A viral flight data recorder for systems biology applications

Erik M Boczko

August 13, 2008

I am interested in self replicating, dynamical systems and the theorems and design principles that underlie their existence, persistence and evolution. I am interested in designing rigorous, direct and quantitative measurements and in utilizing evolved biological objects as new engineering tools at every scale.

Currently, systems biology research is limited not by computing power, nor by algorithmic or logical obstructions but rather by the inability to quantitatively record the chemical and physical transactions inside single living cells. Our modeling efforts have been hampered by this problem and we have turned our sights toward designing rigorous, direct measurements of nanoscale phenomena. The beauty is that each measurement problem brings with it new modeling problems.

I have developed a conceptual strategy to reprogram viral capsid formation and utilize this process as a *flight data recorder* to count arbitrary intracellular transactions in real time. The use of green florescent protein (GFP) or luciferase as a molecular reporters of real time events in single cells has in very large part been responsible for the growth of systems biology research and consciousness. The advantage of these reporters is that they make themselves, and all one has to do is imbue the cell with the proper genetic instruction set. GFP illuminates protein chimeras. Currently it is not possible to illuminate nucleic acids or other macromolecules, *in vivo*, in the same way. The typical materials used in nanotechnology, such as gold or poly dimethyl siloxane(PDMS), are not naturally manufactured or replicated by cells. Fortunately, nature has evolved other objects that we can redesign, reprogram and redeploy to our benefit.

Consider the following facts. Viral capsid proteins spontaneously encapsidate nucleic acid genomes and self organize into nanoparticles. Viral coat proteins have evolved the ability to recognize and bind specific RNA sequences as a prerequisite to capsid nucleation. The capsids of many RNA viruses are formed from a symmetric arrangement of n copies of a single coat protein. A domain of these coat proteins is laden with positively charged amino acids such as arginine. If this domain is not complexed with negatively charged nucleic acid, the symmetric arrangement of positive charges repel each other and disfavor the formation of viral capsid nanoparticles. It has been possible to make coat protein chimeras that display protein fragments on the outer capsid surface. It has also been possible to bring together the two halves of a GFP protein, that were independently expressed, to form a functional GFP molecule. These facts align to yield a flight data recorder.

The process of capsid formation can be co-opted [2]. A single coat protein can be re-engineered in such a way that a capsid forms if and only if a fixed number of molecular transactions have transpired. The coat proteins can be altered in such a way that only the fully formed capsid releases photons. Given a coat protein with these engineered properties we can count the number of completely formed virus particles inside a single cell, as a function of time, to measure the dynamics of the transaction.
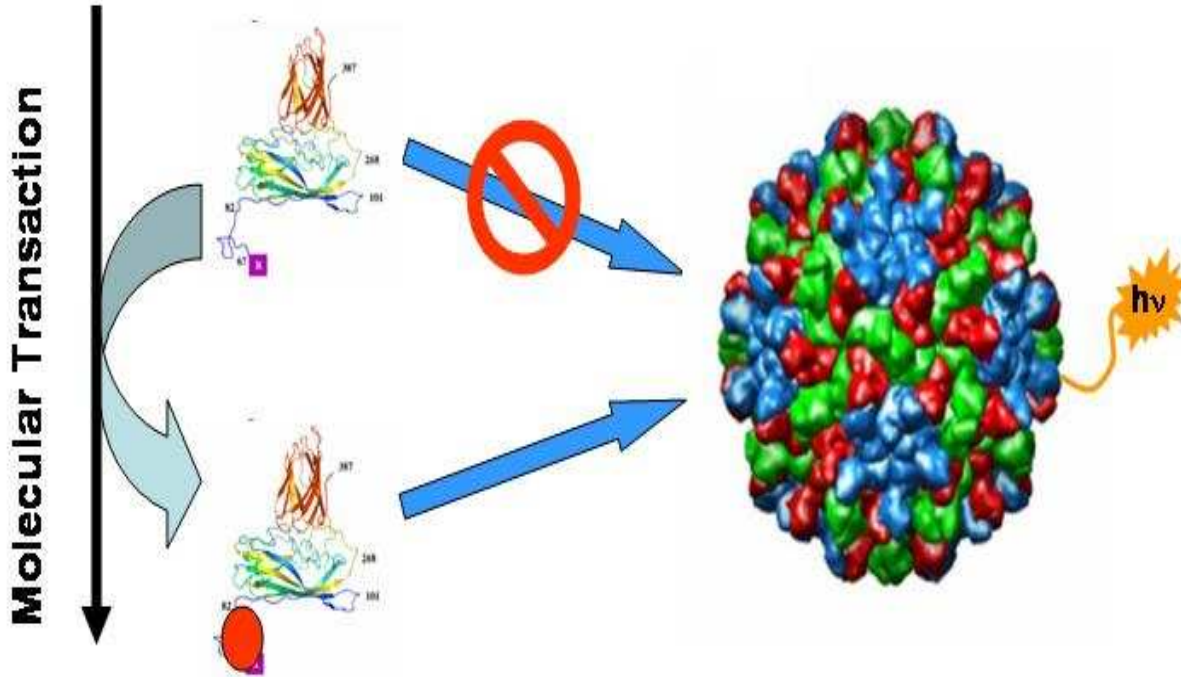
Figure 1: An arbitrary molecular transaction induces the phosphorylation of a viral coat protein. Phosphorylated protein monomers can assemble into a capsid, while unmodified ones cannot. Each fully formed capsid nanoparticle releases photons upon complete assembly. By monitoring the signal in real time from a single cell we can monitor the dynamics of intracellular transactions.

With only slight modification, the TBSV system can be used to follow the dynamics of any specified non-viral mRNA species. Our data, and the success of Wickens' three hybrid system [1] clearly demonstrate the feasibility and potential of the flight data recorder idea for following the dynamics of a 5'-tagged, but otherwise arbitrary, mRNA species. How can arbitrary transactions be tied to capsid formation? A key feature to be exploited is the symmetric arrangement of positive charge. Custom phosphorylation sites within the R-domain would allow the positive charge to be ameliorated not by binding with nucleic acids but through phosphorylation. Phosphorylation is the major form of intracellular signaling, whose explicit purpose is to tie together disparate intracellular transactions. Nature has already created signaling networks for us. We simply reprogram this signaling process to tie together an arbitrary transaction with coat protein phosphorylation. In this way we can create customizable, organic, intracellular data recorders, see Figure 1. This technology to deploy intracellular flight data

recorders could revolutionize systems biology across the board. Molecular Beacons and fluorescence in situ hybridization notwithstanding, the paradigm by itself is worth developing since we see no difference at the abstract level between symbiosis, intracellular symbiont and organic intracellular recording and control devices [3]. The ideas outlined above are the beginning of this larger vision.

References

[1] D.S.Bernstein, N.Buter, C.Stumpf, and M.Wickens, Analyzing mRNA-Protein complexes using a yeast three hybrid system. (2002) *Methods* **26**:123-141

[2] C.Hsu, P.Singh, W.Ochoa, D.J.Manyani, M.Manchester, A.Schneemann and V.S.Reddy, Characterization of polymorphism displayed by the coat protein mutants of tomato bushy stunt virus. (2006) *Virology* **349**:222-229.

[3] K. W. Jeon, Genetic and physiological interactions in the Amoeba-Bacteria symbiosis. (2004) *J. Eukaryot. Microbiol.* **51**:502-508.

[4] T.Keef, C.Michaletti, R.Twarock, Master equation approach to the assembly of viral capsids. (2006) *J. Theor. Biol.* **242**:713-721.

[5] V.S.Reddy, H.A.Giesing, R.T.Morton, A.Kumar, C.B.Post, C.L.Brooks III and J.E.Johnson, Energetics of quasiequivalence:computational analysis of protein-protein interactions in icosohedral viruses. (1998) *Biophys. J.* :546-558.