Neurobiology of Learning and Memory **78,** 565–569 (2002)
doi:10.1006/nlme.2002.4098

Hippocampal Memory Formation, Plasticity, and the Role of Sleep

Matthew A. Wilson

Center for Learning and Memory, Departments of Brain and Cognitive Sciences, and Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The hippocampus is a structure that has long been associated with memory (Squire & Zola, 1998; Tulving, 1983). Our views on this association have been shaped by three seminal observations. The first was the discovery of anterograde memory loss in the patient H.M. who underwent a surgical procedure that led to the bilateral loss of portions of the medial temporal lobes, including large portions of the hippocampus (Scoville & Milner, 1957). The second important observation involved the characterization of the activity of hippocampal neurons during spatial behavior. Second, by placing microelectrodes into the hippocampus of a freely behaving rat, O'Keefe discovered that hippocampal neurons would fire in restricted regions of space (O'Keefe & Dostrovsky, 1971). He termed these spatial receptive fields *place fields* and the cells *place cells.* Third, during that same period Bliss and Lomo (1973) found that electrical stimulation of the afferent fibers leading into the hippocampus could produce long-lasting enhancement of the efficacy of synaptic transmission or LTP. Together these observations link the hippocampus to the cognitive aspects of memory and identify potential neural mechanisms that might be responsible for the formation and maintenance of memory.

The involvement of the hippocampus in the formation of long-term memory has not yet been firmly established, but several lines of evidence suggest that the hippocampus participates in a gradual process, referred to as memory consolidation, in which mnemonic information initially established within the hippocampus gradually becomes incorporated into extrahippocampal sites. As originally noted by O'Keefe, one characteristic of hippocampal spatial activity is that individual cells can show consistent responses within a given environment over long periods of time. Conversely, responses in different environments lead to distinct responses. It is typical to find from 30 to 50% of hippocampal pyramidal cells exhibiting spatial responses in a specific environment. It is clear that representations of distinct locations require the evaluation of responses of neural ensembles.

To determine what attributes of ensemble response might be significant in the formation and representation of mnemonic traces using multiple electrode recording techniques, Wilson and McNaughton (1993) monitored the dynamics of ensemble activity during

Address correspondence and reprint requests to Matthew A. Wilson, Center for Learning and Memory, Departments of Brain and Cognitive Sciences and Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.

1074-7427/02 $35.00

exposure to a novel environment using a dual box apparatus. In this task the animal was familiarized with a 62-cm square box in which the rat performed a random foraging task. The animal was allowed to become familiarized with the space over a 10-day period through daily 10-min exposures. The animal was then given access to an additional box that was attached to the familiar box and was allowed to explore the new space while still being allowed to move between the familiar and novel spaces. In this way characteristics of neural activity patterns in the two spaces could be compared and the evolution of these characteristics could be observed. Three fundamental changes in neural activity were observed in the novel but not the familiar space during the initial 5–10 min of exploration of this augmented environment. First, inhibitory cells recorded in the CA1 pyramidal layer showed a significant reduction in firing rate that was restricted to the novel region of the box. The change in firing rate could occur rapidly—on the order of 1 s—as the animal crossed the implicit dividing line separating the familiar and novel space either from novel to familiar or vice versa. Second, the consistency or temporal covariance of neuronal ensembles was reduced in the novel region of the box. This was assessed by attempting to reconstruct the position of the animal in space by decoding the moment-by-moment neuronal ensemble activity. Reconstruction error was found to be high during the same initial 5–10 min period in the novel space. Third was the observation that the majority of cells that responded in the novel space consisted of cells that had been silent in the familiar space. That is, the representation of the augmented space was an augmented representation of the familiar space. These findings suggested that exposure to novelty engaged mechanisms of plasticity to allow the formation of a representation which minimized the overlap with existing representations.

Clearly, given that in the initial familiar environment, 30–40% of the available neurons are used, continued augmentation requires the reuse of neurons. Several strategies could be employed. Given that neural patterns at single spatial locations appear to activate approximately 5% of the roughly 300,000 neurons in CA1 of the rat, each pattern could be composed of unique configurations of 10–15,000 cells. This yields a combinatorially enormous representational space. While this result did not directly implicate mechanisms of plasticity with the experience-dependent changes that occurred, it suggested that changes in covariance of ensemble activity might be a significant consequence of this plasticity and that manipulation of plasticity might produce a similar phenotype.

In order to examine the contribution of hippocampal plasticity to the structure of place-cell activity, McHugh, Blum, Tsien, Tonegawa, and Wilson (1996) took advantage of new molecular genetic techniques that allowed the knockout of the NMDA receptor largely restricted to hippocampal CA1 pyramidal cells. This manipulation had the effect of disrupting NMDA-mediated plasticity at synapses into CA1 originating in area CA3 and the entorhinal cortex. Animals tested on a spatial reference memory version of the Morris water maze showed a significant impairment in spatial memory. When the place cells in these animals were recorded during exploration of open field and linear track environments, they showed three primary deficits. First, spatial receptive fields were enlarged by approximately 30%. Second, the fields displayed multiple spatial peaks rather than the singular spatially coherent peak typically seen in control animals. Third, the covariance of firing of place cells that had overlapping place fields was significantly reduced and was not significantly different from zero. This indicates that individual cells did not fire consistently during repeated passes through their place fields, and hence, multiple cells would not

covary in their firing at those same locations. The lack of the ability to provide coordinated, or covarying output was argued to be one of the primary sources of the behavioral deficit, with the hippocampus lacking the ability to signal prior experience within the spatial environment through the use of correlated neural ensemble output.

A simple model that captured these characteristics of NMDA-mediated spatial receptive field plasticity was described in which Hebbian plasticity served to enhance response to overlapping inputs and therefore increase signal-to-noise, thus increasing covariance, providing more robust firing in single spatial locations.

While this simple model relied upon the role of coactivity as a mechanism for enhancing synaptic plasticity based on the first-order associative properties of NMDA-mediated LTP in which temporally overlapping inputs would tend to be potentiated, it did not incorporate two critical properties of hippocampal activity and synaptic plasticity. First, on linear tracks or during behavior in which animals systematically followed limited paths, hippocampal neurons discharge in both a spatially and a directionally dependent manner. Second, timing of pre- and postsynaptic activity is known to regulate the direction of synaptic plasticity in a temporally asymmetric fashion such that presynaptic activity that arrives prior to postsynaptic output within a narrow time window on the order of tens of milliseconds leads to synaptic potentiation, while reversing this order leads to synaptic depression (Markram, Lubke, Frotscher, & Sakmann, 1997). These two properties suggest that synaptic plasticity in the hippocampus would lead to spatially asymmetric changes in spatial receptive fields that would reflect the history of behavior in a given environment. This prediction was confirmed when it was found that place fields in area CA1 rapidly take on a spatially asymmetric shape consistent with the history of behavior and the temporally asymmetric properties of NMDA-mediated LTP in this area (Mehta, Quirk, & Wilson, 2000).

This result suggests that a prior state or position can influence the recognition (or signaling of relative familiarity) of a current state or position. During spatial behavior this order dependency would manifest itself as encoding of trajectories. Precisely such responses have been observed during spatial tasks involving multiple paths. Frank, Brown, and Wilson (2000) recorded simultaneously from cells in the superficial and deep layers of the entorhinal cortex as well as area CA1 in the hippocampus. Cells in each of these areas were found to show dependence upon spatial trajectories or paths. The observation of path dependence of cells providing the input to the hippocampus suggests that similar responses might be seen in other cortical areas that both provide the source of input to the hippocampus as well as serve as the potential repository of long-term consolidated mnemonic information and that asymmetric or temporal order dependent receptive fields could be a general property of cortical receptive fields. This would be consistent with the assumption that the neocortex is capable of maintaining mnemonic information which is initially hippocampally dependent, which would require that the neocortex possess similar or compatible encoding capabilities (Teyler & DiScenna, 1986). A further property of neurons in the deep entorhinal cortex was also seen that has implications for our understanding of the use of hippocampal mnemonic information and the role of the neocortex in establishing consolidated memory representations that differ from but depend upon the output of the hippocampus proper. Neurons in the deep entorhinal cortex were observed to produce spatial responses that appeared to generalize across environments, capturing

regularities in spatial behavior that might be used to construct generalized models that are derived from specific unique experiences, but are more applicable to novel circumstances.

The ability to establish dependencies on the temporal order of spatial experience in the response of hippocampal neurons and in neurons of the adjacent cerebral cortex suggests that memory of temporally ordered events might also be maintained within these structures. In order to observe such temporally ordered event memory we examined the activity of hippocampal neurons during periods of sleep. The study of sleep provides an opportunity to identify mnemonic activity that results from behavior in a context in which sensory and behavioral input no longer contributes to that activity. Hence, it can be argued that this activity is a direct reflection of the residual influence of experience on neural substrates and therefore must necessarily be derived from underlying mechanisms of memory.

Early observations confirmed that such residual experience-dependent activity could be observed during periods of non-REM sleep immediately following awake experience (Wilson & McNaughton, 1994). This activity took the form of a correlated discharge of CA1 place cells that had recently been coactive during spatial exploration. This mnemonic reactivation took place predominantly during brief periods of coordinated network discharge known as *ripples* lasting approximately 100 ms and strongly modulated by a 200 Hz oscillatory rhythm (Buzsaki, Horvath, Urioste, Hetke, & Wise, 1992).

Interestingly, suggestions that these events might serve as the initial stage of a process of memory consolidation involving communication between the hippocampus and the neocortex received further support when it was found that these hippocampal discharge events tended to occur in conjunction with another neocortical rhythmic event known as the sleep spindle (Siapas & Wilson, 1998). The coincidence of these two distinctive non-REM sleep events suggested that they reflected a mechanism designed to facilitate hippocampal influence on the formation of memory traces in the neocortex during sleep by providing bias to neocortical neurons that would be synchronized through spindle-modulated activity.

While the observation of non-REM reactivation demonstrated mnemonic content during sleep, it did not provide evidence of preservation of long timescale, sequential event memory which is believed to be a critical function of hippocampal circuits. Subsequent analysis of REM sleep by Louie and Wilson (2001) revealed just such temporal replay. During REM sleep in the rat, hippocampal neurons were found to replay the sequence of activity that had been experienced on a timescale of tens of seconds to minutes. These extended patterns of ensemble response could be directly matched with corresponding patterns that had been recorded during training on a simple spatial behavioral task. Over 40% of REM episodes, each lasting 1 to 2 min, was found to show significant match with the sequential patterns established during awake behavior. Further, the general patterns of theta rhythmic modulation of population response that reflect the state of locomotion of the animal were also recapitulated. The correspondence was sufficiently robust to allow reconstruction of the spatial trajectories being replayed on a second by second basis over the course of an entire REM episode.

Overall, these results indicate a primary role of the hippocampus in establishing recognition of context reflected in the multimodal inputs from the entorhinal cortex through the coordinated discharge of CA1 neurons answering the question, Has this pattern passed through the hippocampus before? The readout would reflect the relative familiarity encoded in the covariance of response. This indication of relative familiarity would be used during

navigation to identify the spatial context in which memory for sequential events could then be accessed. The temporal component of memory established through specific behavior within that environment would further be encoded through temporally asymmetric modifications of synapses within this region allowing paths or trajectories to be incorporated into the hippocampal memory system. The generalization of these paths or sequential events in the neocortex as evidenced by the responses in the deep entorhinal cortex would provide the means to construct models of behavior derived from experience but able to guide the animal under varied conditions. The retrieval and replay of these memories during sleep might provide a mechanism by which this mnemonic information in the hippocampus is gradually incorporated into neocortical circuits.

REFERENCES

Bliss, T. V. P., & Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiology, London* **232,** 331–356.

Buzsaki, G., Horvath, Z., Urioste, R., Hetke, J., & Wise, K. (1992). High-frequency network oscillation in the hippocampus. *Science* **256,** 1025–1027.

Frank, L. M., Brown, E. N., & Wilson, M. A. (2000). Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron* **27,** 169–178.

Louie, K., Wilson, M. A., (2001). Temporally structured REM sleep replay of awake hippocampal ensemble activity. *Neuron* **29,** 145–156.

Markram, H., Lubke, J., Frotscher, M., Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* **275,** 213–215.

McHugh, T. J., Blum, K. E., Tsien, I. Z., Tonegawa, S., & Wilson, M. A. (1996). Impaired hippocampal representation of space in CA1-specific NMDAR 1 knockout mice. *Cell* **87,** 1339–1349.

Mehta, M. R., Quirk, M. C., & Wilson, M. A. (2000). Experience-dependent asymmetric shape of hippocampal receptive fields. *Neuron* **25,** 707–715.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Research* **34,** 171–175.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry* **20,** 11–21.

Siapas, A. G., & Wilson, M. A. (1998). Coordinated interactions between hippocampal ripples and cortical spindles during slow-wave sleep. *Neuron* **21,** 1123–1128.

Squire, L. R., & Zola, S. M. (1998). Commentary: Episodic memory, semantic memory, and amnesia. *Hippocampus* **8,** 205–211.

Teyler, T. I., & DiScenna, P. (1986). The hippocampal memory indexing theory. *Behavioral Neuroscience* **100,** 147–54.

Tulving, E. (1993) *Elements of episodic memory.* Oxford: Clarendon Press.

Wilson, M. A., & McNaughton B. L. (1993). Dynamics of the hippocampal ensemble code for space. *Science* **261,** 1055–1058.

Wilson, M. A., & McNaughton B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science* **65,** 676–679.