

**A proposed function for hippocampal theta rhythm:
Separate phases of encoding and retrieval enhance
reversal of prior learning.**

Michael E. Hasselmo†, Clara Bodelón and Bradley P. Wyble*

*†Department of Psychology, Program in Neuroscience and Center for BioDynamics,
Boston University, 64 Cummington St., Boston, MA 02215, (617) 353-1397, FAX:
(617) 353-1424, hasselmo@berg.bu.edu.*

**Department of Psychology, Harvard University, 33 Kirkland St., Cambridge, MA
02138, wyble@berg.bu.edu*

†Corresponding author.

Abstract.

The theta rhythm appears in the rat hippocampal EEG during exploration and shows phase locking to stimulus acquisition. Lesions which block theta rhythm impair performance in tasks requiring reversal of prior learning, including reversal in a T-maze where associations between one arm location and food reward need to be extinguished in favor of associations between the opposite arm location and food reward. Here, a hippocampal model shows how theta rhythm could be important for reversal in this task by providing separate functional phases during each 100-300 msec cycle, consistent with physiological data. In the model, effective encoding of new associations occurs in the phase when synaptic input from entorhinal cortex is strong, and when long-term potentiation (LTP) of excitatory connections arising from hippocampal region CA3 is strong, but when synaptic currents arising from region CA3 input are weak (to prevent interference from prior learned associations). Retrieval of old associations occurs in the phase when entorhinal input is weak and synaptic input from region CA3 is strong, but when depotentiation occurs at synapses from CA3 (to allow extinction of prior learned associations that do not match current input). These phasic changes require that long-term potentiation (LTP) at synapses arising from region CA3 should be strongest at the phase when synaptic transmission at these synapses is weakest. Consistent with these requirements, our recent data shows that synaptic transmission in stratum radiatum is weakest at the positive peak of local theta, which is when previous data shows induction of LTP is strongest in this layer.

1. Introduction.

The hippocampal theta rhythm is a large amplitude, 3-10 Hz oscillation that appears prominently in the rat hippocampal EEG during locomotion or attention to environmental stimuli, and decreases during immobility or consummatory behaviors such as eating or grooming (Green and Arduini, 1957; Buzsaki et al., 1983). In this paper, we link the extensive data on physiological changes during theta to the specific requirements of behavioral reversal tasks. Theta rhythm is associated with phasic changes in the magnitude of synaptic currents in different layers of the hippocampus, as shown by current source density analysis of hippocampal region CA1 (Buzsaki et al., 1986; Brankack et al., 1993; Bragin et al., 1995). As summarized in Figure 1, at the trough of the theta rhythm recorded at the hippocampal fissure, afferent synaptic input from the entorhinal cortex is strong, synaptic input from region CA3 is weak (Brankack et al., 1993), but long-term potentiation at synapses from region CA3 is strong (Holscher et al., 1997; Wyble et al., 2001). In contrast, at the peak of the fissure theta rhythm, afferent input from entorhinal cortex is weak, the amount of synaptic input from region CA3 is strong (Brankack et al., 1993), but stimulation of synaptic input from region CA3 induces depotentiation (Holscher et al., 1997).

FIGURE 1 ABOUT HERE

1.1. Role of theta rhythm in behavior.

This paper relates these phasic changes in synaptic properties to data showing behavioral effects associated with blockade of theta rhythm. Physiological data suggest that rhythmic input from the medial septum paces theta-frequency oscillations recorded from the hippocampus and entorhinal cortex (Stewart and Fox, 1990; Toth et al., 1997). This regulatory input from the septum enters the hippocampus via the fornix, and

destruction of this fiber pathway attenuates the theta rhythm (Buzsaki et al., 1983). Numerous studies show that fornix lesions cause strong impairments in reversal of previously learned behavior (Numan, 1978; M'Harzi et al., 1987; Whishaw and Tomie, 1995), including reversal of spatial response in a T-maze. The T-maze task is shown schematically in Figure 2. During each trial in a T-maze reversal, a rat starts in the stem of the maze and finds food reward when it runs down the stem and into one arm of the maze (e.g. the left arm). After extensive training of this initial association, the food reward is moved to the opposite arm of the maze (e.g. the right arm), and the rat must extinguish the association between left arm and food, and learn the new association between right arm and food. Rats with fornix lesions make more errors after the reversal -- continuing to visit the old location (left arm) which is no longer rewarded.

FIGURE 2 ABOUT HERE.

In the analysis and simulations presented here, we test the hypothesis that phasic changes in the amount of synaptic input and synaptic modification during cycles of the hippocampal theta rhythm could enhance reversal of prior learning. Specifically, these phasic changes allow new afferent input (from entorhinal cortex) to be strong when synaptic modification is strong, to encode new associations between place and food reward, without interference caused by synaptic input (from region CA3) representing retrieval of old associations. Synaptic input from region CA3 is strong on a separate phase when depotentiation at these synapses can allow extinction of old associations.

1.2. Overview of the model.

In the model, we focus on the strengthening and weakening of associations between location (in left or right arm) and food reward (in left or right arm). These

associations are encoded by modifying synaptic connections between neurons in region CA3 and region CA1 representing location (place cells) and food reward. The place representations are consistent with evidence of hippocampal neurons showing place selective responses in a variety of spatial environments (McNaughton et al., 1981; Skaggs et al., 1996). The food reward representations are consistent with studies showing unit activity selective for the receipt of reward during behavioral tasks (Wiener et al., 1989; Otto and Eichenbaum, 1992; Hampson and Deadwyler, 1999; Young et al., 1997). Units have been shown to be selective for the receipt of reward at a particular location (Wiener et al., 1989), and reward-dependent responses have been shown in both region CA1 (Otto and Eichenbaum, 1992) and entorhinal cortex (Young et al., 1997). This model assumes place cell representations already exist, and does not focus on their formation or properties, which have been modeled previously (e.g. Kali and Dayan, 2000). Instead, the mathematical analysis focuses on encoding and retrieval of associations between place cell activity in each maze arm and the representation of food reward.

The structure and phases in the model are summarized in Figure 1. The analysis presented here suggests that oscillatory changes in magnitude of synaptic transmission, long-term potentiation and post-synaptic depolarization during theta rhythm cause transitions between two functional phases within each theta cycle: 1.) In the encoding phase, entorhinal input is dominant, activating cells in region CA3 and CA1. At this time, excitatory connections arising from region CA3 have decreased transmission, but these same synapses show enhanced long-term potentiation to form associations between sensory events. 2.) In the retrieval phase, entorhinal input is relatively weak, but still brings retrieval cues into the network. At this time, excitatory synapses arising

from region CA3 show strong transmission, allowing retrieval of the previously learned associations. During this retrieval phase, long-term potentiation must be absent in order to prevent encoding of the retrieval activity. Depotentiation or long-term depression (LTD) during this phase allows extinction of prior learned associations. We propose that these two phases appear in a continuous interleaved manner within each 100-300 msec cycle of the theta rhythm.

Section 2 presents mathematical analysis showing how theta oscillations could play an important role in reversal of learned associations. First, we will describe the assumptions for modeling synaptic transmission, synaptic modification and phasic changes in these variables during theta rhythm. Then we describe the assumptions for modeling patterns of activity representing sensory features of the behavioral task in the T-maze. Subsequently, we present a criterion for evaluating the function of the model, in the form of a performance measure based on reversal of learned associations in the T-maze. Finally, we show how specific phase relationships between network variables provide the best performance of the model. These phase relationships correspond to those observed with physiological recording during hippocampal theta rhythm.

2. Mathematical analysis of theta rhythm.

We modeled hippocampal activity when a rat is in one of the two arms of a T-maze during individual trials performed during different stages of a reversal task. Note that a trial refers to the period from the time the rat reaches the end of an arm to its removal from the arm. This period can cover several cycles of theta. A stage of the task refers to multiple trials in which the spatial response of the rat and the food reward location are the same.

We represent the pattern of activation of pyramidal cells in region CA1 with the vector $a_{CA1}(t)$. As shown in Figures 1 and 2, the pattern of activity in CA1 is influenced by two main inputs: 1.) The synaptic input from region CA3, which is a product of the strength of modifiable Schaffer collateral synapses $W_{CA3}(t)$ and the activity of CA3 pyramidal cells $a_{CA3}(t)$. The activity pattern in CA3 remains constant across trials within a single behavioral stage of the task, but changes in different stages depending on the spatial response of the rat. The weight matrix remains constant within each trial, but can change at the end of the trial. 2.) The synaptic input from entorhinal cortex, which is a product of the strength of perforant path synapses W_{EC} (which are not potentiated here) and the activity of projection neurons in layer III of entorhinal cortex $a_{EC}(t)$. The activity pattern in entorhinal cortex also remains constant within each stage of the task, but changes in different stages depending on the location of food reward. The weight matrix W_{EC} does not change in this model.

Note that individual elements of the vectors $a_{CA3}(t)$ and $a_{EC}(t)$ assume only two values, zero or q , to represent presence or absence of spiking activity in each neuron. However, these influences make the activity of neurons in region CA1 $a_{CA1}(t)$ take a range of values. The following equation represents the effects of synaptic transmission on region CA1 activity:

$$a_{CA1}(t) = W_{EC}a_{EC}(t) + W_{CA3}(t)a_{CA3}(t) \quad (1)$$

The structure of the model is summarized in Figures 1 and 3.

2.1. Phasic changes in synaptic input during theta.

This paper focuses on oscillatory modulation of equation 1 to represent phasic changes during theta rhythm. Experiments have shown phasic changes in the amount of synaptic input in stratum radiatum of region CA1 (Ruddell and Fox, 1984; Buzsaki et al., 1986; Brankack et al., 1993; Bragin et al., 1995; Wyble et al., 2000). These changes in amount of synaptic input could depend upon the magnitude of transmitter release, or on other factors such as the firing rate of afferent neurons. Both factors have the same functional influence in the current model, and are represented as follows. The synaptic weights, $W_{CA3}(t)$, are multiplied by a sine wave function, $\theta_{CA3}(t)$, which has phase ϕ_{CA3} and is scaled to vary between 1 and $(1-X)$, as shown in equation 2. The parameter X represents the magnitude of change in synaptic currents in different layers, and was constrained to values between 0 and 1 ($0 < X < 1$).

$$\theta_{CA3}(t) = X / 2 * \sin(t + \phi_{CA3}) + (1 - X / 2) \quad (2)$$

The simulation also includes phasic changes in magnitude of entorhinal input $\theta_{EC}(t)$, as shown in equation 3.

$$\theta_{EC}(t) = X / 2 * \sin(t + \phi_{EC}) + (1 - X / 2) \quad (3)$$

These changes in entorhinal input are consistent with current source density (CSD) data on stratum lacunosum-moleculare (s. l-m) in region CA1 (Brankack et al., 1993; Bragin et al., 1995) and phasic activity of entorhinal neurons (Stewart et al., 1992). The analysis presented here shows how the best performance of the model is obtained with relative phases of these variables that are shown in Figures 1 and 3.

FIGURE 3 ABOUT HERE.

With inclusions of these phasic changes, equation 1 becomes:

$$a_{CA1}(t) = \theta_{EC}(t)W_{EC}a_{EC}(t) + \theta_{CA3}(t)W_{CA3}(t)a_{CA3}(t) \quad (4)$$

2.2. Phasic changes in LTP during theta.

Long-term potentiation in the hippocampus has repeatedly been shown to depend upon the phase of stimulation relative to theta rhythm in the dentate gyrus (Pavrides et al., 1982; Orr et al., 1999) and region CA1 (Huerta and Lisman, 1993; Holscher et al., 1997; Wyble et al., 2001). In region CA1, experiments have shown that LTP can be induced by stimulation on the peak -- but not the trough -- of the theta rhythm recorded in stratum radiatum in slice preparations (Huerta and Lisman, 1993) urethane- anesthetized rats (Holscher et al., 1997) and awake rats (Wyble et al., 2001). Note that theta recorded locally in radiatum will be about 180 degrees out of phase with theta recorded at the fissure. Based on these data, the rate of synaptic modification of the Schaffer collateral synapses ($W_{CA3}(t)$) in the model was modulated in an oscillatory manner given by:

$$\theta_{LTP}(t) = \sin(t + \phi_{LTP}) \quad (5)$$

This long-term modification requires post-synaptic depolarization in region CA1 $a_{CA1}(t)$, and presynaptic spiking activity from region CA3 $a_{CA3}(t)$. For simplicity, we assume that the synaptic weights during each new trial remain constant at their initial values, and only change at the end of the trial. This is consistent with delayed expression of LTP after its initial induction. We also assume that synapses do not grow beyond a maximum strength. We can compute the change in synaptic weight at the end of a single behavioral trial $m+1$, which depends upon the integral of activity in the

network during that trial (from the time at the start of that trial, t_m , to the time at the end of that trial, t_{m+1}). For mathematical simplicity, we integrate across full cycles of theta.

$$\Delta W_{CA3}(t_{m+1}) = \int_{t_m}^{t_{m+1}} \theta_{LTP}(t) a_{CA1}(t) a_{CA3}(t)^T dt \quad (6)$$

The dependence on pre and post-synaptic activity could be described as Hebbian (Gustafsson et al., 1987). However, the additional oscillatory components of the equation go beyond strict Hebbian properties, and can cause the synaptic weight to weaken if pre and post-synaptic activity occurs during a negative phase of the oscillation.

This paper focuses on how the quality of retrieval performance in region CA1 depends upon the phase relationship of the oscillations in synaptic transmission (in equation 4) with the oscillations in synaptic modification (in equation 6). Combining these equations gives us:

$$\Delta W_{CA3}(t_{m+1}) = \int_{t_m}^{t_{m+1}} \theta_{LTP}(t) [\theta_{EC}(t) W_{EC} a_{EC}(t) + \theta_{CA3}(t) W_{CA3}(t_m) a_{CA3}(t)] a_{CA3}(t)^T dt \quad (7)$$

Note that we allow theta phasic modulation of W_{EC} , but we are not modeling Hebbian synaptic modification of W_{EC} . Therefore, after this point we will not include that matrix of connectivity, but will represent patterns of input from entorhinal cortex with single static vectors, as if the synapses were an identity matrix.

2.3. Modeling learning of associations in a reversal task.

The mathematical analysis focuses on the performance of the network in the T-maze reversal task, in which learning of an association between one location and food must be extinguished and replaced by a new association. Our analysis focuses on synaptic changes occurring during initial learning of an association (e.g. between left arm and food), subsequent extinction of this association when it is invalid, and encoding of a new association (e.g. between right arm and food). Thus, we focus on the mnemonic aspects of this task. Changes in the synaptic matrix $W_{CA3}(t)$ include the learning during three sequential stages of the task, whereas for the analysis of performance a fourth stage is also needed. These stages are described as follows:

- 1.) Initial learning – During this period, the rat runs to the left arm of the T-maze and finds food reward in that arm (Figure 2, A1 and B1). Synaptic modification encodes associations between left-arm place (represented by CA3 activity) and left-arm food reward (represented by entorhinal activity) on the n trials with time ranging from $t=0$ to $t=t_n$. Note that during this behavioral stage, activity in CA3 and EC is constant during all n trials. We use the superindex (n) to indicate the activity during this stage in region CA3 ($a_{CA3}^{(n)}$) and in entorhinal cortex ($a_{EC}^{(n)}$).
- 2.) Reversal learning with erroneous responses – After stage 1, the food reward is changed to the right arm. In the initial stage after this reversal, the rat runs to the left arm place but does not find food reward (Figure 2, A2 and B2). This causes extinction or weakening of the associations between left-arm place and left-arm food reward. We denote with the superindex (e) the activity during this stage. The left arm place representation in CA3 is the

same as during initial learning, and for simplicity we assume the CA3 activity vector for initial trials has unity dot product with the activity vector for erroneous reversal trials $a_{CA3}^{(n)T} a_{CA3}^{(e)} = 1$. Because the rat does not encounter food reward on erroneous trials, the EC activity vector representing food reward is zero $a_{EC}^{(e)} = 0$. This period contains e trials and ends at t_{n+e} . This stage could vary in number of trials in individual rats.

- 3.) Reversal learning with correct responses – During this period, the rat runs to the right arm and finds food (Figure 2, A3 and B3). In this stage, the superindex will be (c) . Synaptic modification encodes associations between right arm place $a_{CA3}^{(c)}$ and right arm food reward $a_{EC}^{(c)}$. This period contains c trials and ends at t_{n+e+c} . Note that we make the natural assumption that the CA3 activity representing left arm place cell responses during initial and erroneous trials has a zero dot product with the right arm place cell responses during correct reversal trials $a_{CA3}^{(n)T} a_{CA3}^{(c)} = a_{CA3}^{(e)T} a_{CA3}^{(c)} = 0$. We also assume that the entorhinal input for food in the right arm has zero dot product with the input for food in the left arm $(a_{EC}^{(c)})^T a_{EC}^{(n)} = 0$ and that the dot product of food reward input from entorhinal cortex for the same food location is unity $((a_{EC}^{(c)})^T a_{EC}^{(c)} = (a_{EC}^{(n)})^T a_{EC}^{(n)} = 1)$.

- 4.) Retrieval at choice point on subsequent trials. The analysis of performance in the network focuses on retrieval activity at the choice point of the T-maze (Figure 2, A4 and B4). This stage is described further in section 2.4 below.

The association between left arm place and food reward, learned on the initial n trials (Fig. 2, A1 and B1), is represented by the matrix $W_{CA3}(t_n)$. As noted above, each entry of this matrix does not grow beyond a maximum, which is a factor of K times the pre and post-synaptic activity. This matrix represents the association between left arm place cells activated during these trials in region CA3 ($a_{CA3}^{(n)}$) and food reward representation cells in region CA1 activated by entorhinal input ($a_{CA1} = a_{EC}^{(n)}$). At the end of the n trials, this matrix becomes:

$$W_{CA3}(t_n) = K a_{EC}^{(n)} a_{CA3}^{(n)T} \quad (8)$$

After reversal, the extinction of the association between left arm and reward occurs on erroneous trials (Figure 2, A2 and B2). The learning during each trial is represented by summing up integrals of equation 7 over each trial from $n+1$ to $n+e$. As noted above, the lack of reward is represented by an absence of entorhinal input $a_{EC}^{(e)} = 0$. However, retrieval of prior learning can occur during these erroneous trials due to left arm food representations activated by the spread of activity from left arm place cells in region CA3 over the matrix (equation 8) from the first n trials (recall that $a_{CA3}^{(n)T} a_{CA3}^{(e)} = 1$). This appears in the equation below.

$$W_{CA3}(t_{n+e}) = W_{CA3}(t_n) + \sum_{k=n+1}^{n+e} \int_{t_k}^{t_{k+1}} \theta_{LTP}(t) [\theta_{CA3}(t) W_{CA3}(t_k) a_{CA3}^{(e)}] a_{CA3}^{(e)T} dt \quad (9)$$

In contrast with incorrect trials, correct trials after reversal are indexed with the numbers between $n+e+1$ and $n+e+c$ (Fig 2, A3 and B3), where c represents correct post-reversal trials. The desired learning during each trial c consists of the association

between place cells active in region CA3 when the rat is at the end of the right arm

$a_{CA3}^{(c)T}$, and the food representation activated by entorhinal input $a_{EC}^{(c)}$. As noted above,

$a_{CA3}^{(n)T} a_{CA3}^{(c)} = 0$, so the spread across connections modified in initial learning is

$W_{CA3}(t_n) a_{CA3}^{(c)} = k a_{EC}^{(n)} a_{CA3}^{(n)T} a_{CA3}^{(c)} = 0$ We assume these trials occur sequentially after the

error trials described above. However, with the assumption of non-overlapping place

representations in CA3, even if the erroneous and correct trials were intermixed the

same results would apply. The learning on these correct trials is:

$$W_{CA3}(t_{n+e+c}) = W_{CA3}(t_{n+e}) + \sum_{k=n+e+1}^{n+e+c} \int_{t_k}^{t_{k+1}} \theta_{LTP}(t) [\theta_{EC}(t) a_{EC}^{(c)} + \theta_{CA3}(t) W_{CA3}(t_k) a_{CA3}^{(c)}] a_{CA3}^{(c)T} dt$$

(10)

2.4. Performance measure.

The behavior of the network was analyzed with a performance measure M , which measures how much the network retrieves the new association between right arm and food reward relative to retrieval of the old association between left arm and food reward. We focus on a retrieval trial r , with place cell activity in region CA3, given by $a_{CA3}^{(r)}$, but before the rat has encountered the food reward, so that no food input enters the network $a_{EC}^{(r)} = 0$. For the performance measure, we focus on retrieval of memory when the rat is at the choice point (T-junction). At the choice point, activity spreads along associations between place cells within CA3, such that the place cells in the left and right arm are activated. These place cells can then activate any associated food representation in region CA1 (see Figure 2, A4 and B4). Thus, we use an activity pattern in region CA3 $a_{CA3}^{(r)}$, which overlaps with the region CA3 place activity for both

the left arm $a_{CA3}^{(n)T} a_{CA3}^{(r)} = 1$ and the right arm $a_{CA3}^{(c)T} a_{CA3}^{(r)} = 1$. This assumes that activity dependent upon future expectation can be induced when the rat is still in the stem or at the T-junction. This is consistent with evidence for selectivity of place cells dependent upon future turning response (Wood et al., 2000; Frank et al., 2000).

Note that when the rat is at the choice point on trial r (Figure 2, A4 and B4), the CA1 activity will be represented by an equation which uses the weight matrix shown in equation 10:

$$a_{CA1}(t) = \theta_{EC}(t)a_{EC}^{(r)} + \theta_{CA3}(t)(W_{CA3}(t_{n+e+c}))a_{CA3}^{(r)} \quad (11)$$

The performance measure then takes the retrieval activity $a_{CA1}(t)$ and evaluates its similarity (dot product) with the food representation on the correct reversal trial $a_{EC}^{(c)}$ minus its similarity (dot product) with the now incorrect memory of food from initial trials $a_{EC}^{(n)}$. Due to the oscillations of CA3 input to CA1 during this period, we use the maximum, $\max[\]$, of this measure as an indicator of performance, but similar results would be obtained with integration over time during retrieval.

In the model, the pattern of food reward activity in region CA1, evaluated during the choice point period on this trial r , is taken as guiding the choice on this trial (i.e. the choice will depend upon the relative strength of the activity representing left arm food reward, versus the strength of activity representing right arm food reward). Thus, this equation measures the tendency that the response guided by region CA1 activity on this trial will depend on memory of associations from correct post-reversal trials $a_{EC}^{(c)}$ (left side of equation) versus memory for prior strongly-learned associations on trials $a_{EC}^{(n)}$ (right side of equation).

$$M = \max_{(t \in (t_r, t_r + \epsilon))} [(a_{EC}^{(c)})^T a_{CA1}(t) - (a_{EC}^{(n)})^T a_{CA1}(t)] \quad (12)$$

To see how this measure of behavior changes depending on the phase of oscillatory variables, we can replace the matrices in equation 11 with the components from equations 8, 9 and 10, and obtain the postsynaptic activity for inclusion in equation 12. For the purpose of simplification we will consider retrieval after a single error trial and a single correct trial. This allows removal of the summation signs, and use of the end point of a single error trial ($t_{n+e} = t_{n+1}$) and the end point of a single correct trial ($t_{n+e+c} = t_{n+2}$). We move the constant values outside the integrals. In simplifying this equation, we use the assumptions about activity in different stages itemized at the start of section 2.3 and 2.4. above.

These assumptions allow the performance measure to be reduced to an interaction of the oscillatory terms.

$$M = \max_{(t \in (t_r, t_r + \epsilon))} [\theta_{CA3}(t) \{ \int_{t_{n+e}}^{t_{n+e+c}} \theta_{LTP}(t) \theta_{EC}(t) dt - K - \int_{t_n}^{t_{n+e}} \theta_{LTP}(t) \theta_{CA3}(t) dt \}] \quad (13)$$

Similarly, we can consider the analytical conditions which would allow the previously learned associations to be erased. To accomplish this during the error stage, we desire that the previously strengthened connections should go toward zero, so we

need $\int_{t_k}^{t_{k+1}} \theta_{LTP}(t) [\theta_{CA3}(t) W_{CA3}(t_k) a_{CA3}^{(e)}] a_{CA3}^{(e)T} dt < 0$. If we ignore the constant vectors and

focus on a single theta cycle, we obtain

$$\int_0^{2\pi} \sin(t + \phi_{LTP})(X/2) \sin(t + \phi_{CA3}) dt + \int_0^{2\pi} \sin(t + \phi_{LTP})(1 - X/2) dt. \quad \text{This gives}$$

$$(X/4) \int_0^{2\pi} \cos(\phi_{LTP} - \phi_{CA3}) - \cos(2t + \phi_{CA3} + \phi_{LTP}) dt = (X/2)\pi \cos(\phi_{LTP} - \phi_{CA3}). \text{ This}$$

equation will be negative for $\frac{\pi}{2} < \phi_{LTP} - \phi_{CA3} < \frac{3\pi}{2}$. Thus, for the old synapses to decay, the oscillations of CA3 input and LTP should be out of phase. A similar analysis shows that the synapses strengthened by the correct association in equation 10 will grow for $-\frac{\pi}{2} < \phi_{LTP} - \phi_{EC} < \frac{\pi}{2}$, that is, if the entorhinal input is primarily in phase with synaptic modification. These same techniques can be used to solve equation 13, which over a single theta cycle becomes:

$$M = (X/2)\pi \cos(\phi_{LTP} - \phi_{EC}) - K - (X/2)\pi \cos(\phi_{LTP} - \phi_{CA3}) \quad (14).$$

Figure 4 shows this equation for $X=K=1$. Note that with changes in the value of X (the magnitude of synaptic transmission in equations 2 and 3), the performance measure changes amplitude but has the same overall shape of dependence on the relative phase of network variables.

Increases in the variable M corresponds to more retrieval of correct associations between correct right arm location and food after the reversal, and a decreased retrieval of incorrect associations between left arm location and food. As illustrated in Figure 4, this function is maximal when $\theta_{LTP}(t)$ and $\theta_{CA3}(t)$ have a phase difference of 180 degrees (π), and $\theta_{LTP}(t)$ and $\theta_{EC}(t)$ have a phase difference near zero. In this framework, the peak phase of $\theta_{LTP}(t)$ could be construed as an encoding phase, and the peak phase of $\theta_{CA3}(t)$ could be seen as a retrieval phase. These are the phase relationships shown earlier in Figures 1 and 3. This analysis shows how a particular behavioral constraint –

the ability to reverse a learned behavior – requires specific phase relationships between physiological variables during theta rhythm.

2.5. Relation between theoretical phases and experimental data.

The phase relationship between oscillating variables which provides the best performance in equation 14 (Figure 4) corresponds to the phase relationships observed in physiological experiments on synaptic currents (Buzsaki et al., 1986; Brankack et al., 1993; Bragin et al., 1995) and LTP (Huerta and Lisman, 1993; Holscher et al., 1997; Wyble et al., 2001), as summarized in Figures 1 and 3. In particular, our model suggests that the period of strongest potentiation at excitatory synapses arising from region CA3 should correspond to the period of weakest synaptic input at these same synapses. These synapses terminate in stratum radiatum of region CA1. Experimental data has shown that stimulation of stratum radiatum can best induce LTP at the peak of local theta in urethane anesthetized rats (Holscher et al., 1997). To analyze whether this corresponds to the time of weakest synaptic transmission, we have reanalyzed recent data on the size of evoked synaptic potentials during theta rhythm in urethane anesthetized rats (Wyble et al., 2000), using the slope of stratum radiatum theta measured directly before the evoked potential. As shown in Figure 5, the smallest slope of evoked potentials was noted near the peak of stratum radiatum theta, which is the time of peak LTP induction (Holscher et al., 1997). Consistent with the time of weakest evoked synaptic transmission, previous studies show the smallest synaptic currents in stratum radiatum at this phase of theta (Brankack et al., 1993). The analysis presented here provides a functional rationale for this paradoxical phase difference between

magnitude of synaptic input and magnitude of long-term potentiation in stratum radiatum.

Thus, the requirement for best performance in the model is consistent with physiological data showing that LTP of the CA3 input is strongest when synaptic input at these synapses is weakest. The time of best LTP induction in stratum radiatum is also the time of least pyramidal cell spiking in region CA1 (Fox et al., 1986; Skaggs et al., 1996). The lack of pyramidal cell spiking could result from the fact that the pyramidal cell soma is hyperpolarized during this phase, while the proximal dendrites are simultaneously depolarized (Kamondi et al., 1998). This suggests that synaptic modification at the peak of the theta rhythm in stratum radiatum requires that Hebbian synaptic modification at the Schaffer collaterals should depend upon post-synaptic depolarization caused by perforant pathway input from entorhinal cortex rather than associative input from Schaffer collaterals, but should not require somatic spiking activity. This is consistent with data on NMDA receptor responses to local depolarization and Hebbian long-term potentiation in the absence of spiking (Gustafsson et al., 1987).

Experimental data already show that LTP is not induced at the trough of theta rhythm recorded in stratum radiatum (Huerta and Lisman, 1993; Holscher et al., 1997). In fact, consistent with the process of extinction of prior learning during retrieval, stimulation at the trough of local theta results in LTD (Huerta and Lisman, 1995) or depotentiation (Holscher et al., 1997). These data suggest that some process must suppress Hebbian long-term potentiation in this phase of theta, because this is the phase of strongest Schaffer collateral input (Brankack et al., 1993) and maximal CA1 spiking

activity (Fox et al., 1986; Skaggs et al., 1996). The lack of LTP may result from lack of dendritic depolarization in this phase (Kamondi et al., 1998).

As shown in Figure 6, the loss of theta rhythm after fornix lesions could underlie the impairment of reversal behavior in the T-maze caused by these lesions (M'Harzi et al., 1986). This figure summarizes how separate phases of encoding and retrieval allow effective extinction of the initially learned association and learning of the new association, whereas loss of theta results in retrieval of the initially learned association during encoding, which prevents extinction of this association.

2.6. Changes in postsynaptic membrane potential during theta.

The above results extend previous proposals that separate encoding and retrieval phases depend on strength of synaptic input (Hasselmo et al., 1996; Wallenstein and Hasselmo, 1997; Sohal and Hasselmo, 1998a). This same model can also address the hypothesis that encoding and retrieval can be influenced by phasic changes in the membrane potential of the dendrite (Hasselmo et al., 1996; Paulsen and Moser, 1998) and soma (Paulsen and Moser, 1998). Data shows that during theta rhythm there are phasic changes in inhibition focused at the soma (Fox et al., 1986; Fox, 1989; Kamondi et al., 1998). This phenomenon can be modeled by assuming that the CA1 activity described above represents dendritic membrane potential in pyramidal cells (a_{dCA1}), and a separate variable represents pyramidal cell soma spiking (a_{sCA1}) (Golding and Spruston, 1998). Spiking at the soma is influenced by input from the dendrites (a_{dCA1}), multiplied by oscillatory changes in soma membrane potential (θ_{soma}) induced by shunting inhibition:

$$a_{sCA1} = \theta_{soma} a_{dCA1} \quad (15)$$

If these oscillatory changes in soma membrane potential are in phase with region CA3 input θ_{soma} then this enhances the output of retrieval activity. If they are out of phase with entorhinal input θ_{EC} and LTP θ_{LTP} , this will help prevent interference due to network output during encoding, allowing dendritic depolarization to activate NMDA channels without causing spiking activity at the soma. This corresponds to crossing of a threshold for synaptic modification without crossing a firing threshold for generating output. Use of a modification threshold enhances the influence of even relatively small values for the modulation of synaptic transmission (the parameter X in equations 2, 3 and 14). A modification threshold would select a segment of each cycle during which postsynaptic activity is above threshold. Outside of this segment, the modification rate would be zero. To simplify the analysis of the integrals, we will take the limiting case as the modification threshold approaches 1. In this case, the postsynaptic component of potentiation becomes a delta function with amplitude one and phase corresponding to the time of the positive maximum of the sine wave for synaptic transmission:

$\delta(t - \phi_{CA3} + \pi/2)$. The integral for one component of equation 13 then becomes:

$$\int_0^{2\pi} \sin(t + \phi_{LTP}) \delta(t - \phi_{CA3} + \pi/2) dt = \sin(\pi/2 + \phi_{LTP} - \phi_{CA3}) = \cos(\phi_{LTP} - \phi_{CA3}) \quad \text{This results}$$

in a performance measure very similar to equation 14, with a maximum only slightly smaller than that in equation 14, but with no dependence on magnitude of X :

$$M = \cos(\phi_{LTP} - \phi_{EC}) - K - \cos(\phi_{LTP} - \phi_{CA3}) \quad (16)$$

The disjunction between dendritic and somatic depolarization has been observed experimentally (Golding and Spruston, 1998; Kamondi et al., 1998; Paulsen and Moser, 1998). Phasic changes in disinhibition could underlie phasic changes in population spike induction in stratum pyramidale during theta (Ruddell et al., 1980; Leung, 1984), though population spikes in anesthetized animals appear to depend more on strength of synaptic input (Ruddell and Fox, 1984).

3. Discussion

The model presented here shows how theta rhythm oscillations could enhance the reversal of prior learning by preventing retrieval of previously encoded associations during new encoding. In this paper, we developed expressions for the phasic changes in synaptic input and long-term potentiation during theta rhythm, and for sensory input during different stages of behavior in a T-maze task. A performance measure evaluated memory for recent associations between location and food in this task, by measuring how much retrieval activity resembles the new correct food location, as opposed to the initial (now incorrect) food location. As shown in Figure 4, the best memory for recent associations in this task occurs with the following phase relationships between oscillations of different variables: 1.) sensory input from entorhinal cortex should be in-phase (zero degrees difference) with oscillatory changes in long-term potentiation of the synaptic input from region CA3, and 2.) retrieval activity from region CA3 should be out-of-phase (180 degree difference) with the oscillatory changes in long-term potentiation at these same synapses arising from region CA3. As shown in Figures 1 and 3, this range of best performance appears consistent with physiological data on the

phase dependence of synaptic transmission (Buzsaki et al., 1986; Brankack et al., 1993; Bragin et al., 1995; Wyble et al., 2000) and long-term potentiation (Huerta and Lisman, 1993; Holscher et al., 1997; Wyble et al., 2001).

This model predicts that recordings during tasks requiring repetitive encoding and retrieval of stimuli should show specific phase relationships to theta. These tasks include delayed matching tasks and delayed alternation (Numan et al., 1995; Otto and Eichenbaum, 1992; Hampson et al., 1999). In these tasks, hippocampal neurons should show spiking activity at the peak of local theta in stratum radiatum for information being encoded and spiking activity near the trough of local theta for information being retrieved. This is consistent with theta phase precession (O'Keefe and Recce, 1993; Skaggs et al., 1996), in which the retrieval activity (before the animal enters its place field) appears at different phases of theta than the stimulus driven encoding activity (after the animal has entered the place field).

The encoding of new information will be most rapid if theta rhythm shifts phase to allow encoding dynamics during the time when new information is being processed. This is consistent with evidence that the theta rhythm shows phase-locking to the onset of behaviorally relevant stimuli for stimulus acquisition (Macrides et al., 1982; Semba and Komisaruk, 1984; Givens, 1996). Phase locking could involve medial septal activity shifting the hippocampal theta into an encoding phase during the initial appearance of a behaviorally relevant stimulus, thereby speeding the process of encoding. In tasks requiring access to information encoded within the hippocampus, this information will become accessible during the retrieval phase, which could bias

generation of responses to the retrieval phase. This requirement is consistent with the theta phase-locking of behavioral response times (Buno and Velluti, 1977)

The mechanisms described here could also be important for obtaining relatively homogeneous strength of encoding within a distributed environment (Kali and Dayan, 2000). In particular, the model contains an ongoing interplay between synaptic weakening during retrieval and synaptic strengthening during encoding. Without a prior representation, retrieval does not occur and synaptic weakening is absent, allowing pure strengthening of relevant synapses. However, for highly familiar portions of the environment, considerable retrieval will occur during the trough of theta which will balance the new encoding at the peak. With a change in environment (e.g. insertion of a barrier, or removal of food reward), the retrieval of prior erroneous associations in the model would be no longer balanced by re-encoding of these representations, and the erroneous representations could be unlearned. (If retrieval occurs without external input, a return loop from entorhinal cortex would be necessary to prevent loss of the memory.)

Without the separation of encoding and retrieval on different phases of theta, memory for the response performed during a trial cannot be distinguished from retrieval of a previous memory used to guide performance during that trial. In addition to the model of reversal learning presented above, the same principles could account for impaired performance caused by medial septal lesions in tasks such as spatial delayed alternation (Numan and Quaranta, 1990) and delayed alternation of a go/no-go task (Numan et al., 1995), which place demands on the response memory system. In contrast, medial septal lesions do not impair performance of a task requiring working memory for a sensory stimulus (Numan and Klis, 1992). Fornix lesions also impair delayed nonmatch to sample in a spatial T maze in rats (Markowska et al., 1989) and

monkeys (Murray et al., 1989), where animals must remember their most recent response in the face of interference from multiple previous responses. Performance of a spatial delayed match to sample task is also impaired by hippocampal lesions (Hampson et al., 1999). This analysis can guide further experiments testing the predictions about specific phase relationships required by behavioral constraints in reversal and delayed alternation tasks. These behavioral constraints are ethologically relevant, as they would arise in any foraging behavior where food sites are fixed across some time periods but vary at longer intervals.

References.

- Blum, K.I. and Abbott, L.F. (1993) A model of spatial map formation in the hippocampus of the rat. *Neural Comput* 8(1):85-93
- Bragin A, Jando, G, Nadasdy, Z, Hetke, J, Wise, K, Buzsaki, G. Gamma (40-100 Hz) oscillation in the hippocampus of the behaving rat. *Journal of Neuroscience*, 15: 47-60 (1995).
- Brankack, J, Stewart, M, Fox, SE (1993). Current source density analysis of the hippocampal theta rhythm: associated sustained potentials and candidate synaptic generators. *Brain Research*, 615(2): 310-327
- Buno W Jr, Velluti JC (1977) Relationships of hippocampal theta cycles with bar pressing during self-stimulation. *Physiol Behav* 19(5):615-21
- Buzsaki, G, Czopf, J, Kondakor, I, Kellenyi, L (1986). Laminar distribution of hippocampal rhythmic slow activity (RSA) in the behaving rat: current-source density analysis, effects of urethane and atropine. *Brain Res*, 365(1): 125-37
- Buzsaki G, Leung LW, Vanderwolf CH. (1983) Cellular bases of hippocampal EEG in the behaving rat. *Brain Res*. 287(2):139-71.
- Fox, S. E. (1989). Membrane potential and impedance changes in hippocampal pyramidal cells during theta rhythm. *Exp. Brain Res.*, **77**, 283-294.
- Fox, S. E., Wolfson, S. and Ranck, J. B. J. (1986). Hippocampal theta rhythm and the firing of neurons in walking and urethane anesthetized rats. *Exp. Brain Res.*, **62**, 495-508.
- Frank LM, Brown EN, Wilson M. (2000) Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron*. 27(1):169-78.

- Givens B (1996) Stimulus-evoked resetting of the dentate theta rhythm: relation to working memory. *Neuroreport* 8: 159-163.
- Golding NL, Spruston N. (1998) Dendritic sodium spikes are variable triggers of axonal action potentials in hippocampal CA1 pyramidal neurons. *Neuron*. 21(5):1189-200.
- Gustafsson B, Wigstrom H, Abraham WC, Huang YY (1987) Long-term potentiation in the hippocampus using depolarizing current pulses as the conditioning stimulus to single volley synaptic potentials. *J Neurosci*. 7(3):774-80
- Hampson RE, Jarrard LE, Deadwyler SA (1999) Effects of ibotenate hippocampal and extrahippocampal destruction on delayed-match and -nonmatch-to-sample behavior in rats. *J Neurosci*. 19(4):1492-507.
- Hampson RE, Simeral JD, Deadwyler SA (1999) Distribution of spatial and nonspatial information in dorsal hippocampus. *Nature* 402(6762):610-4
- Hasselmo, M.E., Wyble, B.P. and Wallenstein, G.V. (1996) Encoding and retrieval of episodic memories: Role of cholinergic and GABAergic modulation in the hippocampus. *Hippocampus*, 6(6): 693-708.
- Holscher, C, Anwyl, R, Rowan, MJ (1997). Stimulation on the positive phase of hippocampal theta rhythm induces long-term potentiation that can be depotentiated by stimulation on the negative phase in area CA1 in vivo. *J Neurosci*, 17(16): 6470-6477.
- Hopfield JJ (1995) Pattern recognition computation using action potential timing for stimulus representation. *Nature* 376(6535):33-36.
- Huerta, PT, Lisman, JE (1993). Heightened synaptic plasticity of hippocampal CA1 neurons during a cholinergically induced rhythmic state. *Nature*, 364: 723-725
- Huerta PT, Lisman JE (1995). Bidirectional synaptic plasticity induced by a single burst during cholinergic theta oscillation in CA1 in vitro. *Neuron*, 15(5): 1053-1063.

Jensen O, Lisman JE (1996) Novel lists of 7 +/- 2 known items can be reliably stored in an oscillatory short-term memory network: interaction with long-term memory. *Learn Mem.* 3(2-3):257-63.

Kali, S., Dayan, P. (2000) The involvement of recurrent connections in area CA3 in establishing the properties of place fields: a model. *J. Neurosci.* 20(19):7463-7477.

Kamondi, A, Acsady, L, Wang, XJ, Buzsaki, G (1998). Theta oscillations in somata and dendrites of hippocampal pyramidal cells in vivo: activity-dependent phase-precession of action potentials. *Hippocampus*, 8(3): 244-61

Leung, L.-W. S. (1984) Model of gradual phase shift of theta rhythm in the rat. *J. Neurophysiol.* 52: 1051-1065.

Macrides F., Eichenbaum, HB, Forbes, W.B (1982) Temporal relationship between sniffing and the limbic theta rhythm during odor discrimination reversal learning. *J. Neurosci.* 2: 1705-1717.

Markowska AL, Olton DS, Murray EA, Gaffan D (1989) A comparative analysis of the role of fornix and cingulate cortex in memory: rats. *Exp Brain Res.* 74(1):187-201.

M'Harzi, M., Palacios, A., Monmaur P, Willig F, Houcine O and Delacour J. (1987) Effects of selective lesions of fimbria-fornix on learning set in the rat. *Physiol. & Behav.* 40: 181-188.

Murray EA, Davidson M, Gaffan D, Olton DS, Suomi S. (1989) Effects of fornix transection and cingulate cortical ablation on spatial memory in rhesus monkeys. *Exp Brain Res.* 74(1):173-86

Numan, R., Feloney, M.P., Pham, K.H. and Tieber, L.M. (1995) Effects of medial septal lesions on an operant go/no-go delayed response alternation task in rats. *Physiol. Behav.* 58: 1263-1271.

Numan R, Klis D (1992) Effects of medial septal lesions on an operant delayed go/no-go discrimination in rats. *Brain Res. Bull.* 29: 643-650.

Numan R, Quaranta, JR Jr. (1990) Effects of medial septal lesions on operant delayed alternation in rats. *Brain Res.* 531: 232-241.

O'Keefe, J. and Recce, M.L. (1993) Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3: 317-330.

Orr G, Rao G, Stevenson GD, Barnes CA, McNaughton BL (1999) Hippocampal synaptic plasticity is modulated by the theta rhythm in the fascia dentata of freely behaving rats. *Soc. Neurosci. Abstr.* 25: 2165 (864.14).

Otto T, Eichenbaum H (1992) Neuronal activity in the hippocampus during delayed non-match to sample performance in rats: evidence for hippocampal processing in recognition memory. *Hippocampus* 2(3):323-34.

Paulsen O, Moser EI. (1998) A model of hippocampal memory encoding and retrieval: GABAergic control of synaptic plasticity. *Trends Neurosci.* 21(7):273-8.

Pavrides C, Greenstein YJ, Grudman M, Winson J. (1988) Long-term potentiation in the dentate gyrus is induced preferentially on the positive phase of theta-rhythm. *Brain Res.* 439(1-2):383-7.

Rao, G, Orr, G, Barnes, CA, and McNaughton BL (1998) Theta rhythm modulation of CA1 evoked potential size during locomotion. *Soc. Neurosci. Abstr.* 24: 933 (367.16).

Rudell, A. P. and Fox, S. E. (1984). Hippocampal excitability related to the phase of the theta rhythm in urethanized rats. *Brain Res.*, **294**, 350-353

Rudell, A. P., Fox, S. E. and Ranck, J. B. J. (1980). Hippocampal excitability phase-locked to the theta rhythm in walking rats. *Exp. Neurol.*, **68**, 87-96.

Semba K, Komisaruk BR (1984) Neural substrates of two different rhythmical vibrissal movements in the rat. *Neuroscience* 12(3):761-74

Skaggs, W.E., McNaughton, B.L., Wilson, M.A. and Barnes, C.A. (1996) Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* 6, 149-172.

Sohal, V.S. and Hasselmo, M.E. (1998a) Changes in GABAB modulation during a theta cycle may be analogous to the fall of temperature during annealing. *Neural Computation* 10: 889-902.

Sohal, V.S. and Hasselmo, M.E. (1998) GABAB modulation improves sequence disambiguation in computational models of hippocampal region CA3. *Hippocampus* 8: 171-193.

Stewart, M., Quirk, G.J., Barry, M., Fox, S.E. (1992) Firing relations of medial entorhinal neurons to the hippocampal theta rhythm in urethane anesthetized and walking rats. *Exp Brain Res.* 90(1):21-8.

Wallenstein, G.V. and Hasselmo, M.E. (1997a) GABAergic modulation of hippocampal population activity: Sequence learning, place field development and the phase precession effect. *J. Neurophysiol.* 78(1): 393-408.

Wiener SI, Paul CA, Eichenbaum H. (1989) Spatial and behavioral correlates of hippocampal neuronal activity. *J Neurosci.* 9(8):2737-63.

Whishaw IQ, Tomie JA (1997) Perseveration on place reversals in spatial swimming pool tasks: further evidence for place learning in hippocampal rats. *Hippocampus* 1997;7(4):361-70

Wood ER, Dudchenko PA, Robitsek RJ, Eichenbaum H. (2000) Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron*. 27(3):623-33.

Wyble BP, Linster C, Hasselmo ME. (2000) Size of CA1-evoked synaptic potentials is related to theta rhythm phase in rat hippocampus. *J. Neurophysiol.* 83(4):2138-2144.

Wyble BP, Hyman JM, Goyal V, Hasselmo ME (2001) Phase relationship of LTP induction and behavior to theta rhythm in the rat hippocampus. *Soc. Neurosci. Abstr.* 27.

Young, B.J., Otto, T., Fox, G.D. and Eichenbaum, H. (1997) Memory representation within the parahippocampal region. *J. Neurosci.* 17: 5183-5195

ACKNOWLEDGEMENTS: We appreciate the assistance on experiments of Christiane Linster and Bradley Molyneaux. I appreciate critical comments by Marc Howard, James Hyman, Terry Kremin, Ana Nathe, Anatoli Gorchetchnikov, Norbert Forin and Seth Ramus. In memory of Patricia Tillberg Hasselmo. This work was supported by NIH MH61495, NIH MH60013, and NSF IBN9996177.

FIGURE LEGENDS:

Figure 1. Schematic representation of the change in dynamics during hippocampal theta rhythm oscillations. LEFT: appropriate dynamics for encoding. At the trough of the theta rhythm in the EEG recorded at the hippocampal fissure, synaptic transmission arising from entorhinal cortex is strong. Synaptic transmission arising from region CA3 is weak, but these same synapses show a strong capacity for long-term potentiation. This allows the afferent input from entorhinal cortex to set patterns to be encoded, while preventing interference from previously encoded patterns on the excitatory synapses arising from region CA3. RIGHT: appropriate dynamics for retrieval. At the peak of the theta rhythm, synaptic transmission arising from entorhinal cortex is relatively weak (though strong enough to provide retrieval cues). In contrast, the synaptic transmission arising from region CA3 is strong, allowing effective retrieval of previously encoded sequences. At this phase, synapses undergo depotentiation rather than LTP, preventing further encoding of retrieval activity and allowing forgetting of incorrect retrieval.

Figure 2. A. Schematic T-maze illustrating behavioral response during different stages in the task. Location of food reward is marked with the letter *F*. A1. Initial learning of left turn response. A2. Error after reversal. No food reward is found in the left arm of the maze. A3. Correct after reversal – After a correct response, food reward is found in the right arm of the maze. A4. Retrieval choice. Analysis focuses on performance at the choice point in a retrieval trial (*r*), when network activity reflects memory for food location. B. Functional components of the network during different trials in a behavioral task. B1. Initial encoding trials (*n*). When in the left arm, the rat encodes associations between the left arm CA3 place cell activity (*L*) and left arm food reward

activity (F). Thick lines represents strengthened synapses. B2. Performance of an error after reversal (e). When in the left arm (L), no food reward is found. This results in weakening of the previous association (thinner line) between left arm place and left arm food reward. (No food reward activity appears because rat does not find reward in error trials). B3. Encoding of correct reversal trial (c). When in the right arm, food reward is found. This results in encoding of an association between right arm CA3 place cell activity (R) and right arm food reward (F) activated by entorhinal input. B4. Retrieval of correct reversal trial (r). This test focuses on activity observed when the rat is at the choice point (Ch), allowing retrieval of both arm place representations (L and R), and potential retrieval of their food reward associations.

Figure 3. Summary of experimental data on changes in synaptic potentials and long-term potentiation at different phases of the theta rhythm. Dotted lines indicate the phase of theta with best dynamics for encoding (E) and the phase with best dynamics for retrieval (R). Entorhinal input: Synaptic input from entorhinal cortex in stratum lacunosum-moleculare is strongest at the peak of theta recorded in stratum radiatum -- equivalent to the trough of theta recorded at the hippocampal fissure (Buzsaki et al., 1986; Branckak et al., 1993; Bragin et al., 1995). Soma: Intracellular recording in vivo shows the depolarization of the soma in phase with CA3 input (Kamondi et al., 1998). However, dendritic depolarization is in phase with entorhinal input (Kamondi et al., 1998). LTP: Long-term potentiation of synaptic potentials in stratum radiatum is strongest at the peak of local stratum radiatum theta rhythm in slice preparations (Huerta and Lisman, 1993) and urethane anesthetized rats (Holscher et al., 1997). Strong entorhinal input and strong LTP in stratum radiatum is associated with the encoding

phase (E). CA3 input: Synaptic potentials in stratum radiatum are strongest at the trough of the local stratum radiatum theta rhythm as demonstrated with current source density analysis (Buzsáki et al. 1986; Branckak et al., 1993; Bragin et al., 1995) and studies of evoked potentials (Ruddell and Fox, 1984; Wyble et al., 2000). Strong synaptic transmission in stratum radiatum is present during the retrieval phase (R).

Figure 4. A. Performance of the network depends on the relative phase of oscillations in synaptic transmission and long-term potentiation. This graph shows how performance in T-maze reversal changes as a function of the phase difference between LTP and synaptic input from entorhinal cortex (EC phase) and of the phase difference between LTP and synaptic input from region CA3 (CA3 phase). This is a graphical representation of equation 14. Best performance occurs when EC input is in phase (zero degrees) with LTP, but when region CA3 synaptic input is 180 degrees out of phase with LTP.

Figure 5. Data showing the magnitude of initial slope of evoked synaptic potentials (EPSP slope) in stratum radiatum of hippocampal region CA1 relative to the slope of the local theta field potential (Local EEG slope) recorded from the same electrode just before recording of the evoked potential. This data is reanalyzed from a previous study correlating slope of potentials with hippocampal fissure theta (Wyble et al., 2000). This data demonstrates that the initial slope of evoked potentials is smallest near the peak of local theta – the same phase when induction of long-term potentiation is maximal (Holscher et al., 1997; Wyble et al., 2001).

Figure 6. Schematic representation of the encoding and extinction of associations in the model during normal theta (with theta) and after loss of theta due to fornix lesions (no theta). Circles represent populations of units in region CA3 representing arm location (left = L, right = R), and populations in region CA1 activated by food location. Arrows labeled with “F” represent sensory input about food reward arriving from entorhinal cortex. Filled circles represent populations activated by afferent input from entorhinal cortex, or by retrieval activity spreading from region CA3. Initial learning: during initial encoding trials, simultaneous activity of left arm place cells and food reward input causes strengthening of connections between CA3 and CA1 (thicker line). This association can be retrieved in both normal and fornix lesioned rats. Error after reversal: with normal theta, when food reward is not present, the region CA1 neuron is not activated during encoding, so no strengthening occurs, and during the retrieval phase depotentiation causes the connection storing the association between left arm and food to weaken (thinner line). In contrast, without theta, the activity spreading along the strengthened connection causes post-synaptic activity which allows strengthening and prevents weakening of the association between left arm and food. Correct reversal: with normal theta, a new stronger association is formed between right arm and food. Without theta, the new and old associations are similar in strength. Retrieval choice: at the choice point on subsequent trials, retrieval of the association between right arm and food dominates with normal theta, whereas without theta the associations have similar strength.

Glossary of mathematical terms:

$a_{EC}^{(n)}$ = entorhinal activity (This is a fixed vector during each stage of task. The stage designated by superscript: n=initial learning, e=error after reversal, c=correct after reversal, r=retrieval test)

$a_{CA3}^{(n)}$ = region CA3 activity (superscript designates which stage of task)

$a_{CA1}(t)$ = region CA1 activity (varies with time, t_n = end of initial period, t_{n+e} = end of erroneous trials, t_{n+e+c} = end of correct trials)

$W_{CA3}(t)$ = synaptic input matrix from region CA3

W_{EC} = synaptic input matrix from entorhinal cortex

$\theta_{CA3}(t)$ = oscillatory modulation of synaptic transmission from CA3 during theta rhythm, with phase = ϕ_{CA3}

$\theta_{EC}(t)$ = oscillatory modulation of synaptic transmission from entorhinal cortex during theta rhythm, with phase = ϕ_{EC}

$\theta_{LTP}(t)$ = oscillatory modulation of rate of synaptic modification during theta rhythm, with phase = ϕ_{LTP}

M = performance measure in reversal task dependent on relative phase of oscillatory modulation.

Figure 1