Cell, Vol. S116, S73–S74, January 23, 2004 Copyright ©2004 by Cell Press

# Building an Artificial Regulatory System to Understand a Natural One

Commentary

**Roger Brent***
The Molecular Sciences Institute
Berkeley, California 94704

It was my ninth year in Mark Ptashne's lab. I had learned from outstanding students, post-docs, and the great soul at the top to identify important problems and explore them by experiment.

The problem was "action at a distance." A site on DNA *here* causes transcription to initiate at a promoter *there*, here and there being hundreds or thousands of nucleotides apart. Mysterious. Fascinating. In prokaryotes, this phenomenon had an inglorious history and unsatisfying mechanisms to explain it. Perhaps proteins bound to regulatory sites transmitted changes in DNA structure ("telestability"; Burd et al., 1975) to the transcription start site. Perhaps RNA polymerase bound a site, then slid without transcribing to the transcription start (this idea, marvelously termed "polymerase drift," arose first from EM studies of polymerase binding sites and promoters in phage λ [Wollwieder and Szybalski, 1978]). The problem resurfaced in the 1980s, due to the discovery of distant positively acting sites in higher eukaryotes (enhancers; Banerji et al., 1981) and in yeast (upstream activating sequences [UAS]; Guarente et al., 1982). One UAS, upstream of the yeast GAL1 and GAL10 genes, required Gal4 protein for it to work. The putative sites of Gal4 action in this UAS were 2-fold rotationally symmetric, so the active protein was probably a dimer.

I built on graduate work on LexA, an *E. coli* repressor that controls genes induced after DNA damage. Ptashne had tolerated, then supported the work, even before the protein turned out to be λ-repressor-like. Since spring 1984, I had been putting LexA into yeast. The idea was that the prokaryotic LexA might be cleaved after DNA damage in yeast as it was in *E. coli*, and, if so, that might provide path into the response to DNA damage in eukaryotes. I worried that I lacked sufficient control over the new things I was doing to the organism. My comrade Pam Silver and I had a term for the worry, "the great fear," that we might be missing something basic about yeast biology and that ignorance would render our (or at least my) experiments worthless.

The DNA damage experiments failed, but fear-powered technical control proved to be useful. Now there were yeast I was certain produced a bacterial repressor, and I could ask if the repressor repressed in eukaryotes. It did: LexA repressed reporter genes with LexA binding sites inserted in their promoters (Brent and Ptashne, 1984). This first use of "effector" and "reporter" constructs together in eukaryotes was inspired, again, by work in λ: the marvelous dual-component control circuit, *lac* promoter reading λ repressor repressing λ promoter reading *lacZ* gene *E. coli*, constructed by Russ Maurer (1978). These LexA-repression-in yeast experiments argued against "telestability" (see Figure 1). More importantly, they provided both a proven experimental setup and license to ignore aspects of eukaryotic biology one might reasonably have thought relevant, including nuclear localization of regulatory proteins, nucleosomes, and any assertions expressed in declarative sentences containing the word "chromatin."

These results enabled a leap in the dark. It might be that Gal4 could help LexA's DNA binding 87 amino terminal residues dimerize, and that the resulting chimeric protein would bind LexA sites. Since DNA binding by native LexA did not activate transcription (but rather repressed it), if the chimeric protein activated, activation could not be due to a change in DNA structure upon DNA binding by the LexA moiety.

So, I built things. I made *lacZ* reporter plasmids that lacked UASs and carried LexA binding sites instead. I made a fusion gene to make a fusion protein, paying great attention to make sure that the LexA DNA binding region might be connected to Gal4 via a flexible hinge. I tested whether the chimera bound LexA binding sites by recycling graduate work and showing that, when I expressed LexA-Gal4 in *E. coli*, it repressed the bacterial response to DNA damage.

Then, in late 1984, I took yeast that contained both artificial reporter and artificial protein and smeared said yeast onto an indicator plate with the broad end of a toothpick. I was hoping for some hint of blue color, indicating reporter activation, in the next two days. I came back to the plate a half hour later to take a peek. The streak was already dark blue. LexA-Gal4 activated gene expression and did so strongly. Native LexA did not. Native Gal4 did not. The experiment worked. Simplification worked. My major emotion was profound relief.

From experiment to stable interpretation took about two months. Interpretation grew over a series of conversations with Ptashne that stretched over about week before Mark was off to a meeting to describe the results. In those, we locked in on the following view. If one can swap pieces, and those pieces are sufficient for function, then those pieces are *domains* or *modules*, the protein is *modular*, and the experiment was a *domain swap*. For me, this process demonstrated that truth can be revealed by experiment, but that finding the right words may be needed to complete its creation

The interpretation left a number of loose ends. Even though we had established that the C-terminal 807 amino acids of Gal4 comprised an "activation domain," the swap left open whether activation was a complex enzymatic activity or something simple, whether it required the entire Gal4 moiety or only pieces of it, etc. But one consequence of the concept of modularity was to enable arbitrary proteins, including quite crude deletions of known activators, to be stuck to defined sites on DNA upstream of reporters where their activation and repression could be measured. There was clearly to be a great wave of these experiments during the late 1980s, but I had decided to stop studying transcription activation. Rather, I wanted to use activation in yeast as a phenotype to study processes in higher eukaryotes, a quest that eventually led me to protein-interaction genomic approaches.

*Correspondence: brent@molsci.org

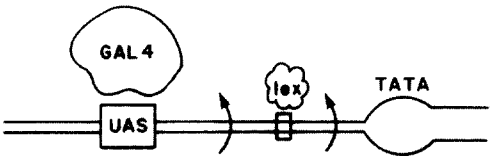


Figure 1. It Can't Work This Way.

If Gal4 activated transcription by unwinding DNA near it, LexA bound between Gal4 and the transcription start shouldn't repress gene activation. Slide from "Puffball series," November, 1984.

Another consequence was another license, to set aside many considerations from protein biochemistry and structural biology. I drew the LexA and Gal4 figures for the paper as horizontal rectangles on one of the early Macintosh computers, a machine that facilitated the cutting and pasting of graphics. Although it wasn't the cleverest concept in the history of molecular biology, cutting and pasting did turn out to be surprisingly good way to learn about proteins. The most significant next step in the jettisoning of structural considerations was probably taken by Paul Godowski in Keith Yamamoto's lab (1988), who swapped LexA's DNA binding region into the middle of the glucocorticoid receptor, without prior worry about the amino acid sequence that would be generated at the junctions. As easy to simply do the experiment.

Another consequence was some contribution to the rise of a unifying concept, that changes in intracellular function arise from changes in localization, which often arise from noncovalent interactions with other molecules (now called recruitment, see the Ptashne essay in this supplement). This idea received impetus from the picture of cellular function generated by application of the two-hybrid method (Fields and Song, 1989) and of follow on means to perform wholesale surveys of protein-protein interactions (Finley and Brent, 1994).

This burst of knowledge of interactions in turn helped bring about the current semi-impasse. Knowing interactions provides real information to load into databases, but it doesn't provide satisfying understanding of how proteins work together to create cellular outcomes. Now, the challenge is different. If anatomy is comprised both of parts and their connections, then, now that we know these, we are obliged to turn anatomy into physiology. This physiological understanding will need to be quantitative and predictive.

### References

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299–308.

Brent, R., and Ptashne, M. (1984). A bacterial repressor protein or a yeast transcriptional activator can block upstream activation of a yeast gene. Nature *312*, 612–615.

Burd, J.F., Wartell, R.M., Dodgson, J.B., and Wells, R.D. (1975). Transmission of stability (telestability) in deoxyribonuclic acid. Physical and enzymatic studies on the duplex block polymer d(C15A15)-d(T15G15). J. Biol. Chem. *250*, 5109–5111.

Fields, S., and Song, O. (1989). A genetic system to detect protein-protein interactions. Nature *340*, 245–246.

Finley, R.J., and Brent, R. (1994). Interaction mating reveals binary and ternary connections between *Drosophila* cell cycle regulatory proteins. Proc. Natl. Acad. Sci. USA *91*, 12980–12984.

Godowski, P.J., Picard, D., and Yamamoto, K.R. (1988). Signal transduction and transcriptional regulation by gluocorticoid receptor-LexA fusion proteins. Science *241*, 812–816.

Guarente, L., Yocum, R.R., and Gifford, P. (1982). A GAL10–CYC1 hybrid yeast promoter identifies the GAL4 regulatory region as an upstream site. Proc. Natl. Acad. Sci. USA *79*, 7410–7414.

Maurer, R. (1978). Regulation and autoregulation by the bacteriophage λ repressor protein. Ph.D. thesis, Harvard University, Cambridge Massachusetts

Pabo, C.O., Sauer, R.T., Sturtevant, J.M., and Ptashne, M. (1979). The lambda repressor contains two domains. Proc. Natl. Acad. Sci. USA *76*, 1608–1612.

Vollenweider, H.J., and Szybalski, W. (1978). Electron microscopic mapping of RNA polymerase binding to coliphage λ DNA. J. Mol. Biol. *123*, 485–498.