A Mechanism for Stochastic Decision Making by Bacteria

Zach Hensel and Jie Xiao*[a]

Nature presents organisms of every scale with a constantly shifting set of challenges. The correct response to these challenges is critical to an organism's survival. Churchill once remarked, "True genius resides in the capacity for evaluation of uncertain, hazardous, and conflicting information", and this ability often determines whether an organism survives, as well as who prevails in war. In contrast to humans, who make rational decisions after careful evaluation of their situation, bacterial cells sometimes appear to randomly choose their fates. Choi et al. recently reported that an *E. coli* cell's decision to induce or not to induce the uptake of a metabolite hinges on the stochastic action of a single molecule.[1]

Stochastically driven cellular choices have been increasingly observed in biology. In recent years, studies have found that the decisions by *Bacillus subtilis* to become competent for DNA uptake[2] and by *E. coli* to enter a quiescent state to survive antibiotic exposure[3] are stochastically controlled. Similar phenomena have also been observed in higher eukaryotes; early-stage hematopoietic stem cell differentiation[4] and sensory neuron differentiation[5] appear to involve stochastic mechanisms as well.

Why would a cell make its choices stochastically instead of programmatically? Stochasticity often leads to phenotypic variations in an isogenic population. In a fluctuating environment, this can be advantageous as cells can be in different states to contend with many future possibilities. Recent work has found that phenotypic variation can confer advantage in dynamic environments.[6, 7] In contrast, committing to a single fate across

a population, although advantageous in a stable environment, is dangerous in the event of change. Additionally, programming every cell in a population to deterministically respond to any conceivable environmental cue can be unbearably costly or even impossible.[8]

While the benefit of stochastic cellular choices under certain circumstances is often evident, the exact cause of such stochasticity may differ in each case. For example, noise in gene expression has been shown to be a major contributor to phenotypic heterogeneity.[2, 9] Now, Choi et al. have revealed a new mechanism: they have determined that stochastic molecular interactions between the *lac* repressor and its operators underlie the transition of an *E. coli* cell to induction.

The *lac* operon of *E. coli* has served as the first paradigm for gene regulation.[10] Induction of the *lac* operon occurs when the *lac* repressor unbinds from the *lac* operator in the presence of lactose or a nonmetabolizable inducer, such as methyl-β-$_D$-thiogalactoside (TMG), thus enabling transcription of the *lacZYA* genes. Once expressed, lactose permease, encoded by *lacY*, imports more lactose and creates a positive feedback loop. Landmark work by Novick and Weiner showed that induction of the *lac* operon was all-or-none, and it was postulated that the generation of the first lactose permease molecule was the rate-limiting step in induction.[11] Once it occurred, full induction followed due to the positive feedback of lactose permease. However, Choi et al. have now shown that one permease molecule is not enough. Instead, a few hundred permease molecules are needed before full induction can take place.

This claim by the Xie group relies on the ability to detect single permease molecules fused to a yellow fluorescent protein. The authors found that many uninduced cells have one or more per-

mease molecules, irrespective of the TMG concentration (Figure 1 A). This indicates that one permease molecule is insufficient to induce a cell.

If one permease molecule is not enough, then how many are needed? To measure this threshold, Choi et al. used a clever strategy to prepare cells with a wide range of initial numbers of permease molecules. They first fully induced cells, and then grew cells in media lacking TMG to dilute the initial permeases through multiple cell divisions. They quantified the number of permease molecules in these cells, added TMG back to the cells, and observed whether induction occurred. If the initial number of permease molecules in one cell is high enough when TMG is added back, a sufficient amount of inducer will be imported into the cell to actively "pull" off bound *lac* repressor from its operators, leading to induction. They found that the probability of a cell switching to induction as a function of the initial permease numbers is well described by a Hill function (Figure 1 B). The threshold, defined as the number of permease molecules leading to a 50% probability of induction, is approximately 375 molecules at 40 μM TMG. The presence of such a threshold indicates a nonlinear response of the induction process, probably resulting from the cooperativity of *lac* repressor binding to DNA at two operators through a DNA loop. Positive cooperativity has been shown to be vital in achieving a binary response.[12]

How are hundreds of permease molecules generated if the cell is initially at the uninduced state? Previous work from the Xie group demonstrated that, in the absence of an inducer, the *lac* promoter fires infrequently; transcription occurs approximately once per cell cycle, generating on average one mRNA transcript and a handful of protein molecules.[13, 14] A very similar expression profile is observed here for the uninduced

[a] *Z. Hensel, Dr. J. Xiao*
Department of Biophysics and Biophysical Chemistry
Johns Hopkins University School of Medicine
725 N Wolfe Street, Baltimore, MD 21205 (USA)
Fax: (+1)410-502-7221
E-mail: xiao@jhmi.edu

fraction of cells; this implies that this type of expression is incapable of generating enough permease molecules for induction. When the authors followed real-time protein production in the presence of 200 μm TMG, they observed that about 1% of expression bursts generated more than 100 protein molecules. This indicates that these rare, large expression bursts could provide the initial permease molecules necessary for induction.

But what is the molecular mechanism underlying these rare, large expression bursts? Choi et al. proposed that it could be the stochastic, complete dissociation of the tetrameric *lac* repressor from the looped DNA formed between the operator O_1 and one of its auxiliary operators O_2 or O_3. Binding of tetrameric *lac* repressor at O_1 has been shown to be strengthened by cooperative looping with O_2 or O_3, which originally explained why *lac* repression was more resilient than expected against mutations at O_1.[15,16] The relatively frequent, small bursts observed in uninduced cells are likely due to the partial dissociation of the repressor from O_1 but not the other operators, which is usually followed by rapid rebinding.

To test their hypothesis, Choi et al. generated a mutant strain in which DNA looping is abolished by deleting O_2 and O_3. This mutant displayed uniform, monotonic induction irrespective of the initial permease numbers in a cell, and no induction threshold was observed (Figure 1B). In addition, they found that, without looping, the large bursts occur much more frequently than with looping and burst size increases with increasing TMG concentration, further proving that large bursts are the direct result of complete repressor dissociation.

From these experiments, a stochastic mechanism of *lac* operon induction, illustrated in Figure 2, becomes clear: while *lac* is repressed, small expression bursts arise from single transcription events when the *lac* repressor unbinds O_1, but quickly rebinds due to its association with other operators. Induction occurs only when a rare, large expression burst resulting from the complete dissociation of *lac* repressor from DNA produces enough permease molecules to cross
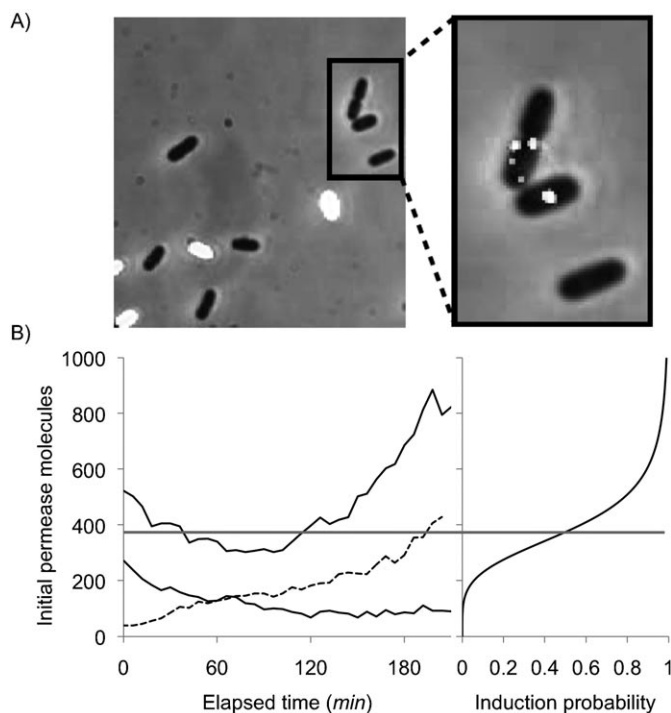

Figure 1. A) After incubation with the inducer TMG, *E. coli* cells with fluorescently labeled lactose permease molecules are observed to undergo all-or-none induction. Upon closer inspection with single-molecule resolution (inset), many uninduced cells are found to have a handful of permease molecules. B) Time courses of induction of the wild-type *lac* operon (——) upon the addition of 40 μm TMG show a threshold of initial permease molecules required for induction (——). When DNA looping is knocked out by eliminating auxiliary operator sites (----), induction occurs independent of the initial permease number. The probability of induction as a function of the initial number of permease molecules is well fit by a Hill function (right). A) From Choi et al.,[1] reprinted with permission. Copyright AAAS 2008.
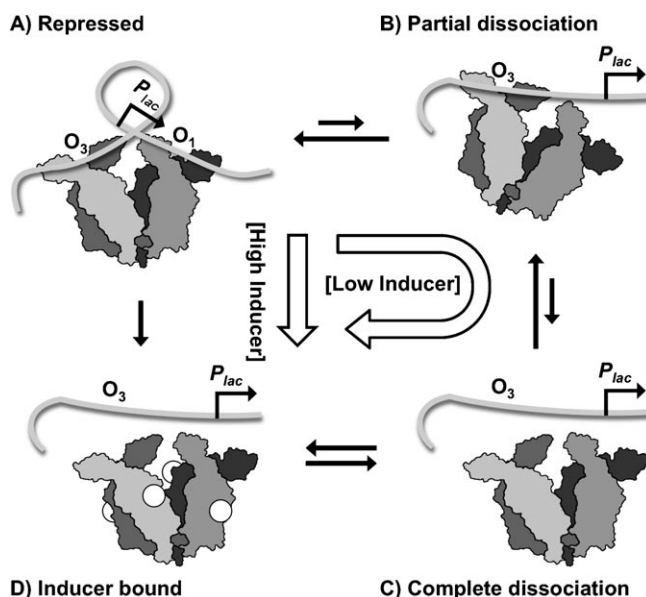

Figure 2. The phenomena observed by Choi et al. point to a model for induction of the *lac* operon in which DNA looping (to O_2 or O_3) stabilizes the repressed promoter state (A). Random partial dissociations of the tetrameric *lac* repressor from O_1 (B) lead to small bursts of expression, usually generating only a single mRNA molecule and a handful of proteins. Very rare complete dissociation (C) can lead to the transcription of multiple mRNAs and the production of the hundreds of lactose permease molecules required for induction if inducer is present to sequester *lac* repressor (D). The *lac* operon can be directly induced if inducer concentrations are high enough to bind and overcome a reduced affinity for DNA-bound *lac* repressor. Modeled by using VMD[21] and an all-atom repressor model.[22]

the induction threshold. However, it should be noted that complete dissociation alone might not be sufficient to establish induction. According to Elf et al., it takes at most a few minutes for a free *lac* repressor to find its binding site inside an *E. coli* cell,[17] and those few minutes might not be long enough to generate the few hundred permease molecules required. Only when the *lac* repressor is prevented from rebinding long enough in the presence of an inducer can many transcripts and hundreds of protein molecules be produced. Additionally, positive feedback from lactose permease should not be underestimated, as it likely extends the duration of large bursts: Choi et al. show large bursts lasting many tens of minutes, during which newly expressed permeases would further increase intracellular inducer concentration, facilitating induction.

One important feature of this model is that the cell's commitment to induction stems from a single, stochastic dissociation event of *lac* repressor. This is because large bursts occur very rarely (once in many cell cycles) and are uncorrelated with each other. Permease molecules generated during bursts falling short of the induction threshold do not accumulate as they are diluted by cell growth. Induction only becomes likely when the protein generated in a single burst crosses the threshold. This interpretation introduces a novel mechanism for stochastic decision making in bacteria. Although recent work casts doubt as to whether the *lac* operon exhibits all-or-none induction in response to its natural inducer lactose instead of a nonmetabolizable inducer,[18, 19] the stochastic molecular interactions revealed in this work still hold.

The finding by Choi et al. has obvious implications for those developing mathematical models of *lac* induction. Current models treat *lac*-repressor binding in equilibrium.[20] The *lac* repressor is present at very low concentrations (ca. 10 molecules per cell), so explicit accounting of stochastic noise in repressor binding cannot be ignored here. Future models of the *lac* operon should aim to reproduce the induction threshold observed by Choi et al. as well as the stochastic gene-expression bursts that give rise to induction.

The discovery that *lac* operon induction by TMG arises from a stochastic single molecular event has the potential to change our view of how a cell functions. Many biologically significant molecules, such as regulatory proteins, DNAs, and mRNAs, exist in small copy numbers with inherently noisy interactions between them. How do cells deal with this? Choi et al. have shown a new way in which cells can actually exploit stochasticity to drive phenotypic change. Is this a commonly used mechanism to govern decision making in cells? And more broadly, when are such stochastic mechanisms preferred and when are they eschewed? Clearly, this will depend upon the task cells have at hand. More work will allow us to answer these questions, and the ability to probe the dynamics of single molecules in real time in a living cell will provide unprecedented insight. One thing that is clear is that unlike Hamlet, who was torn between life and death, it is apparently not difficult for an *E. coli* cell to choose to become induced or not. It just closes its eyes and rolls the dice.

[1] P. J. Choi, L. Cai, K. Frieda, X. S. Xie, *Science* **2008**, *322*, 442–446.

[2] H. Maamar, A. Raj, D. Dubnau, *Science* **2007**, *317*, 526–529.

[3] E. Kussell, R. Kishony, N. Q. Balaban, S. Leibler, *Genetics* **2005**, *169*, 1807–1814.

[4] T. Enver, C. M. Heyworth, T. M. Dexter, *Blood* **1998**, *92*, 348–351; discussion p. 352.

[5] R. J. Johnston, Jr., C. Desplan, *Curr. Opin. Neurobiol.* **2008**, *18*, 20–27.

[6] M. Thattai, A. van Oudenaarden, *Genetics* **2004**, *167*, 523–530.

[7] M. Acar, J. T. Mettetal, A. van Oudenaarden, *Nat. Genet.* **2008**, *40*, 471–475.

[8] E. Kussell, S. Leibler, *Science* **2005**, *309*, 2075–2078.

[9] M. Ackermann, B. Stecher, N. E. Freed, P. Songhet, W.-D. Hardt, M. Doebeli, *Nature* **2008**, *454*, 987–990.

[10] F. Jacob, J. Monod, *J. Mol. Biol.* **1961**, *3*, 318–356.

[11] A. Novick, M. Weiner, *Proc. Natl. Acad. Sci. USA* **1957**, *43*, 553–566.

[12] A. Becskei, B. Seraphin, L. Serrano, *EMBO J.* **2001**, *20*, 2528–2535.

[13] L. Cai, N. Friedman, X. S. Xie, *Nature* **2006**, *440*, 358–362.

[14] J. Yu, J. Xiao, X. Ren, K. Lao, X. S. Xie, *Science* **2006**, *311*, 1600–1603.

[15] S. Oehler, E. R. Eismann, H. Kramer, B. Muller-Hill, *EMBO J.* **1990**, *9*, 973–979.

[16] R. Schleif, *Annu. Rev. Biochem.* **1992**, *61*, 199–223.

[17] J. Elf, G. W. Li, X. S. Xie, *Science* **2007**, *316*, 1191–1194.

[18] E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, A. van Oudenaarden, *Nature* **2004**, *427*, 737–740.

[19] M. J. A. van Hoek, P. Hogeweg, *Biophys. J.* **2006**, *91*, 2833–2843.

[20] M. Santillán, M. C. Mackey, *J. R. Soc. Interface* **2008**, *5 Suppl. 1*, S29–39.

[21] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33–38, 27–28.

[22] E. Villa, A. Balaeff, K. Schulten, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6783–6788.