A glimpse on the designing principle of hierarchical recruitment chains

Chu Chen

Backgrounds

A common scenario in signal transductions is the assembly of protein machineries which are capable of producing/activating a large number of cytosolic soluble ligands/proteins to amplify the signal. The architecture of these machineries consists of platforms/scaffolds, adaptors and effectors. In the case of the spindle assembly checkpoint (SAC) signaling complex, through which mitotic cells sense the proper attachment of microtubules onto kinetochores, the platform is KNL1, the adaptor is the BUB1-BUB3 heterodimer and the effector is the MAD1-MAD2 heterotetramer (London and Biggins, 2014). If microtubules are not properly attached to the kinetochore, the MELT repeats of KNL1 will be phosphorylated by the kinetochore-anchored kinase MPS1. Then the BUB1-BUB3 complex will recognize the phosphorylated MELT repeats, bind to KNL1 and then be phosphorylated by the kinetochore-anchored MPS1. This further recruits the MAD1-MAD2 complex, which facilitates generation of closed MAD2 through a disputable mechanism. Besides the SAC signaling complex, there are other hierarchical recruitment chains that feature the platform-adaptor-effector architecture, like the ones of inflammasomes (Lamkanfi and Dixit, 2012) and the death-inducing signaling complex/DISC (Ashkenazi and Salvesen, 2014).

If we look at the functions of individual components of these recruitment chains, the platform/ scaffold might help to increase the efficiency of interaction and provide regulation to signal transductions (Good *et al*., 2011), and the effector is essential to generate outputs. Our current understandings about the functions of adaptors seem to be limited to offering binding sites in a flexible manner and additional possibility for regulations (Schechtman and Mochly-Rosen, 2001; Samelson, 2002). However, for some well-characterized hierarchical recruitment chains like the one of SAC signaling complex, binding flexibility does not seem to be the case since the downstream effector is just the MAD1-MAD2 complex. Of course, we cannot rule out the possibility that there are still some unknown regulations at the level of the BUB1-BUB3 adaptor. But even in this case, any recombination event that fuse the adaptor and effector or any mutations that allow them to form a stable complex before being recruited to the platform (concerning that both BUB1-BUB3 and MAD1-MAD2 are stable complex in the cytosol) would not necessarily affect the organism. However, we see that the 3-level hierarchical recruitment architecture of SAC signaling complex is conserved from yeast to human.

Here it is proposed that the existence of adaptor layer(s) might be critical in decrease the sensitivity of the recruitment chain to nonspecific (not signaling-driven) phosphorylation/ dephosphorylation, thereby making the switching from assembly to disassembly more robust.

The toy model

Assume that there are a total number of N levels in the recruitment chain and each level could be occupied by one protein molecule at most. Level 1 is the platform/scaffold and the reference frame in this model. Once activated*, level-n protein could recruit level-(n + 1) protein quickly; oppositely, level-(n + 1) protein and all higher-level proteins would also quickly fall off from the chain† (if not specified, a "protein" indicates either a protein or a multi-protein complex in the following context). Only when level N, the effector, is recruited could output be generated. In addition, phosphorylation is precisely regulated and could only take place after the protein is recruited to the chain, but phosphatases are soluble in the cytosol and dephosphorylates recruited proteins at a low velocity as well as all unrecruited proteins. This implies that any newly recruited molecule is initially in its unphosphorylated form. To simplify the model, phosphorylation/dephosphorylation kinetics of all of the proteins are set to be identical [or substrate-independent in the case of SAC signaling complex, as the platform and the adaptor are both phosphorylated by MPS1 and dephosphorylated by PP1 (London *et al*., 2012)]. Molecular interactions are facilitated by allosteric conformational changes induced by phosphorylation. This means that molecular interactions do not mask phosphorylation sites so that phosphorylation/dephosphorylation kinetics stay the same regardless of binding/unbinding of the next level in the recruitment chain.

Under these assumptions, phosphorylation and dephosphorylation are not in equilibrium, given that the newly recruited proteins are always unphosphorylated initially and that unbinding is fast (see supplementary materials for verification). Therefore, the master equations which describe the transitioning of the highest level occupied are considered. We denote the probability to observe level n ($0 \le n \le N$) as the highest level occupied at time t as $P_n(t)$. The time derivative of P_n ($P_0 \equiv 0$) is:

$$\dot{P}_n = \begin{cases} -[(n-1)k_- + k_+]P_n + k_+ P_{n-1} + k_- \sum_{i=n+1}^{N} P_i = -[(n-1)k_- + k_+]P_n + k_+ P_{n-1} + k_-(1 - \sum_{i=1}^{n} P_i), 1 \le n < N \\ -(N-1)k_- P_N + k_+ P_{N-1}, n = N \end{cases}$$

* In the context of SAC and some other signaling complexes, activation means phosphorylation and deactivation means dephosphorylation. Because of this, we will not distinguish the terminologies in the following context. In the case of SAC, all level shares a common kinase - the MPS1 kinase. In some other signaling complexes (like the yeast yeast mating response MAPK cascade), it is the level n that activates level (n + 1). This mechanistic difference does not affect the model here.

† This is not true in the case of the yeast yeast mating response MAPK cascade, where proteins all bind to the level 1 scaffold.

The (n - 1) coefficient before k_- is due to the fact that for a n-level recruitment chain, there are (n - 1) pairs of interactions, any one of which might be torn apart due to the allosteric effect of dephosphorylation‡.

At long run, the initial state of the system at $t = 0$ would have little influence on the current state. Therefore,

$$\dot{P}_n = 0, \forall n : 1 \le n \le N$$

It is easy to prove by mathematical induction that:

$$P_n = \begin{cases} nk_-k_+^{n-1}\prod_{i=1}^{n}(ik_- + k_+)^{-1}, 1 \le n < N \\ k_+^{N-1}\prod_{i=1}^{N-1}(ik_- + k_+)^{-1}, n = N \end{cases} \quad (※)$$

To verify this theoretical result, discrete-event simulation is performed using MATLAB. In the case of SAC signaling complex, a previous biochemical study reported the k_{cat}/K_m of mammalian MPS1 as 3.0×10^4 $(M \cdot s)^{-1}$ at pH 7.4 and room temperature, via an *in vitro* assay using Mps1-NTD as the substrate (Sun *et al*., 2010). Another enzymologic study reported the k_{cat}/K_m of PP1γ (the catalytic domain of PP1) as (1850 ± 500) $(M \cdot s)^{-1}$ at pH 6 ~ 7 and 25 °C, via an *in vitro* assay using 4-nitrophenyl phosphate as the substrate (McWhirter *et al*., 2008). These measurements serve as rough estimations about the relative catalytic velocities of these two enzymes, since the substrates and temperature (37 °C, physiological) in this model are different from the ones in the literatures. In this simulation, a ten fold difference between the velocities of phosphorylation ($k_+ = 0.1$) and dephosphorylation ($k_- = 0.01$) is assumed, as well as fast binding/unbinding kinetics (set as $k = 1$ for both)§.

In the case of SAC signaling complex, once proper attachment of microtubules are secured, the recruitment chain is shut down by physically blocking the access of MPS1 to KNL1

---

‡ The dephosphorylation of protein m induces a conformational change to itself, after which protein (m + 1) (and all levels above) will be dissociated from the chain. The timing of dephosphorylation (t) for a certain phosphorylated protein follows the exponential distribution: $t \sim \text{Exp}(k_-)$. Therefore, for a short time period Δt, the probability that dephosphorylation does not take place is $p = 1 - [1 - \text{Exp}(-k_-\Delta t)] = \text{Exp}(-k_-\Delta t)$. Since dephosphorylation events are mutually independent, the probability that dephosphorylation does not take place at any level from 1 to (n - 1) is $p^{n-1} = \text{Exp}[-(n-1)k_-\Delta t] \approx 1 - (n-1)k_-\Delta t$. Thus, the probability that a dephosphorylation event takes place in Δt is $(n-1)k_-\Delta t$.

§ Incorporating correct absolute values into the model will be ideal but not a necessity. Any one of these kinetics parameters defines the timescale and as long as the simulation is run for a long enough time and the ratios between velocities are consistent with reality, the time fractions for each level being the highest occupied level would not be different. This is also revealed by (※), which could be written into a function of k_-/k_+.

(Aravamudhan *et al*., 2015). As in the model here, this corresponds to an increase in K_m (or a decrease in $k_+ = k_{cat}/K_m$) of phosphorylation. There is no direct measurement on this off-state kinetics parameter so far, so different extents of decrease ($k_+' = 0.01$ or 0.001) are tried out in the simulation, which produce qualitatively similar results.

Simulation results

A reasonable range of numbers ($2 \leq N \leq 7$; 2 means no adaptor) is tested for total number of levels in the hierarchical recruitment chain and the time fraction for each level n ($1 \leq n \leq N$) being the highest level occupied is calculated in each case. Since the simulation time is long enough, the time fraction should be approximately the same as the probability given by (※). The results are plotted in Fig 1 and supplementary figure 1.

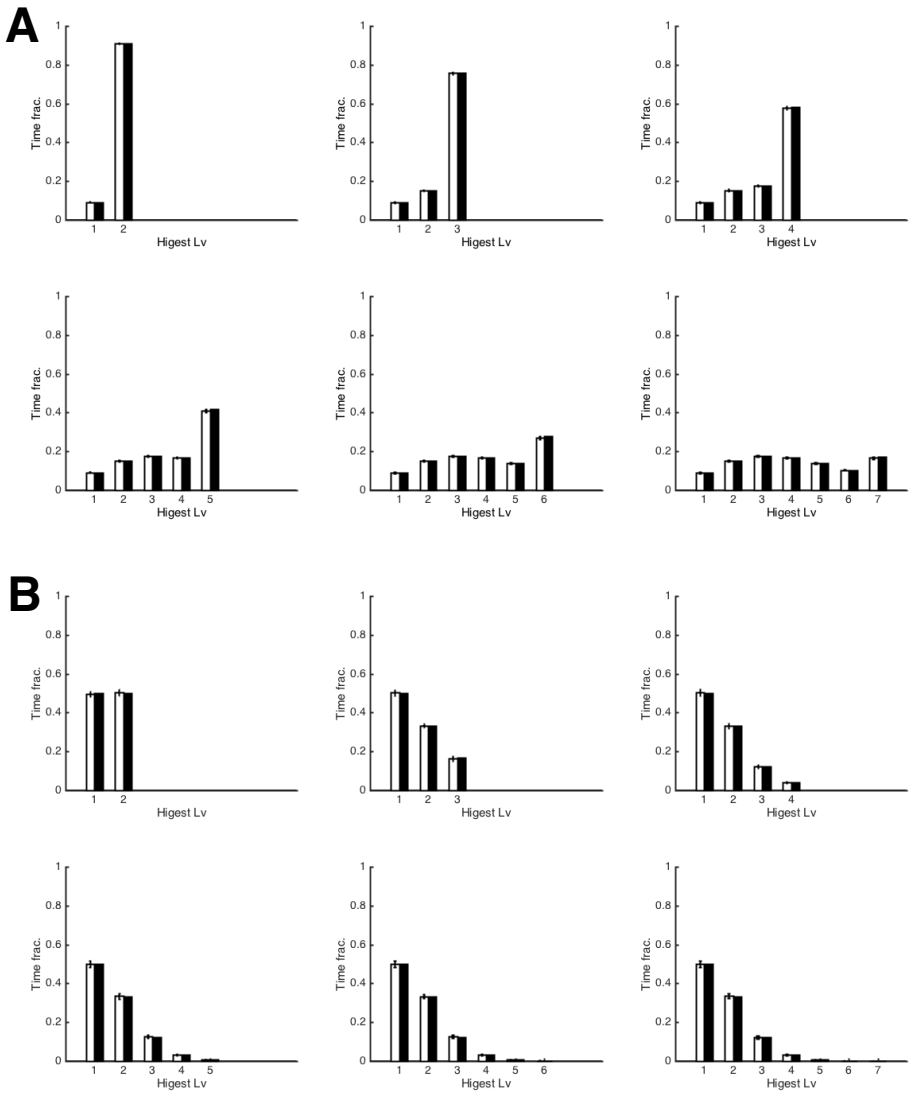


Fig 1. Time fractions for each level being the highest level occupied. Total number of levels in the chain: 2 (top left panel) - 7 (bottom right panel). (A) on-state: $k_+ = 0.1$, $k_- = 0.01$. (B) off-state: $k_+ = 0.01$, $k_- = 0.01$. White tiles represent average values from 100 simulations, with error bars showing standard deviations. Black tiles represent probabilities calculated from (※).

If we plot the time fraction of full recruitment versus N (Fig 2), we can see that as N increases, the probability of observing full recruitment drops down both before and after proper attachment of microtubules. If a switch-like behavior is required, the total number of levels should not be either too large or too small. If it is too large, the random dephosphorylation will make it hard to achieve full assembly at on-state. If it is too small (like $N = 2$ which corresponds to no adaptor layer), then the random phosphorylation will make it hard to avoid full assembly.

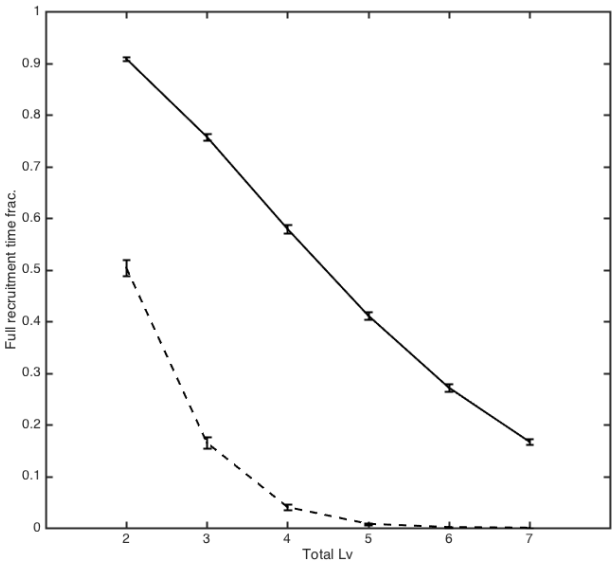


Fig 2. Probability of full assembly goes down as the total number of levels increases. Solid line: $k_+ = 0.1$, $k_- = 0.01$ (on-state). dashed line: $k_+ = 0.01$, $k_- = 0.01$ (off-state). Showing average values from 100 simulations ± standard deviations.

Discussions

There is another common feature of the recruitment chains mentioned previously. It has been shown that the numbers of MELT repeats in KNL1 of different species are different, which can be as many as 35-40 in some species (Tromer et al., 2015). Inflammasome has a wheel structure with a C_7-symmetry, while the DISC has a trimeric architecture. It is well-known that oligomerization provides the possibility for cooperativity (Segel, 1991) and that cooperativity

might shape input-output curves in signal transductions. The model discussed here is is robust to noise in nonspecific (not signaling-driven) binding and unbinding, opposite to the ultrasensitivity of MAPK signaling. What would be the overall outcome if multiple recruitment sites and cooperativity are taken into account is a worthwhile issue for further investigation.

References

London, N., and Biggins, S. "Signalling dynamics in the spindle checkpoint response." *Nature Reviews Molecular Cell Biology* 15.11 (2014): 736-748.

London, N., *et al*. "Phosphoregulation of Spc105 by Mps1 and PP1 regulates Bub1 localization to kinetochores." *Current Biology* 22.10 (2012): 900-906.

Sun, T., *et al.* "Cellular abundance of Mps1 and the role of its carboxyl terminal tail in substrate recruitment." *Journal of Biological Chemistry* 285.49 (2010): 38730-38739.

McWhirter, C., *et al*. "Mechanistic study of protein phosphatase-1 (PP1), a catalytically promiscuous enzyme." *Journal of the American Chemical Society* 130.41 (2008): 13673-13682.

Aravamudhan, P., *et al*. "The kinetochore encodes a mechanical switch to disrupt spindle assembly checkpoint signalling." *Nature Cell Biology* 17.7 (2015): 868-879.

Lamkanfi, M., and Dixit, V.M. "Inflammasomes and their roles in health and disease." *Annual Review of Cell and Developmental Biology* 28 (2012): 137-161.

Ashkenazi, A., and Salvesen, G. "Regulated cell death: signaling and mechanisms." *Annual Review of Cell and Developmental Biology* 30 (2014): 337-356.

Segel, L.A. *Biological kinetics*. Cambridge: Cambridge University Press, 1991.

Good, M.C., *et al*. "Scaffold proteins: hubs for controlling the flow of cellular information." *Science* 332.6030 (2011): 680-686.

Schechtman, D., and Mochly-Rosen, D. "Adaptor proteins in protein kinase C-mediated signal transduction." *Oncogene* 20.44 (2001).

Samelson, L.E. "Signal Transduction Mediated by the T Cell Antigen Receptor: The Role of Adapter Proteins." *Annual Review of Immunology* 20.1 (2002): 371-394.

Tromer, E., *et al*. "Widespread recurrent patterns of rapid repeat evolution in the kinetochore scaffold KNL1." *Genome Biology and Evolution* 7.8 (2015): 2383–2393.

Supplementary materials for
A glimpse on the designing principle of hierarchical recruitment chains

Chu Chen

We denote the event "level n is occupied by the corresponding protein" as A_n, and the probability (or the time fraction in the long run) to observe the phosphorylated form of the protein at level i as r_i. Since phosphorylation is the switch of fast binding and unbinding, this leads to the following recurrence relations:

$$P(A_n) = \begin{cases} 1, n = 1 \\ P(A_n, A_{n-1}) = P(A_n \mid A_{n-1})P(A_{n-1}) = r_{n-1}P(A_{n-1}), n > 1 \end{cases}$$

It is easy to derive:

$$P_n = \begin{cases} P(A_1) - P(A_2) = 1 - r_1, n = 1 \\ P(A_n) - P(A_{n+1}) = (1 - r_n)\prod_{i=1}^{n-1} r_i, 1 < n < N \ (*) \\ P(A_N) = \prod_{i=1}^{N-1} r_i, n = N \end{cases}$$

If we assume that phosphorylation and dephosphorylation are in equilibrium, the probabilities to observe the phosphorylated/unphosphorylated forms of a protein would satisfy:

$$k_- P^{\text{eq}}_{\text{phosphorylated}} = k_+ P^{\text{eq}}_{\text{unphosphorylated}}$$

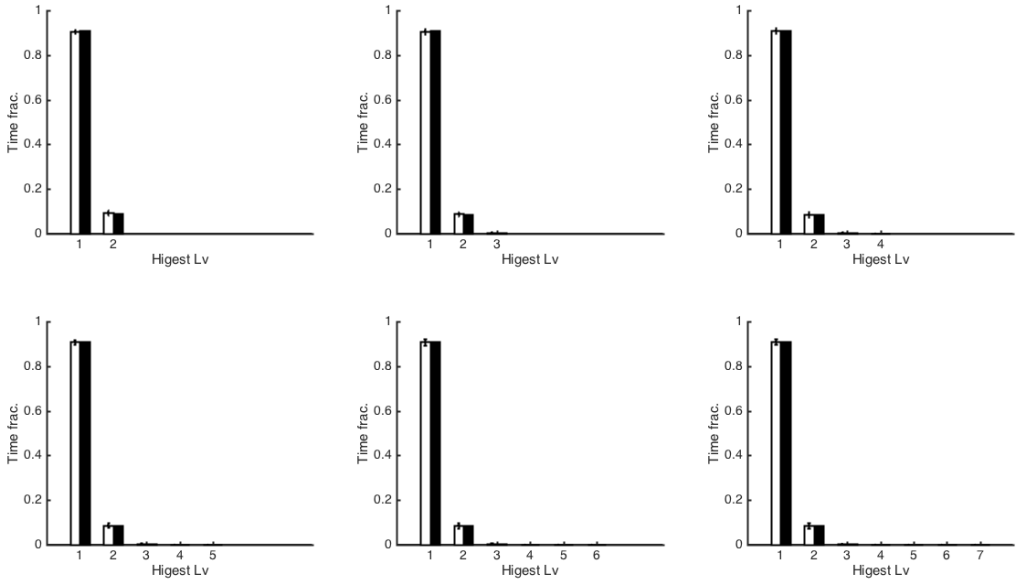
Thus, we have:

$$r_i = P^{\text{eq}}_{\text{phosphorylated}} = \frac{k_+}{k_- + k_+}$$

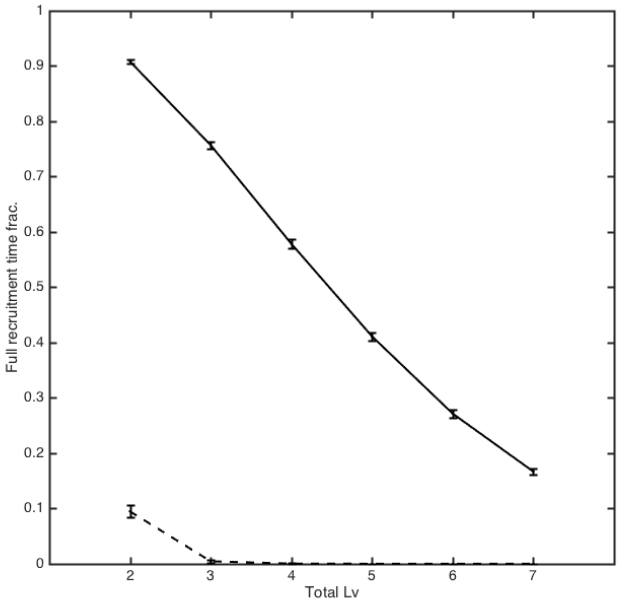
Supplementary table 1 shows that only r_1 is close to $k_+/(k_- + k_+)$, which is as expected since level 1 is the only level where corresponding protein would never fall off from the chain and whose phosphorylation and dephosphorylation could be considered as in equilibrium in the long run. If we take the values in the table as r_i's and plug them into (*), the resulting P_n's (not shown) would be consistent with the results of simulations.

Supplementary table 1. Mean values of r_i's ± standard deviations from 100 simulations (k_+ = 0.1, k_- = 0.01, $k_+/(k_- + k_+)$ = 10 / 11 ≈ 0.909).

Total levels\level	1	2	3	4	5
2	0.909 ± 0.004	0	NA	NA	NA
3	0.909 ± 0.004	0.840 ± 0.005	0	NA	NA
4	0.910 ± 0.004	0.841 ± 0.004	0.781 ± 0.006	0	NA
5	0.909 ± 0.004	0.840 ± 0.005	0.781 ± 0.006	0.730 ± 0.007	0



Supplementary figure 1. Time fractions for each level being the highest level occupied. Total number of levels in the chain = 2 - 7, k_+ = 0.001, k_- = 0.01. White tiles represent average values from 100 simulations, with error bars showing standard deviations. Black tiles represent probabilities calculated from (※). The probability of full assembly could be ignored when the total number of levels exceeds 2 under this off-state parameter setting.



Supplementary figure 2. Probability of full assembly goes down as the total number of levels increases. Solid line: $k_+ = 0.1$, $k_- = 0.01$ (on-state). dashed line: $k_+ = 0.001$, $k_- = 0.01$ (off-state). Showing average values from 100 simulations ± standard deviations.