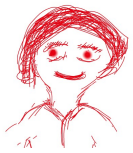
Switches in the brain?

A potential mechanism for long-term memory storage

Dilawar Singh

We forget often. But we remember some memories as long as we live. This means that our brain is capable of protecting memories for years. This is a remarkable feat given that the *biochemical hardware* involved in creating new memories is a hostile place for its storage. What are the challenges involved? And what type of biochemical mechanisms may overcome them? This article explores a major hypothesis that molecular switches may be behind our remarkable ability to remember for a lifetime.



Dilawar Singh is currently a graduate student at National Center for Biological Sciences (NCBS), Bengaluru. His hobby is to convince people to move to open-source softwares to live happily ever after.

1. Introduction

Our brain is made up of roughly 100 billion neurons, joined together with over 100 trillion connections called **synapse**s. Each neuron on average makes 1000 connections. It is now widely accepted that memories are created by changing these connections.

Let's label these synapses as $s_1, s_2, \ldots s_n$. A subset of these synapses participate in a specific memory formation. For example, my memory of being chased by a ferocious street dog named *Lalu* (lets call it M_{Lalu}) is represented by the set of synapses $M_{\text{Lalu}} = (s_{10}, s_{21}, s_{12}, \ldots, s_{331})$ i.e., these connections were changed during my troubling encounter with Lalu. I sometimes recall this memory whenever I see a similar looking dog.

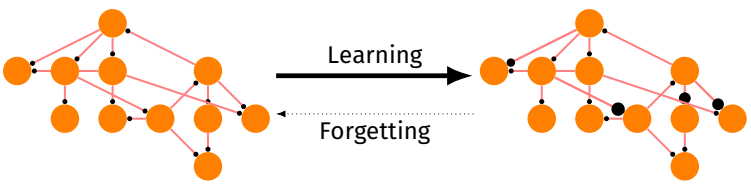
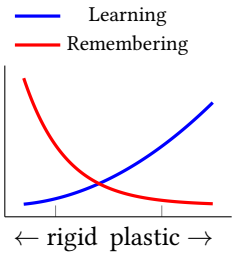


Figure 1. Memory formation and forgetting. During formation of a memory, some synapses become stronger. The longer you can maintain these connections, the longer you can hold on to this memory.

I can recall an experience as long as the set of synapses in which the particular experience was stored remains *intact*. Therefore, our ability to remember is contingent on our brain's ability to keep its connections intact. But on the other hand, our ability to learn depends on our brain's ability to change its connections. And here is the first challenge!

1.1 *Learning quickly v/s forgetting slowly, a zero-sum game*



Ability to change (plasticity) is good for learning, but bad for remembering.

For M_{Lalu} (or any other memory) to remain intact, each of its components ($s_{10}, s_{21}, s_{12} \ldots$) should also remain intact. The longer a synapse can keep itself unchanged, the better it will be at remembering the memory. Let's assume that somehow I create a synapse which maintains its state for a very long time (i.e., a rigid synapse). This synapse will not *forget* easily, but it causes another problem. Rigid synapse will not participate in any memory formation anymore since learning requires change, and a rigid syanpse can't change. It behaves like a read-only compact disk. On the other hand, if I create a synapse which is easily changeable (i.e., a plastic synaspe), it will be good at learning new experiences but won't be able to retain them for long. Plastic synapse forgets easily. We know that not only we remember for long time, we are capable of learning quickly too. For example, honey bees can learn the location of food after one encounter with flowers. Indeed, a good memory system is the one which learns quickly from new experiences and forgets old information as slowly as possible. Forgetting and remembering are two sides of the same coin. They are conflicting demands i.e., improving one will always deteriorate the other – a zero-sum game. The challenge is to strike a balance.

2. Hopfield network – associative memory network

Memory storage and retrieval is trivially done by a computer. It will be helpful to compare memory storage in the computer and the brain. In the computer, we always know the address of every stored memory and we access it by providing this address. The

file icon on your desktop is a graphical way of encoding this addressing scheme. This process is very similar to looking up the index page in a reference book to find a chapter. Our brain, on the other hand, is very unlikely to have such an indexing mechanism.

We recall when we are provided with *cues*. For example, when you see some part of a familiar person in a wedding album – while the rest of the person may be hidden behind others – you could easily identify the person. And many other memories of that person will also be recalled. A famous class of recurrent neural network popularly known as Hopfield network can do just the same as shown in *Figure 2*.

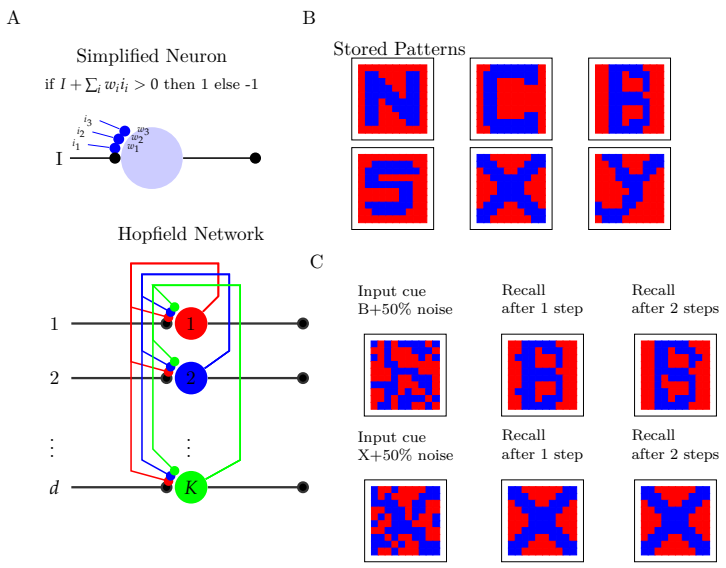


Figure 2. Hopfield network with 100 spiking neurons. These *recurrent* configurations give rise to interesting brain-like computation. **(B)** 6 patterns (memory) i.e. NCBSXY are stored in this network. **(C)** When a very distorted *cue* is applied to the network input, it *fetches* one of stored pattern which is the *closest* to the applied cue.

How does this recurrent network work is beyond the scope of article. Readers are encouraged to explore more by themselves. *"How well can we explain biological memory by these networks?"* is an active research area. Though these networks are extremely successful in accomplishing various *brain-like* computations (*a la* machine learning), I would like to advise the reader to be skeptical by noting the following:

- Neurons used in these networks are highly simplified. *Real* neu-

rons are not this simple. Even though these simplified neurons capture the essential *all-or-none* (electrical spike) way of communication and learning by changing synaptic connections, they do ignore rich local computations which can be accomplished by branches of these neurons called *dendrites*.

- There is no evidence that neurons make such dense recurrent connections. However some studies have shown that Hopfield network can work with very sparse recurrent connections as well.

- Activity in these networks does not match usually observed activity in the primate brain during memory-recall experiments.

Solutions contributed by other disciplines are helpful for comparison and contrast and often provide very useful insight. But in the end, these solutions must be tested under the constraints imposed by biology.

Nonetheless, these networks provide us with a framework to concretely think about the problem of memory storage and recall. We learn a great deal about these problems by pointing out the limitations and failures of these models. Hopfield network has properties which will sound very natural to us. Can you store as many memories as you like in these networks? No. There is an upper limit. Adding more patterns over maximum limit causes distortion in memories. When a cue is given, the network no longer fetches the right pattern. It often fetches a pattern which was not even stored; the retrieved pattern instead resemble some mixture of stored patterns. When too many memories are stored, they corrupt each other by mixing up. One can ask more questions. When connections decay in these networks and memories start disappearing, which memory disappear first: the weakest or the newest? And, when a new memory is added, what happens to the old memories?

After this necessary detour, lets go back to the main theme: how do synapses maintain their state?

3. How does a synapse maintain its state?

Very complex biochemistry plays out during learning that changes the synapse. Surprisingly, the net effect of this complex biochemistry can be summarised by a simple mathematical expression.

Ah, *the unreasonable effectiveness of mathematics*[4]! Let's assume that synaptic strength w is tightly correlated with a chemical species X found at synapse i.e. w changes with X. The problem of maintenance of w can be rephrased as the problem of maintenance of the activity of X. Therefore, the problem of "*synapse maintaining its state*" becomes the problem of "*molecule X maintaining its state*" – a more concretely defined problem.

Let's assume that X is converted to its active form X^* by adding a phosphoryl group (PO_3^{2-}). The phosphoryl group is removed by a phosphatase and X^* is turned back into inactive X. The phosphorylation and its counterpart dephosphorylation are a very common motif for controlling various chemical reactions by activating and inactivating protein molecules. Once most if not all X has been turned into X^* during memory formation, how do we make sure that X^* does not turn back into X (lose memory)?

$$X \longrightarrow X^*$$
$$X^* + P \rightleftharpoons X^*P \longrightarrow X + P$$

Let's mull over a solution to this problem of long term maintenance of X^*. Here is one potential solution.

1. **(Amplification)** X^* **auto-phosphorylates** itself i.e. $X \xrightarrow{\curvearrowright} X^*$. If we manage to get sufficient X^* somehow, it will act as a catalyst to its own production. X^* will always remain high.

2. Dephosphorylation of X^* is minimized by controlling the number of P or reducing the reaction rate.

Both (1) and (2) help in making X^* highly stable. Problem solved? No. Now we have constructed a rigid synapse. Recall the *rigid* v/s *plastic* synpase dilemma discussed previously (section **1.1**). This synapse will definitely remember for longer but it will not participate in any new learning anymore.

Can you think of other set of hypotheses? It must conform to laws of chemistry!

As long as we are in the realm of theory, let's propose a solution to this problem. We add another reaction say $P' + X^* \rightarrow P'X \rightarrow P' + X$ which deactivates X^* when the *need* arises. Phosphatase P' is different from P. This adds another layer of control to an already complicated problem i.e. forgetting is now controlled by another process. This requires one more explanation: how does this new

mechanism controlling *forgetting* work? And philosophically – if you care about it – it violates the principle of **parsimony** which recommends to pick the simplest explanation.

We still have two big problems hiding underneath. We have not considered the underlying biological hardware i.e. synapse in any detail where this biochemical network suppose to function. The first problem is chemical noise . For a biochemical system operating in very small volumes, effect of chemical noise can be very strong. There are over 200 types of protein molecules in a typical synapse. Indeed, most of these protein molecules have few tens of molecules. The brain is always active and the chemical noise caused by the brain's background activity will surely turn some molecules of X into X^*. Then due to auto-phosphorylation, sooner than later, all of the X will be turned into X^*. We have created a very stable memory of nothing but background noise. This is highly undesirable!

The volume of a typical synapse is 1×10^{-20} m^3. At this volume, 1 μM concentration is roughly equal to 6 molecules.

The second problem is *turnover* i.e. old molecules are continuously degraded and being replaced by newly minted molecules. Let's assume that at the time of memory formation, we had 100 molecules of X^* in the synapse. And let's also assume that on average, every day one new formed molecule (i.e. X) replaces an old one (X^*). After 50 days, half of the synaptic strength is gone! To counter this, we must have a *refresh* mechanism by which the newly added molecule quickly changes its state according to the state of synapse i.e. newborn X becomes X^* if most molecules at synapses are X^*.

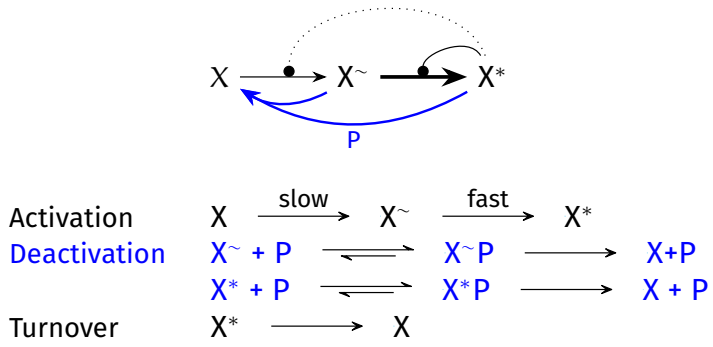
Effectively, we want a stable `ACTIVE` state (all X are X^*) which is immune to turnover. We also don't want chemical noise to turn X into X^*. We want a switch like behaviour i.e. if it is *OFF* or `ON` it tends to stay *OFF* or `ON`. If a few X are turned into X^* by background noise, we want them to be quickly turned back into X by the phosphatase P. And if during memory formation, a significant portion of X has been turned into X^* then we want that any X^* deactivated into X is quickly activated again into X^*. This system should operate like a switch which does not flip unless significant force is applied. These are called **bistable switch**es

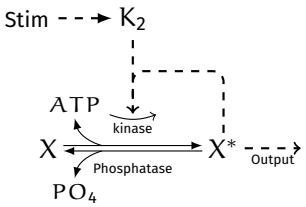
(See Box 1 for overview of bistability).

Is there any evidence that bistable chemical reaction networks exist? Do they occur at all in living cells? Bistability (and its close relative oscillations) are very common in biology (Why? What is the reason for their *selection* by evolution?)); from cellular level to population levels. So it won't be surprising if we discover bistable switches operating at syanpse as well. Indeed, various studies have shown that calcium/calmodulin dependent protein Kinase II (CaMKII) *may* form a bistable switch in synapse.

4. Molecular bistable switch at synapse

John Lisman hypothesised that a kinase and a phosphatase together can form a bistable switch immune to *turnover*. CaMKII and its phosphatase Protein phosphatase 1 (PP1) were identified as the hypothesised kinase and phosphatase. This chemical system has been extensively studied using computational models for over two decades [3]. There is some evidence that CaMKII is bistable *in vitro* conditions. Whether CaMKII is bistable inside a living neuron is still an open question (see Box 2 for overview of CaMKII properties at synapse.).





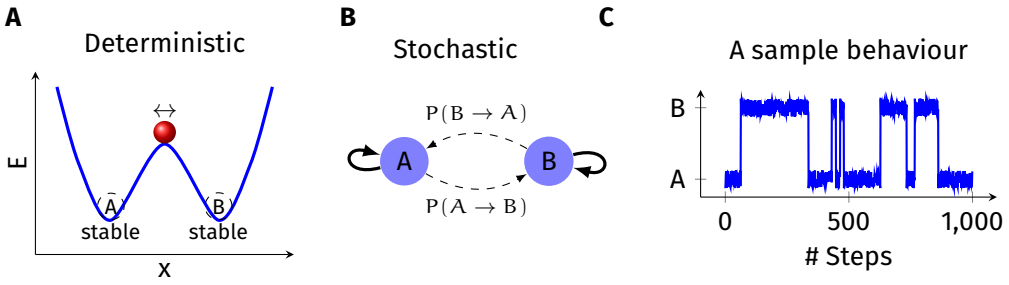
Reaction in a bistable switch proposed by John Lisman. Modified from [1].

Figure 3. A hypothetical network which can solve the problem of chemical noise and turnover with suitable parameters (how?). The activation step is divided into slow and fast components such that fluctuations caused by background noise do not cause the system to activate itself. X^* also partially activate X to $X^{\sim}$ to overcome *turnover.*

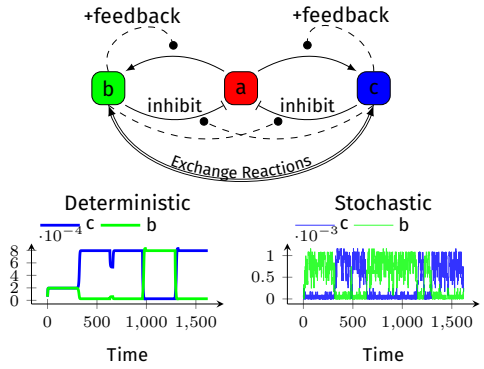
Box 1. Bistability in reaction network

In a world full of fluctuations, stability is indeed very useful property. Life is remarkably stable even though it is made up of inherently noisy components. Bistable chemical networks are ubiquitous in biology. In bistable chemical networks, noisy components acts together to give rise to highly (bi)stable behaviour.

A bistable system – as its name implies – has two stable states. From the point of view of an experimentalist, if you obverse a bistable system for very long time, you would almost always find it in either of its two stable states. Just like an electrical switch which you would almost always find in either `ON` or `OFF` state (and rarely in transition between `ON` and `OFF`).



In the figure above, **A** show a conceptual model of bistability. The energy landscape of a bistable system has two minima ($\widehat{A}$ and $\widehat{B}$), therefore, the system always ends up in either of them. It is a deterministic bistable system. **B** shows a conceptual representation of stochastic bistable system as state transition diagram. The arrows depict the probability of transition from one state to the other. In **C**, we simulate the system described in **B** with probability of transition from A to B ($P(A \rightarrow B)$), and from B to A ($P(B \rightarrow A)$) set equal to 0.01. Such bistable motif are commonplace in biology [6], especially found in situations where cell has to make a decision. One such motif is show below in both deterministic and stochastic settings. There are three necessary conditions a network must have to be able to show bistability: positive feedback, mechanism to filter out small stimuli, and a mechanism to prevent explosions[5].

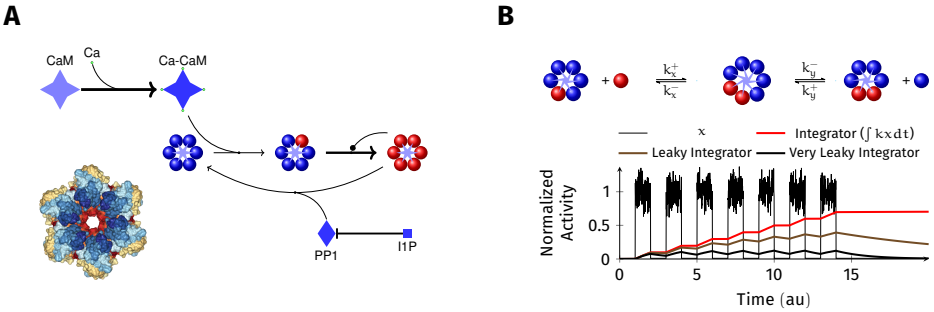


A bistable chemical system adapted from [6]. Species b and c catalyze their production (positive feedback) and inhibits production of a. **(below)** Two trajectories are shown in different settings. (Left) System is deterministic; (right) System is stochastic.

Box 2. CaMKII at synapse: A brief overview of its properties and function

Among more than 2000 proteins found in brain, CaMKII constitutes roughly 2% of them all. Furthermore, it is enriched in Hippocampus – a brain structure essential for memory formation. Indeed, CaMKII is known to play essential role in memory formation. In experiments involving mice, deactivating CaMKII in any way has always resulted in impairment of memory formation and learning.

CaMKII molecule has interesting properties which makes it an attractive candidate for storing information. 12 or 14 subunits of CaMKII make up one holoenzyme, usually arranged in dodecameric (top view) or tetradecameric form (top view). CaMKII structure is shown in (**A**) inset below (from Protein Data Bank (https://www.rcbs.org)).



(A) Summary of CaMKII pathway of. Upon its influx into synapse, calcium (Ca^{++}) binds to calmodulin (CaM) and create a complex calcium/calmodulin complex (Ca^{++}/CaM). Ca^{++}/CaM binds to CaMKII and phosphorylate it. CaMKII is dephosphorylated by PP1. **(B)** Subunit exchange between two holoenzymes i.e. a fully active CaMKII holoenzyme can loose an *ACTIVE* subunit which can be picked up by another holoenzyme which becomes partially active.

Note that first step is very slow for it requires binding of two Ca^{++}/CaM simultaneously. Once a subunit has been activated, phosphorylation of its clockwise neighbour requires binding of only one Ca^{++}/CaM and therefore further phosphorylation proceeds at much faster rate. Note that the first very slow step of CaMKII phosphorylation can be overcome by subunit exchange when an inactive holoenzyme picks up an active subunit released by other holoenzyme. Now this holoenzyme requires binding of only one Ca^{++}/CaM for further phosphorylation. Therefore subunit exchange helps in spreading CaMKII activation.

CaMKII integrates Ca^{++} signal A single CaMKII holoenzyme acts as a leaky integrator of Ca^{++} activity i.e. it sums up Ca^{++} activity in time and also decays with a time-constant (leaky). See **(B)** to compare behaviour of three integrators of calcium activity (x): a non-leaky integrator and two leaky integrators with small and large amount of leakage. Mathematically, it is similar to a leaky capacitor if we model Ca^{++} activity as applied current and CaMKII activity as voltage across capacitor. Integrators are useful when you want accumulate *enough* information about x before taking a decision e.g. a plant can *decide* to flower or shed leaves only if integration of moisture in air and/or temperature during the day crosses a threshold value.

In our computational study of this pathway, we explored the ef-

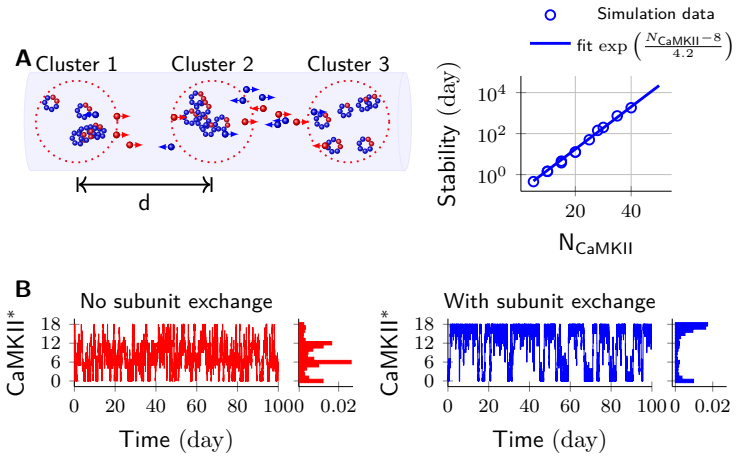


Figure 4. **(A)** (left) Three clusters of CaMKII in a synapse . (Right) This system ability to hold memories increases exponentially with system size. 52 holoenzymes can keep the memory intact roughly for 100 years! **(B)** Subunit exchange improves CaMKII ability to store information. **(Red)** 3 individual bistable switches independently without subunit exchange. **(Blue)** When subunit exchange is enabled, these independent switches synchronize. Notice that the now system spends more time in its stable states.

fect of subunit exchange on CaMKII pathway [7]. We found that the subunit exchange improves information retention capacity of CaMKII. This add weightage to the idea that CaMKII is *the memory melecule*. Individual CaMKII acts as leaky integrator of Ca^{++} activity (see Box 2.). With subunit exchange, the leaky integrator becomes less leaky and hold information for longer time. When many CaMKII holoenzymes come together and form a cluster then they give rise to a bistable switch (*Figure 4*). Such clusters have been observed in experiments. We found subunit exchange synchronizes bistable activity of distributed clusters. That is multiple bistable switches act as a single but much more stable bistable switch. To summarise, we found the subunit-exchange makes CaMKII molecule better at retaining information.

5. Conclusion

In this article, we have discussed why bistable motif is an attractive candidate for storing biological memories. Because bistable switches maintain their states for long time (remembering). Most support for this idea has came from computational studies. To really prove it, we need experimental data supporting this hypothesis. There is now a growing experimental evidence that synapse change in *all-or-none* manner, a finding which is consistent with this idea. Some other studies claim that changes are graded i.e.

synapse changes in step-wise manner much like a **multi-stable** synapse. A multi-stable synapse is an ensemble of many bistable components. Whether CaMKII is bistable in synapse (or in some part of it) is still an open question. So far there is no concrete evidence that it is. There could be other still unknown mechanisms which can give rise to bistability. Given that bistable chemical motifs are widespread, it is reasonable to believe that there are indeed switches in our brain – much like flip-flops in your pen-drives and memory cards – which are evolved to keep our memories safe from the onslaught of time and noise.

Acknowledgements

I'd like to thank Somya Mani for helpful suggestions on the manuscript.

Address for Correspondence
Bhalla Lab, National Center for Biological Sciences, Bengaluru
GKVK Campus, Bellary Road
Bengaluru - 560065.
Email: dilawars@ncbs.res.in

Suggested Reading

[1] **Lisman J. E., *A mechanism for memory storage insensitive to molecular turnover: a bistable autophosphorylating kinase*. Proc. Natl. Acad. Sci. USA, May 1985**

[2] **Christof Koch *Biophysics of computations*. Oxford University Press, 1999**

[3] **Malin Sandstorm, *Models of CaMKII activation*, Master Thesis, Royal Institute Of Technology Sweden**

[4] **Eugene Wigner, *The Unreasonable Effectiveness of Mathematics in the Natural Sciences*, Communications in Pure and Applied Mathematics, vol. 13, No. I (February 1960)**

[5] **Thoman Wilhelm, *The smallest chemical reaction system with bistability*, BMC Systems Biology, vol. 3, No 1 (Sep 2009)**

[6] **Naren Ramakrishnan, Upinder S. Bhalla *Memory Switches in Chemical Reaction Space*, PLOS Computational Biology, No 7 vol. 4 (July 2008)**

[7] **Dilawar Singh, Upinder S Bhalla, *Subunit exchange enhances information retention by CaMKII in dendritic spines*, biorxiv, doi:10.1101/372748 (2018)**