Stochastic model validation for the Pap (pili) epigenetic switch

Brian Munsky[1], Brooke Trinh[2], David Low[2], and Mustafa Khammash[3]

***Short Abstract* — The biochemical regulation and epigenetic control of pyelonephrytis-associated pili (Pap) in uropathogenic *E. coli* is a stochastic, but well-characterized process. This study introduces a detailed numerical model that improves this qualitative understanding to allow for accurate quantitative predictions. Through integrated modeling and experimental design, model parameters have been directly measured or numerically identified using *in vitro* Mobility Shift Assays and *in vivo* Florescence Activated Cell Sorting. The result is highly predictive Markov chain model, which can now be extended to discover new behaviors in Pap and other similarly controlled genetic regulatory mechanisms.**

***Keywords* — Pyelonephrytis-Associated Pili, Genetic Regulatory Network, Predictive Model, Stochastic Switch, Epigenetic Control**

Pyelonephrytis-associated pili (Pap) are small hair-like structures that cover and increase the infectiousness of uropathogenic *E. coli*. Over many generations, the Pap system switches randomly, between ON (piliated) and OFF (unpiliated) states, each of which is stabilized by positive feedbacks and epigenetic control via heritable DNA methylation patterns. At its core, the Pap switch is primarily controlled by leucine responsive regulatory protein (Lrp), a global regulator that also affects many other *E. coli* genes. When Lrp binds proximal to the *pap* operon, RNAP is blocked from the gene and the system cannot produce Pap proteins. Conversely, when Lrp binds further upstream from the promoter, RNAP is effectively recruited and the gene turns ON. Lrp binding, and therefore the Pap switch, is indirectly controlled by other regulatory features including: (i) a local regulatory protein, PapI, interacts with and changes the strength of certain Lrp-DNA complexes; (ii) DNA adenine methylase (Dam) mediates methylation of GATC sequences within Lrp binding regions and affects Lrp affinity in those regions; and (iii) Bound Lrp molecules induce DNA conformations that affect Lrp binding at other sites. Each of these effects has been quantified and all are combined in this paper to provide a comprehensive and extensively validated bottom-up model of the Pap epigenetic switch.

The most critical parameters in the model are Lrp binding and unbinding rates, which can be quantified from *in vitro* electrophoretic Mobility Shift Analyses (EMSA) with disjoined distal and proximal DNA fragments [1]. These rates have been determined for various PapI concentrations and DNA methylation patterns and have been fit with a simple cooperative binding model (see representative example in Fig. 1A).

The pap gene has been fused to GFP reporter gene and *in vivo* cell populations are examined with florescence activated cell sorting (FACS). For model calibration of extrinsic nose, baseline florescence of OFF cells is measured by sorting cells with an under-induced GFP promoter. The resulting measurements are incorporated into the model as an additive continuous valued Gaussian signal. The model quantitatively matches the measured histograms in a wide array of circumstances (see representative example Fig. 1B).

In the numerical and biological experiments, cells are defined ON when florescence exceeds a certain threshold; otherwise they are OFF. The ON probability has been measured after ten cell generations with concentrations of Dam, Lrp and PapI varied under arabinose or Lac control. The model captures the quantitative effects of variation in theses three concentrations (see, for example, Fig. 1C).

This rigorously validated Pap model provides one of the most detailed and accurate representations of any stochastic gene regulatory switch, and it substantially increases the biological understanding of the Pap system. Further, the model enables testable quantitative predictions on *in vivo* Lrp binding and Dam methylation patterns under numerous attainable conditions (not shown).

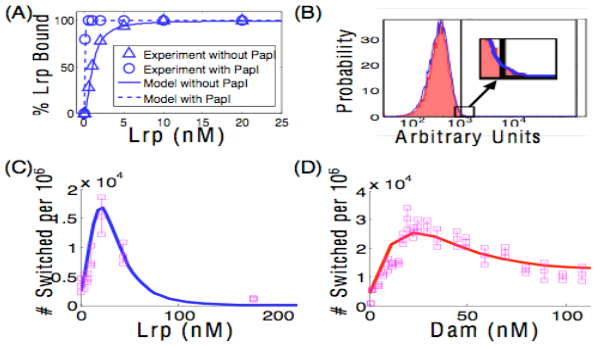

Figure 1. Validation of the Pap Model with identified parameter set. (A) Analytical (lines) determination of Lrp affinities for the distal Lrp binding location at 0nM and 5nM PapI fit to *in vitro* EMSA data (markers). (B) Quantitative analysis (blue line) of *in vivo* FACS data (red histogram) for a particular value of Dam and Lrp concentrations. Cells above the threshold are ON—see inset. (C) Matching *in vivo* proportion of ON cells as function of Lrp. (D) Predicted switch rates and measured *in vivo* Pap protein levels as functions of Dam induction.

References

[1] Hernday A, Braaten B, and Low D *Mol. Cell*, **12**, 947-957.

[1]Center for NonLinear Studies and CCS-3, Los Alamos National Lab. E-mail: brian.munsky@gmail.com

[2]Department of Molecular, Cellular and Developmental Biology, University of California at Santa Barbara.

[3]Department of Mechanical Engineering, University of California at Santa Barbara.