Reprint Series
4 August 1995, Volume 269, pp. 650–656

SCIENCE

Circuit Simulation of Genetic Networks

Harley H. McAdams and Lucy Shapiro

Copyright © 1995 by the American Association for the Advancement of Science

Circuit Simulation of Genetic Networks

Harley H. McAdams and Lucy Shapiro

Genetic networks with tens to hundreds of genes are difficult to analyze with currently available techniques. Because of the many parallels in the function of these biochemically based genetic circuits and electrical circuits, a hybrid modeling approach is proposed that integrates conventional biochemical kinetic modeling within the framework of a circuit simulation. The circuit diagram of the bacteriophage lambda lysis-lysogeny decision circuit represents connectivity in signal paths of the biochemical components. A key feature of the lambda genetic circuit is that operons function as active integrated logic components and introduce signal time delays essential for the in vivo behavior of phage lambda.

Genetic networks that include many genes and many signal pathways are rapidly becoming defined in prokaryotes and eukaryotes. As network size increases, intuitive analysis of feedback effects is increasingly difficult and error prone. Electrical engineers routinely analyze circuits with thousands of interconnected complex components. Electromechanical devices switch in small fractions of a second, and common transistor circuits can operate at more than 10^8 cycles per second. In contrast, the protein signal–controlled switching rate in genetic circuits is around 10^{-2} per second. Although there is a great disparity in time scales between genetic switching circuits and electrical switching circuits, there are many parallels in their function. These similarities lead to the question: Which electrical engineering circuit analysis techniques are applicable to genetic circuits that comprise tens to hundreds of genes?

The conducting pathways between components determine the connectivity of electrical circuits. The connectivity of genetic circuits is determined by the connection between the source of a protein signal and its site of action established by its site-specific biochemical address. Site-specific biochemical addressing permits many genetic circuits to operate in parallel within the same small cell volume. Thus, the cell can achieve high computational density in terms of operations per second per cubic centimeter and the instantaneous amount of genetic computation within any living organism is enormous, in spite of slow switching rates.

Electrical circuits are typically described by circuit diagrams and characterized by simulation models. The simulation provides a calculating tool for predicting time behavior of the interconnected system. The circuit diagram shows the overall organization of the circuit and the detailed interconnectivity between components. We now apply this perspective to the genetic circuit used by bacteriophage lambda (λ) to choose between lysis and lysogeny (*1–5*). After infecting a host *Escherichia coli* cell, phage λ either propagates as a prophage integrated into the host DNA (lysogeny) or becomes an actively replicating virus (lysis). The λ decision circuit, which controls one phase of a single phage life cycle and operates synergistically within a single *E. coli* cell, is perhaps the most completely characterized complex genetic network. Many genetic circuits have broader scope. For example, elements of several cells may be components of a circuit that controls the life cycle of the different cell types (*6*) and cell generations (*7*). Symbiotic relations between bacteria and higher organisms involve genetic circuits that cross species boundaries (*8*).

Biochemical feedback plays an essential role in cellular regulation (*9*), and algebraic formalisms for analysis of biological networks represented as asynchronous automata have been proposed (*10*). As with electrical networks, however, the algebraic approach quickly becomes obtuse for all but the simplest networks. Short-term dynamics of small-scale biochemical reaction networks that control physiological mechanisms in bacteria have been simulated (*11*), and modeling of metabolic biochemical reaction networks is well advanced (*12*). A hybrid approach is needed to integrate conventional biochemical kinetic modeling with models of control and delay mechanisms in large genetic circuits.

Signal Timing in Genetic Networks

Electrical switching circuits are frequently characterized as networks of idealized switching devices; that is, devices with instantaneous transitions between states at precise times. However, practical electrical devices exhibit finite transition times and transient responses. The idealization of practical electronic devices permits simplified characterization of switching circuits based on Boolean logic while retaining the observed behavior of the system. In a parallel manner, a Boolean gate representation can characterize biochemical repression or activation of transcriptional promoter elements when the switching action is definitive and relatively fast. (The term "gate" refers to a circuit element that outputs a signal when its required input conditions are satisfied. Boolean algebra treats relations between logical variables with the values TRUE and FALSE.) Transcription elongation and translation control mechanisms augment promoter-based logic to determine expression of specific genes. Biochemical mechanisms described below that determine the dynamic balance between protein production and decay, and thus determine signal levels, are important parameters in genetic circuit logic. Time delay mechanisms, especially transcription delays and signal accumulation delays, are central to the correct function of the circuits.

Figure 1A shows the interplay of these mechanisms in a hypothetical genetic circuit with two promoters, three genes, and a termination site. The time evolution of signals in this circuit (Fig. 1B) is determined

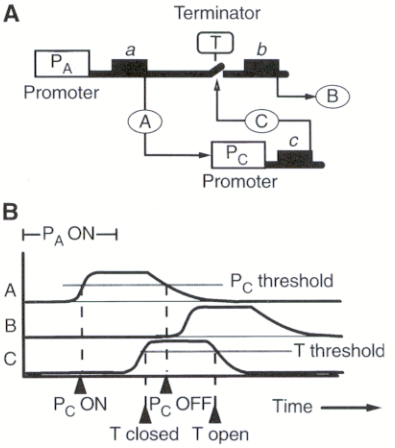

Fig. 1. (**A**) Circuit diagram of a hypothetical genetic network. Bold line indicates the RNAP path on the DNA after binding at P_A. Promoter P_C is ON when protein signal A is ON; that is, at effective concentration at its site of action. Switch T, which is usually open but closes to permit RNAP passage in the presence of the protein signal C, models terminator site function. (**B**) Timing diagram of signals A, B, and C. Time lags depend on the gene spacing on the DNA, transcription rates, the time required for accumulation of an effective signal concentration, and protein decay rate constants.

H. H. McAdams is in Palo Alto, CA 94305, USA. L. Shapiro is in the Department of Developmental Biology, Beckman Center, Stanford University School of Medicine, Stanford, CA 94305, USA.

by transcription time delays resulting from the rate of transcription, from delays while transcription is blocked at the terminator site, and from protein signal accumulation delays. Transcription of the operon containing genes *a* and *b* is initiated at time zero when RNA polymerase (RNAP) binds at promoter P_A, an open complex is formed, and RNAP transcribes to and through the gene (*a*) encoding protein A. Time delays in this operon result from (i) the rise and fall times of the protein signals A and C, and (ii) the time required for RNAP to transcribe the operons initiated at P_A and P_C. The time from P_A ON to P_C ON, for example, is the sum of these two times. The initial rate of signal protein production is proportional to the product of the transcript production rate under the prevailing repression or activation condition of the promoter, the number of proteins translated from each transcript, and the multiplicity of infection (MOI). After open complex formation at the promoter site, the rate and extent of subsequent transcription are actively controlled by pause sites and by termination sites. Transcription delays can range from a few seconds to several minutes, depending on the distance and the average transcription rate (*13*).

After appearance of the first gene *a* transcript, an additional delay is required to achieve a concentration of A sufficient to turn on P_C at an effective rate. Signal protein C controls terminator switch T. When T closes (is antiterminated), transcription can continue through *b* to produce signal protein B. The time from P_C ON to the closing of T is the sum of the delay time attributable to RNAP movement plus the rise time of the C protein signal. Determinants of the time from initiation of binding at P_A to initiation of B production are more complex: Both the RNAP travel time along the DNA from P_A to the end of *b* and the delaying effect of events that influence terminator switch T must be included.

When the controlling signal concentrations change and a promoter turns off, a pipeline of RNAPs must be cleared before transcription ceases. This latent signal capacity, which continues until transcript translation ends, must be considered when modeling the circuit.

Steady-state signal protein concentrations are determined by the dynamic balance between protein production and degradation (Fig. 2A, Eq. 1). [Equation numbers refer to equations for signal dynamics in Fig. 2A (*14*).] A short signal protein half-life results in low steady-state signal levels and a short time to steady state (Fig. 2A, Eqs. 7, 9, and 10). Cells actively control protein signal degradation rates and thus steady-state signal levels. For example, active control of degradation is central to the control of protein CII.

Feedback Circuit Dynamics

Two feedback circuits involving proteins CI and Cro, with differing configurations and dynamics, are critical design elements of the λ decision circuit. The CI feedback loop shown in Fig. 3 is a self-regulating circuit. When conditions are favorable for the lysogenic path, CII increases to a level that turns on promoter P_{RE} (Fig. 3A). There is a transcription delay before production of the first *cI* transcript. Translation of this and successive transcripts adds to the cumulative CI signal; simultaneously, protein degradation reduces the signal. The instantaneous signal change rate, $\Delta CI(t)$, is determined by these processes, and the current signal level, $CI(t)$, is the time integral of $\Delta CI(t)$. Initially, when the CI signal level is low, P_{RM} does not initiate transcripts, but there is a high level of transcription from P_{RE}. The CI concentration rapidly increases to a level sufficient for P_{RM} to initiate transcription (Fig. 3C) and to sustain CI production after P_{RE} becomes inactive when the CII signal decreases. At the steady-state CI concentration [~140 to 200 molecules per cell (*15*, *16*)], the rate of transcription from P_{RM} results in a rate of CI production equal to that of CI degradation. This steady-state CI concentration represses P_R and P_L (Fig. 3B) to prevent induction of the lysogen and to prevent transcription from additional λ phages that may infect the host promoter. Promoter P_R is repressed at a CI concentration much less than the steady-state concentration and, consequently, is switched off rapidly. P_L dynamics are similar. The switching of P_L and P_R is definitive and rapid, and, hence, Boolean in character.

When the lytic path is favored, negative feedback loop mediated by Cro repression of P_R controls signal levels and timing. The sustaining level of transcription in the CI loop described above produces only CI molecules. However, when the Cro loop is active, the sustaining transcripts initiated at P_R extend beyond *cro* to other genes (Fig. 4). The dynamics of the Cro feedback loop as calculated from Eqs. 5, 7, 10, and 11 in Fig. 2A are shown in Fig. 2, B to E. The steady-state signal level is lower with feedback than without (Fig. 2B). The steady-state signal protein level and transcript initiation rate increase with MOI and there is an initial transient burst of transcription for higher MOI (Fig. 2, C to E). The steady-

A

Signal protein concentration is determined by

$$\text{protein concentration change} = (\text{protein production}) - (\text{protein loss}) \quad (1)$$

With no repression by protein produced (no feedback), Eq. 1 becomes

$$dL_N/dt = k_p - k_d L_N(t) \quad (2)$$

$$= \text{MOI}\ A_{GENE}\ OC_{MAX} - [\ln(2)/T_{1/2}]\ L_N(t) \quad (3)$$

With repression by the protein produced (through feedback) Eq. 1 becomes

$$dL_{FB}/dt = k_p\,[1 - Y(t)] - k_d\,L_{FB}(t) \quad (4)$$

$$= \text{MOI}\ A_{GENE}\ OC_{MAX}\ \{1/[1 + L_{FB}(t)/A_g\ V_{cell}\ K_E] - [\ln(2)/T_{1/2}]\ L_{FB}(t)\} \quad (5)$$

Eqs. 4 and 5 must be integrated numerically. With $L(0) = 0$, Eqs. 2 and 3 integrate to

$$L_N(t) = (k_p/k_d)\,[1 - \exp(k_d\,t)] \quad (6)$$

$$= [\text{MOI}\ A_{GENE}\ OC_{MAX}\ T_{1/2}/\ln(2)]\ \{1 - \exp[-\ln(2)\,t/T_{1/2}]\} \quad (7)$$

Solving for Eqs. 2, 3, and 5 for steady-state concentration when $dL/dt = 0$,

$$L_{NSS} = k_p/k_d \quad (8)$$

$$= (\text{MOI}\ A_{GENE}\ OC_{MAX}\ T_{1/2})/\ln(2) \quad (9)$$

$$L_{FBSS} = (A_g\ V_{cell}\ K_E/2)$$

$$(-1 + \{1 + [4\ \text{MOI}\ A_{GENE}\ OC_{MAX}\ T_{1/2}]/[K_E\ A_g\ V_{cell}\ \ln(2)]\}^{1/2}) \quad (10)$$

With feedback, the steady-state average rate of transcript initiation (OC_{FBSS}) is

$$OC_{FBSS} = [\ln(2)\ L_{FBSS}]/(A_{GENE}\ T_{1/2}) \quad (11)$$

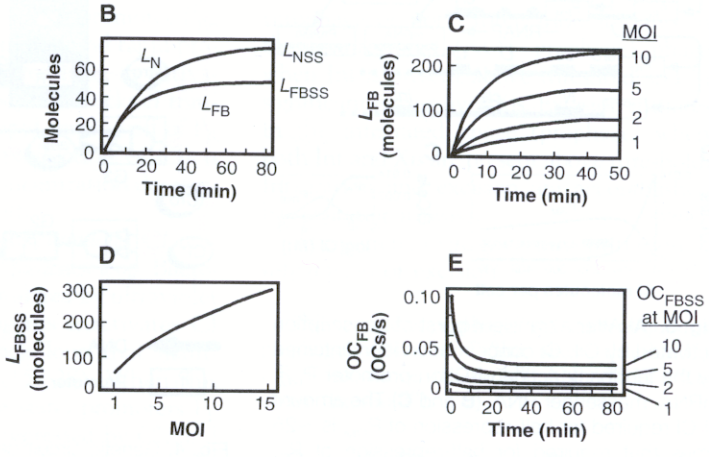

Fig. 2. (**A**) Equations for signal dynamics in simple genetic circuits. A_g, Avogadro's number; V_{cell}, *E. coli* cell volume; OC_{MAX}, unrepressed open complex formation rate; A_{GENE}, average number of proteins per transcript; K_E, equilibrium constant for repressing protein-promoter interaction; $T_{1/2}$, protein half-life; k_p (= MOI $OC_{MAX}A_{GENE}$), unrepressed protein synthesis rate; k_d [= $\ln(2)/T_{1/2}$], protein degradation rate; $L_N(t)$, protein concentration with no feedback; $L_{FB}(t)$, protein concentration with feedback (L_{NSS} and L_{FBSS} are steady-state values); $Y(t)\{= L(t)/[L(t) + K_E]\}$, fractional repressor occupancy (*37*). (**B** to **E**) Circuit dynamics with parameters approximating the Cro feedback loop: $V_{cell} = 1.41 \times 10^{-15}$ liters; $OC_{MAX} = 0.1$ per second; $A_{GENE} = 6$; $K_E = 1.2 \times 10^{-7}$ M; MOI = 1, unless otherwise indicated; $T_{1/2} = 15$ min.

state protein concentration from each downstream gene in the P_R operon will be proportional to the product of the characteristic proteins/transcript ratio of each gene and the half-life of the resulting protein (Fig. 2A, Eq. 9). This product can vary widely from gene to gene in the same operon. Promoters P_L and P_{RM} are switched rapidly with a design similar to that of P_R and P_L switching by the CI loop. These two feedback circuits support both rapid Boolean switching (of P_R, P_L, and P_{RM}) and controlled, sustained synthesis from several genes (from P_R on the lytic path) when needed.

Time Delays

The transcription delay between a promoter and each gene is $NT_{PG}/\langle R_T \rangle$, where NT_{PG} is the nucleotide count from the promoter to the end of the gene and $\langle R_T \rangle$ is the average rate of transcript elongation. The initial rate of signal protein increase is linear and proportional to MOI (Fig. 2A, Eqs. 3 and 5; Fig. 2, B and C). Thus, the total time delay (T_{delay}) from promoter ON to effectiveness at the site of action of a gene is approximately given by

$$T_{delay} = NT_{PG}/\langle R_T \rangle + (K_s A_g V_{cell})/(\text{MOI } A_{GENE}\, OC_{MAX})$$

where K_s is the equilibrium constant for the signal protein–site of action interaction (see Fig. 2A for the definition of other terms). This equation applies when K_s is much smaller than the steady-state signal concentration. The second term can be relatively small at higher MOIs (Fig. 2C) or for switching configurations designed for speed (for example, P_R and P_L control by CII in Fig. 3). A transcription delay mechanism assures time delays necessary for correct circuit function even at high MOI. For example, gene *Q* is ~6500 nucleotides from P_R, requiring several minutes for transcription (*6*). In the second term, the K_s/A_{GENE} ratio can differ widely for genes in the same operon, contributing to the flexibility of the operon structure as a genetic circuit control element.

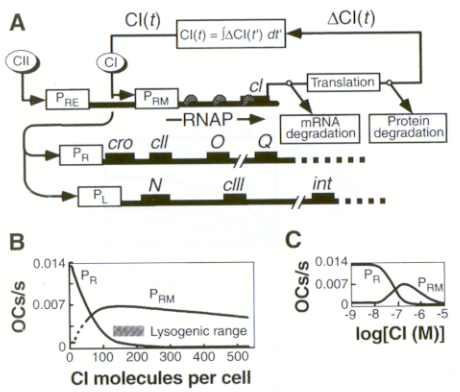

Fig. 3. (**A**) After a transient burst of transcription activated by CII, CI concentration is maintained by the feedback action of CI on promoter P_{RM}. mRNA, messenger RNA. (**B** and **C**) The amount of CI required for half-repression of P_{RM} is ~25 times that required for half-repression of P_R. Promoter P_L is also repressed at a low CI concentration. Because the initial rate of CI production is high and both P_L and P_R are repressed at well below the steady-state CI concentration, this circuit design produces rapid and definitive switching at P_L and P_R.

Lambda Decision Circuit

The λ genetic circuit that determines the course of the phage infection is shown in Fig. 4 (*17*). A switch selects between the two stable configurations determined by the CI- or Cro-based negative feedback loops (*4*). The physiological state of the bacterial host and the MOI together bias the switch toward latching in one state if conditions favor lysogeny and the other if conditions favor lysis (Fig. 4, gate G1).

Electrical engineers avoid "races" (simultaneously changing signals along two different, but interacting, signal paths) in switching circuit designs; they especially avoid "critical" races (a race condition in which the outcome differs depending on which path completes first). A central element of the λ decision circuit is the critical race created by the competitive buildup of CI and Cro in which the outcome determines whether the phage will integrate into the host DNA or begin replication, resulting in lysis. The phage circuit design creates a "fuzzy logic" mechanism by integrating dependence on internal health (state of nutrition) and extracellular environment (MOI) into the stability, rate of growth, and steady-state concentration of CII and CIII that determines the logical outcome at gate G8.

(G1) P_R and P_{RM} promoters control two-state switch. After λ infection of *E. coli*, transcription initiates at P_R and P_L (*1*, *2*, *4*). Translation of the P_R transcript produces Cro. Promoter P_{RM} is initially OFF. Transcripts initiated at P_R and P_L induce a cascade of events that result in rapid production of CI if environmental conditions favor CII stability, and low or no CI production if not (gate G8). Cro and CI bind competitively and in sequence, but in opposing order, to three sites (OR_1, OR_2, and OR_3) on the λ DNA. As the CI concentration increases, P_R is repressed by CI at OR_1, after which P_{RM} is stimulated by CI at OR_2 and repressed by CI at OR_3 (Fig. 2C). As the Cro concentration increases, P_{RM} is repressed by Cro at OR_3, and P_R is then repressed by Cro at OR_2 or OR_1. Eventually, either the CI or the Cro feedback loop is locked on determining the lytic or lysogenic path choice. The overall effect is of an integrated logic component providing a bistable two-state switch. The switch set-

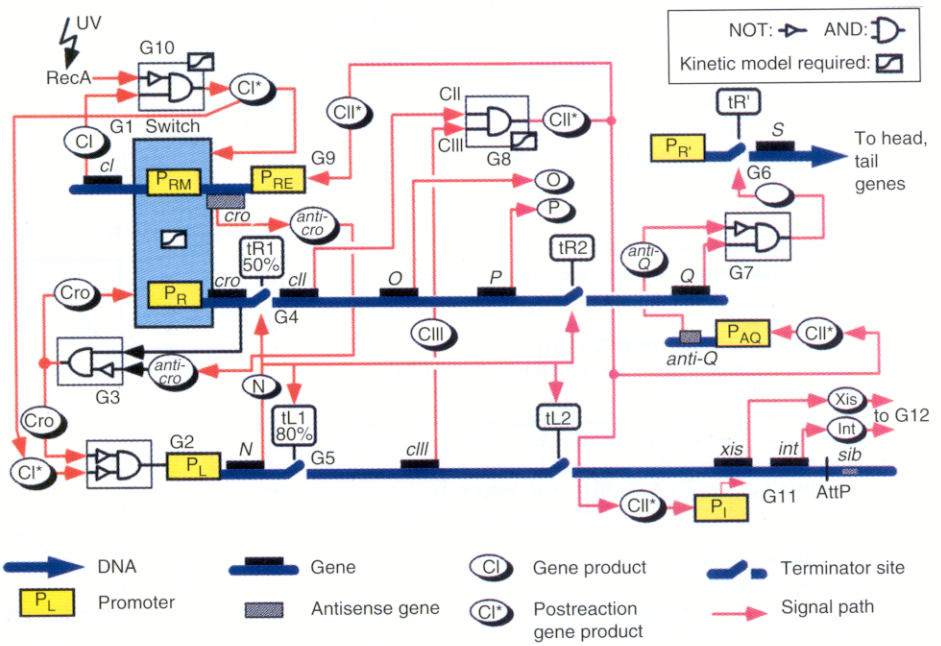

Fig. 4. Genetic circuit determining the phage λ lysis-lysogeny decision. The early λ genes are shown configured in operons relative to controlling promoters. The right- or leftward orientation of the operons does not necessarily correspond to the orientation on the chromosome. Protein signal paths connect gene product source and site of action. Intermediate modulating reactions (at G8, for example) and the associated logic are shown. The rectangles enclosing only Boolean logic symbols identify control logic amenable to approximation as Boolean logic. The sigmoid curve–in-rectangle symbol identifies control logic requiring kinetic modeling. UV, ultraviolet; cI* and cII*, effective signal levels at site of action.

ting is established unambiguously if either Cro or CI is present in sufficient excess. On either path, early genes are suppressed after the path decision.

(G2) Negative feedback through P_L promoter. Promoter P_L is repressed by either CI* or Cro after the path choice decision.

(G3) Cro production is reduced by transcription from P_{RE}. Transcription from P_{RE} reduces Cro production either because antisense *cro* sequences quench *cro* messenger RNA (mRNA) translation (*4*, p. 58) or because convergent transcription reduces P_R transcription (*18*). This reduction in Cro production favors establishment of lysogeny.

(G4) Terminators tR1 and tR2 control transcripts initiated at P_R. Antitermination of tR1 allows transcription initiated at P_R to continue well past *cro* (*19*). Initially, the bypass permits ~50% of the incident RNAPs to pass the terminator site and continue transcription to tR2, which reduces, but does not eliminate, P_R-initiated O and *P* transcripts. The N antiterminator protein produced after P_L is ON interacts with a site on the mRNA to convert the RNAP to a termination-resistant form that remains termination resistant 5 to 10 kb downstream (*20*). Transcription is definitely blocked at tR2 after the path decision is made and P_L is OFF. The tR2 termination site then blocks RNAPs in the pipeline, accelerating cessation of production of Q.

(G5) Terminators tL1 and tL2 control transcripts initiated at P_L. Terminator tL1 modulates transcription through *cIII*, *xis*, *int*, and beyond (*19*). In the absence of the N antitermination protein, transcription is 80% blocked at tL1 and completely blocked at tL2. The antiterminated form of RNAP produced by interaction with N also transcribes through a termination site located beyond *int* in the unintegrated viral DNA. The P_I-initiated transcript terminates at that point. This difference between P_L- and P_I-initiated RNAPs is crucial to the regulation of *int* and *xis* as described below (gate G11).

(G6) Q protein antiterminates tR′ to allow transcription of lysis, head, and tail genes. The Q protein antiterminates terminator tR′ so that the $P_{R'}$-initiated transcript will continue through *S* and downstream genes that encode cell lysis proteins and head and tail coat proteins (*21*).

(G7) Q production is quenched by antisense mRNA from anti-Q. Transcription of *anti-Q* is activated by CII (*22*). Promoter P_{AQ} is located within gene Q and initiates reversed transcription of the Q gene. The resulting antisense Q mRNA quenches translation of the Q mRNA to prevent antitermination of tR′ by Q, to block expression of *S* and downstream genes (gate G6), and to assure establishment of lysogeny.

(G8) CIII and CII dependence on MOI and active protein stability control determine lysogeny choice. CII activates three promoters essential for lysogeny: P_{RE} (gate G9), P_I (part of gate G11), and P_{AQ} (gate G7) (*2*, *23*–*25*). Because CI production is locked on by the feedback loop through P_{RM}, CII is necessary for establishment, but not for maintenance, of lysogeny. CIII protects CII from degradation by host proteases; the half-life of CII is ~5 min in the presence of CIII and <1 min in its absence (*23*). Thus, the absence or presence of CIII affects the concentration of CII by a factor of 5 (Fig. 2A, Eq. 9).

An MOI of ≤2 is required for lysogen production in exponentially growing cells, and the percentage lysogeny increases markedly with increasing MOI up to an MOI of ~7 (*26*). At an MOI of 1, both CII and CIII are produced at a low level and the steady-state concentration of CII is apparently too low to activate P_{RE}. (In an induced lysogen, the MOI is 1, so CII does not activate P_{RE} to restore the CI feedback loop after induction.) The production rate of both CII and CIII is higher with a higher MOI (Fig. 2A, Eq. 9; Fig. 2E), and the half-life of CII also increases as a result of the higher CIII concentration. Thus, because of the multiplicative effect of MOI and half-life on the steady-state concentration (Fig. 2A, Eq. 9), the steady-state concentration of CII at an MOI of 10 should be 40 to 50 times that at an MOI of 1. The characterization of the CII-CIII interaction as an AND gate in Fig. 4 is only valid at high MOI; full characterization requires a kinetic model.

Infection of a cell population at some average MOI produces a distribution of MOIs across the population. Because signal growth rates, steady-state concentrations, and timing depend markedly on MOI (Fig. 2A), the result is a distribution of observed lysis-lysogeny outcomes.

(G9) Promoter P_{RE} activates cI transcription. CII activates the strong promoter P_{RE} to initiate transcription of *cI* and rapid production of CI (Fig. 3), resulting in activation of the feedback loop that locks *cI* transcription on and the rapid turning off of P_R and P_L (*1*, *27*). The production of CII is then halted, the CII concentration decreases, and P_{RE} is turned off.

(G10) Ultraviolet light stimulates CI proteases, breaking the P_{RM} feedback loop. Bacterial RecA protease stimulated by ultraviolet radiation and other agents inactivates CI and breaks the feedback loop that maintains *cI* transcription (*15*, *28*). When this occurs in the prophage state, promoters P_L and P_R are no longer repressed by CI, and events that result in excision, phage replication, and lysis are initiated.

(G11) State-dependent logic controls Int and Xis production. The control of *int* and *xis* is dependent on whether the phage is integrated into the host bacterial DNA (*2*). The control of Int and Xis is mediated in part by a topologically determined mechanism that depends on the state of the phage DNA (prophage or not). In the unintegrated phage DNA, the *sib* region located downstream of *int* is transcribed by the N-antiterminated RNAP initiated at P_L. The *sib* portion of this mRNA facilitates nuclease attack followed by sequential destruction of *int* progressing back toward *N* (retroregulation), so that Xis is preferentially produced by the P_L-initiated transcript. In contrast, the P_I-initiated transcript does not transcribe a complete *xis* mRNA and terminates before *sib*, thus producing only Int. The attachment point (AttP) of the phage is between *int* and *sib*, so that, during λ induction, P_L-initiated RNAP transcribes through *int*, but the resulting mRNA does not contain the *sib* region. The location of *xis*, well separated from P_L on the DNA, assures a time delay to permit execution of the lysogenic logic when conditions are correct for that decision.

The genetic logic produced by these biochemical mechanisms that govern Int and Xis production is

"Int is produced" if "CII* is above threshold" OR ("state is prophage" AND "P_L-initiated RNAP is present").

"Xis is produced" if "P_L-initiated RNAP is present" AND NOT "CII* is present."

(G12) State-dependent logic based on Int and Xis controls integration and excision. Relative concentrations of Xis and Int control phage integration and excision (Fig. 5) (*29*–*31*). Integration requires only the Int protein among the phage-specified proteins; Xis inhibits integration. Excision requires both Int and Xis; less Int is required for excision than for integration. As in gate G11, the control logic depends on the state of the phage DNA (prophage or not). Integration initiates in the unintegrated (NOT prophage) phage when Int is present above threshold and in significant excess over Xis. Excision initiates in the integrated phage (prophage) when both Int and Xis are present above threshold. Integration and excision are accomplished by separate pathways.

The genetic logic governing initiation of integration or excision is

"Next state is prophage" if ("current state is prophage" AND NOT "Int above threshold") OR ("Int above threshold" AND NOT "Xis above threshold").

Genetic Circuits as Sequential Logic Circuits

The logic circuit diagram in Fig. 5A is an equivalent symbolic representation of the genetic logic statements for gates G11 and

G12 given above. The logic controlling integration and excision depends on the state of the phage (prophage or not). Thus, the λ prophage state functions as a long-term memory mechanism in the logic circuit. Time delays in the circuit provide short-term memory.

A sequential circuit is defined as a circuit in which the output of the circuit depends not only on current inputs, but also on the stored state of the circuit consistent with the model in Fig. 5B. Comparison of Figs. 5A and 5B shows that the logic configuration of G11 and G12 fits the sequential circuit model. Sequential circuits in which the memory results from time delays are asynchronous sequential circuits. The analysis and design of a large fraction of all electrical logic circuits, including virtually all digital computation circuits, are based on elaboration of the sequential circuit paradigm.

The checkpoint phenomena observed in cell cycles (*32*) represent another parallel between well-known electrical logic circuit phenomena and genetic circuits. When the external input to a sequential circuit changes, the combined input vector (stored plus external input values) to the combinational logic results in a new output vector (new external outputs plus next-state values). The next-state values are stored and, after a time delay, the output values from the memory representing the present state change to the next-state values (Fig. 5B). The new combined input to the logic may result in another change in the outputs; this cycle will continue until the circuit settles into a stable state in which the current state equals the next state. A sequential circuit may pass through several of these transitional states before reaching a stable state. Initiation of a transition out of a stable state depends on receipt of changed inputs. When the changed inputs reflect completion of necessary precursor actions, this sequential circuit mechanism of operation is analogous to the checkpoint control mechanism.

Across the lambdoid phages, completely different proteins, with different mechanisms of action, perform analogous circuit functions (*31*, *33*). In these instances, genome organization and the pattern of gene grouping into modules with similar function is highly conserved (*34*). For electrical circuits, there are always many equivalent alternative implementations of any logic function based on alternative choices of components or design details. Engineers select among these implementations on the basis of criteria such as cost, reliability, or power consumption. Similarly, evolution selects among alternative biochemical implementations of logic functions, but the selection criteria relate to survival value. We conjecture that the phage logic circuit design is the most conserved element that defines the lambdoid phage "species," with each specific phage representing an alternative implementation of the lambda logic design—"the lambda algorithm." The biochemical design of any individual lambdoid phage then represents optimization of the implementation of the common lambda logical design for a specific host environment.

Genetic circuits exhibit hierarchical organization: Regulons control operons, which control gene groupings. Electronic circuit designers structure complex systems as hierarchical structures to facilitate reuse of modular functions and simplified control by a few signals. The multigene genetic subfunctions in the hierarchy are points of high leverage for evolutionary adaptability because a single mutation in circuit logic can change the control of a large genetic cascade, thereby amplifying evolutionary consequences. The evolutionary consequences of rearrangement of modular functions in the genome by homologous recombination are well documented (*34*). Relocation of a single gene that encodes a controlling signal protein can change connectivity and, hence, the circuit logic, resulting in a radical effect on timing, duration of effect, or sequencing of the controlled subcircuits. Identification of the circuit-level organization of genetic circuits, as in Figs. 4 and 5, together with established methods for logic circuit analysis, will provide a functional framework for analysis of such large-scale reorganization of genetic logic.

Verification of Decision Circuit Logic

Conventionally, biochemical simulations have emphasized modeling coupled kinetic equations. Electrical circuit simulations emphasize the circuit connectivity and the functionality of the circuit components, such as resistors, capacitors, and transistors. Consideration of electrical circuit simulations suggests a hybrid approach to genetic circuit modeling that integrates the following ideas with kinetic models: (i) identify the circuit connectivity and model point-to-point signal paths, (ii) simplify transcription control logic by treating it as Boolean logic when justified, and (iii) model the functionality of complex or nonlinear control elements in specialized subroutines.

In many instances, the signal path dynamics can be modeled (Fig. 2) with four parameters: (i) rate of signal protein transcript production, (ii) average proteins/transcript, (iii) signal protein half-life, and (iv) the equilibrium constant for the signal protein–site of action interaction. Past experimental effort has not been directed toward characterizing the point-to-point

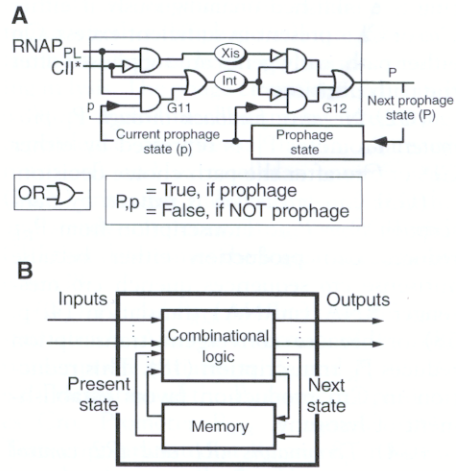

Fig. 5. (**A**) The biochemical logic controlling phage λ integration into or excision from host DNA depends on the current state of the phage (whether the phage is an integrated prophage or not). RNAP$_{PL}$, antiterminated RNAP initiated at P$_L$. (**B**) Sequential circuit model. By definition, sequential logic circuits are configured from logic and memory elements. Sequential circuit outputs depend on both the value of current external inputs and stored values of past outputs. The decision circuit of phage λ includes biochemically based sequential subcircuits.

links, per se, under controlled conditions, because these parameters were not viewed in the circuit context presented here. As a result, we find solid data for only two of the four needed parameters for most links in Fig. 4, and three of four for a few (*35*). We expect that this situation will improve as the importance of understanding the signal paths and timing in complex genetic networks is widely realized. Because overall genetic circuit functions are highly interdependent, an integrated simulation will allow inference of missing protein parameters from a combined consideration of the circuit design, the known parameters, and the overall timing and outcome of the circuit logic under varying conditions. This capability to exploit the interdependence of circuit elements will be a major benefit from simulating the circuits.

Operons function as key signal-generating and control components in genetic circuits and are candidates for modeling with specialized subroutines. The operon model and associated routines must treat (i) promoter control logic, (ii) promoter activity (transcript initiation rate), reflecting repressor or activator kinetics, (iii) operon layout (gene location), (iv) any elongation control (for example, antitermination mechanisms), (iv) posttranscriptional controls, and (v) translation efficiency (for example, average proteins/transcript) for each gene. Common control of one or more genes by several promoters [for example, common control of *cI* by P$_{RE}$ and P$_{RM}$ (Fig.

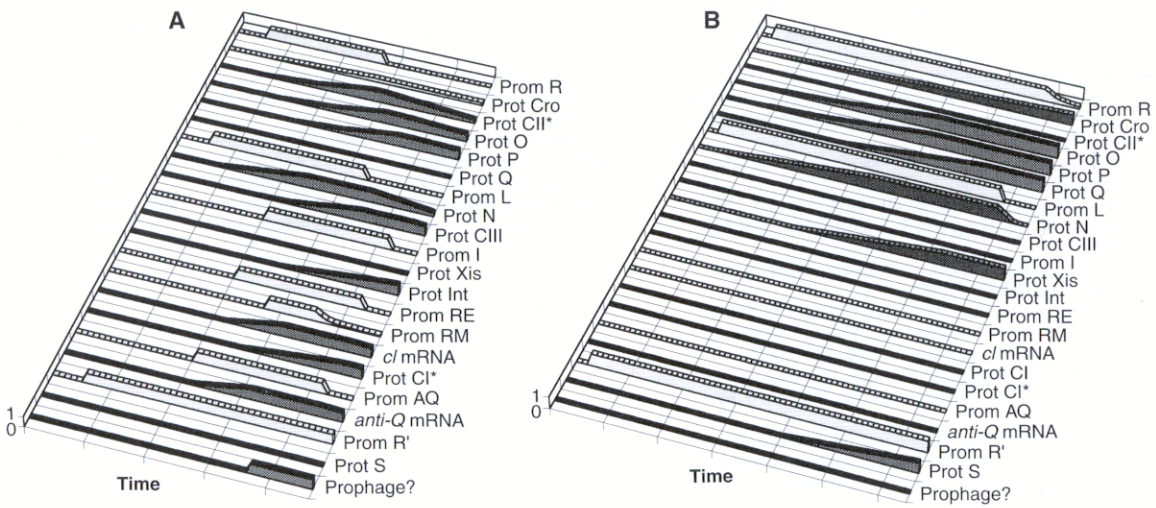

Fig. 6. Timing diagrams of promoter and protein signal sequencing for (**A**) lysogeny decision at an MOI of 10 and (**B**) lysis decision at an MOI of 1. Signals are normalized from 0 (no signal) to 1 (maximum signal). Prom and light shading identify promoters. Prot and dark shading identify signal molecules. The curve labeled Prophage? is 1 if the phage DNA is integrated as a prophage and 0 otherwise.

3A)] should be incorporated into one operon model. Our approach is to define a software operon "object." This object is comprised of a data structure that specifies the operon configuration plus software procedures for operating on the structure. The operon is viewed as a sequence of multinucleotide segments represented by array elements in the data structure. A parameter representing the instantaneous distribution of RNAP transcription complexes in each DNA segment is stored in each array element. Two other parallel arrays reflect any termination sites and antitermination; one contains the fraction of RNAP molecules that moves forward along the DNA at each segment (for example, 0, if blocked; 1, if not; intermediate, if partially blocked), the other carries a flag along indicating whether the RNAP in that segment has been antiterminated. A procedure shifts the RNAP distribution along the array at each time step (taking account of effects at termination sites) to model transcript elongation. This operon object construct models transcription time delays to different genes on the operon and the pipelines of RNAP molecules that continue transcription after the promoter is off. It also provides a simple technique for modeling RNAP flow control by transcription termination sites. Within a circuit simulation, when the promoter control logic determines the promoter is ON, RNAP molecules are injected into one end of the array and then shifted along the array automatically at each time step. At any time, the program can determine the level of transcription of a gene on an operon object by querying the object regarding the rate of RNAP molecules traversing the location of that gene.

In the limits of high MOI (10, favoring high CII and, hence, lysogeny) and low MOI (1, low CII and, thus, lysis), the action of the λ switch (36) and of the CII-CIII interaction is definitive and control of the other promoters fits the Boolean approximation. Thus, we can check the circuit logic in Figs. 4 and 5 by assessing whether correct outcomes are achieved in these limiting cases. The logic validation program incorporates the signal path connectivity and the promoter control logic in Figs. 4 and 5. Transcription time delays, the first term in the T_{delay} equation above, can be modeled by operon objects defined as described above for the five operons headed by P_R, P_L, P_{RE}, P_{AQ}, and $P_{R'}$ (Fig. 4). Because data to model signal path dynamics are not available, signal accumulation delays were estimated for the high and low MOI cases with the second term in the T_{delay} equation and plausible parameters. Signal time delays (not full signal dynamics) were modeled with a signal proxy constrained to the range 0 to 1.0 and using linear growth when the signal gene is being transcribed and translated, and linear decay otherwise. Growth and decay slopes were chosen to produce the required delays.

Figure 6A (high MOI) shows the correct sequencing of promoters, the delays during signal protein buildup, the onset of regulation by CI, and initiation of integration of the lysogen. After integration, the CI feedback loop decreases P_{RM} activity to the low level sufficient to maintain the CI repressor protein concentration necessary to prevent excision, CI concentration falls to its steady-state value, production of other λ proteins ceases, and the proteins are degraded. As predicted, the repression of Cro production (gate G3) and *anti-Q* mRNA regulation (gate G7) stabilize the lysogen path. Figure 6B (low MOI) shows establishment of the lytic path. For the MOI = 1 case, CII was set to zero because it never rises above threshold (gate G8). Then, promoters P_{RM}, P_I, and P_{AQ} remain OFF and the circuit proceeds to events that result in production of S and products of other late genes on the lytic path. The time delays are essential to correct sequencing of circuit functions. If the controlling signal paths in the circuit are eliminated or connected incorrectly, the circuit does not operate correctly. In summary, the order of observed promoter activation, gene expression, and decision outcomes validates the circuit logic shown in Figs. 4 and 5.

We conclude from experience with the λ decision circuit that construction of a simulation model of a genetic circuit that is hypothesized to explain experimental observation provides a powerful test of the hypothesis. The simulation forces identification of connectivity and explicit accounting for timing and sequencing of events. Because intuitive analysis of systems with time lags and feedback is notoriously difficult and error prone, the simulation calculations provide a check on the intuitive understanding. If the simulation does not replicate observed behavior, then the hypothesized circuit is incorrect unless the deviation can be explained by modeling approximations. For example, we initially had a simpler conception of the λ circuit. We were led to detailed examination of transcription time delays and the MOI dependence of logic involving CII and CIII by simulation results. We are optimistic that libraries of generic object-oriented software models of common genetic mechanisms can be developed to provide geneticists the type of user-friendly simulation tools that electrical circuit analysts now take for granted.

REFERENCES AND NOTES

1. I. Herskowitz and D. Hagen, *Annu. Rev. Genet.* **14**, 399 (1980).
2. H. Echols and G. Guarneros, in *Lambda II*, R. Hendrix, J. W. Roberts, F. W. Stahl, R. A. Weisberg, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1983), pp. 75–92.
3. R. W. Hendrix, J. W. Roberts, F. W. Stahl, R. A. Weisberg, Eds., *Lambda II* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1983).

4. M. Ptashne, *A Genetic Switch: Phage λ and Higher Organisms* (Cell Press and Blackwell Scientific Publications, Cambridge, MA, ed. 2, 1992).
5. D. L. Friedman, *Curr. Opin. Genet. Dev.* **2**, 727 (1992).
6. H. R. Horvitz and I. Herskowitz, *Cell* **68**, 237 (1992); S. K. Kim and D. Kaiser, *Annu. Rev. Microbiol.* **46**, 117 (1992).
7. Y. Brun, G. Marczynski, L. Shapiro, *Annu. Rev. Biochem.* **63**, 419 (1994).
8. M. J. McFall-Ngai and E. G. Ruby, *Science* **254**, 1491 (1991); S. Long and B. J. Staskawicz, *Cell* **73**, 921 (1993).
9. F. Jacob and J. Monod, *Cold Spring Harbor Symp. Quant. Biol.* **26**, 389 (1961).
10. R. Thomas, *J. Theor. Biol.* **153**, 1 (1991); M. Kaufman, J. Urbain, R. Thomas, *ibid.* **114**, 527 (1985); L. Glass and S. A. Kauffman, *ibid.* **39**, 103 (1973).
11. D. Bray, R. B. Bourret, M. I. Simon, *Mol. Biol. Cell* **4**, 469 (1993); D. C. Hauri and J. Ross, *Biophys. J.* **68**, 708 (1995).
12. A. Arkin and J. Ross, *Biophys. J.* **67**, 560 (1994).
13. The in vivo transcription rate in *E. coli* has been directly observed at 42 nucleotides (nt) per second [S. L. Gotta, O. L. Miller Jr., S. L. French, *J. Bacteriol.* **173**, 6647 (1991)]. The overall in vivo rate is 12 to 19 nt/s [A. Kornberg and T. A. Baker, *DNA Replication* (W. H. Freeman, New York, 1992), p. 246]. Higher rates, in the range of 50 to 60 nt/s, are also observed, particularly in vitro. See also U. Vogel and K. F. Jensen, *J. Bacteriol.* **176**, 2807 (1994).
14. Steady-state concentrations of all cellular proteins are determined by the dynamic balance between production and losses. Assumptions for the equations in Fig. 2A include (i) transcription initiation is the rate-limiting step; (ii) the repressing protein-promoter interaction has first-order kinetics (higher order kinetics enter through the repression term, $Y(t)$, in Eq. 5); (iii) transcription and translation are coupled so that the number of proteins produced per transcript is a constant (A_{GENE}) independent of transcription rate but differing from gene to gene (*16*) [O. Yarchuk, N. Jacques, J. Guillerez, M. Dreyfus, *J. Mol. Biol.* **226**, 581 (1992)]; and (iv) for low (<10?) MOI, protein production is proportional to MOI (RNAP, ribosome, or substrate availability is limiting at high MOI) [K. Luk and K. Mark, *J. Virol.* **64**, 183 (1983); G. P. Zambetti and R. C. Shuster, *Mol. Gen. Genet.* **193**, 322 (1984)]. In vivo experiments at an average MOI result in a Poisson distribution of MOIs within the experimental cell population, with the exception of induced monolysogens for which the initial MOI is 1. We neglect (i) additional concentration reduction attributable to dilution in growing cells and (ii) additional production from newly replicated viral DNA.
15. A. Levine, A. Bailone, R. Devoret, *J. Mol. Biol.* **131**, 655 (1979).
16. M. A. Shea and G. K. Akers, *ibid.* **181**, 211 (1985).
17. The circuit in Fig. 4 includes the functions essential to the path choice decision. Functions such as the OOP antisense RNA whose role in λ regulation appears to be nonessential for the path choice are not included (*1*) [L. Krinke and D. L. Wulff, *Genes Dev.* **1**, 1005 (1987); *ibid.* **4**, 2223 (1990)].
18. U. Schmeissner, D. Court, H. Shimatake, M. Rosenberg, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3191 (1980).
19. A. Das, *J. Bacteriol.* **174**, 6711 (1992).
20. S. W. Mason, J. Li, J. Greenblatt, *J. Biol. Chem.* **267**, 19418 (1992); J. Greenblatt *et al.*, *Nature* **364**, 401 (1993).
21. J. W. Roberts, *Cell* **52**, 5 (1988).
22. B. C. Hoopes and W. R. McClure, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3134 (1985); Y. S. Ho and M. Rosenberg, *J. Biol. Chem.* **260**, 11838 (1985); R. W. Simons and N. Kleckner, *Annul Rev. Genet.* **22**, 567 (1988).
23. M. A. Hoyt, D. M. Knight, A. Das, H. I. Miller, H. Echols, *Cell* **31**, 565 (1982).
24. C. Herman *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10861 (1993).
25. J. A. Noble *et al.*, *ibid.*, p. 10866.
26. P. Kourilsky, *Mol. Gen. Genet.* **122**, 183 (1973).
27. D. I. Wulff and M. Rosenberg, in (*3*), pp. 53–74.
28. J. W. Roberts and R. Devoret, in (*3*), pp. 123–144.
29. L. M. De Vargas and A. Landy, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 588 (1991).
30. A. Landy, *Annu. Rev. Biochem.* **58**, 913 (1989); *Curr. Opin. Genet. Dev.* **3**, 699 (1993).
31. J. M. Leong *et al.*, *J. Mol. Biol.* **189**, 603 (1986).
32. L. H. Hartwell and T. A. Weinart, *Science* **246**, 629 (1989); L. H. Hartwell, *Genetics* **129**, 975 (1991).
33. A. Campbell and D. Botstein, in (*3*), pp. 365–380.
34. S. Casjens, G. Hatfull, R. Hendrix, *Semin. Virol.* **3**, 383 (1992).
35. Measurements of the signal protein half-lives are available. The rate of signal protein transcript production can be estimated for the critical operons, but with diminishing certainty for downstream genes on P_R and P_L operons as a result of polarity effects. The kinetics of the P_R, P_L, P_{RM}, and P_{RE} promoters have been extensively studied. The average proteins/transcript parameter is not available.
36. For a kinetic model of the switch, see (*16*).
37. Descriptive of Cro repression of P_R, P_L, and P_{RM}. Descriptive of CI repression of P_R and P_L. CI interaction with P_{RM} is more complex (*4, 16*) [B. J. Meyer, R. Maurer, M. Ptashne, *J. Mol. Biol.* **139**, 163 (1980); B. J. Meyer and M. Ptashne, *ibid.*, p. 195].
38. We thank I. Hershowitz, D. Botstein, D. Kaiser, S. Kustu, and A. Campbell for valuable suggestions.