Review

Tracking the flow of hippocampal computation: Pattern separation, pattern completion, and attractor dynamics

James J. Knierim,⇑, Joshua P. Neunuebel

Krieger Mind/Brain Institute and Solomon H. Snyder Department of Neuroscience, Johns Hopkins University, United States
Dept. of Psychological and Brain Sciences, University of Delaware, United States

Abstract

Classic computational theories of the mnemonic functions of the hippocampus ascribe the processes of pattern separation to the dentate gyrus (DG) and pattern completion to the CA3 region. Until the last decade, the large majority of single-unit studies of the hippocampus in behaving animals were from the CA1 region. The lack of data from the DG, CA3, and the entorhinal inputs to the hippocampus severely hampered the ability to test these theories with neurophysiological techniques. The past ten years have seen a major increase in the recordings from the CA3 region and the medial entorhinal cortex (MEC), with an increasing (but still limited) number of experiments from the lateral entorhinal cortex (LEC) and DG. This paper reviews a series of studies in a local–global cue mismatch (double-rotation) experiment in which recordings were made from cells in the anterior thalamus, MEC, LEC, DG, CA3, and CA1 regions. Compared to the standard cue environment, the change in the DG representation of the cue-mismatch environment was greater than the changes in its entorhinal inputs, providing support for the theory of pattern separation in the DG. In contrast, the change in the CA3 representation of the cue-mismatch environment was less than the changes in its entorhinal and DG inputs, providing support for a pattern completion/error correction function of CA3. The results are interpreted in terms of continuous attractor network models of the hippocampus and the relationship of these models to pattern separation and pattern completion theories. Whereas DG may perform an automatic pattern separation function, the attractor dynamics of CA3 allow it to perform a pattern separation or pattern completion function, depending on the nature of its inputs and the relative strength of the internal attractor dynamics.

Keywords:
Place cells
Attractors
Pattern completion
Pattern separation
Dentate gyrus
CA3

1. Introduction

Although most computational theories of the mnemonic functions of the hippocampus have focused on the CA3 and dentate gyrus (DG) regions, the large majority of single-unit studies of the hippocampus have been recordings from the CA1 region. This emphasis on CA1 is understandable for both functional and practical reasons. CA1 is the region that primarily transmits the output of DG/CA3 processing to the rest of the cerebrum (Witter & Amaral, 2004). Thus, it can serve as a functional readout of the information provided by the hippocampus to other brain areas that are involved in hippocampus-dependent cognition and behavior. Moreover, CA1 is the first cell layer encountered in the rodent hippocampus when an electrode is advanced from the dorsal surface of the brain, and it is by far the easiest layer of the hippocampus to record large ensembles of well-isolated units.

Although recordings from CA1 can illuminate the types of information and representations being sent to other brain regions, these recordings in isolation can inform little about the nature of the information processing that occurs within the hippocampal circuitry. For example, one may know that CA1 place cells form independent spatial representations of two distinct environments (Bostock, Muller, & Kubie, 1991). However, this knowledge alone tells us little about the computational processing that creates these two representations, and whether that processing occurs within CA1, in upstream hippocampal regions like DG or CA3, or even in regions entirely afferent to the hippocampus. To understand the neural computations of the hippocampus, it is necessary to understand the information represented in hippocampal inputs, in its internal processing stages, and in its outputs, as well as the transformations that occur as information is processed through these circuits.

⇑ Corresponding author at: Krieger Mind/Brain Institute, Johns Hopkins University, 338 Krieger Hall, 3400 N. Charles Street, Baltimore, MD 21218, United States.
E-mail address: jknierim@jhu.edu (J.J. Knierim).

http://dx.doi.org/10.1016/j.nlm.2015.10.008
1074-7427/© 2015 Elsevier Inc. All rights reserved.
This article will review a series of studies from our lab over the past decade in which we recorded the activity of hippocampal input regions and output regions, as well as the intrahippocampal processing in the DG and CA3 regions. To induce controlled, parametric changes to the animal’s environment, we used a “double rotation” manipulation, in which both the local reference frame of the behavioral track was placed in varying degrees of conflict with the global reference frame of the laboratory environment (Knierim, 2002; Shapiro, Tanila, & Eichenbaum, 1997). We investigated how neural populations in the hippocampal system respond to these alterations in order to deduce the neural representations and computations associated with the different regions. In particular, we addressed the questions of whether we can interpret (1) DG responses as evidence for its proposed role of performing pattern separation on its inputs and (2) CA3 responses as evidence for its proposed role of performing pattern completion (or the related concepts of error correction and generalization) on its inputs. We begin with a brief history of the computational theories of pattern separation and completion.

2. Classic theories of DG function: Pattern separation in DG vs. pattern completion in CA3

The most prominent theory of DG function is the pattern separation theory (Kesner, Gilbert, & Wallenstein, 2000; McNaughton & Morris, 1987; McNaughton & Nadel, 1990; Rolls & Treves, 1998; Yassa & Stark, 2011), which originated in David Marr’s theory of the cerebellum (Marr, 1969). Marr proposed that the cerebellar granule layer created a very sparse representation of incoming sensorimotor input by an expansion recoding strategy; that is, highly overlapping representations encoded by populations of pontine mossy fibers, which formed synapses onto an enormous layer of cerebellar granule cells, were transformed into extremely sparse, independent representations in the granule layer. This transformation allowed the storage, with minimal interference, of the countless contexts in which a movement occurred. Investigators elaborated Marr’s subsequent theory of the hippocampus (Marr, 1971) and ascribed a similar function to the DG granule cell layer (McNaughton & Morris, 1987; McNaughton & Nadel, 1990; Myers & Scharffman, 2009, 2011; O’Reilly & McClelland, 1994; Rolls & Treves, 1998). If the entorhinal cortex (EC) represented two experiences with a high degree of overlap in the population of active cells, an expansion recoding mechanism caused the hippocampal granule layer to create representations that overlapped considerably less than its EC inputs. This process is illustrated in Fig. 1, which shows a hypothetical relationship between two input patterns and two output patterns. Each sphere represents a cell in the population, and the colored spheres represent the currently active neurons. The input patterns share a great overlap, in that each pattern shares 2 of the 3 active neurons. With expansion recoding onto a larger output population, there is no overlap in the sets of active cells.

Pattern separation may be graphically illustrated in terms of the input–output curves of the DG (Fig. 2A). Consider how the DG and the EC represent two different events. The x-axis represents the difference in the combined medial entorhinal cortex (MEC) and lateral entorhinal cortex (LEC) representations of the two events. This difference may be quantified by the correlation of neural population activity between the representations of each event (a high ΔInput corresponds to a low correlation between the two representations). The y-axis represents the difference in the corresponding DG representations. According to the pattern separation theory, the ΔOutput of the DG is greater than the ΔInput from the EC, quickly reaching a maximum difference (i.e., the DG representations of the two experiences are completely decorrelated, even when the ΔInput is still fairly small). In support of the theory that the DG performs such an operation, rats with selective DG lesions were impaired in tasks thought to rely on spatial pattern separation (Gilbert, Kesner, & Lee, 2001); genetic knockout of the NR1 subunit of the NMDA receptor in DG impaired the ability of mice to discriminate similar environments in a fear conditioning experiment (McHugh et al., 2007); and CA3 place fields in these mice were impaired in the discrimination of two environments (McHugh et al., 2007). Human imaging studies have supported a role of the hippocampus in pattern separation, but these studies were limited by an inability to resolve the DG and CA3 regions (Bakker, Kirwan, Miller, & Stark, 2008). (For more detailed reviews, see Rolls & Kesner, 2006; Santoro, 2013; Yassa & Stark, 2011.)

The complementary role of pattern completion is commonly ascribed to the CA3 region. Strictly defined, pattern completion refers to the ability of a network to retrieve from memory a complete pattern of activity when presented with incomplete or degraded input patterns; for example, a network may output a complete phone number when presented with only a fraction of the number as a retrieval cue. This process is thought to be the result of the recurrent collateral circuitry providing an anatomical basis for an autoassociative network, that is, a network in which activity patterns can become associated with themselves. Such a network might also display attractor dynamics. In general terms, an attractor refers to a set of stable states in a system that have the property that they “attract” neighboring states to move toward the stable states, like a magnet attracts nearby metal particles or like a ball rolls down the slope of a hill toward a valley (Knierim & Zhang, 2012). Attractor dynamics in CA3, in concert with pattern...
separation processing in its DG inputs, may cause the CA3 network to display outputs that show both pattern separation and pattern completion.

Fig. 2B shows the sigmoidal relationship between pattern input similarity (from combined MEC, LEC, and DG inputs) and pattern output similarity that arises from these attractor dynamics (Guzowski, Knierim, & Moser, 2004; McClelland & Goddard, 1996; Rolls & Treves, 1998). When two input representations are similar (Δinput = 0), the output representations are also very similar (Δoutput = 0). As the representations diverge (Δinput > 0, but still small), the output representations of CA3 maintain a higher degree of similarity than the inputs (Δoutput < Δinput). This resistance to changes in the input is caused by the putative attractor dynamics of CA3 (Knierim & Zhang, 2012). The inputs cause CA3 cells to fire within the basin of attraction set up by the rat’s previous experiences. The attractor dynamics cause the CA3 representation to move closer to the Δoutput = 0 state, although the influences from the external input prevent the two representations from being completely correlated (i.e., Δoutput > 0). As the Δinput increases further, eventually the attractor dynamics cause a non-linear change in the Δoutput, causing the output representations to be less similar than the inputs (pattern separation: Δoutput > Δinput). This can be interpreted as the input representations of each experience causing the CA3 representations to fall into two distinct attractor states that are mutually inhibited.

Note that the recurrent collaterals and putative attractor dynamics of CA3 do not just reflect pattern completion. When Δinput is small, CA3 reflects pattern completion (Δoutput < Δinput). When Δinput is large, CA3 output reflects pattern separation (Δoutput > Δinput). Note also that in this model, pattern completion, error correction, and generalization are all different instances of the same underlying computational dynamics. Outputs can be made more similar than the inputs when the changes to the input (Δinput) are due to missing or degraded parts of the representation (pattern completion), small errors in the input representation (error correction), or small differences in the representations of legitimately similar inputs (generalization). Although these different cases might result in variations in the shape of the sigmoid (e.g., making it shallower or sharper; moving the inflection point along the x-axis), the attractor dynamics will endow the CA3 with qualitatively similar, nonlinear, input–output curves. Thus, rather than performing pattern separation per se, the CA3 output can be thought of as the final arbiter between the pattern separation processes of the DG inputs and the pattern completion processes of the recurrent collateral circuitry.

Fig. 2C shows the hypothesized relationship between input changes and output changes for CA1 (Guzowski et al., 2004).

Lacking both the numerical expansion between the input and output layers of DG and the strong, recurrent collateral system of CA3, CA1 is thought to show a more linear relationship between its inputs and outputs (as Δinput increases, Δoutput increases by approximately the same amount). Note that this linear relationship does not predict that CA1 will necessarily vary its inputs in linear proportion to changes in the external environment or other experimental manipulations. The Δinput to CA1 includes not only the entorhinal representations, but also the input from CA3. Thus, CA1 might show nonlinear changes in its place cell outputs as a function of experimental manipulations, but this might reflect nonlinear changes inherited from CA3, rather than nonlinear processing in CA1 itself. In other words, if one could measure precisely the changes in the CA1 inputs and the changes in the CA1 outputs, the input–output curve might show a linear change as in Fig 2C, even if the output of CA1 plotted relative to changes in the experimental manipulations is highly nonlinear.

3. Double rotation experiments

With these considerations in mind, we will now review a series of neurophysiological recording experiments from hippocampal afferent regions (the MEC, LEC, and anterior thalamus), intrahippocampal regions (DG and CA3), and the hippocampal output layer (CA1). In these experiments, rats ran clockwise on a circular track (Knierim, 2002) (Fig. 3). The track was divided into 4 quadrants, each with a distinct visual and tactile texture. The track was centered in a room with a circular, black curtain at the periphery. Along the curtain were 6 salient global cues. After many days of training to run laps on the track for irregularly placed food reward, the rats experienced over 4 days a sequence of alternating sessions in which the cues were arranged in the standard configuration (standard session) or in a mismatched configuration (mismatch session). The mismatches were generated by rotating the local cues on the track in a counterclockwise (CCW) direction and the cues along the curtains by an equal amount in a clockwise (CW) direction (double rotation). The net mismatch between the local and global reference frames was 45°, 90°, 135°, or 180°.

3.1. Superficial-layer MEC spatial representations and thalamic head direction representations are controlled by the global cues

The superficial layers of MEC provide the major spatial information to the hippocampus, in the form of grid cells, boundary cells, and head direction cells (Hafting, Fyhn, Molden, Moser, & Moser, 2005; Sargolini et al., 2006; Savelli, Yoganarasimha, & Knierim, 2008; Solstad, Boccara, Kropff, Moser, & Moser, 2008; Zhang...
et al., 2013). MEC cells are modulated by the theta rhythm, just like place cells (Brun et al., 2008; Hafting, Fyhn, Bonnevie, Moser, & Moser, 2008; Ranck, 1973; Stewart, Quirk, Barry, & Fox, 1992). MEC cells, CA1 place cells, and head direction cells appear to be tightly coupled to each other: When head direction cells rotate their preferred firing directions in an environment, both MEC cells and place cells rotate their spatial tuning profiles by equal amounts (Hargreaves, Yoganarasimha, & Knierim, 2007; Knierim, Kudrimoti, & McNaughton, 1995, 1998). Under most circumstances, head direction cells are controlled by the most peripheral landmarks in an environment (Taube, Muller, & Ranck, 1990; Zugaro, Berthoz, & Wiener, 2001). Thus, as predicted, head direction cells of the thalamus changed their preferred firing directions to follow the global cues in the double rotation experiment (Fig. 4). Whenever more than one head direction cell was recorded simultaneously (range 2–7 cells), all of the head direction cells rotated by the same amount. This result provided convincing evidence that the head direction system formed a tightly coupled network, consistent with ring attractor models of head direction cells (see below for further descriptions of ring attractors) (Skaggs, Knierim, Kudrimoti, & McNaughton, 1995; Zhang, 1996; Blair, 1996; Redish, Elga, & Touretzky, 1996; Song & Wang, 2005; Knierim & Zhang, 2012).

Because of the close anatomical association between areas with head direction cells and the MEC (Witter & Amaral, 2004), along with the presence of head direction cells in MEC (Sargolini et al., 2006) and the previously described coupling between head direction cells and MEC spatial firing (Hargreaves et al., 2007), one would predict that the MEC spatial representations would be controlled strongly by the global cues in the mismatch sessions. As predicted, the firing fields of MEC cells rotated in alignment with the global cues (Neunuebel, Yoganarasimha, Rao, & Knierim, 2013). Fig. 5A shows population correlation matrices between two standard sessions (left) and between a standard and mismatch session (right). Each pixel in the matrix represents the correlation between the population vectors representing the location of the rat (in degrees of angle along the track) in one session and the next. The band of high correlation (bright white stripe) along the main diagonal on the left matrix indicates that the MEC population representation of location was stable between the two standard sessions (i.e., the population vector was highly correlated between the same locations on the track in the two sessions and poorly correlated between different locations). During the mismatch session (right), the correlation band was maintained, but shifted above the main diagonal (red dashed line), indicating that the representations followed the rotation of the global cues in the mismatch session. We converted the 2-dimensional correlation matrices into 1-dimensional polar plots by calculating the average correlation of all pixels along the diagonals of the matrix (Fig. 5B). Fig. 5C shows a well-formed tuning curve for the STD1–STD2 correlation (gray) and a similarly well-formed tuning curve for the STD1–MISMATCH correlation (purple), rotated 45° clockwise to follow the global cue set.

The responses to all mismatch angles are represented as polar plots in Fig. 6A. As the cue-mismatch increased in magnitude, the population correlations between the standard and mismatch ses-

---

**Fig. 3.** Double rotation experiment. The inner circle denotes the behavioral track with different local textures on the 4 quadrants. The black, outer ring denotes the circular curtain at the perimeter, with 6 distinct landmarks that constitute the global cue set. Shown here is a typical sequence of recordings, with 3 standard sessions interleaved with 2 mismatch sessions (180° and 45°). In the mismatch sessions, the local cues are rotated CCW and the global cues are rotated CW by the same amount. From Neunuebel et al. (2013).

**Fig. 4.** Head direction cell responses to the double rotation. In these circular histograms, 0° is at the top of the plot. Each dot indicates the rotation angle of the directional tuning curve of a head direction cell between the standard session and a mismatch session. For example, if the preferred direction of a cell rotated 90° between sessions, a dot is added to the circle corresponding to 90° (i.e., at the 3 o’clock position)”L” indicates the rotation of the local cues and ”G” indicates the rotation of the global cues. Almost all head direction cells rotated by approximately the same amount as the global cues. Modified from Yoganarasimha, Yu, and Knierim (2006).
sessions decreased, indicating that the MEC representations continuously degraded with increasing local–global conflict. This decorrelation may reflect a combination of CW rotations of the underlying grid cells and other spatial cells (following the global cue set) as well as potential shifts of the grid in the horizontal plane, as the representations become decoupled from the external landmarks in the environment. In general, though, the MEC representations followed the global cues, with little evidence of a strong, local-cue driven component to the MEC population response.

3.2. Superficial-layer LEC representation is controlled by the local cues

In contrast to the MEC, grid cells and head direction cells appear to be absent in LEC (Hargreaves, Rao, Lee, & Knierim, 2005; Knierim, Neunuebel, & Deshmukh, 2013; Yoganarasimha, Rao, & Knierim, 2011), and LEC cells are much more weakly modulated by theta compared to MEC and hippocampal cells (Deshmukh, Yoganarasimha, Voicu, & Knierim, 2010). Superficial-layer LEC cells fire in the presence of discrete objects in an environment, and under these conditions a small minority show spatial tuning profiles that resemble the robust place fields of the hippocampus, even when the firing location is away from the objects (Deshmukh & Knierim, 2011). Thus, one might predict that LEC cells would be more strongly controlled by local cues on the track, rather than the global landmarks on the curtains. Consistent with this prediction, Neunuebel et al. (2013) demonstrated that, even though the LEC did not have a robust spatial representation of the track, there was a significant correlation between the LEC representations of the standard and mismatch sessions, and this representation was clearly controlled by the local cues (Fig. 6B). (Note that we only analyzed firing when the rat's head was on the track and the rat was moving forward at a minimal speed. We did not analyze neural activity when the rat was looking off the track, performing so-called “head scanning” behavior and potentially attending to individual cues on the curtains (Monaco, Rao, Roth, & Knierim, 2014)).

3.3. DG performs pattern separation on its EC inputs

Given that the MEC and LEC are the major inputs to the DG, we can begin to measure the \( \Delta \text{input–} \Delta \text{output} \) curve of the DG (Neunuebel & Knierim, 2014). As the cue mismatch increased, the MEC representation gradually degraded (\( \Delta \text{input} \) increased) and was controlled by the global cues (Fig. 6A). In contrast, the LEC representation was not strongly correlated between the standard and mismatch sessions for any cue-mismatch angle, but a weak, local-cue-controlled signal was detectable (Fig. 6B). The DG representation of the 45° mismatch environment was very similar to the standard environment, as the two representations were highly correlated (although less correlated than the standard ses-
Fig. 6. Responses of the MEC, LEC, DG, and CA3 populations to the double rotation. Each polar plot denotes the correlation between the population representations of two standard sessions (gray plots) or between a standard and a mismatch session (colored plots). The correlation value is plotted along the radial dimension and the angular dimension indicates the amount that the population vectors were rotated relative to each other. The dark tick marks at the outer ring indicate the rotation extent of the global cues and the light tick marks indicate the rotation extent of the local cues, as in Fig. 5C. High population correlations are indicated by well-formed tuning curves with larger peak correlations; low population correlations are indicated by poorly formed tuning curves with smaller peak correlations. See original reports for more details. (A) MEC population correlations degraded gracefully with increasing mismatch angles and were controlled by the global (G) cues. From Neunuebel et al. (2013). (B) LEC population correlations were poorly tuned in all standard-standard and standard-mismatch comparisons, but the peak correlations were controlled by the local (L) cues. From Neunuebel et al. (2013). (C) DG population correlations were well-formed for the 45° mismatch, but became almost completely decorrelated with the larger mismatch sessions. The population of DG cells with significant spatial information is plotted. Modified from Supplementary Data in Neunuebel and Knierim (2014). (D) CA3 population correlations were well-formed for all mismatch angles and were controlled by the local (L) cues. The population of CA3 cells with significant spatial information is plotted. Modified from Supplementary Data in Neunuebel and Knierim (2014). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.4. CA3 performs pattern completion/error correction/generalization on its DG/EC inputs

Lee, Yoganarasimha, Rao, and Knierim (2004) showed that the CA3 population maintained a relatively coherent representation of the standard environment in the mismatch sessions. As the mismatch angle increased, the correlation slowly degraded, similar to the MEC input. Whereas the graceful degradation in MEC may reflect an increasing shift of the spatial representations of MEC relative to the track (e.g., grid cell firing patterns may shift in the horizontal plane, causing some grid vertices to “fall off” the track and other to “climb on”; Neunuebel et al., 2013), the degradation in CA3 may result from a combination of such a shift as well as an increase in partial remapping of CA3 cells. Critically, in contrast...
to the MEC, the CA3 representation was controlled by the local cues, and Neunuebel and Knierim (2014) independently verified this finding (Fig. 6D). Thus, armed with the information about how the MEC, LEC, and DG inputs change with the increasing mismatch amount, we can begin to measure the Ainput–Aoutput curve for CA3. The MEC maintains a high degree of correlation between the standard and mismatch sessions, but the representation is controlled by the global cues (Fig. 6A). Thus, the MEC representation cannot explain the local-cue-controlled, coherent representations retrieved by the CA3 network. The LEC has weakly correlated representations between all standard and mismatch sessions (Fig. 6B), and the DG has weakly correlated representations between standard sessions and mismatch sessions >45° (Fig. 6C). Since the only cues that CA3 receives about the local cues appear to come from these weakly correlated inputs, this provides strong evidence for pattern completion in CA3 (i.e., a large Ainput results in a smaller Aoutput).

Importantly, in this experiment we do not appear to explore the parameter space in which the Ainput is large enough to detect the pattern separation functions of the CA3 attractor (i.e., the rightward extreme of the sigmoid curve of Fig. 2). However, experiments from other laboratories have shown evidence that the CA3 representations are completely independent under more extreme conditions that would presumably generate a larger Ainput from EC, (Guzowski et al., 2004; Leutgeb, Leutgeb, Treves, Moser, & Moser, 2004; Vazdarjanova & Guzowski, 2004). Leutgeb et al. (2004) recorded the activity of CA3 and CA1 place cells as rats explored similar enclosures in the same location or in completely different rooms. Under these conditions, the CA3 representations of the different rooms were completely orthogonal to each other, whereas the CA1 representations maintained some degree of similarity that presumably reflected the common aspects of the two rooms. Vazdarjanova and Guzowski (2004) imaged immediate early gene activity when changes to an environment were small (individual environmental cues were altered) or large (the rat was moved to a new room). With the small manipulations, the CA3 ensemble representation changed less than the CA1 representation; conversely, with the large change to a new environment, the CA3 representation changed more than the CA1 representation. Taken together with our double rotation data, these experiments suggest that CA3 displays the sigmoidal Ainput–Aoutput relationship predicted by theory (Guzowski et al., 2004; McClelland & Goddard, 1996; Rolls & Treves, 1998).

3.5. CA1 representations reflect a (linear?) combination of CA3 and EC inputs

Lee et al. (2004) demonstrated that CA1 populations created a split representation in the mismatch sessions, whereas CA3 simultaneously produced a more coherent representation controlled by the local cues. Some CA1 place fields rotated with the local cues and a similar number rotated with the global cues (some place fields even split in half, with one subfield rotating with the local cues and the other with the global cues). Like CA3, other CA1 cells showed a remapping response, either losing their firing fields or gaining a field.

Because the major input to CA1 is from CA3, the CA1 result seems counterintuitive. If the CA3 network goes through all the trouble of completing the input pattern of the altered environment, why does CA1 “ignore” this CA3 input and instead send the “corrupted” signal to the rest of the brain? It is important to realize that different parts of the transverse axis of CA1 receive input from different parts of the CA3 transverse axis as well as direct inputs from layer III of the EC. Proximal CA1 receives input from distal CA3 and from MEC, whereas distal CA1 receives input from proximal CA3 and LEC (Witter & Amaral, 2004; Witter, Wouterlood, Naber, & Van Haeften, 2000). The recordings from Lee et al. (2004) were primarily from proximal CA1 and from distal CA3. Thus, we can explain the split representation of CA1 by these anatomical projection patterns (Fig. 7). Proximal CA1 receives a local-cue-driven representation from distal CA3 and a global-cue-driven representation from MEC. Lacking a strong recurrent collateral system, it appears that CA1 performs a linear operation on its inputs: the inputs convey separate local and global representations, so the CA1 output forms a split representation that mimics the dual nature of its inputs. This split representation may reflect the comparator function often ascribed to CA1, that it compares the EC representation about the current state of the world with the mnemonic representations stored in CA3 of expected events predicted by the current inputs (Hasselmo, 2005; Levy, 1996; Lisman & Otamakhova, 2001; Vinogradova, 2001). Thus, the competition between pattern separation and pattern completion in the DG-CA3 processing loop retrieves from CA3 the most likely stored representation based on the current input, and this representation is then compared to the current input in CA1. Analogously, we would expect that distal CA1 would show responses that reflect a relatively linear combination of its inputs from proximal CA3 and LEC. Although such a comparator function was not explicitly demonstrated in our analyses, a finer-grained analysis (perhaps segregating EC-related firing from CA3-related firing based on variables such as theta phase or firing coherence at beta, low gamma, and high gamma frequency bands; Bieri, Bobbitt, & Colgin, 2014; Colgin et al., 2009; Hasselmo, Bodelon, & Wyble, 2002; Igarashi, Lu, Colgin, Moser, & Moser, 2014) might reveal intriguing neural dynamics between the CA1 local and global representations that could underlie such a comparison.

Fig. 8 summarizes the relationship between the LEC, MEC, and DG spatial input patterns to CA3 and the responses of these cell populations to the double rotation experiment. For clarity of presentation, the CA3 population activity on the circular track is depicted as a ring attractor, although the activity of place cells is more accurately modeled as a 2-dimensional sheet attractor (Knierim & Zhang, 2012; Samsonovich & McNaughton, 1997). The cyan circles represent cells with place fields at the corresponding locations on the track. The diameters of the circles indicate the current firing rate of each cell, with larger diameters indicating higher firing rates. The green lines indicate excitatory connections among cells with neighboring place fields and the dashed black lines indicate inhibitory connections to all other cells. Although the excitatory and inhibitory connections are shown only for cells at the top of the ring, all cells have this connectivity pattern, resulting in a continuous ring attractor in which every location on the track is a stable state. With appropriate weights, a single bump of activity will form on this ring even in the absence of external inputs. When external inputs are added to the network, the location of the activity bump can be set by these inputs. The colored rings below demonstrate the activity levels of cells in the LEC, MEC, and DG inputs to CA3 when the rat is located at the north location on the track (left column). It is important to note that these rings indicate the activity of cells that represent locations on the track; they do not represent the place fields on the track. A topography of projections is assumed, such that cells that represent a certain location on the track project to cells in CA3 that represent the same location. The orange LEC ring shows a noisy representation with only a weak bias for higher activity at the north representation in the standard session. The purple MEC ring shows a well-formed spatial representation at the current location of the rat, from grid cells and other spatially modulated cells. The red DG ring also shows a well-formed spatial representation of the current location of the rat from DG place cells. In the familiar, standard environment, these 3 representations provide a coherent input to CA3 that drives the CA3 attractor.
Fig. 7. Proximal CA1 (CA1p) receives input from the MEC that is controlled by the global cues. It also receives input from distal CA3 (CA3d) that is controlled primarily by local cues. The CA1 output back to the deep layers of MEC is an apparent linear split representation of the conflicting local–global inputs it receives from its afferent inputs. That is, CA1 receives conflicting reference frame signals from its inputs, and its outputs faithfully reflect this conflict. Adapted from Neunuebel et al. (2013).

Fig. 8. Putative attractor dynamics of CA3. CA3 population activity is represented schematically as a bump in a continuous ring attractor. Excitatory connections are shown in green and inhibitory connections in black (dashed lines). Inputs from DG, MEC, and LEC are indicated as input rings of activity. In the standard session (left column), the active inputs when the rat is at the north location of the track drive the CA3 activity bump at the north location of the ring. G1 and G2 refer to the locations of two global cues that control MEC activity, and L1 and L2 refer to the locations of two local cues that control LEC activity. When the rat is on the west location of the track (middle column), the corresponding cells controlled by the G2 and L2 cues cause the CA3 activity bump at the west location of the attractor ring. The right column shows the active cells during a mismatch session when the rat is back at the north location. For clarity of illustration, we show a 90° CW rotation of the global cues alone, which, in the absence of any other reference frame, is equivalent to the 90° mismatch of the double rotation. When the global cues are rotated 90° CW, the DG remaps and the LEC continues to fire in a weak spatial representation controlled by the local cues (L1). However, because the G2 global cues have been rotated to the north, the MEC cells that formerly encoded the west location defined by the G2 global cues are now active when the rat is at north. The CA3 attractor dynamics cause a relatively coherent activity bump to form at the local-cue-predicted location.
as they will fire only when the rat travels to the east location (where G1 is now visible). The MEC cells active at the initial G2 location (west in the standard session; middle column) are active instead, sending excitatory drive to the west location of the CA3 ring. Finally, the DG representation has almost entirely remapped (pattern separation), with a weak bias for the cells at the north location to fire. Other DG cells that encode random locations on the track are also active, indicating the global remapping that occurs in the DG.

Given these inputs, the attractor dynamics in CA3 determine how the CA3 place cells fire in the mismatch session. The data clearly indicate that CA3 is controlled by the local cues. Thus, the combined LEC and DG input apparently form the “seed” that biases the attractor bump (i.e., the active cells in Fig. 8) to fire in the locale-predicted location. Although the MEC provides a strong input for CA3 cells that represent the global cues (G2) to fire, the inhibition from the attractor bump appears to prevent strong activity at that location of the ring. (The CA3 data show that there are a small number of CA3 cells controlled by the global cues, suggesting that the strong external drive can overcome the attractor dynamics to some degree (Lee et al., 2004; Neunuebel & Knierim, 2014)). Other active cells outside the bump can generate alterations in the attractor dynamics, such as the formation of spurious attractor states (Knierim & Zhang, 2012), as the result of the conflict between the external inputs. Thus, given the altered input patterns of it inputs, we see that the CA3 attractor dynamics allow the input “errors” to be partially corrected and a single, coherent output representation to form. When the rat is at the north, the inputs to the north CA3 cells are severely weakened and degraded by the absence of the drive from MEC and DG. However, the ability of the network to form a strong activity bump at this location, given the weak inputs, can be considered a demonstration of pattern completion.

It is important to note that these diagrams illustrate only the sets of active cells that represent the standard environment. Not depicted are other DG and CA3 cells that were silent in the standard session but that became active in the mismatch session. In cases of “global remapping,” completely different ensembles of neurons (active sets) are activated as the result of large environmental changes (Kubie & Ranck, 1983). The different active sets correspond to activity on different “charts,” to use the terminology of Samsonovich and McNaughton (1997). Such global remapping in CA3, such as that produced by bringing the rat to a completely different environment (where G1 is now visible). The DG cells active at the north location to fire. Other DG cells that encode random locations on the track are also active, indicating the global remapping that occurs in the DG.

What occurs when the animal enters a completely different environment? The EC inputs are presumably very different in the new environment (where G1 is now visible), and the DG imposes a completely novel pattern of activity on the CA3 network, creating a new attractor basin in CA3 that is associated with the EC inputs. The interesting dynamics occur when the new environment is only partly different. If the change to the EC inputs is small, the EC input will place the activity pattern of CA3 within the basin of attraction of the initial pattern. The DG input pattern will be less similar to the initial pattern, due to the pattern separation, but if this does not drive the CA3 pattern far enough away from the attractor basin, then the system will retrieve a pattern similar to the initial pattern (pattern completion/generalization). As the EC input changes even further, the DG will continue to drive the CA3 pattern away from the original pattern, but the EC will drive it to fire within the basin. Eventually, the combined input of both EC and DG will cause CA3 to fire in a pattern outside the initial attractor, and this will then cause a new attractor basin to emerge. Under this scenario, the DG is always providing an input to drive CA3 away from a stored attractor basin (driving CA3 toward pattern separation). Whether CA3 actually performs pattern completion or pattern separation depends on the relative balance of EC inputs and DG inputs, as well as on the strength of the attractors stored within the CA3 network and other modulatory inputs that may alter the relative weights of the DG/EC inputs. These ideas might be explicitly tested with optogenetics or other techniques that could selectively enhance/suppress either the EC or DG inputs to CA3 and thereby causally tip the balance in favor of pattern separation or pattern completion in ambiguous environments.

4. Final comments and caveats

With this review we hope to have demonstrated how a series of experimental studies from our laboratory over the past decade have provided direct, physiological evidence in favor of classic models of hippocampal computation and its relationship to memory. We have taken the approach from the computational literature that the concepts of pattern separation and pattern completion can only be studied directly in terms of input–output transformations of neural representations (McClelland & Goddard, 1996; O’Reilly and McClelland, 1994; Santoro, 2013; Treves & Rolls, 1994). Thus, we have investigated the responses of 6 specific processing stages in the hippocampal circuit (anterior thalamus, LEC, MEC, DG, CA3, and CA1) to a manipulation designed to introduce graded changes to the hippocampal inputs. By measuring the correlations between the representations of the standard and cue-mismatch sessions, we provided evidence that conforms well to the models’ predictions that the DG performs pattern separation and that the CA3 contains attractor dynamics that can support pattern completion.

A number of caveats are in order, however. The data presented in this review capture the essence of the results, but the entire picture is much more complex, and the interested reader is directed to the original research papers to find further details. Moreover, the hippocampal anatomy is more complex than that described here. Although the MEC and LEC are the major inputs into the hippocampus, there are other inputs (such as from the septum, perihinal cortex, and brainstem) that we have not accounted for in our
that person. Reactivating neocortical representations via CA1 output to construct a full recollection of the event would constitute an episodic memory. However, upon greeting the familiar-looking person, one may realize that this is a complete stranger when they do not return the greeting. Under these conditions, the pattern separation mechanisms of the DG may impose a new attractor in the CA3 region to create a completely distinct representation of the stranger, such that in the future one does not repeat the error. Even in situations in which the memory retrieval was correct, one still would wish to remember the original episode in which the person was previously encountered as well as storing a new memory of the current meeting. Perhaps this dual function explains why the CA3 region displays a topographical organization along its transverse axis, in which the proximal part of CA3 is biased toward pattern separation (to form a new, distinct memory of the current episode) whereas the distal part of CA3 is simultaneously biased toward pattern completion (to retrieve the memory of the prior episode). Experiments utilizing high-resolution imaging with human subjects might be able to test these ideas and bridge the neurophysiological results from rodents with the neuropsychological literature on human episodic memory, with both approaches interpreted in the framework of memory processing from the computational literature.

Acknowledgments

We thank I. Lee, D. Yoganarasimha, F. Savelli, and G. Rao for help in data collection for some of the figures shown in this manuscript. The experiments reviewed here were supported by Public Health Service grants NS039456 and MH094146 and by the Johns Hopkins University Brain Sciences Institute. The funding agency had no role in the design, data collection, analysis, or writing of the paper.

References


