The Building of a Cellular Detector

Haider Riaz,* David Herrera,* and Paul Francois*

Department of Physics, McGill University, Montreal

E-mail: haider.riaz@mail.mcgill.ca; david.herrera@mail.mcgill.ca; paulf@physics.mcgill.ca

Abstract

The generic detection problem in molecular biology is the quest for the mechanism by which a cell can detect a small quantity of a specific molecule in a sea of heterogenous molecules (*Needle in a haystack*). This problem is particulalrly central to the early adaptive immune reponse where T-Cells have been selected by evolution to search their surrounding environment for foreign peptides amidst a huge number of self peptides. In this study, we propose an improved cellular detector model that builds upon the paradigmatic model (KPR) by adding positive and negative feedback mechanisms.

1- Introduction

The molecular detection problem involves the ability of the detector cells to differentiate between self and foreign ligands (ligands are cell markers). Detection of foreign ligands causes the cell to activate and potentially trigger a response. It is assumed that the ligands differ by a single kinetic parameter: the dissociation time (τ). This is known as the "life-time" dogma [1]. Self ligands have a lower value for the dissociation time compared to foreign ligands. Therefore, any successful model must exploit the difference in dissociation times to trigger the immune response with a high sensitivity and specificity.

*To whom correspondence should be addressed

We assume that there are two types of ligands in the vicinity of the detector cell, the "correct" and "incorrect" ligands, with respective concentrations L_c and L_i and respective dissocaition times τ_c and τ_i. We also presume that the "life-time" dogma holds and that $\tau_c > \tau_i$ (the correct ligand binds more strongly to the T-Cell receptor than the incorrect ligand).

It is also assumed that the binding of the ligand to the receptor initiates a downstream signalling cascade that transmits information to the cell. The output of the cascade is assumed to be a single molecule. The steady state output in the presence of both ligand types is denoted by $O(L_i; L_c)$ whereas $O(L_x)$ denotes the output only in the presence of ligand $x \in \{i, c\}$ only. We let Θ be the output threshold that corresponds to activation of the detector cell.

The problem is then to have $O(L_i; L_c) \gg \Theta$ (detection in presence of correct ligands), but $O(L_i) \ll \Theta$ (no detection in presence of only incorrect ligands). A schematic diagram of the immune detection problem is schematically illustrated in Figure 1. The purpose of this project is to determine generic properties of the signalling pathways that lets a detector cell solve the problem even for large concentrations of incorrect ligands.

Kinetic Proofreading (KPR) is the paradigmatic approach to solving this problem [2]. KPR introduces reaction intermediates that are more likely to irreversibly exit the pathway if they lead to an incorrect product and less likely to exit if they lead to the correct product. The large number of proofreading steps required to distinguish foreign and self ligands accurately reduces sensitivity. Therefore, KPR has a sensitivity/tolerance trade-off for high incorrect ligand concentrations. This trade-off is more acute for ligands with similar dissociation times. In the next section, we explore more thoroughly the inadequacy of KPR in solving the generic detection problem.

2- Kinetic Proofreading Falls Short

A schematic diagram of the KPR chemical system is depicted below. C_N is the ouput that triggers cell activation after a certain threshold level is reached. C_i ($\forall\, i \in [0, N-1]$) is an

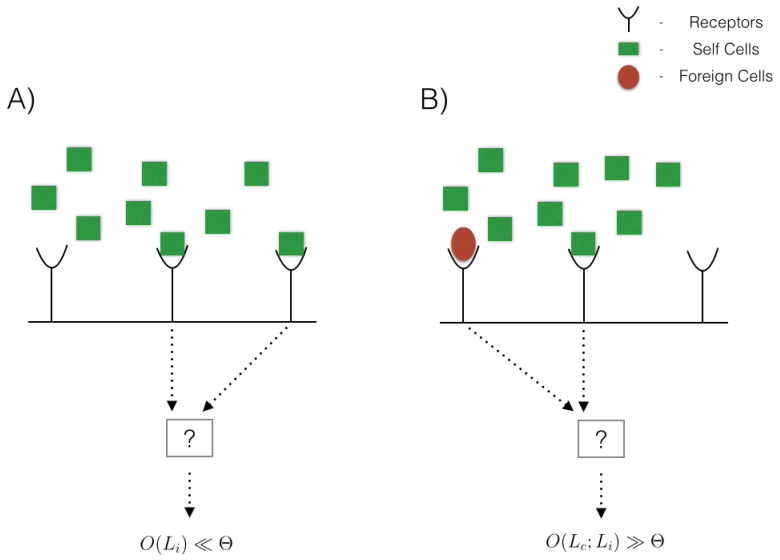

Figure 1: A schematic diagram of the detection problem. a) The incorrect ligands bind to the receptors but do not produce enough output ($O(L_i) \ll \Theta$) to illicit the activation of the cell. b) The presence of small amounts of correct ligands in the midst of a large number of incorrect ligands should trigger cell activation ($O(L_i; L_c) \gg \Theta$).

intermediate product in the KPR cascade. L^f and R^f are the number of free ligands and receptors respectively. It is assumed that the total concentration of receptors is much bigger than the number of ligands present.

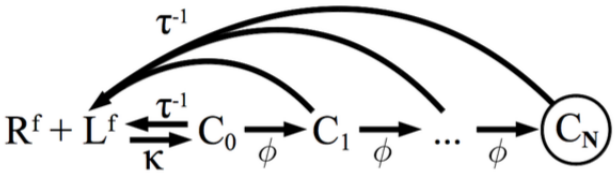

Figure 2: Visual depiction of the KPR cascade. The complexes dissociate with a rate of τ^{-1}. The output is C_N (circled).

The rate equations for the species in our model form a system of coupled differential equations (for single ligand species):

$$\dot{R}^f = \dot{L}^f = -\kappa R^f L^f + \tau^{-1} \sum_{i=0}^{N} C_i$$

$$\dot{C}_0 = \kappa R^f L^f - (\phi + \tau^{-1})$$

$$\dot{C}_n = \phi(C_{n-1} - C_n) - \tau^{-1} C_n \text{ , where } 1 \leq n \leq N-1$$

$$\dot{C}_N = \phi C_{N-1} - \tau^{-1} C_N$$

For analytical convenience, we set $\kappa R = \phi$.We solve the system exactly for the steady state solution:

$$C_N = \left(\frac{\phi\tau}{1+\phi\tau} \right)^{N+1} \quad L = \gamma^{N+1} L \tag{1}$$

Where $\gamma = (\frac{\phi\tau}{1+\phi\tau})$. Since we assumed $R \gg L$ (no receptor saturation), we do not get any interaction between ligand types so that $O(L_i; L_c) = O(L_i) + O(L_c)$.

The steady state solution for KPR is proportional to τ. Therefore, ligands with longer dissociative times will have a higher output for a particular concentraion compared to ligands with smaller dissocative times. Since $\tau_c > \tau_i$, this is desirable. But another important feature of the steady state solution is that it is linear with ligand concentration. It should be immediately apparent that this will cause a loss in specificity at high incorrect ligand concentrations.

The linear behavior of the output for different ligands is depicted in Figure 2. Observe that at high enough concentration, even ligands with lower τ cross the detection threshold. At that point it becomes impossible to discriminate ligands using pure KPR.

3- Adaptive Sorting Saves Specificity

A new model by Lalanne and Francois, called *Adaptive Sorting* circumvents the problem of low specificity at high concentrations by adding a negative feedforward loop [3]. Specifically,

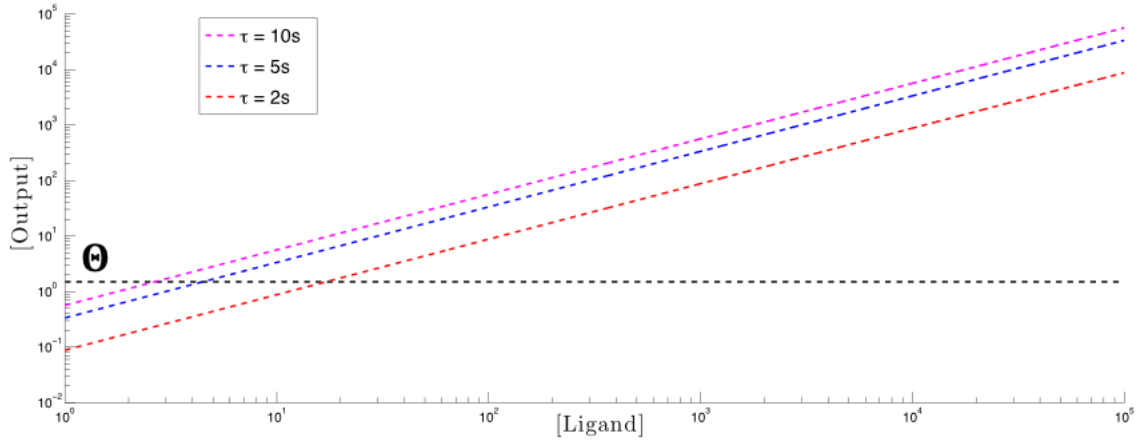

Figure 3: Output versus input for three different ligand types. The linear relationship between output and input is evident in all three plots. The output is greater for ligands with higher τ for a specific concentration compared to that of a ligand with a lower τ. Note that at high ligand concentrations, even the ligand with low dissociation time ($\tau = 2s$) triggers activation. $\phi = 1$, $N = 10$.

one complex of the cascade (C_m) inhibits a kinase K which is required to produce the output C_N. The addition of the Kinase K and its inhibition by C_m is a novel shift from pure KPR. The ouput becomes independent of L at large concentrations as a result of the incoherent negative feedforward loop.

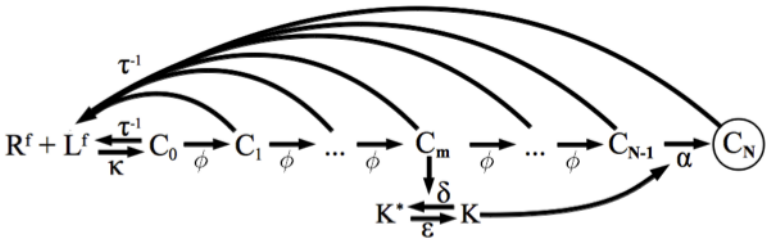

Figure 4: The adaptive sorting system adds a negative feedforward loop to the KPR backbone. The kinase K regulates the production of the output C_N. The complex C_m inhibits the production of the kinase K and therefore inhibits the production of the output C_N.

The equations for adaptive sorting are identical to those of the kinetic proofreading system, except for the equations C_{N-1}, C_N and K:

$$\dot{C}_{N-1} = \phi C_{N-2} - (\alpha K + \tau^{-1}) C_{N-1}$$

$$\dot{C}_N = \alpha K C_{N-1} - \tau^{-1} C_N$$

$$\dot{K} = -\delta K C_m + \epsilon (K_T - K) \text{ , where } K_T = \text{Total Kinase}$$

We assume for simplicity that $\alpha K_T = \phi$ and proceed to solve for the steady state solution:

$$C_N = \left(\frac{C_* (\phi\tau)^2 \gamma^{N-2}}{C_* (1 + \phi\tau) + \gamma^m C_0} \right) C_0 \text{ ,where } C_0 = \frac{\gamma L}{1 + \phi\tau} \tag{2}$$

Where $C_* = \epsilon\delta^{-1}$ is inversely proportional to the strength of the kinase inhibition. The output at large L is independent of the input:

$$C_N \to C_* (\phi\tau)^2 \gamma^{N-m-2} \text{ for large } L$$

By placing the threshold (Θ) above the maximum output of the incorrect ligand, adaptive sorting gives infinite tolerance (no amount of incorrect ligand concentration can trigger detection). This is depicted in Figure 5. The dashed lines represent pure KPR whereas the solid lines represent the output of adaptive sorting. Incorrect ligands do not cross the threshold even at high concentrations in adaptive sorting.

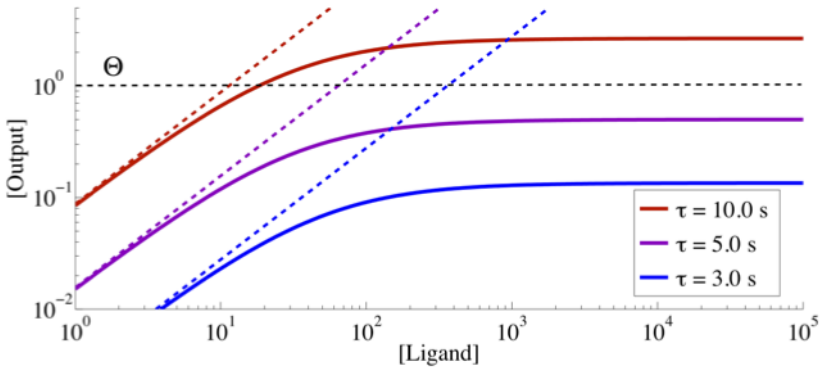

Figure 5: Output versus input for adaptive sorting. Full lines have negative feedforward, dashed lines are pure kinetic proofreading solutions. $N = 5$, $m = 2$, $C_* = 1$, $\phi = \kappa R = \alpha K_T = 0.2$, $\Theta = 1.5$.

There is however one other major difference between adapive sorting and KPR. The regulating kinase in KPR leads to interaction between the different ligands and we no longer have: $O(L_i; L_c) = O(L_i) + O(L_c)$. The adaptive sorting model with two ligands present is illustrated in Figure 6.

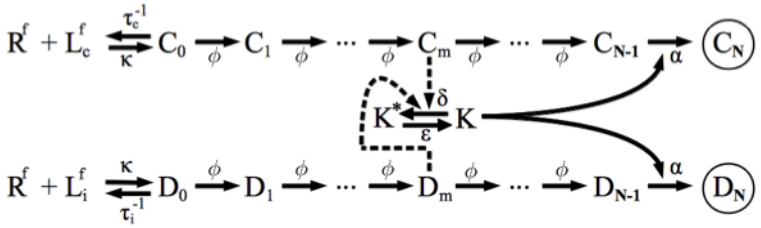

Figure 6: Adaptive Sorting model for two ligand species. The output is now $C_N + D_N$. The kinase is now inhibited by both C_m and D_m simultaneously.

The output of the system is $C_N + D_N$. But now notice that large numbers of incorrect ligands (L_i) lead to strong inhibition of the kinase K (and therefore the output). This causes a severe loss in sensitivity (for high incorrect ligand concentrations) as depicted in Figure 7.

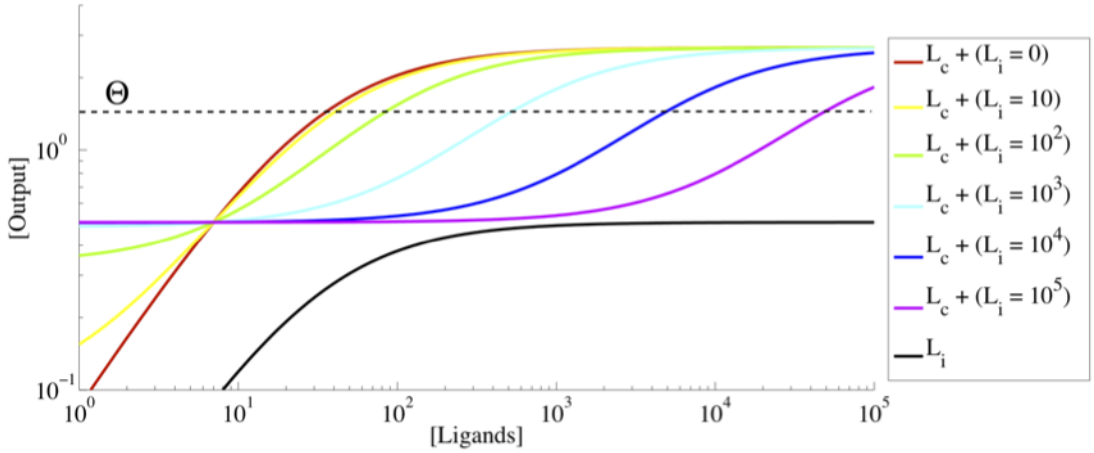

Figure 7: The interaction between ligands leads to loss in sensitivity. Shown above is the output versus ligand concentration for L_c with different L_i (coloured lines) and L_i (black line). Notice how a larger L_i decreases the sensitivity of detection. $\tau_c = 10s$, $\tau_i = 2s$, $\epsilon = 0.5$,$N = 5$,$m = 2$,$C_* = 1$, $\phi = \kappa R = \alpha K_T = 0.2$, $\Theta = 1.5$.

4- A More Balanced Model

The sensitivity can be restored by adding several modifications to the adaptive sorting model. The first change is that C_N now positively regulates the production of the kinase K (positive feedback loop). This modification on its own, however, will simply lead back to KPR (the negative feedforward and positive feedback cancel each other effects). Therefore the negative feedforward loop is also changed from containing one inhibiting complex to two complexes that are necessary for inhibition of the kinase (cooperative negative feedback).

We define M to be the product of complexes that contribute to the negative feedforward: $M = C_i * C_j$. The rate equations for our new model are then as follows:

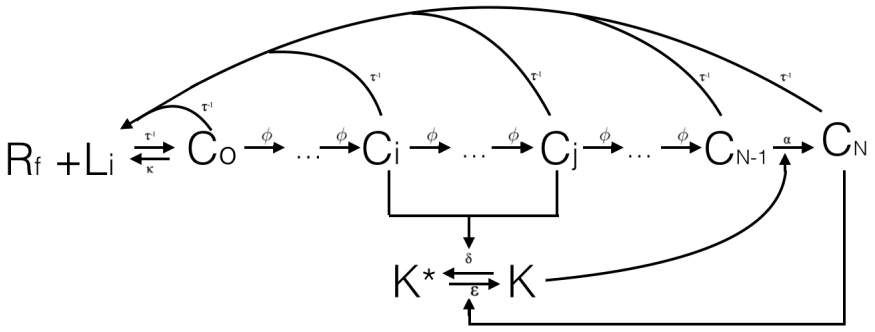

Figure 8: A modified version of adaptive sorting. The output C_N positively regulates the kinase production. The negative feedforward now requires two complexes: C_i and C_j.

$$\dot{C}_n = \phi(C_{n-1} - C_n) - \tau^{-1}C_n \text{ , where } 1 \leq n \leq N-1$$

$$\dot{C}_{N-1} = \phi C_{N-2} - (\alpha K + \tau^{-1})C_{N-1}$$

$$\dot{C}_N = \alpha K C_{N-1} - \tau^{-1}C_N$$

$$\dot{K} = -\delta K M + \epsilon C_N(K_T - K) \text{ , where } K_T = \text{Total Kinase}$$

Solving the system for the steady state solution as before gives us:

$$C_N = \frac{\alpha K_T \gamma^N L}{\alpha K_T \tau + 1} - \frac{\delta M}{\epsilon(\alpha K_T \tau + 1)} \text{ , where } M \propto L^2 \tag{3}$$

The output of this system is linear initally (the positive term dominates) but starts to fall off in a quadratic fashion (the negative term starts dominating) at higher concentrations. Therefore like adaptive sorting, the model has a high specificity (the output does not continue growing with ligand concentration). This is portrayed graphically for single ligand species in Figure 9.

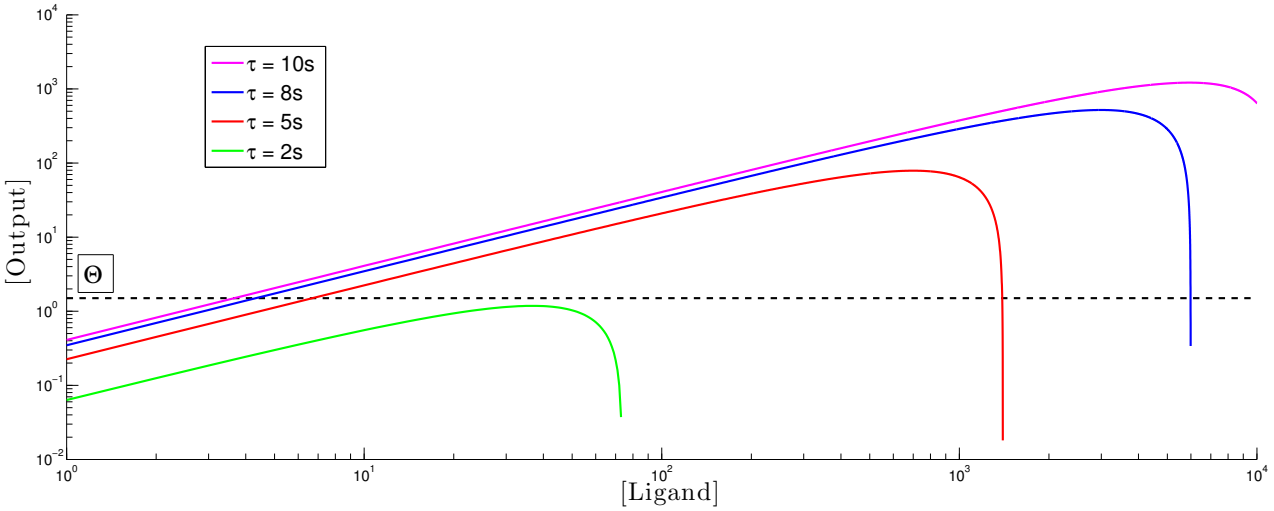

Figure 9: Output versus input for the new model. The initial behavior of the output is similar to that of KPR but at higher concentration the ouput plateaus and drops down. $\epsilon = 10$,$N = 10$,$i = 4$, $j = 5$, $\phi = \kappa R = 5$, $\alpha K_T = 0.2$, $\delta = 0.5$,$\Theta = 1.5$.

We now turn to the case of two ligand species present in the system. As in adaptive sorting, the two ligand species interact with each other through the common kinase. The output of the system is now: $C_N + D_N$ and the differential equations describing the system read:

$$\dot{C}_N = (\alpha K \gamma_c^N L_c)\,(\alpha K + \tau_c^{-1})^{-1}$$

$$\dot{D}_N = (\alpha K \gamma_i^N L_i)\,(\alpha K + \tau_i^{-1})^{-1}$$

$$\dot{K} = -\delta K(M_c + M_j) + \epsilon(C_N + D_N)(K_T - K)\ \text{, where } K_T = \text{Total Kinase}$$

We solve for the steady solution and as was done for adaptive sorting, we plot the output $(C_N + D_N)$ as a function of the correct ligand concentration for various incorrect ligand concentrations in Figure 10.

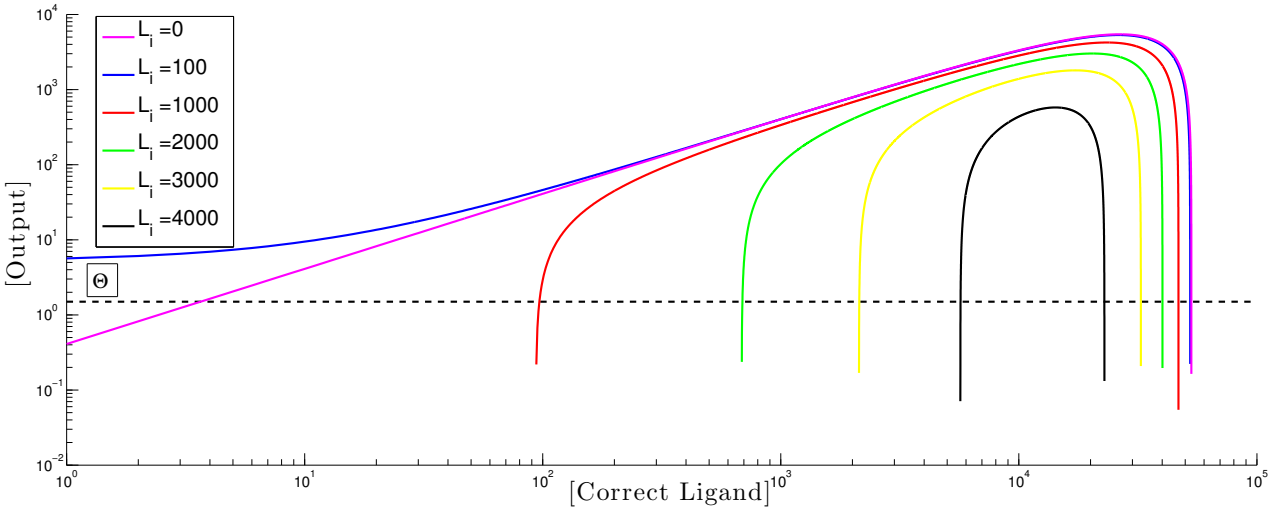

Figure 10: Output versus input for the new model in the presence of both incorrect and correct ligands. The suppressive behavior seen in simple adaptive sorting due to incorrect ligands is not seen for concentrations of up to a 1000 (therefore improving sensitivity). For concentrations higher than 1000 the sensitivity deteriorates again and the beahvior of the output follows that of a skewed normal distribution. $\epsilon = 10$,$N = 10$,$i = 4$, $j = 5$, $\phi = \kappa R = 5$, $\alpha K_T = 0.2$, $\delta = 0.5$,$\Theta = 1.5$.

There is a marked improvement in sensitivity compared to simple adaptive sorting. For incorrect ligand concentrations up to 10^3, the sensitivity improves by at least half an order of magnitude. For incorrect concentrations higher than 10^3, the sensitivity deteriorates sharply due to the effect of the cooperative negative feedback by the incorrect ligand complexes. The output mimicks the behavior of a phase transition at these high incorrect ligand concentrations. There is severe suppression up until a critical concentration value, after which it shoots up rapidly.

A more subtle advantage of this model over adaptive sorting is that the output grows linearly with concentration for correct ligands (with only no or little incorrect ligands present) until saturation ($L_c = 10^5$). This means that information about the system is not lost as in adaptive sorting where the output is constant at high concentrations. The constant output provides no information on the changing ligand concentrations.

5- Looking Ahead

We have shown that by building on the KPR model, one can move closer to solving the generic detection problem in molecular biology. Our enhanced model reatins the high tolerance of adaptive sorting whilst improving sensitivity for environments with incorrect concentrations below a certain threshold. The model's parameter can perhaps be fine tuned further and that can lead to improved performance. As noted before, another advantage of the model over adaptive sorting is that information about the ligand landscape is not lost at high input concentrations.

This work is being done at a time when experiments are being done that link theoretical discussions to the real phenomena of life and allow for falsification of theoretical models. We hope to have conveyed that the new detection model is suitable for the role of detection and brings us a step closer to the correct description of molecular detectors.

We have found it of paramount importance that in order to progress we must recognize our ignorance and leave room for doubt. Scientific knowledge is a body of statements of varying degrees of certainty some most unsure, some nearly sure, but none absolutely certain.
- Richard Feynman

Acknowledgement

We would like to thank our supervisor Paul Francois for introducing us to the problem of designing cellular detectors. We would also like to acknowlege the work of Jean-Benoit Lallane on this problem from which we benefitted immensely.

References

(1) O. Feinerman, R. N. Germain, and G. Altan-Bonnet, "Quantitative challenges in understanding ligand discrimination by $\alpha\beta$ T cells", Molecular immunology 45, 619 (2008).

(2) J. J. Hopfield, "Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity", Proc. Natl. Acad. Sci. U.S.A. 71, 4135 (1974).

(3) Jean-Benoit Lalanne and Paul Francois, "Principles of Adaptive Sorting Revealed by In Silico Evolution", PRL 110, 218102 (2013).