In 1953, Lwoff et al. discovered that a strain of *E. Coli* when exposed to UV light lyse spewing forth λ viruses. Some of the newly infected *E. Coli* would soon lyse while others grow and divide normally until exposed to UV light. In other words, some cells follow a *lysis* pathway while other followed a *lysogeny* pathway. The decision between the lysis and the lysogeny developmental pathway is made by a fairly simple genetic circuit.

**The OR Operator**

**CI, the λ Repressor**
Nothing in biology is clear-cut.

- Without CI bound to O_{R2}, RNAP can still bind to P_{RM} and initiate transcription of CI at a reduced basal rate.
- With CI bound to O_{R2}, transcription occurs at enhanced activated rate.
Cro₂ Bound to O₃ Turns off P₉M

Low Concentrations of CI₂

High Concentrations of CI₂

Cooperativity Aids CI₂ Binding to O₉₂

Low Concentrations of Cro₂
Moderate Concentrations of Cro_2

High Concentrations of Cro_2

Cooperativity of CI_2 Binding

Cooperativity of CI_2 Binding
Effect of Cooperativity

![Graph showing effect of cooperativity](image)

- At low to moderate concentrations of CI and Cro, there are three common configurations:
  - No molecules bound to $O_R$, Cro produced at full rate and CI produced at low basal rate.
  - CI bound to $O_R1$ and $O_R2$, Cro production repressed, and CI activated.
  - Cro bound to $O_R3$, CI cannot be produced, Cro is produced.

Configurations of $O_R$

- Feedback of the products as transcription factors coupled with affinities makes $O_R$ behave as a bistable switch.
- In lysis state, Cro produced locking out production of CI.
- In lysogeny state, CI produced locking out production of Cro.

Immunity

- In lysogeny state, cell is immune to further infection.
- $cro$ genes on DNA inserted by further infections are shut off by CI2 molecules from first infection.
- Once cell commits to lysogeny, it becomes very stable and does not easily change over to the lysis pathway.

Recognition of the Operators

- CI2 and Cro2 bind to operator sites that are 17 base pairs long.
- How do these proteins locate these sequences from amongst the millions within the bacteria?
- Observing from midpoint, a strand on one side is nearly symmetric with complimentary strand on other side.

Near Symmetry in the Operator Sequences

<table>
<thead>
<tr>
<th>Operator sequences</th>
<th>Operator half sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_R1$</td>
<td>T A T C A C C G C C A G G A G T A T A A C A C C G C</td>
</tr>
<tr>
<td>$O_R2$</td>
<td>T A T C A C C G C C A G G A G T A T A A C A C C G C</td>
</tr>
<tr>
<td>$O_{L1}$</td>
<td>T A T C A C C G C C A G G A G T A T A A C A C C G C</td>
</tr>
<tr>
<td>$O_{L2}$</td>
<td>T A T C A C C G C C A G G A G T A T A A C A C C G C</td>
</tr>
<tr>
<td>$O_{L3}$</td>
<td>T A T C A C C G C C A G G A G T A T A A C A C C G C</td>
</tr>
</tbody>
</table>

Induction

![Diagram showing induction](image)
Consensus Sequence

- The consensus sequence is as follows:
  
  TATCACCGcCGGTGATA
  ATAGTGGCgGCCACTAT

- Many entries are highly preserved.
- Differences exist that cause the differences in affinity for CI$_2$ and Cro$_2$ for the different operators.
- Notice that the first half of the operator sites $O_{R1}$ and $O_{R3}$ agree perfectly with the consensus sequence while second half has several differences.

Amino Acid-Base Pair Interactions

$O_{R1}$

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & C & C & T & T \\
\end{array}
\]

$O_{R2}$

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

$O_{R3}$

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{gly} & \text{asn} & \text{leu} & \text{asn} & \text{val} \\
T & A & T & G & A & G \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{gln} & \text{ser} & \text{ala} & \text{ala} & \text{his} \\
C & A & A & C & C & G \\
\end{array}
\]
Early Events

Retroregulation of Int

Retroregulation of Int

Lysis/Lysogeny Decision

Lysis/Lysogeny Decision (cont)

- Key to lysis/lysogeny decision is the protein CII.
  - CII activates $P_{RE}$ which jump-starts production of CI.
  - After jump-start, positive feedback in $P_{RM}$ increases CI and locks out Cro production resulting in lysogeny decision.
  - Activity of CII is determined by environmental factors.
  - Bacterial proteases attack and destroy CII.
  - Growth in a rich medium activates these proteases whereas starvation has the opposite effect.
  - Thus, $\lambda$, tends to lyosenize starved cells.
  - CIII protein protects CII from degradation promoting lysogeny.
  - Production of CII and CIII enhanced by anti-terminator, N.
  - Higher multiplicity of infection means more N, cII, and cIII gene copies leading to higher probability of lysogeny.

- High protease levels and low gene counts
  - CI produced slowly and degrades rapidly
  - little CI is synthesized
  - Q and Cro are synthesized
  - lysis decision
- Low protease levels and high gene counts
  - more N, CII, and CIII are produced
  - CI and Int are made from $P_{RE}$ and $P_I$
  - Int integrates phage chromosome and CI turns off all genes except CI
  - lysogeny decision
Late Lytic Events

Late Lysis

Chris J. Myers (Lecture 4: Phage λ)

Late Lysogenic Events

Late Lysogeny

Chris J. Myers (Lecture 4: Phage λ)

Integration and Induction

attP
attB
sibint xis cIII N cI

Chemical Reaction Network Model

Variety of ways to model a system using chemical reactions.
This model includes:
Five genes: cI, cro, N, cII, and cIII.
Four promoters: P_{RM}, P_{R}, P_{RE}, and P_{L}.
Model somewhat simplified (see appendix at end of Chapter 2).

Phage λ Decision Circuit

Model of the Promoter $P_{RE}$

$P_{RE} + \text{RNAP}$
$K_{PRE2} \rightarrow P_{RE} \cdot \text{RNAP}$
$P_{RE} + \text{CII}$
$K_{PRE3} \rightarrow P_{RE} \cdot \text{CII}$
$P_{RE} + \text{CII} + \text{RNAP}$
$K_{PRE4} \rightarrow P_{RE} \cdot \text{CII} \cdot \text{RNAP}$
$P_{RE} \cdot \text{RNAP}$
$K_{PRE5} \rightarrow P_{RE} \cdot \text{RNAP} + \text{cII}$
$P_{RE} \cdot \text{CII} \cdot \text{RNAP}$
$K_{PRE6} \rightarrow P_{RE} \cdot \text{CII} \cdot \text{RNAP} + \text{cII}$

Constant | Value | Constant | Value
---|---|---|---
$K_{PRE2}$ | 0.01 M$^{-1}$ | $K_{PRE3}$ | 0.00726 M$^{-1}$
$K_{PRE4}$ | 0.00161 M$^{-1}$ | $K_{PRE5}$ | 0.000004 sec$^{-1}$
$K_{PRE6}$ | 0.015 sec$^{-1}$ | $n$ | 10
Model of the $O_R$ Operator

- Modeling method just described requires $(n - 1)$ chemical reactions where $n$ is number of potential configurations of transcription factors and RNAP bound to the operator and promoter sites.
- Also requires $m$ reactions for configurations leading to transcription.
- $O_R$ operator has $40$ possible configurations with $13$ leading to transcription.
  - $O_R3$ has four states (empty, CI2, Cro2, and RNAP).
  - $O_R2$ and $O_R1$ also can bind all transcription factors but bind to RNAP jointly (i.e., $10$ possibilities).

Simplified Model of the $O_R$ Operator (1 Site Occupied)

\[
\begin{align*}
O_R + CI & \overset{k_{ORI}}{\longrightarrow} O_R - CI \\
O_R + Cro & \overset{k_{O'R}}{\longrightarrow} O_R - Cro \\
O_R + RNAP & \overset{k_{ORPR}}{\longrightarrow} O_R - RNAP \\
O_R + Cro + RNAP & \overset{k_{ORPR}}{\longrightarrow} O_R - Cro - RNAP \\
O_R + 2RNAP & \overset{k_{ORPR}}{\longrightarrow} O_R - 2RNAP \\
O_R + Cro + RNAP & \overset{k_{ORPR}}{\longrightarrow} O_R - Cro - RNAP \\
O_R - RNAP & \overset{k_{ORPR}}{\longrightarrow} O_R - RNAP + nCI \\
O_R - RNAP - 2Cro & \overset{k_{ORPR}}{\longrightarrow} O_R - RNAP + nCro
\end{align*}
\]

\[
\begin{array}{ccc}
\text{Constant} & \text{Value} & \text{Constant} & \text{Value} \\
K_{ORR1} & 0.06568 M^{-1} & k_{PRM} & 0.001 sec^{-1} \\
K_{ORR6} & 0.09455 M^{-1} & k_{PRM} & 0.011 sec^{-1} \\
K_{ORR7} & 0.1779 M^{-1} & k_{PR} & 0.014 sec^{-1} \\
K_{ORR8} & 0.25123 M^{-1}
\end{array}
\]

Simplified Model of the $O_R$ Operator (2 Sites Occupied)

\[
\begin{align*}
O_R + 2CI & \overset{k_{ORBI}}{\longrightarrow} O_R - 2CI \\
O_R + CI + Cro & \overset{k_{ORB}}{\longrightarrow} O_R - CI + Cro \\
O_R + Cro + RNAP & \overset{k_{ORPR}}{\longrightarrow} O_R - Cro - RNAP \\
O_R + 2RNAP & \overset{k_{ORPR}}{\longrightarrow} O_R - 2RNAP \\
O_R + Cro + RNAP & \overset{k_{ORPR}}{\longrightarrow} O_R - Cro - RNAP \\
O_R - 2RNAP & \overset{k_{ORPR}}{\longrightarrow} O_R - 2RNAP + nCI \\
O_R - RNAP + nCro & \overset{k_{ORPR}}{\longrightarrow} O_R - RNAP + nCro
\end{align*}
\]

CI and Cro Dimerization and Degradation

\[
\begin{align*}
2CI & \overset{k_{C1}}{\longrightarrow} CI2 & \text{Constant} & \text{Value} \\
2Cro & \overset{k_{C2}}{\longrightarrow} Cro2 & K_2 & 0.0007 sec^{-1} \\
CI & \overset{k_{C3}}{\longrightarrow} & K_3 & 0.1 M^{-1} \\
Cro & \overset{k_{C4}}{\longrightarrow} & K_4 & 0.0025 sec^{-1} \\
\text{N} & \overset{(n)}{\longrightarrow} & K_5 & 0.1 M^{-1}
\end{align*}
\]

$N$ Production from promoter $P_L$ and $N$ Degradation

\[
\begin{align*}
P_L + Cro & \overset{k_{PRL}}{\longrightarrow} P_L - Cro \\
P_L + CI & \overset{k_{PRL}}{\longrightarrow} P_L - CI \\
P_L + 2Cro & \overset{k_{PRL}}{\longrightarrow} P_L - 2Cro \\
P_L + RNAP & \overset{k_{PRL}}{\longrightarrow} P_L - RNAP \\
P_L - RNAP + nN & \overset{(n)}{\longrightarrow} P_L - RNAP + nN
\end{align*}
\]

\[
\begin{array}{ccc}
\text{Constant} & \text{Value} & \text{Constant} & \text{Value} \\
K_{PRL2} & 0.0132 M^{-1} & K_{PRL8} & 0.014 M^{-1} \\
K_{PRL3} & 0.1234 M^{-1} & K_{PRL9} & 0.058 M^{-1} \\
K_{PRL4} & 0.0245 M^{-1} & K_{PRL10} & 0.058 M^{-1} \\
K_{PRL5} & 0.0656 M^{-1} & K_{PRLN} & 0.011 sec^{-1} \\
K_{PRL6} & 0.0256 M^{-1} & K_{PRL7} & 0.00231 sec^{-1} \\
K_{PRL7} & 0.0156 M^{-1} & K_{PRL10} & 0.058 M^{-1}
\end{array}
\]
**CIII and CII Production**

\[
\begin{align*}
NUT_L + N & \stackrel{k_{NUT}}{\longrightarrow} NUT_L \cdot N \\
P_L \cdot RNAP + NUT_L & \stackrel{0.2k_{PL}}{\longrightarrow} P_L \cdot RNAP + NUT_L \cdot N \\
P_L \cdot RNAP + NUT_L \cdot N & \stackrel{k_{P1}}{\longrightarrow} P_L \cdot RNAP + NUT_L \cdot N + nCIII \\
NUT_R + N & \stackrel{k_{NUT}}{\longrightarrow} NUT_R \cdot N \\
O\alpha_{12} \cdot RNAP + NUT_R & \stackrel{0.5k_{PR}}{\longrightarrow} O\alpha_{12} \cdot RNAP + NUT_R \cdot N + nCII \\
O\alpha_{12} \cdot RNAP + NUT_R \cdot N & \stackrel{k_{O12}}{\longrightarrow} O\alpha_{12} \cdot RNAP + NUT_R \cdot N + nCII
\end{align*}
\]

Note that \( k_{NUT} \) is 0.2 M\(^{-1}\).

**Model for CII and CIII Degradation**

\[
\begin{align*}
CII + P1 & \stackrel{k_{CII}}{\longrightarrow} CII \cdot P1 \\
CII \cdot P1 & \stackrel{k_{10}}{\longrightarrow} P1 \\
CIII + P1 & \stackrel{k_{11}}{\longrightarrow} CIII \cdot P1 \\
CIII \cdot P1 & \stackrel{k_{13}}{\longrightarrow} P1
\end{align*}
\]

**Why Study Phage λ?**

- Bacteria and their phages multiply quickly making it easier to analyze gene regulation with them than higher organisms.
- Phage λ has been the subject of study for over 50 years now.
- It is one of, if not the best, understood genetic circuit.
- Excellent illustration of a circuit that analyzes its environment makes decision between two competing pathways.
- Similarities with bacteria that must respond to stress and circuits involved in development and cell differentiation.
- Genes from phage λ are used in synthetic biology where DNA is produced to perform particular functions.
- Phage λ is an excellent testbed for trying new ideas.
- Virtually every modeling method has been applied to phage λ.
- This course also uses it as a running example throughout.

**Sources**

- Excellent history in paper by Gottesman and Weisberg (2004)
- Mark Ptashne - A Genetic Switch
- [http://www.mun.ca/biochem/courses/4103/](http://www.mun.ca/biochem/courses/4103/)
- Wikipedia entry for phage λ
- Model inspired by Arkin et al.'s phage λ model