of Neurobiology, The University of Texas at Austin, Austin, Texas 78712, and <sup>3</sup>Department of Neurobiology and Anatomy, University of Texas Medical School at Houston, Houston, Texas 77030

The temporally specific learning displayed by the cerebellum facilitates mechanistic analysis of neural timing and temporal coding. We report evidence for a subtraction-like mechanism of temporal coding in cerebellar cortex in which activity in a subset of granule cells specifically codes the interval between the offset of two mossy fiber inputs. In a large-scale cerebellar simulation, cessation of one of two ongoing mossy fiber inputs produces a robust temporal code in the population of granule cells. This activity supports simulation learning in response to temporal patterns of stimuli, even when those same stimuli do not support learning when presented individually. Using stimulation of mossy fiber inputs to the cerebellum as training stimuli in rabbits, we confirmed these unusual predictions in a cerebellum-dependent form of learning. Analysis of the simulations reveals a specific working hypothesis for this temporal subtraction process that involves interactions between granule cells and the inhibitory Golgi cells. The results suggest how feedforward inhibition, such as that present in the cerebellar cortex, can contribute to temporal coding.

## Introduction

Timing is an important aspect of behavior, and thus the mechanisms of timing and temporal coding are fundamental to our understanding of neural function (Mauk and Buonomano, 2004). The cerebellum displays a capacity for temporally specific learning that, for motor function, contributes to adaptation and coordination of movement (Raymond et al., 1996; Medina and Mauk, 2000; Ohyama et al., 2003). Adaptation of smooth pursuit eye movements (Medina et al., 2005; Medina and Lisberger, 2008) and conditioning of eyelid responses (Perrett et al., 1993; Medina et al., 2000) are particularly clear and experimentally tractable examples. In particular, the relatively direct way in which eyelid conditioning engages the cerebellum (Fig. 1A) makes it a fruitful system for analyzing the mechanisms of learning, information processing, and timing. Together, these factors also facilitate the application of large-scale computer simulations in such studies (Medina and Mauk, 2000; Ohyama et al., 2003). Here we present results from cerebellar simulations and eyelid conditioning experiments that provide evidence for a novel, subtraction-like mechanism of temporal coding in the cerebellum.

In eyelid conditioning, repeatedly pairing a tone conditioned stimulus (CS) with a reinforcing unconditioned stimulus (US) (stimulation near the eye) promotes acquisition of precisely timed conditioned responses: the eyelids close in response to the tone CS, and these responses are timed to peak near US onset

(Schneiderman and Gormezano, 1964; Mauk and Ruiz, 1992; Perrett et al., 1993). Evidence from reversible lesions, recordings of Purkinje cells in cerebellar cortex, and from simulations suggests that the learned timing of cerebellar responses is mediated by temporal coding mechanisms in the cerebellar cortex (Medina et al., 2000; Jirenhed et al., 2007; Medina and Lisberger, 2008; Rasmussen et al., 2008).

Using large-scale computer simulations of the cerebellum, we find that the cessation of one of two ongoing mossy fiber (MF) inputs produces a robust temporal code in a population of granule cells. This code, which lasts hundreds of milliseconds, involves different subsets of granule cells becoming active at different times in the interval between the offsets of the two mossy fiber inputs. Because this activity is unique to the time period when one input is present and the second input has just been removed (subtracted), we refer to this as temporal subtraction. For convenience, we refer to the interval between stimulus offsets as the “offset interval.” Feedforward inhibition acting on granule cells makes this granule cell temporal code unique to the offset interval. Because of these processes, the simulations display robust, well-timed learning for patterns of mossy fiber inputs that are otherwise too long in duration to support learning when presented individually (Schneiderman and Gormezano, 1964). After characterizing the properties of this learning in the simulation, we implemented the same unusual training trials using direct electrical stimulation of mossy fiber inputs to the cerebellum as the CS in eyelid conditioning experiments. The properties of this learning in the rabbit match the simulation results.

An analysis of the simulation suggests that learning under these circumstances relies on feedforward inhibition of granule cells by the population of cerebellar Golgi cells. In addition to revealing a novel mechanism of temporal coding in the cerebellar

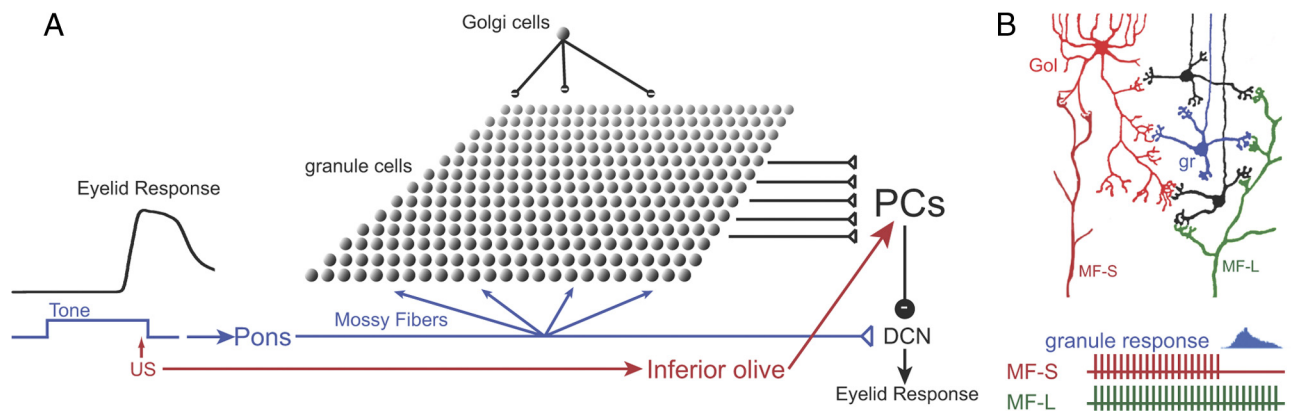
Received Aug. 11, 2010; revised Nov. 10, 2010; accepted Nov. 22, 2010.

This work was supported by National Institutes of Health Grants MH74006 and MH46904. We thank Brenda Houck for her help with the rabbit experiments.

Correspondence should be addressed to either Brian E. Kalmbach or Michael D. Mauk, Center for Learning and Memory, The University of Texas at Austin, 1 University Station, Stop C7000, Austin, TX 78712-0805. E-mail: brian@clm.utexas.edu or mike@clm.utexas.edu.

DOI:10.1523/JNEUROSCI.4212-10.2011

Copyright © 2011 the authors 0270-6474/11/312025-10\$15.00/0



**Figure 1.** A schematic representation of cerebellar pathways and their relevance to delay eyelid conditioning. **A**, In delay eyelid conditioning, the tone conditioned stimulus is conveyed to the cerebellum via activation of mossy fibers that originate in the pons (blue pathway). The US is conveyed to the cerebellum via activation of climbing fibers originating in the inferior olive (red pathway). Mossy fibers excite both granule and Golgi cells in the cerebellar cortex, which provide input to the Purkinje cells (PCs). Plasticity at the granule-to-Purkinje cell synapses and the mossy fiber-to-DCN synapses mediate learning and expression of conditioned eyelid responses. **B**, Connectivity in the cerebellar cortex that suggests the possibility for a subtraction-like mechanism of temporal coding. Mossy fibers excite both granule (gr) and Golgi (Gol) cells, and Golgi cells inhibit granule cells. With the presentation of two mossy fiber inputs, one short (MF-S, red) and one long (MF-L, green), there may be granule cells (blue) that are excited by the long input and are inhibited by the short input via activating a Golgi cell (light red) that inhibits the granule cell. The hypothesis is that this would activate such granule cells at times during the interval between the offsets of the two inputs (the offset interval, blue histogram).

cortex, these data suggest how feedforward inhibition can enhance the generation of temporal codes.

## Materials and Methods

**Subjects and surgical procedures.** Data were obtained from six male New Zealand albino rabbits (*Oryctolagus cuniculus*; Myrtle's Rabbitry). Treatment of animals and surgical procedures were in accordance with National Institutes of Health guidelines and an institutionally approved animal welfare protocol.

Animals were surgically implanted with two laterally spaced (1 mm) tungsten stimulating electrodes (tip exposed to obtain impedance of  $\sim 100$ – $200$  k $\Omega$ ; A-M Systems) in the left middle cerebellar peduncle (5.5 mm lateral, 16 mm ventral, and 3 mm anterior from lambda). Before surgery, a preanesthetic (40 ketamine and 5 mg/kg acepromazine) was injected subdermally, and each rabbit was positioned in a stereotaxic restrainer such that lambda was 1.5 mm ventral to bregma. General anesthesia was maintained with 2% isofluorene mixed in oxygen. Stimulating electrodes, a bolt to attach the eyelid detector to the skull, and anchor and ground screws were secured in place with dental acrylic. Two stainless steel stimulating electrodes were implanted in the peri-orbital muscles of the eye. Subjects were given postoperative analgesics and antibiotics for 2 d after surgery and were allowed to recover for 1 week before experiments began.

**Conditioning.** Subjects were trained in custom-designed, well-ventilated and sound-attenuating chambers. Chambers were equipped with isolated pulse stimulators (model 2100; A-M Systems) used to deliver trains of electrical pulses through the periorbital electrodes as the US. Separate stimulators were used to time the delivery of constant-current pulses by stimulus isolators (model A360; World Precision Instruments) connected to the electrodes implanted in the middle cerebellar peduncle. To measure eyelid position, an infrared emitter/detector was attached to the head stage to record movements of the left external eyelid. These detectors measure eyelid position by measuring the amount of infrared light reflected back to the detector, which increases as the eye closes. At the start of each daily training/test session, the detector was calibrated by delivering the US to elicit maximum eye closure. The amplification of the signal was adjusted to match an assumed maximum eye closure of 6 mm.

Stimulus presentation was controlled by custom-designed software. Subjects were given daily eyelid conditioning sessions that comprised 12 blocks of 9 trials (one CS alone trial and eight paired trials per block). The CS consisted of trains of constant-current cathodal stimulation (50 Hz, 40  $\mu$ s pulse width, 100  $\mu$ A) delivered through the electrodes implanted in the middle cerebellar peduncle. Four rabbits were trained with a subtraction

pattern of stimulation through the implanted electrodes. For this training, the stimulus train lasted 3000 ms through one electrode (MF-short) and lasted 3500 ms through the other electrode (MF-long). The train duration delivered through each electrode was counterbalanced across subjects: two rabbits were trained with MF-long delivered through the medial electrode and two rabbits were trained with MF-long delivered through the lateral electrode. For two rabbits, the duration of the trains delivered through each electrode was 3500 ms. The US was a 50 ms train of constant-current pulses (100 Hz, 1 ms pulse width, 2–3 mA) presented during the last 50 ms of the 3500 ms train. We stimulated mossy fibers in the middle cerebellar peduncle rather than their cell bodies to avoid unwanted activation of fibers of passage or antidromic activation of forebrain regions. Rabbits trained with the subtraction pattern of input were trained 1 additional day after the day in which they first reached 60% responding in a session. Criterion was reached during day 2 in two rabbits and day 4 in two rabbits. Rabbits trained with stimulation through both electrodes lasting 3500 ms were trained for 5 d.

After initial training sessions, rabbits were administered test sessions in which an aspect of the CS was manipulated during non-reinforced probe trials. The first test session involved probe trials in which stimulation through one electrode was omitted. In the second test session, the duration of MF-short was either shortened to 2700 ms or lengthened to 3300 ms while the duration of MF-long was kept constant.

Rabbits trained using the subtraction pattern of inputs were then retrained using a different offset interval. This involved keeping the duration of MF-short constant (3000 ms) and varying the duration of MF-long. Rabbits were trained for 2 d with a MF-long duration of 3800 ms (750 ms offset interval) before switching to training with a MF-short duration of 3300 ms (250 ms offset interval) for 2 d.

Finally, electrode sites were marked by passing 200  $\mu$ A, anodal current for 15 s and then switching polarity for 15 s through electrodes. Animals were killed with an overdose of sodium pentobarbital and were perfused through the heart with 10% Formalin. Brains were embedded in gelatin and sectioned at 80  $\mu$ m using a freezing microtome. Sections were mounted and stained with cresyl violet. Examination of these histological slides revealed that all electrodes were placed in the middle cerebellar peduncle.

**Data analysis.** For each trial, eyelid responses were digitized (200 Hz, 12 bit resolution) and stored for subsequent offline analysis using custom software. Analysis involved the 200 ms before CS onset and the next 5 s. Trials in which eyelid movements greater than 0.3 mm were made in the 200 ms before CS onset were automatically excluded from analysis by the analysis software. A conditioned response was defined as an eyelid movement of a least 0.3 mm within the interstimulus interval. Several response

**Table 1. Connectivity of simulation**

Cell type ( <i>n</i> )	Inputs, source (+/−): number	Number of outputs
Mossy fibers (600)	N/A	Granule cells (+): 80
Climbing fibers (4)	Deep nucleus (−): 2	Deep nucleus (−): 2
		Purkinje (+): 2
Granule (12,000)	Mossy fibers (+): 4	Purkinje (+): 1
	Golgi cells (−): 4	Basket (+): 1
		Stellate (+): 1
Golgi (900)	Mossy fibers (+): 4	Granule (−): 53
	Granule cells (+): 64	
Basket (96)	Granule cells (+): 125	Purkinje (−): 4
Stellate (240)	Granule cells (+): 50	Purkinje (−): 1
Purkinje (24)	Granule cells (+): 500	Deep nucleus (−): 2
	Stellate cells (−): 10	
	Basket cells (−): 16	
	Climbing fiber (+): 1	
Deep nucleus (8)	Mossy fibers (+): 75 <sup>a</sup>	Climbing fibers (−): 1
	Purkinje cells (−): 6	

<sup>a</sup>At simulation inception, these synapses are very weak. They are essentially the potential for a synapse that can be created (strengthened). This is in line with evidence that excitatory synapses are formed in the deep nuclei during eyelid conditioning (Kleim et al., 2002). Also, because strengthened synapses are sparse, the ratio of excitatory-to-inhibitory synapses onto the deep nucleus cells approximates 1:1, which is more in line with known connectivity.

measures were captured by the analysis software, including the following: latency to criterion (latency to a response of 0.3 mm), latency to onset (latency to first inflection in sweep of a conditioned response), and amplitude. *t* tests (two-tailed) and ANOVA followed by Tukey's *post hoc* comparisons were used to test for differences.

**Computer simulations.** The details of the simulation have been published previously (Medina and Mauk, 2000; Medina et al., 2000, 2001, 2002; Mauk and Ohyama, 2004). The full source code for the simulation can be obtained by contacting the senior author (M.D.M.). Briefly, the construction of the simulation emulates the known synaptic organization and physiology of the cerebellum (Eccles et al., 1967). In accordance with the ratio of cell types within the cerebellum, the simulation contained leaky integrate-and-fire representations of 600 mossy fibers, 12,000 granule cells, 900 Golgi cells, 96 basket cells, 240 stellate cells, 24 Purkinje cells, 8 deep cerebellar nuclei (DCN) cells, and 4 climbing fibers. With these cell ratios and the connectivity between the Purkinje and basket cells, the stimulation is intended to represent a parasagittal strip in which all Purkinje cells receive input from climbing fibers of the same type (i.e., those activated by an eyelid stimulus). The change in membrane potential, each 1 ms time step, was calculated according to synaptic and leak conductances and to membrane capacitance. Spikes were initiated when membrane potential exceeded threshold. Threshold increased for each spike and returned exponentially to the baseline value. Synaptic conductances were based on time constants derived from *in vitro* studies.

The simulation captures the geometric relationships and divergence and convergence ratios of synaptic connections within the cerebellum. It accomplishes this by generating a two-dimensional grid of granule cells, Golgi cells, and mossy fiber glomeruli. Within this grid are specified areas over which a simulated neuron can contact a member of the population of target neurons. The criteria for these specifications are based on published accounts of the spatial relationships of connections within the cerebellum. For example, because the axons of granule cells run transversely through the cerebellar cortex, the contact area for granule cells was a narrow rectangle. Although this area constrained the range over which cells could make connections, the connections were determined randomly in a way that obeyed known divergence and convergence ratios (for example, each granule cell could only receive four mossy fiber inputs) (Table 1).

The granule cell-to-Purkinje cell synapse and the mossy fiber-to-DCN synapse were modifiable. A granule cell-to-Purkinje cell synapse underwent long-term depression (LTD) or long-term potentiation (LTP) depending on whether or not its activity fell within a 100 ms time window preceding a climbing fiber input. Synapses active within this window underwent LTD, whereas synapses active outside underwent LTP. Mossy fiber-to-DCN synapses active within a time window of an abrupt pause

in Purkinje cell activity underwent LTP, whereas those active during strong Purkinje activity underwent LTP.

Eyeblink conditioning was simulated by generating mossy fiber and climbing fiber activity in accordance with empirical data. Each of the 600 mossy fibers was assigned a background firing rate between 1 and 40 Hz. To mimic activation of mossy fibers as a MF-short and MF-long, no more than 3% of (18) mossy fibers were randomly selected to represent MF-short and (no more than) another 18 as MF-long (for simulations used in the figures: MF-long, 12; MF-short, 17). The firing rate of these cells was elevated to an average of 80 Hz during the relevant stimulus period. All mossy fiber activity was stochastic with the target firing rate for any given time used to determine the probability of activating an excitatory conductance in ways that made the actual, noisy activity match on average the target (e.g., 80 Hz for the mossy fibers encoding the MF-long). The activation of an excitatory conductance for the four climbing fibers served to mimic the presentation of the US (each Purkinje cell received US-driven climbing fiber input). The averaged and smoothed activity of the eight deep nucleus neurons was used to represent the output of the simulation and the predicted “eyelid response” of the simulation.

## Results

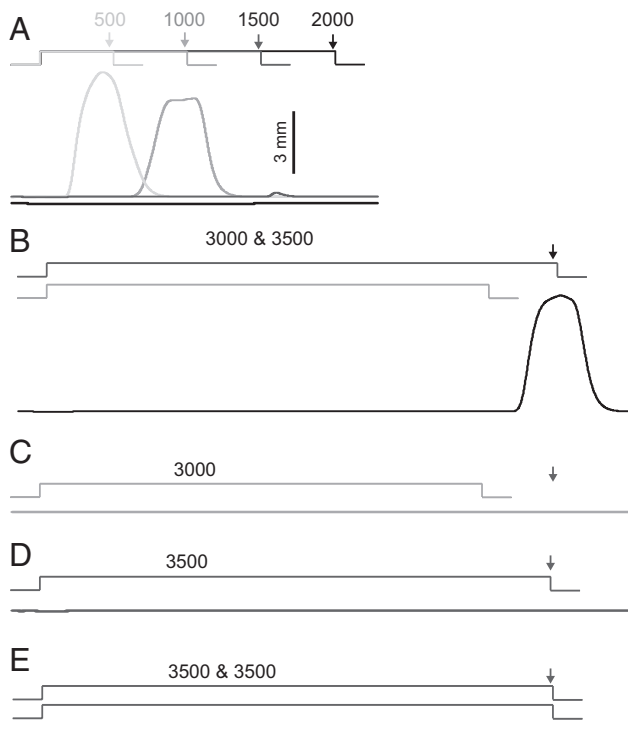
The temporal subtraction hypothesis is based on cerebellar cortex circuitry that influences the transformation of mossy fiber inputs to the activation of granule cells (Fig. 1*B*). The typical granule cell receives excitatory inputs from four different mossy fibers and receives no more than four inhibitory inputs from Golgi cells. The Golgi cells themselves also receive excitatory inputs from mossy fibers (not shown in Fig. 1*A* for clarity but schematized in Fig. 1*B*) in addition to their excitatory inputs from granule cells (Palkovits et al., 1971, 1972). Therefore, these inputs implement both feedforward and feedback inhibition of granule cells, respectively (Eccles et al., 1967; Ito, 1984). Figure 1*B* schematically represents a plausible combination of inputs for some small percentage of granule cells given two mossy fiber inputs presented in a temporal subtraction pattern (i.e., two subsets of active mossy fibers with one terminating before the other). For convenience, we will refer to such mossy fiber inputs as MF-short and MF-long. In this schematic representation, the blue granule cell receives excitatory input from MF-long and inhibitory input from a Golgi cell that is excited by MF-short. For such a granule cell, there would be both excitation and inhibition during the period when both mossy fibers are active but only excitation after the offset of MF-short. In principle, this circuitry suggests the possibility that (1) there will be granule cells activated at specific times during the offset interval (the time interval between the offsets of MF-short and MF-long), and (2) this activity could support learning of timed responses when a US is presented after the offset of MF-short.

## Temporal subtraction in a computer simulation of the cerebellum

We began by testing this prediction in a large-scale computer simulation of the cerebellum that can replicate most of the behavioral properties of eyelid conditioning (Medina and Mauk, 2000; Medina et al., 2000, 2001, 2002; Ohyama et al., 2002; Mauk and Ohyama, 2004). The simulation was presented with two mossy fiber inputs in which the MF-long lasted 500 ms longer than the MF-short. The simulation was trained by pairing this pattern of mossy fiber inputs with the US (brief excitatory conductance applied to the climbing fibers), in which the US was delivered at the end of the MF-long.

It is essential to ensure that any learning promoted in this way is specific to the temporal subtraction pattern and is not simply





**Figure 2.** Results from a computer simulation of the cerebellum during inputs that mimic delay eyelid conditioning. **A**, The decline in simulated conditioning as the duration of the conditioned stimulus (conveyed by mossy fiber inputs) increases. Identical simulations were trained with four intervals (500, 1000, 1500, 2000 ms) between the onset of a mossy fiber input and the presentation of a climbing fiber input that mimics the US (arrows). Whereas training with intervals of 500 or 1000 ms promotes robust learning by the simulation (sample responses after training shown below), training at an interval of 1500 ms barely promotes detectable learning and there is no learning at all for an interval of 2000 ms). **B**, Learning in the simulation during “subtraction training.” When the same simulation is trained with two mossy fiber inputs, one lasting 3000 ms and the other lasting 3500 ms, with the US presented at the end of the 3500 ms input (arrow), robust learning of a well-timed response is promoted. A sample conditioned response is shown below. **C–E**, Various control conditions showing that the learning in **B** is specific to the subtraction pattern of mossy fiber inputs. **C**, No learning was produced in the simulation with the 3000 ms input paired with the US at 3500 ms. **D**, There is also no learning with the 3500 ms input alone paired with the US. **E**, Finally, there is no learning with both inputs presented together for 3500 ms and paired with the US at 3500 ms.

the result of pairing the mossy fiber inputs with the US. To do so, we used mossy fiber stimulus durations that were too long to support learning if simply paired with the US (Schneiderman and Gormezano, 1964). Figure 2A shows results from training identical simulations using increasingly long mossy fiber inputs that are paired with the US. As in rabbit eyelid conditioning (Schneiderman and Gormezano, 1964), there is no detectable learning with mossy fiber durations of 3000 ms. Thus, to test the temporal subtraction hypothesis, the cerebellar simulation was trained with a MF-short that lasted 3000 ms and a MF-long that lasted 3500 ms. As shown in Figure 2B, this temporal subtraction pattern of training resulted in robust learning by the simulation of well-timed responses whose peaks occurred near the time of US presentation. This learning was specific to the short–long temporal pattern because training in the control condition in which each input lasted 3500 ms did not promote learning (Fig. 2E), nor did learning occur in two additional controls in which either the MF-short or the MF-long was omitted (Fig. 2C,D). Sample responses from these simulations for the subtraction pattern of training are shown in Figure 3.

Next, we addressed whether these inputs produced a true temporal code sufficient to mediate learned responses whose timing was determined by the time of US presentation. To test this, two identical simulations were trained with the same MF-short/MF-long input but paired with the US either 250 or 750 ms after MF-short offset (instead of 500 ms as above). The timing of the learned responses from these three conditions was significantly different ( $F_{(2,321)} = 419.99$ ,  $p < 0.0001$ ), and, in each case, the responses were timed to peak near US onset (Fig. 3). Together, these simulation results predict that a temporal subtraction-like process may be implemented by the synaptic organization of the cerebellar cortex. This process would make learning possible for temporal patterns of two inputs under conditions in which learning is not possible for either input alone or with both inputs presented without the subtraction pattern. Because these same eyelid conditioning procedures are possible in rabbits, the next step was to subject these simulation-derived predictions to empirical test.

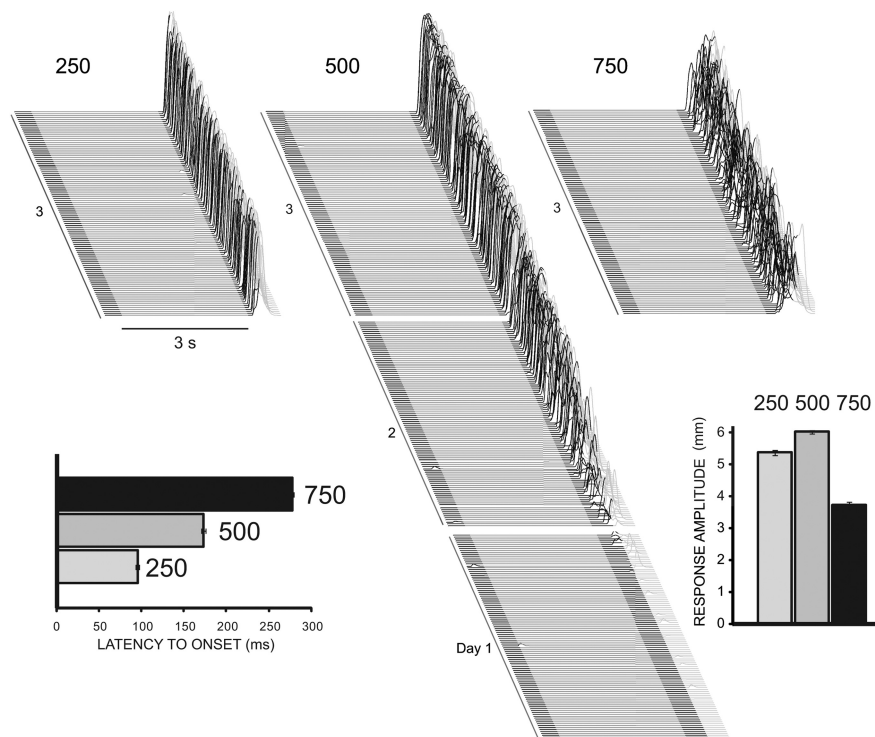
### Temporal subtraction tests in rabbit eyelid conditioning

To implement these same tests in rabbit eyelid conditioning requires substitution of direct electrical stimulation of mossy fiber inputs to the cerebellum for the normal tone CS. This is necessary to obtain precise control over the timing of the inputs and to avoid unwanted activation of forebrain inputs arriving at different times during the trial (Kalmbach et al., 2010b). Two stimulating electrodes were implanted in the middle cerebellar peduncle of six rabbits. The electrodes were laterally spaced 1 mm to activate different subsets of mossy fibers with each. Stimulation of mossy fibers with electrodes placed in the middle cerebellar peduncle, rather than at their primary source in the pontine nuclei, reduces the chances for unwanted activation of axons other than mossy fibers.

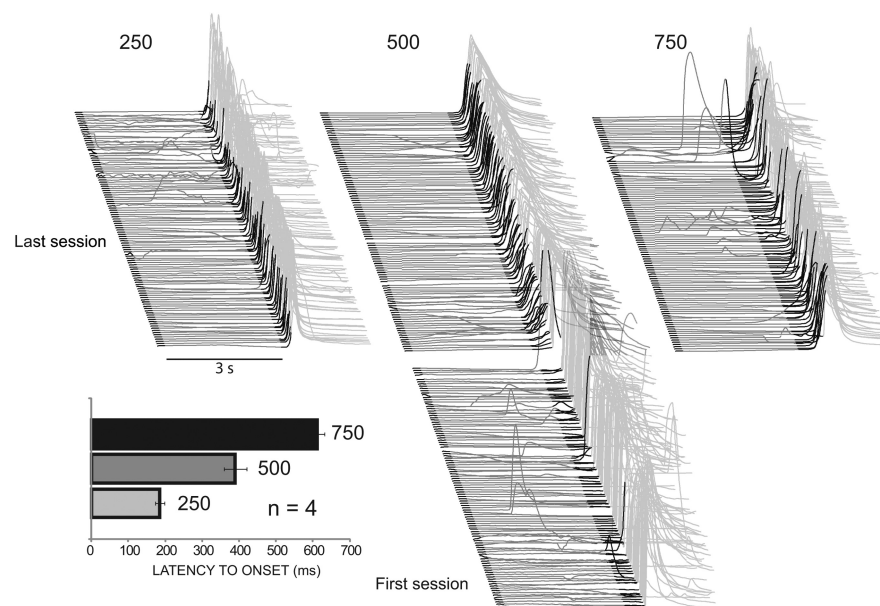
Using stimulation through these electrodes, rabbits were trained with procedures that paralleled those used for the simulations. As shown in Figure 4, eyelid conditioning procedures using a MF-short of 3000 ms and a MF-long of 3500 ms promoted the acquisition of robust learned responses ( $65.6 \pm 12.4\%$  response rate on the last day of training), whereas there was no detectable learning in the control condition in which each stimulus lasted 3500 ms ( $1.0 \pm 1.0\%$  response rate on the last day of training;  $t_{(4)} = 3.58$ ,  $p = 0.02$ ; control condition not shown). To test the learned timing of the responses (Mauk and Ruiz, 1992), each rabbit was retrained with the US presented either 250 or 750 ms after MF-short offset. As with the simulations, there was clear evidence for learned timing: comparing the latencies to onset for the three conditions (US presented 250, 500, or 750 ms after MF-short offset) again showed significant differences, with the responses in each condition tending to peak near the time of the US ( $F_{(2,6)} = 31.07$ ,  $p < 0.001$ ) (Fig. 4).

### Inputs controlling the responses of simulations and rabbits

The ability to control the two mossy fiber inputs in both the simulations and the rabbits makes it possible to present probe trials to evaluate the relative contributions of the two inputs to learning. In both rabbits and simulations, we observed responses to probe presentations of MF-long alone ( $51.39 \pm 22.50\%$  for rabbits; 58.3% for simulation) but not MF-short alone ( $2.78 \pm 1.6\%$  for rabbits; 0 for simulation) (Fig. 5). However, response likelihood to presentations of MF-long alone was highly variable, with some subjects responding robustly and others rarely. In the simulation and in rabbits that responded robustly to presentations of MF-long, we observed that the responses to MF-long



**Figure 3.** Analysis of the learning mediated by the simulation during temporal subtraction training. Sample raw “responses” of the simulation over three sessions of training are shown down the middle, with conditioned responses appearing at the beginning of the second session. For these responses, the pre-input period is shown in black, the 3000 ms in which both mossy fiber inputs are active is shown in light gray, and the 500 ms period in which only the long mossy fiber input is active is shown again in black. Responses during this period are conditioned responses mediated by the subtraction pattern. When similar simulations were trained with the US presented either 250 or 750 ms after the offset of the short mossy fiber input, the resulting conditioned responses were timed to peak near US onset. Note the latencies to onset of the three sets of responses shown across the top of the figure. The differences in response latencies for the three conditions (US presented at 250, 500, and 750 ms) is quantified in the leftmost graph. The graph at the right quantifies the amplitudes of the responses in the three conditions.



**Figure 4.** Temporal subtraction training in rabbit eyelid conditioning. Display of the data and the groups are the same as Figure 3. The training sessions shown in the middle are the first and third sessions of a rabbit trained with a 3000 ms mossy fiber input and a 3500 ms mossy fiber input. As with Figure 3, the responses to the left and to the right are from the same subject after retraining in which the US was presented either 250 or 750 ms after the offset of the 3000 ms mossy fiber input. Note again the differences in the timing of the responses in these three conditions. These differences are quantified for all four subjects on the graph at the bottom left.

alone had significantly shorter onset latencies than responses to presentations of MF-long and MF-short together (Fig. 5) ( $t_{(1)} = 43.77$ ,  $p = 0.01$ ). Finally, presenting probe trials with different durations of MF-short, while keeping the duration of MF-long fixed, resulted in responses whose onset varied with the duration of MF-short ( $F_{(2,6)} = 99.90$ ,  $p < 0.001$  for rabbits;  $F_{(2,41)} = 111.66$ ,  $p < 0.001$  for simulation). These results, which are presented in Figure 6, show that, although responses can be evoked in well-trained rabbits or simulations to presentation of only MF-long, both inputs are required for learning and for the expression of properly timed responses. Because the timing of conditioned responses was time locked to the offset of MF-short during probe trials, the learning appears to be supported primarily by granule cells that become active after the offset of MF-short. These unusual observations support the simulation-derived hypothesis that the cerebellar cortex uses the offset of MF-short to create a temporal code that is used in learning. Because the mechanisms within the cerebellar cortex that generate this code are not known, we returned to an analysis of the simulation to develop a working hypothesis.

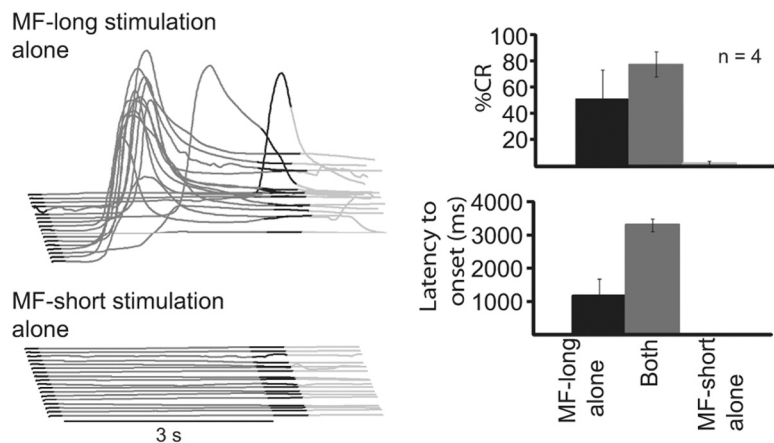
### Simulation-suggested mechanisms of temporal subtraction

As noted above, the initial mechanistic rationale for the temporal subtraction hypothesis was the potential for granule cells that receive excitation from MF-long and feedforward inhibition from MF-short. Because these connections were not explicitly built into the simulation (see Materials and Methods), we turned to an analysis of the simulation to (1) determine whether this connectivity accounts for the ability of the simulation to learn and (2) test whether this connectivity can in principle account for our empirical observations. The simulation is useful in this regard because the activity of each mossy fiber and Golgi cell input to a given granule cell can be analyzed, and the activity of each simulated mossy fiber and granule cell input to a given Golgi cell can similarly be analyzed.

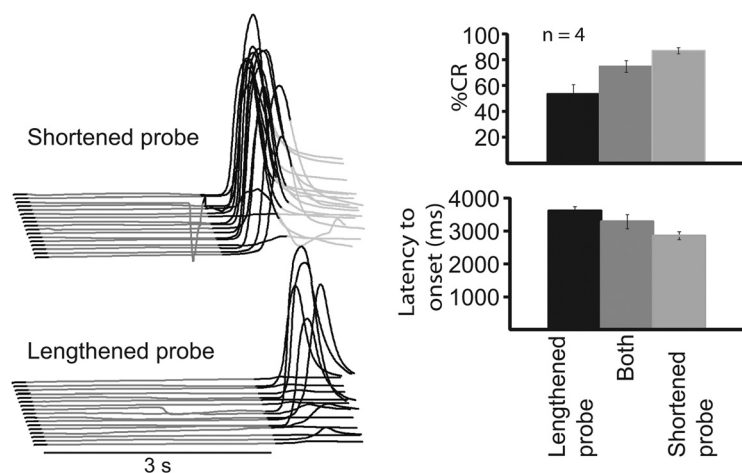
We focused analysis on the 120 granule cells (1% of population) whose synapse onto a Purkinje cell underwent the most LTD (Ito and Kano, 1982; Linden, 1994; Ito, 2001) during training with an MF-short of 3000 ms and an MF-long of 3500 ms. Because Purkinje cells inhibit the output of the cerebellum (Ito, 1984), these synapses would contribute most to the

cessation of ongoing Purkinje cell activity that is necessary for the expression of the learned responses (Perrett et al., 1993; Hesslow, 1994; Jirenhed et al., 2007; Medina and Lisberger, 2008; Rasmussen et al., 2008; Kalmbach et al., 2010a). Of these granule cells, 115 (96%) were excited by at least one mossy fiber activated by MF-long and 113 (94%) were inhibited by Golgi cells that were robustly activated by MF-short. Overall, 108 (90%) were excited by at least one MF-long mossy fiber and inhibited by at least one MF-short Golgi cell. In contrast, for a randomly selected granule cell, there is a 9.5% probability that it receives input from at least one MF-long, a 37.4% probability that it receives input from at least one Golgi cell that receives input from MF-short, and a 3.6% probability that it receives input from at least one of each. Figure 7*B* illustrates this by showing representative peristimulus histograms of granule cells from this sample of 120, and Figure 7*A* shows the four mossy fiber and Golgi cell inputs of a representative example. Figure 7*C* shows two examples of the activity produced in Purkinje cells of a well-trained simulation. Of the remaining 10% of this sample, five cells (4%) were inhibited by a Golgi cell driven by MF-short and were excited by at least one mossy fiber with high spontaneous activity (rather than by a mossy fiber activated by MF-long). The remaining seven cells had quite different peristimulus histograms. This analysis indicates that the ability of the simulation to learn responses for the subtraction pattern of inputs arises from the presence of a small subset of granule cells that are activated by MF-long and receive feedforward inhibition from Golgi cells activated by MF-short (as schematized in Fig. 1*B*).

Finally, we used analysis of the simulation to develop a working hypothesis to address two related questions: (1) why does learning decline with long mossy fiber inputs, and (2) why does the temporal subtraction pattern rescue learning despite the long duration of each MF input? For this analysis, the granule cells were sorted according to their activity in the 200 ms preceding the US, with those most active during this period being the cells whose synapses onto Purkinje cells are most eligible to undergo LTD during learning (Chen and Thompson, 1995; Raymond and Lisberger, 1998; Wang et al., 2000; Safa and Regehr, 2008). Figure 8*A* shows peristimulus histograms for 12 of these cells in a simulation trained with the subtraction protocol (MF-long, 3500; MF-short, 3000). Figure 8*B* shows 12 histograms from the identical simulation trained with the control condition in which both mossy fiber inputs lasted 3500 ms. In both cases, the US was presented 3500 ms after stimulus onset. The cells from the subtraction-trained simulation have robust activity that is specific to the offset interval, whereas the activity in the control simulation occurs over the entire stimulus duration. In addition, the activity in the control simulation showed greater variability



**Figure 5.** Responses of rabbits to probe trials in which either MF-short or MF-long was omitted. Left, Sample sweeps from a representative subject. In most animals, presenting MF-long elicited conditioned responses (CR) with much shorter latencies to onset. Presentation of MF-short did not elicit responses. Right, Comparisons of the likelihood of responding (top) and latency to onset (bottom) for the probe stimuli ( $n = 4$ ).

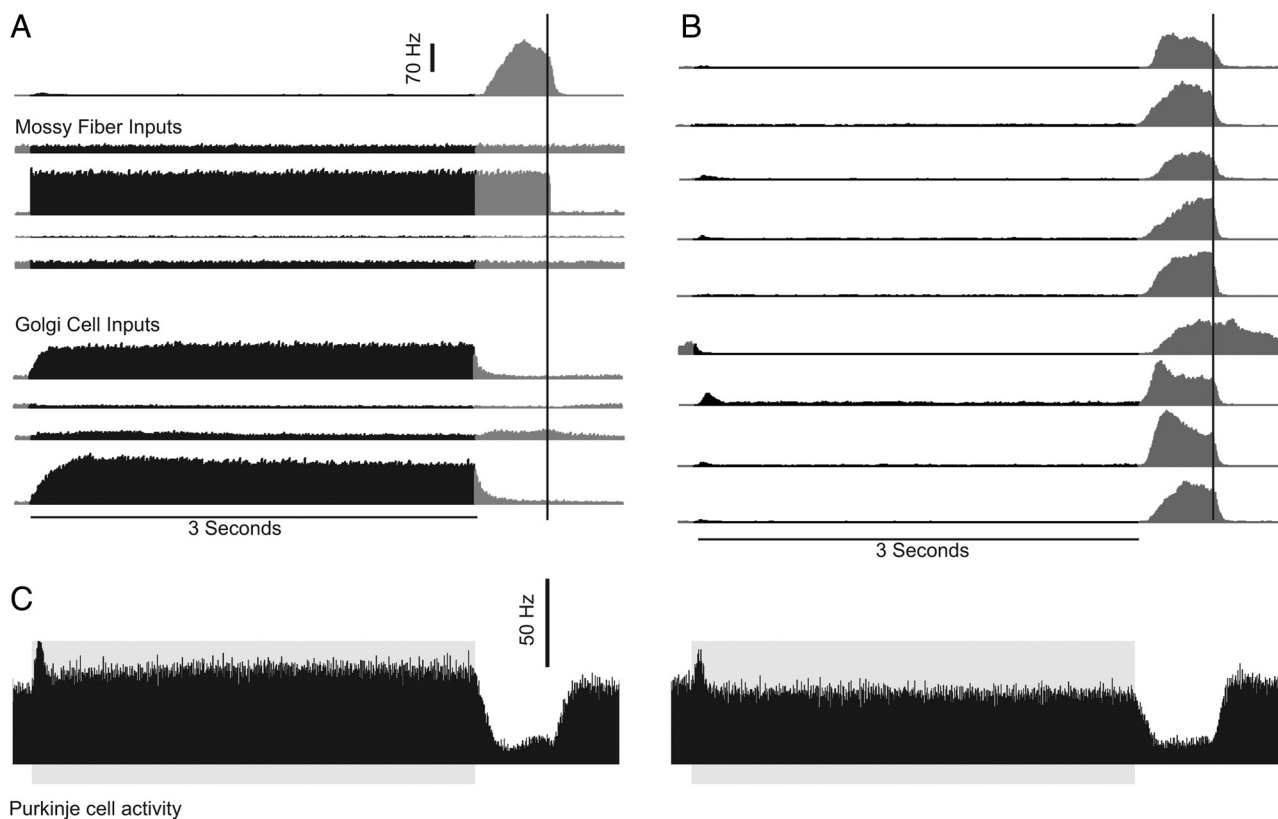


**Figure 6.** Responses of rabbits to probe trials in which MF-short was either lengthened or shortened. Left, Sample sweeps from a representative subject. Top, Trials in which MF-short was shortened from 3000 to 2700 ms. Bottom, Trials in which MF-short was lengthened from 3000 to 3300 ms. Right, Comparisons of the likelihood of responding (top) and latency to onset (bottom) for the probe stimuli ( $n = 4$ ). CR, Conditioned response.

across trials. Although both sets of cells have robust activity during the interval when the US presentation could induce LTD (indicated by gray rectangles), the synapses from the subtraction-trained simulations underwent robust net LTD, whereas synapses from the control simulation did not. Figure 8*C* shows the average synaptic weights for the 120 synapses in each case.

Although the average activity of the subtraction-trained cells is greater during the LTD period, this disparity cannot explain the large difference in the effect of training on synaptic weights. The key factor is the temporal specificity of activity in each group. Granule cell-to-Purkinje cell synapses that are active in the absence of climbing fiber input undergo LTP (Lev-Ram et al., 2003). The activation of these synapses would decrease cerebellar output by increasing Purkinje cell activity. For the control simulation, because the cells that are most active just before the US are also quite active earlier in the same stimulus, LTP induction can compete with induction of LTD. This is quantified in Figure 8*D*, in which the difference between possible LTD and LTP events are shown for the target 120 cells from the two simulations. There were more total LTD events for the subtraction group





**Figure 7.** Sample responses of simulated granule cells during subtraction training with a 3000 ms MF-short and a 3500 ms MF-long. **A**, A peristimulus histogram of a granule cell from a simulation trained as stated. The black portion of each histogram denotes the 3000 ms in which both MF-short and MF-long were present. The vertical line running through each histogram denotes the expected time of US presentation. The top trace is the response of the granule cell, which begins after offset of MF-short and peaks at ~400 ms later. The next four histograms represent the peristimulus histograms of the four mossy fiber inputs to this granule cell. One of the four belonged to the MF-long input. The bottom four peristimulus histograms show the responses of the four Golgi cells that inhibit this granule cell. Two of these four Golgi cells were strongly activated by MF-short. **B**, Nine granule cells from the same simulation that had responses after the offset of MF-short. Note the range of latencies to peak for these responses. **C**, Peristimulus histograms from two representative Purkinje cells in a well-trained simulation. All Purkinje cells showed a similar decrease in activity during the offset interval.

(3500\_3000) compared with the control (3500\_3500). In addition, the LTD for the control groups was counteracted by the larger number of LTP events at the same synapses, which yielded net LTD for the subtraction group and net LTP for the control group. Thus, in addition to trial-to-trial variability in granule cell activity during the control simulation, competition between LTP and LTD appears to contribute to the inability to learn during long-duration mossy fiber presentations (Mauk and Donegan, 1997).

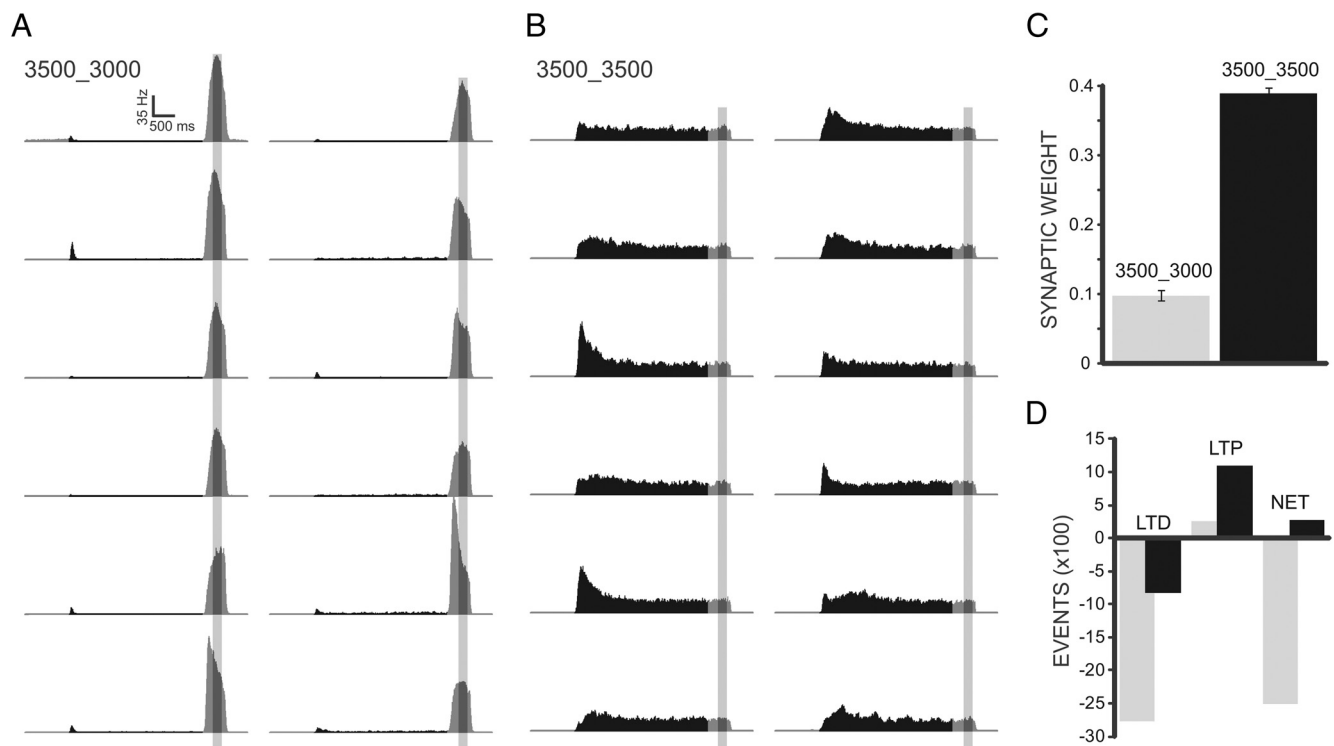
These simulation results suggest a three-component working hypothesis for the effectiveness of temporal subtraction patterns to promote learning. First, cerebellar learning declines as the duration of a mossy fiber input increases as a result of the increasing opportunities for LTP events to counteract the LTD induced by presentation of the US. Second, this competition favors LTP at long stimulus durations in part because the granule cells that are active just before US onset tend also to be active at earlier times during the same stimulus, in other words, from a lack of temporal specificity. Third, the subtraction pattern of mossy fiber inputs promotes robust activation of granule cells that is specific to the offset period, providing the opportunity for LTD events that are not counteracted by LTP induction and thus rescuing learning despite the long duration of the stimuli.

## Discussion

The ability to learn properly timed outputs is central to cerebellar function. We have used a large-scale computer simulation of the

cerebellum and eyelid conditioning experiments to demonstrate and characterize temporal subtraction, a network mechanism of temporal coding in cerebellar cortex. We showed that the cessation of one of two ongoing mossy fiber inputs to a cerebellar simulation produces a robust and specific temporal code in the population of granule cells. Because this granule cell code is unique to the interval after the cessation of the shorter mossy fiber input, i.e., the offset interval, it predicts precisely timed learning for mossy fiber inputs that are otherwise too long to normally support learning. We showed that cerebellar-dependent learning in rabbits supports these unusual predictions of the simulation. We then characterized how each input contributes to learning under these atypical conditions. Finally, we showed how feedforward inhibition present in the synaptic organization of the cerebellum (Eccles et al., 1967; Palkovits et al., 1971, 1972; Ito, 1984) may contribute to this enhanced temporal code and learning of precisely timed responses.

Temporal subtraction via feedforward inhibition adds to the list of connectivity-based mechanisms that can contribute to the generation of temporal codes, which includes recurrent inhibition, feedback inhibition, and feedforward excitation (Durstewitz et al., 2000; Ikegaya et al., 2004; Mauk and Buonomano, 2004; Buhusi and Meck, 2005; D'Angelo et al., 2009). Recurrent inhibition has been invoked as a means to generate oscillations in the activity of individual cells that could be used in timing (Bu-



**Figure 8.** An analysis of why the cerebellar simulation does not learn with long mossy fiber inputs (two inputs, both 3500 ms) and why the subtraction pattern of training (MF-short, 3000; MF-long, 3500) rescues learning. **A**, Peristimulus histograms of the 12 simulated granule cells that were most active during the time when the US-activated climbing fiber input could induce LTD (gray shaded area). These granule cells were not only robustly active, but their activity was relatively specific to the time period in which LTD could be induced. **B**, Peristimulus histograms from the 12 granule cells that were most active during the time period in which the US could induce LTD during training with both mossy fiber inputs of 3500 ms. The activity of these cells is not specific to the time of LTD induction. This creates opportunities for the induction of LTP (and thus the suppression of conditioned responses) during the activity earlier in the trial. **C**, The average synaptic weights from the same two simulations as in **A** and **B** for the top 120 granule cells in each case in terms of level of activity when the US was presented. The synapses in the simulation trained with 3000\_3500 had undergone net LTD, whereas the synapses in the 3500\_3500 simulation were at approximately their beginning strength. **D**, A breakdown of LTD and LTP events for the two groups of synapses shown in **C**. Both sets of synapses underwent a large number of LTD events, but the synapses in the 3500\_500-trained simulation also underwent a large number of LTP events. This led to net LTD for the 3000\_3500 simulation and to net LTP for the 3500\_3500 simulation.

husi and Meck, 2005; D'Angelo et al., 2009). Similarly, feedforward excitation has been hypothesized as a means to generate delay lines or synfire chains (Durstewitz et al., 2000; Ikegaya et al., 2004). In this case, a series of connected neurons generates temporal codes by activating their downstream targets in a highly regulated chain. Previous simulations of the cerebellum have also highlighted the ability of sparse (non-recurrent) feedback inhibition to generate time-varying activity (Medina et al., 2000). Although feedback inhibition and feedforward excitation certainly operate in this simulation of the cerebellum, the present results also illustrate how the feedforward inhibition that is present in many brain regions (Lawrence and McBain, 2003; Swadlow, 2003; Mauk and Buonomano, 2004; Tepper et al., 2004) can contribute to and enhance temporal coding for patterns that involve one input ending before the other.

Temporal subtraction requires only that one of two ongoing mossy fiber inputs terminate, and thus it may represent an often engaged computational principle in the cerebellum. Subtraction may occur during the presentation of peripheral stimuli, such as tones, in which certain mossy fibers respond phasically to the onset of the tone and others respond for the duration of the tone (Boyd and Aitkin, 1976; Aitkin and Boyd, 1978). The stimuli used in eye movement studies produce variety in the temporal pattern of mossy fiber input (Chubb et al., 1984; Lisberger and Pavelko, 1986; Ramachandran and Lisberger, 2006), which may also give opportunity for subtraction. Similarly, the subtraction pattern of

inputs is reminiscent of the pattern of inputs that the cerebellum appears to be presented with during trace eyelid conditioning, in which the offset of the CS and onset of the US is separated by a stimulus-free trace interval. Specifically, our results are consistent with evidence that the cerebellum uses mossy fiber input driven by the forebrain that persists to the US and tone-driven input that is active only for the duration of the tone (e.g., a subtraction pattern of input) to generate granule cell activity during the trace interval (Kalmbach et al., 2009, 2010b). Furthermore, they suggest that this coding during trace eyelid conditioning depends on feedforward inhibition that is driven by the tone-driven mossy fiber input.

These data underscore the importance of a precise temporal code for cerebellar learning. LTD of the granule cell-to-Purkinje cell synapse is a form of synaptic plasticity implicated in cerebellar learning (Ito and Kano, 1982; Linden, 1994; Ito, 2001). The induction of LTD requires a specific temporal relationship between the activation of granule cell and climbing fiber synapses onto Purkinje cells; granule cell synapses must be active within ~300 ms of a climbing fiber input to undergo LTD (Chen and Thompson, 1995; Wang et al., 2000; Safo and Regehr, 2008). Conversely, granule cell synapses active in the absence of climbing fiber input undergo LTP (Lev-Ram et al., 2003). Because the activation of weak synapses would tend to increase cerebellar output, synapses that have undergone LTD would be important for the initiation of responses. Although we observed granule cell activity within the window for LTD induction during training



with both inputs presented for 3.5 s, this activity was not unique to the LTD induction window. These granule cells were active in the time window for the induction of both LTD (because they were active within  $\sim 300$  ms of a climbing fiber input) and LTP (because they were active for seconds in the absence of climbing fiber input). As such, the strength of these synapses either did not change or even increased. This interpretation may explain why learning declines as the length of the CS used in conditioning increases (Schneiderman and Gormezano, 1964; Ohyama et al., 2003): longer CSs are incapable of generating granule cell activity that is specific to the time interval for LTD induction. Conversely, during presentation of mossy fiber input in the subtraction pattern to our cerebellar simulation, there were synapses active precisely in the window for the induction of LTD. This observation suggests that the cessation of one of two ongoing mossy fiber inputs enables learning by creating activity in a subset of granule cells that is unique to the period for LTD induction.

Our results also suggest how the unipolar brush cells (UBCs) found within cerebellar cortex may contribute to temporal coding. UBCs receive a single mossy fiber input and contact granule cells, Golgi cells, and other UBCs (Nunzi et al., 2001). *In vitro* evidence suggests that UBCs are capable of converting a phasic mossy fiber input into a tonic train of action potentials lasting hundreds of milliseconds (Diño et al., 2000; Nunzi et al., 2001). Thus, although these cells were not included in the cerebellar simulation, they may be well suited to transform phasic mossy fiber input into prolonged “mossy fiber inputs” with variable firing durations, which could in turn engage the subtraction mechanism and enhance temporal coding and learning. In this way, UBCs may act as a source of time-varying “intrinsic” mossy fiber input.

In summary, these data suggest how feedforward inhibition present in the synaptic organization of the cerebellar cortex contributes to the generation of precisely timed motor movements by contributing to the generation of a precise temporal code in a population of granule cells. Feedforward inhibition present in other brain regions may similarly function to enhance temporal coding.

## References

- Aitkin LM, Boyd J (1978) Acoustic input to the lateral pontine nuclei. *Hear Res* 1:67–77.
- Boyd J, Aitkin L (1976) Responses of single units in the pontine nuclei of the cat to acoustic stimulation. *Neurosci Lett* 3:259–263.
- Buhusi CV, Meck WH (2005) What makes us tick? Functional and neural mechanisms of interval timing. *Nat Rev Neurosci* 6:755–765.
- Chen C, Thompson RF (1995) Temporal specificity of long-term depression in parallel fiber–Purkinje synapses in rat cerebellar slice. *Learn Mem* 2:185–198.
- Chubb MC, Fuchs AF, Scudder CA (1984) Neuron activity in monkey vestibular nuclei during vertical vestibular stimulation and eye movements. *J Neurophysiol* 52:724–742.
- D’Angelo E, Koekkoek SK, Lombardo P, Solinas S, Ros E, Garrido J, Schoneville M, De Zeeuw CI (2009) Timing in the cerebellum: oscillations and resonance in the granular layer. *Neuroscience* 162:805–815.
- Diño MR, Schuerger RJ, Liu Y, Slater NT, Mugnaini E (2000) Unipolar brush cell: a potential feedforward excitatory interneuron of the cerebellum. *Neuroscience* 98:625–636.
- Durstewitz D, Seamans JK, Sejnowski TJ (2000) Neurocomputational models of working memory. *Nat Neurosci [Suppl]* 3:1184–1191.
- Eccles JC, Ito M, Szentagóthai J (1967) The cerebellum as a neuronal machine. New York: Springer.
- Hesslow G (1994) Inhibition of classically conditioned eyeblink responses by stimulation of the cerebellar cortex in the decerebrate cat. *J Physiol* 476:245–256.
- Ikegaya Y, Aaron G, Cossart R, Aronov D, Lampl I, Ferster D, Yuste R (2004) Synfire chains and cortical songs: temporal modules of cortical activity. *Science* 304:559–564.
- Ito M (1984) The cerebellum and neural control. New York: Raven.
- Ito M (2001) Cerebellar long-term depression: characterization, signal transduction, and functional roles. *Physiol Rev* 81:1143–1195.
- Ito M, Kano M (1982) Long-lasting depression of parallel fiber–Purkinje cell transmission induced by conjunctive stimulation of parallel fibers and climbing fibers in the cerebellar cortex. *Neurosci Lett* 33:253–258.
- Jirenhed DA, Bengtsson F, Hesslow G (2007) Acquisition, extinction, and reacquisition of a cerebellar cortical memory trace. *J Neurosci* 27:2493–2502.
- Kalmbach BE, Ohyama T, Kreider JC, Riusech F, Mauk MD (2009) Interactions between prefrontal cortex and cerebellum revealed by trace eyelid conditioning. *Learn Mem* 16:86–95.
- Kalmbach BE, Davis T, Ohyama T, Riusech F, Noreis WL, Mauk MD (2010a) Cerebellar cortex contributions to the expression and timing of conditioned eyelid responses. *J Neurophysiol* 103:2039–2049.
- Kalmbach BE, Ohyama T, Mauk MD (2010b) Temporal patterns of inputs to cerebellum necessary and sufficient for trace eyelid conditioning. *J Neurophysiol* 104:627–640.
- Klein JA, Freeman JH Jr, Bruneau R, Nolan BC, Cooper NR, Zook A, Walters D (2002) Synapse formation is associated with memory storage in the cerebellum. *Proc Natl Acad Sci U S A* 99:13228–13231.
- Lawrence JJ, McBain CJ (2003) Interneuron diversity series: containing the detonation—feedforward inhibition in the CA3 hippocampus. *Trends Neurosci* 26:631–640.
- Lev-Ram V, Mehta SB, Kleinfeld D, Tsien RY (2003) Reversing cerebellar long-term depression. *Proc Natl Acad Sci U S A* 100:15989–15993.
- Linden DJ (1994) Long-term synaptic depression in the mammalian brain. *Neuron* 12:457–472.
- Lisberger SG, Pavelko TA (1986) Vestibular signals carried by pathways subserving plasticity of the vestibulo-ocular reflex in monkeys. *J Neurosci* 6:346–354.
- Mauk MD, Buonomano DV (2004) The neural basis of temporal processing. *Annu Rev Neurosci* 27:307–340.
- Mauk MD, Donegan NH (1997) A model of Pavlovian eyelid conditioning based on the synaptic organization of the cerebellum. *Learn Mem* 4:130–158.
- Mauk MD, Ohyama T (2004) Extinction as new learning versus unlearning: considerations from a computer simulation of the cerebellum. *Learn Mem* 11:566–571.
- Mauk MD, Ruiz BP (1992) Learning-dependent timing of Pavlovian eyelid responses: differential conditioning using multiple interstimulus intervals. *Behav Neurosci* 106:666–681.
- Medina JF, Lisberger SG (2008) Links from complex spikes to local plasticity and motor learning in the cerebellum of awake-behaving monkeys. *Nat Neurosci* 11:1185–1192.
- Medina JF, Mauk MD (2000) Computer simulation of cerebellar information processing. *Nat Neurosci [Suppl]* 3:1205–1211.
- Medina JF, Garcia KS, Noreis WL, Taylor NM, Mauk MD (2000) Timing mechanisms in the cerebellum: testing predictions of a large-scale computer simulation. *J Neurosci* 20:5516–5525.
- Medina JF, Garcia KS, Mauk MD (2001) A mechanism for savings in the cerebellum. *J Neurosci* 21:4081–4089.
- Medina JF, Noreis WL, Mauk MD (2002) Inhibition of climbing fibres is a signal for the extinction of conditioned eyelid responses. *Nature* 416:330–333.
- Medina JF, Carey MR, Lisberger SG (2005) The representation of time for motor learning. *Neuron* 45:157–167.
- Nunzi MG, Birnstiel S, Bhattacharyya BJ, Slater NT, Mugnaini E (2001) Unipolar brush cells form a glutamatergic projection system within the mouse cerebellar cortex. *J Comp Neurol* 434:329–341.
- Ohyama T, Medina JF, Noreis WL, Mauk MD (2002) Trying to understand the cerebellum well enough to build one. *Ann N Y Acad Sci* 978:425–438.
- Ohyama T, Noreis WL, Murphy M, Mauk MD (2003) What the cerebellum computes. *Trends Neurosci* 26:222–227.
- Palkovits M, Magyar P, Szentagóthai J (1971) Quantitative histological analysis of the cerebellar cortex in the cat. 3. Structural organization of the molecular layer. *Brain Res* 34:1–18.

- Palkovits M, Magyar P, Szentágothai J (1972) Quantitative histological analysis of the cerebellar cortex in the cat. IV. Mossy fiber-Purkinje cell numerical transfer. *Brain Res* 45:15–29.
- Perrett SP, Ruiz BP, Mauk MD (1993) Cerebellar cortex lesions disrupt learning-dependent timing of conditioned eyelid responses. *J Neurosci* 13:1708–1718.
- Ramachandran R, Lisberger SG (2006) Transformation of vestibular signals into motor commands in the vestibuloocular reflex pathways of monkeys. *J Neurophysiol* 96:1061–1074.
- Rasmussen A, Jirenhed DA, Hesslow G (2008) Simple and complex spike firing patterns in Purkinje cells during classical conditioning. *Cerebellum* 7:563–566.
- Raymond JL, Lisberger SG (1998) Neural learning rules for the vestibulo-ocular reflex. *J Neurosci* 18:9112–9129.
- Raymond JL, Lisberger SG, Mauk MD (1996) The cerebellum: a neuronal learning machine? *Science* 272:1126–1131.
- Safo P, Regehr WG (2008) Timing dependence of the induction of cerebellar LTD. *Neuropharmacology* 54:213–218.
- Schneiderman N, Gormezano I (1964) Conditioning of the nictitating membrane of the rabbit as a function of CS-US interval. *J Comp Physiol Psychol* 57:188–195.
- Swadlow HA (2003) Fast-spike interneurons and feedforward inhibition in awake sensory neocortex. *Cereb Cortex* 13:25–32.
- Tepper JM, Koós T, Wilson CJ (2004) GABAergic microcircuits in the neostriatum. *Trends Neurosci* 27:662–669.
- Wang SS, Denk W, Häusser M (2000) Coincidence detection in single dendritic spines mediated by calcium release. *Nat Neurosci* 3:1266–1273.