PHAGE STRATEGIES

3.12.1 Introduction

Key Terms

Lytic infection of a bacterium by a phage ends in the destruction of the bacterium with release of progeny phage.

Lysis describes the death of bacteria at the end of a phage infective cycle when they burst open to release the progeny of an infecting phage (because phage enzymes disrupt the bacterium's cytoplasmic membrane or cell wall). The same term also applies to eukaryotic cells; for example, when infected cells are attacked by the immune system.

Prophage is a phage genome covalently integrated as a linear part of the bacterial chromosome.

Lysogeny describes the ability of a phage to survive in a bacterium as a stable prophage component of the bacterial genome.

Integration of viral or another DNA sequence describes its insertion into a host genome as a region covalently linked on either side to the host sequences.

Induction of prophage describes its entry into the lytic (infective) cycle as a result of destruction of the lysogenic repressor, which leads to excision of free phage DNA from the bacterial chromosome.

The **excision** of phage or episome or other sequence describes its release from the host chromosome as an autonomous DNA molecule.

A **plasmid** is a circular, extrachromosomal DNA. It is autonomous and can replicate itself.

An **extrachromosomal genome** in a bacterium is a self-replicating set of genes that is not part of the bacterial chromosome. In many cases, the genes are necessary for bacterial growth under certain environmental conditions.

An **episome** is a plasmid able to integrate into bacterial DNA.

Immunity in phages refers to the ability of a prophage to prevent another phage of the same type from infecting a cell. It results from the synthesis of phage repressor by the prophage genome.

Immunity in plasmids describes the ability of a plasmid to prevent another of the same type from becoming established in a cell. It results usually from interference with the ability to replicate.

Some phages have only a single strategy for survival. On infecting a susceptible host, they subvert its functions to the purpose of producing a large number of progeny phage particles. As the result of this **lytic infection**, the host bacterium dies. In the typical lytic cycle, the phage DNA (or RNA) enters the host bacterium, its genes are transcribed in a set order, the phage genetic material is replicated, and the protein components of the phage particle are produced. Finally, the host bacterium is broken open (*lysed*) to release the assembled progeny particles by the process of **lysis**.

Other phages have a dual existence. They are able to perpetuate themselves via the same sort of lytic cycle in what amounts to an open strategy for producing as many copies of the phage as rapidly as possible. But they also have an alternative form of existence, in which the phage genome is present in the bacterium in a latent form known as **prophage**. This form of propagation is called **lysogeny**.

In a lysogenic bacterium, the prophage is inserted into the bacterial genome, and is inherited in the same way as bacterial genes. The process by which it is converted from an independent phage genome into a prophage that is a linear part of the bacterial genome is described as **integration**. By virtue of its possession of a prophage, a lysogenic bacterium has **immunity** against infection by further phage particles of the same type. Immunity is established by a single integrated prophage, so usually a bacterial genome contains only one copy of a prophage of any particular type.

Transitions occur between the lysogenic and lytic modes of existence. **Figure 12.1** shows that when a phage produced by a lytic cycle enters a new bacterial host cell, it either repeats the lytic cycle or enters the lysogenic state. The outcome depends on the conditions of infection and the genotypes of phage and bacterium.

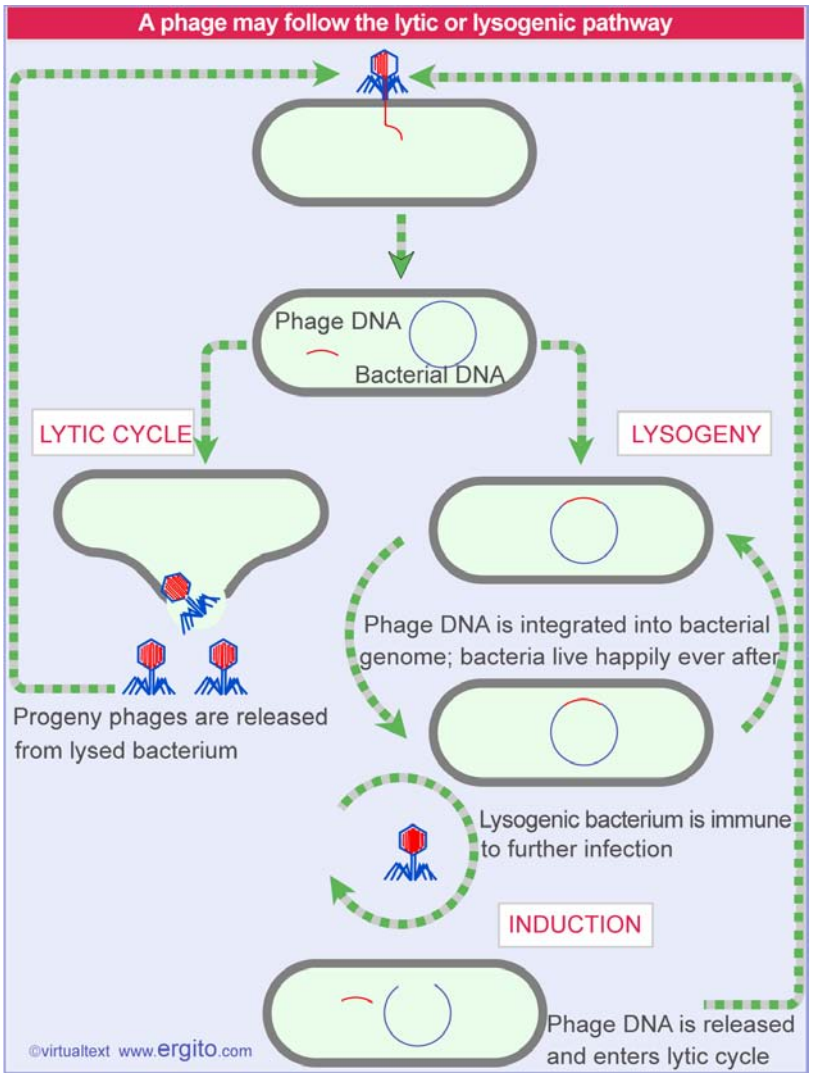

Figure 12.1 Lytic development involves the reproduction of phage particles with destruction of the host bacterium, but lysogenic existence allows the phage genome to be carried as part of the bacterial genetic information.

A prophage is freed from the restrictions of lysogeny by the process called **induction**. First the phage DNA is released from the bacterial chromosome by **excision**; then the free DNA proceeds through the lytic pathway.

The alternative forms in which these phages are propagated are determined by the regulation of transcription. Lysogeny is maintained by the interaction of a phage repressor with an operator. The lytic cycle requires a cascade of transcriptional controls. And the transition between the two life-styles is accomplished by the establishment of repression (lytic cycle to lysogeny) or by the relief of repression (induction of lysogen to lytic phage).

Another type of existence within bacteria is represented by **plasmids**. These are autonomous units that exist in the cell as **extrachromosomal genomes**. Plasmids are self-replicating circular molecules of DNA that are maintained in the cell in a stable

and characteristic number of copies; that is, the number remains constant from generation to generation.

Some plasmids also have alternative life-styles. They can exist either in the autonomous extrachromosomal state; or they can be inserted into the bacterial chromosome, and then are carried as part of it like any other sequence. Such units are properly called **episomes** (but the terms "plasmid" and "episome" are sometimes used loosely as though interchangeable).

Like lysogenic phages, plasmids and episomes maintain a selfish possession of their bacterium and often make it impossible for another element of the same type to become established. This effect also is called **immunity**, although the basis for plasmid immunity is different from lysogenic immunity. (We discuss the control of plasmid perpetuation in *Molecular Biology 4.13 The replicon*.)

Figure 12.2 summarizes the types of genetic units that can be propagated in bacteria as independent genomes. Lytic phages may have genomes of any type of nucleic acid; they transfer between cells by release of infective particles. Lysogenic phages have double-stranded DNA genomes, as do plasmids and episomes. Some plasmids and episomes transfer between cells by a conjugative process (involving direct contact between donor and recipient cells). A feature of the transfer process in both cases is that on occasion some bacterial host genes are transferred with the phage or plasmid DNA, so these events play a role in allowing exchange of genetic information between bacteria.

Phages and plasmids live in bacteria

Type of Unit	Genome Structure	Mode of Propagation	Consequences
Lytic phage	ds- or ss-DNA or RNA linear or circular	Infects susceptible host	Usually kills host
Lysogenic phage	ds-DNA	Linear sequence in host chromosome	Immunity to infection
Plasmid	ds-DNA circle	Replicates at defined copy number May be transmissible	Immunity to plasmids in same group
Episome	ds-DNA circle	Free circle or linear integrated	May transfer host DNA

Figure 12.2 Several types of independent genetic units exist in bacteria.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.3.12.1

PHAGE STRATEGIES

3.12.2 Lytic development is divided into two periods

Key Terms

Early infection is the part of the phage lytic cycle between entry and replication of the phage DNA. During this time, the phage synthesizes the enzymes needed to replicate its DNA.

Late infection is the part of the phage lytic cycle from DNA replication to lysis of the cell. During this time, the DNA is replicated and structural components of the phage particle are synthesized.

Key Concepts

- A phage infective cycle is divided into the early period (before replication) and the late period (after the onset of replication).
- A phage infection generates a pool of progeny phage genomes that replicate and recombine.

Phage genomes of necessity are small. As with all viruses, they are restricted by the need to package the nucleic acid within the protein coat. This limitation dictates many of the viral strategies for reproduction. Typically a virus takes over the apparatus of the host cell, which then replicates and expresses phage genes instead of the bacterial genes.

Usually the phage includes genes whose function is to ensure preferential replication of phage DNA. These genes are concerned with the initiation of replication and may even include a new DNA polymerase. Changes are introduced in the capacity of the host cell to engage in transcription. They involve replacing the RNA polymerase or modifying its capacity for initiation or termination. The result is always the same: phage mRNAs are preferentially transcribed. So far as protein synthesis is concerned, usually the phage is content to use the host apparatus, redirecting its activities principally by replacing bacterial mRNA with phage mRNA.

Lytic development is accomplished by a pathway in which the phage genes are expressed in a particular order. This ensures that the right amount of each component is present at the appropriate time. The cycle can be divided into the two general parts illustrated in **Figure 12.3**:

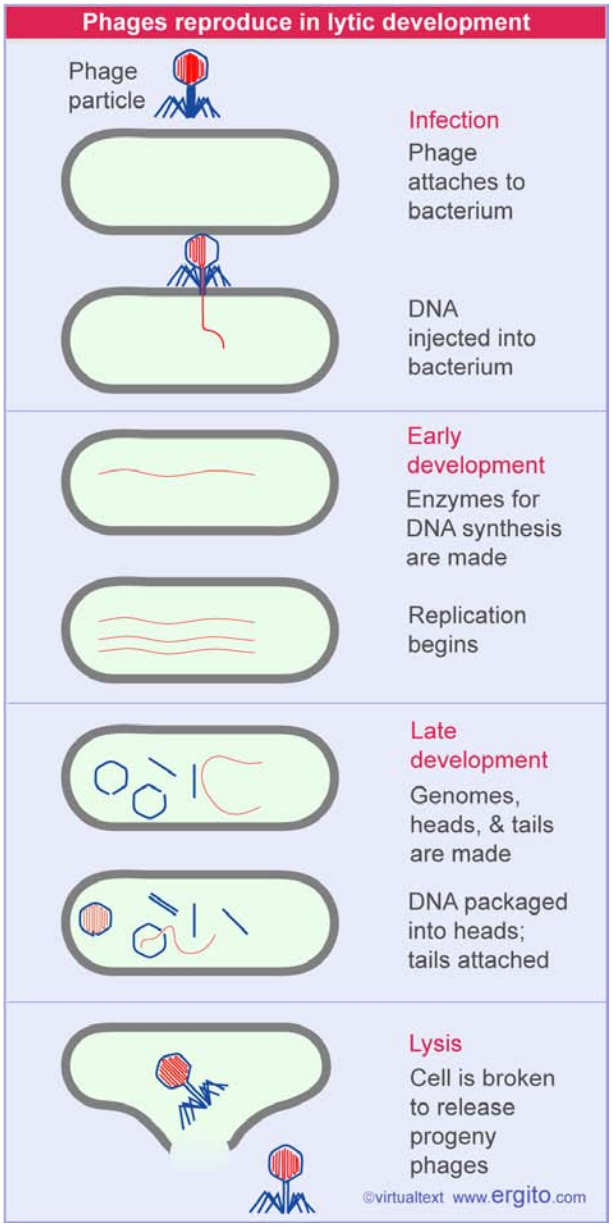

Figure 12.3 Lytic development takes place by producing phage genomes and protein particles that are assembled into progeny phages.

- **Early infection** describes the period from entry of the DNA to the start of its replication.
- **Late infection** defines the period from the start of replication to the final step of lysing the bacterial cell to release progeny phage particles.

The early phase is devoted to the production of enzymes involved in the reproduction of DNA. These include the enzymes concerned with DNA synthesis, recombination, and sometimes modification. Their activities cause a *pool* of phage genomes to

accumulate. In this pool, genomes are continually replicating and recombining, so that *the events of a single lytic cycle concern a population of phage genomes*.

During the late phase, the protein components of the phage particle are synthesized. Often many different proteins are needed to make up head and tail structures, so the largest part of the phage genome consists of late functions. In addition to the structural proteins, "assembly proteins" are needed to help construct the particle, although they are not themselves incorporated into it. By the time the structural components are assembling into heads and tails, replication of DNA has reached its maximum rate. The genomes then are inserted into the empty protein heads, tails are added, and the host cell is lysed to allow release of new viral particles.

PHAGE STRATEGIES

3.12.3 Lytic development is controlled by a cascade

Key Terms

A **cascade** is a sequence of events, each of which is stimulated by the previous one. In transcriptional regulation, as seen in sporulation and phage lytic development, it means that regulation is divided into stages, and at each stage, one of the genes that are expressed codes for a regulator needed to express the genes of the next stage.

Early genes are transcribed before the replication of phage DNA. They code for regulators and other proteins needed for later stages of infection.

Immediate early phage genes in phage lambda are equivalent to the early class of other phages. They are transcribed immediately upon infection by the host RNA polymerase.

Delayed early genes in phage lambda are equivalent to the middle genes of other phages. They cannot be transcribed until regulator protein(s) coded by the immediate early genes have been synthesized.

Middle genes are phage genes that are regulated by the proteins coded by early genes. Some proteins coded by middle genes catalyze replication of the phage DNA; others regulate the expression of a later set of genes.

Late genes are transcribed when phage DNA is being replicated. They code for components of the phage particle.

Key Concepts

- The early genes transcribed by host RNA polymerase following infection include or comprise regulators required for expression of the middle set of phage genes.
- The middle group of genes include regulators to transcribe the late genes.
- This results in the ordered expression of groups of genes during phage infection.

The organization of the phage genetic map often reflects the sequence of lytic development. The concept of the operon is taken to somewhat of an extreme, in which the genes coding for proteins with related functions are clustered to allow their control with the maximum economy. This allows the pathway of lytic development to be controlled with a small number of regulatory switches.

The lytic cycle is under positive control, so that each group of phage genes can be expressed only when an appropriate signal is given. **Figure 12.4** shows that the regulatory genes function in a **cascade**, in which a gene expressed at one stage is necessary for synthesis of the genes that are expressed at the next stage.

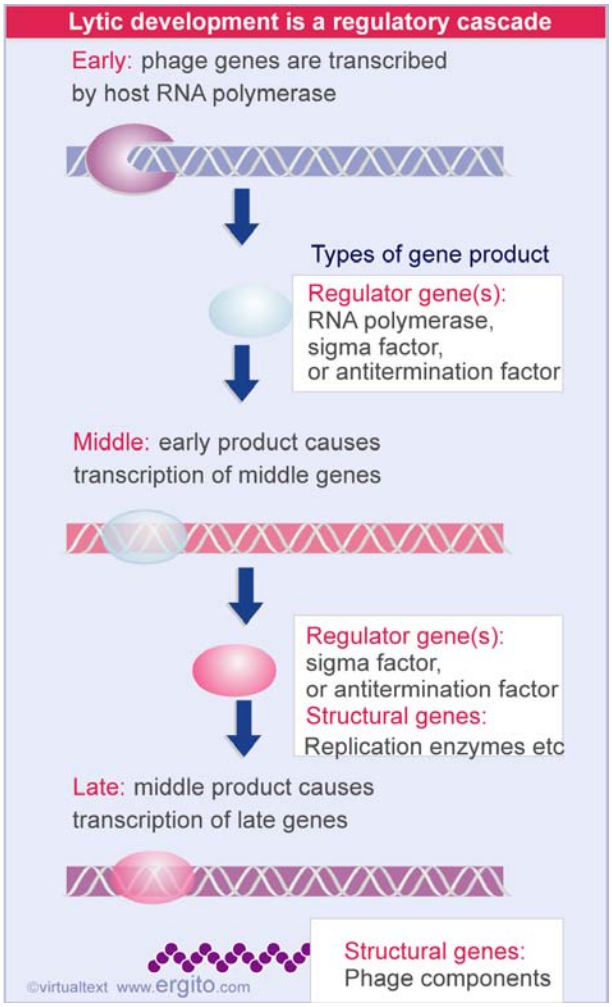

Figure 12.4 Phage lytic development proceeds by a regulatory cascade, in which a gene product at each stage is needed for expression of the genes at the next stage.

The first stage of gene expression necessarily relies on the transcription apparatus of the host cell. Usually only a few genes are expressed at this stage. Their promoters are indistinguishable from those of host genes. The name of this class of genes depends on the phage. In most cases, they are known as the **early genes**. In phage lambda, they are given the evocative description of **immediate early**. Irrespective of the name, they constitute only a preliminary, representing just the initial part of the early period. Sometimes they are exclusively occupied with the transition to the next period. At all events, *one of these genes always codes for a protein that is necessary for transcription of the next class of genes.*

This second class of genes is known variously as the **delayed early** or **middle gene** group. Its expression typically starts as soon as the regulator protein coded by the early gene(s) is available. Depending on the nature of the control circuit, the initial set of early genes may or may not continue to be expressed at this stage. If control is at initiation, the two events are independent (see **Figure 12.5**), and early genes can be switched off when middle genes are transcribed. If control is at termination, the early genes must continue to be expressed (see **Figure 12.6**). Often the expression of

host genes is reduced. Together the two sets of early genes account for all necessary phage functions except those needed to assemble the particle coat itself and to lyse the cell.

When the replication of phage DNA begins, it is time for the **late genes** to be expressed. Their transcription at this stage usually is arranged by embedding a further regulator gene within the previous (delayed early or middle) set of genes. This regulator may be another antitermination factor (as in lambda) or it may be another sigma factor (as in SPO1).

A lytic infection often falls into three stages, as shown in **Figure 12.4**. The first stage consists of early genes transcribed by host RNA polymerase (sometimes the regulators are the only products at this stage). The second stage consists of genes transcribed under direction of the regulator produced in the first stage (most of these genes code for enzymes needed for replication of phage DNA). The final stage consists of genes for phage components, transcribed under direction of a regulator synthesized in the second stage.

*The use of these successive controls, in which each set of genes contains a regulator that is necessary for expression of the next set, creates a cascade in which groups of genes are turned on (and sometimes off) at particular times.*The means used to construct each phage cascade are different, but the results are similar, as the following sections show.

PHAGE STRATEGIES

3.12.4 Two types of regulatory event control the lytic cascade

Key Concepts

- Regulator proteins used in phage cascades may sponsor initiation at new (phage) promoters or cause the host polymerase to read through transcription terminators.

At every stage of phage expression, one or more of the active genes is a regulator that is needed for the subsequent stage. The regulator may take the form of a new RNA polymerase, a sigma factor that redirects the specificity of the host RNA polymerase (see *Molecular Biology 3.9.18 Sigma factors may be organized into cascades*), or an antitermination factor that allows it to read a new group of genes (see *Molecular Biology 3.9.23 Antitermination is a regulatory event*). The next two figures compare the use of switching at initiation or termination to control gene expression.

One mechanism for recognizing new phage promoters is to replace the sigma factor of the host enzyme with another factor that redirects its specificity in initiation (see **Figure 9.31**). An alternative mechanism is to synthesize a new phage RNA polymerase. In either case, the critical feature that distinguishes the new set of genes is their possession of *different promoters from those originally recognized by host RNA polymerase*.**Figure 12.5** shows that the two sets of transcripts are independent; as a consequence, early gene expression can cease after the new sigma factor or polymerase has been produced.

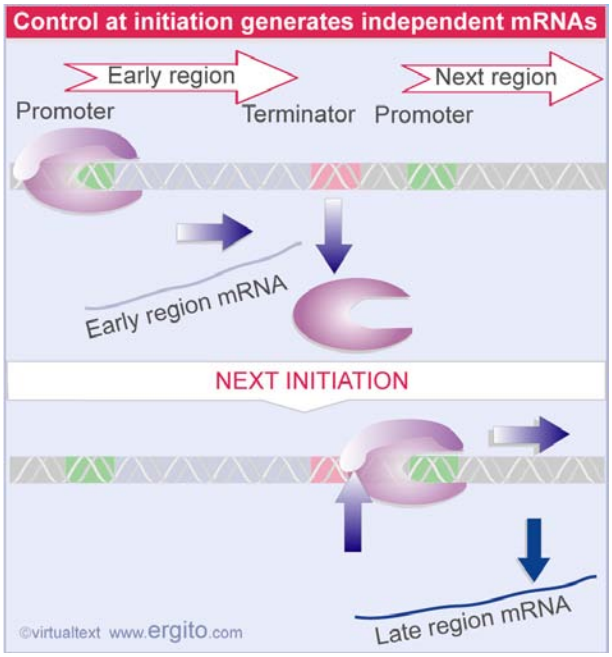

Figure 12.5 Control at initiation utilizes independent transcription units, each with its own promoter and terminator, which produce independent mRNAs. The transcription units need not be located near one another.

Antitermination provides an alternative mechanism for phages to control the switch from early genes to the next stage of expression. The use of antitermination depends on a particular arrangement of genes. **Figure 12.6** shows that the early genes lie adjacent to the genes that are to be expressed next, but are separated from them by terminator sites. *If termination is prevented at these sites, the polymerase reads through into the genes on the other side.* So in antitermination, the *same promoters* continue to be recognized by RNA polymerase. So the new genes are expressed only by extending the RNA chain to form molecules that contain the early gene sequences at the 5′ end and the new gene sequences at the 3′ end. Since the two types of sequence remain linked, early gene expression inevitably continues (for review see 102).

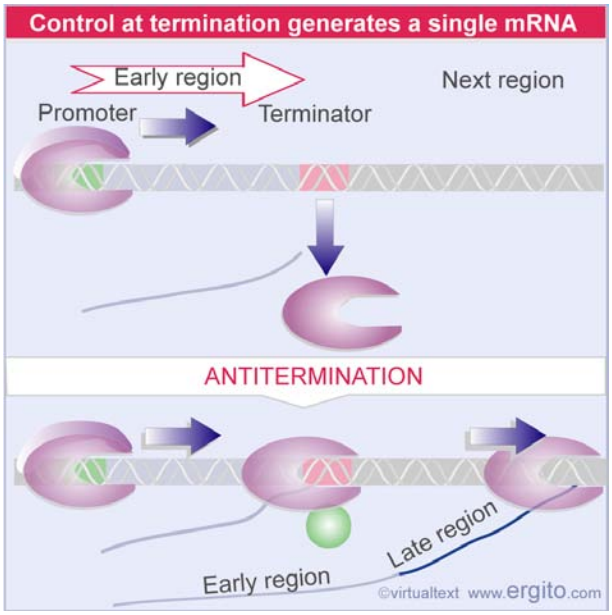

Figure 12.6 Control at termination requires adjacent units, so that transcription can read from the first gene into the next gene. This produces a single mRNA that contains both sets of genes.

The regulator gene that controls the switch from immediate early to delayed early expression in phage lambda is identified by mutations in gene *N* that can transcribe *only* the immediate early genes; they proceed no further into the infective cycle (see **Figure 9.53**). The same effect is seen when gene *28* of phage SPO1 is mutated to prevent the production of σ^{gp28} (see **Figure 9.40**). From the genetic point of view, the mechanisms of new initiation and antitermination are similar. *Both are positive controls in which an early gene product must be made by the phage in order to express the next set of genes.* By employing either sigma factors or antitermination proteins with different specificities, a cascade for gene expression can be constructed

Reviews

102. Greenblatt, J., Nodwell, J. R., and Mason, S. W. (1993). *Transcriptional antitermination*. Nature 364, 401-406.

PHAGE STRATEGIES

3.12.5 The T7 and T4 genomes show functional clustering

Key Concepts

- Genes concerned with related functions are often clustered.
- Phages T7 and T4 are examples of regulatory cascades in which phage infection is divided into three periods.

The genome of phage T7 has three classes of genes, each constituting a group of adjacent loci. As **Figure 12.7** shows, the class I genes are the immediate early type, expressed by host RNA polymerase as soon as the phage DNA enters the cell. Among the products of these genes are a phage RNA polymerase and enzymes that interfere with host gene expression. The phage RNA polymerase is responsible for expressing the class II genes (concerned principally with DNA synthesis functions) and the class III genes (concerned with assembling the mature phage particle).

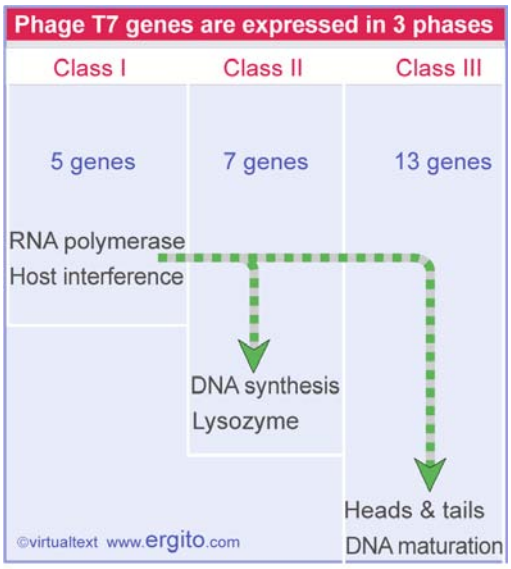

Figure 12.7 Phage T7 contains three classes of genes that are expressed sequentially. The genome is ~38 kb.

T4 has one of the larger phage genomes (165 kb), organized with extensive functional grouping of genes. **Figure 12.8** presents the genetic map. *Essential genes* are numbered: a mutation in any one of these loci prevents successful completion of the lytic cycle. *Nonessential genes* are indicated by three-letter abbreviations. (They are defined as nonessential under the usual conditions of infection. We do not really understand the inclusion of many nonessential genes, but presumably they confer a selective advantage in some of T4's habitats. In smaller phage genomes, most or all

of the genes are essential.)

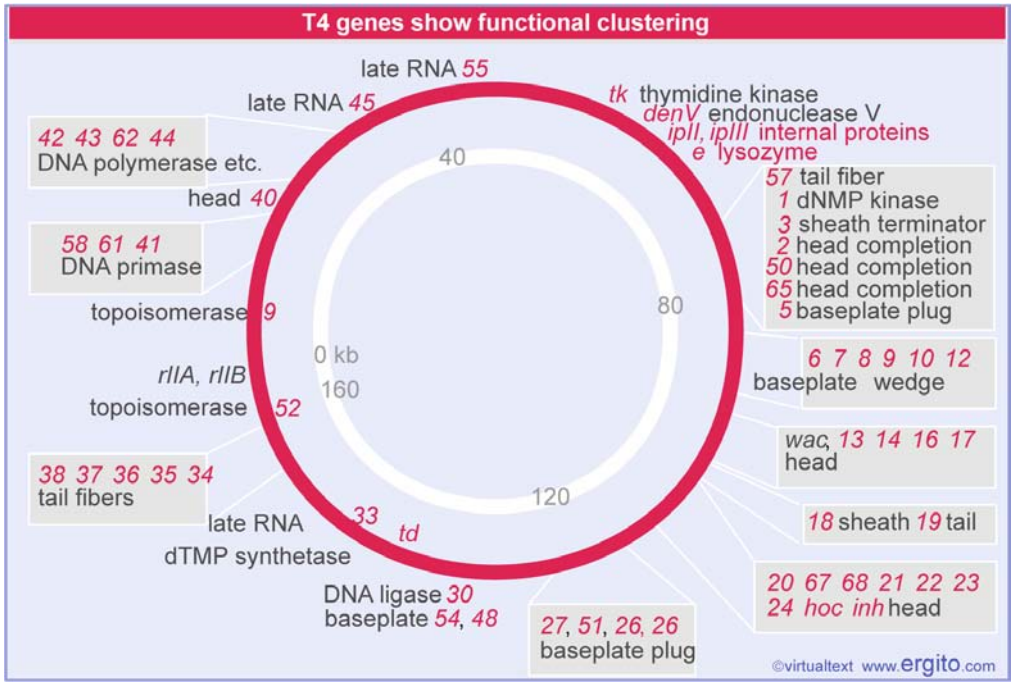

Figure 12.8 The map of T4 is circular. There is extensive clustering of genes coding for components of the phage and processes such as DNA replication, but there is also dispersion of genes coding for a variety of enzymatic and other functions. Essential genes are indicated by numbers. Nonessential genes are identified by letters. Only some representative T4 genes are shown on the map.

There are three phases of gene expression. A summary of the functions of the genes expressed at each stage is given in **Figure 12.9**. The early genes are transcribed by host RNA polymerase. The middle genes are also transcribed by host RNA polymerase, but two phage-encoded products, MotA and AsiA, are also required. The middle promoters lack a consensus –30 sequence, and instead have a binding sequence for MotA. The phage protein is an activator that compensates for the deficiency in the promoter by assisting host RNA polymerase to bind. (This is similar to a mechanism employed by phage lambda, which is illustrated later in **Figure 12.28**.) The early and middle genes account for virtually all of the phage functions concerned with the synthesis of DNA, modifying cell structure, and transcribing and translating phage genes.

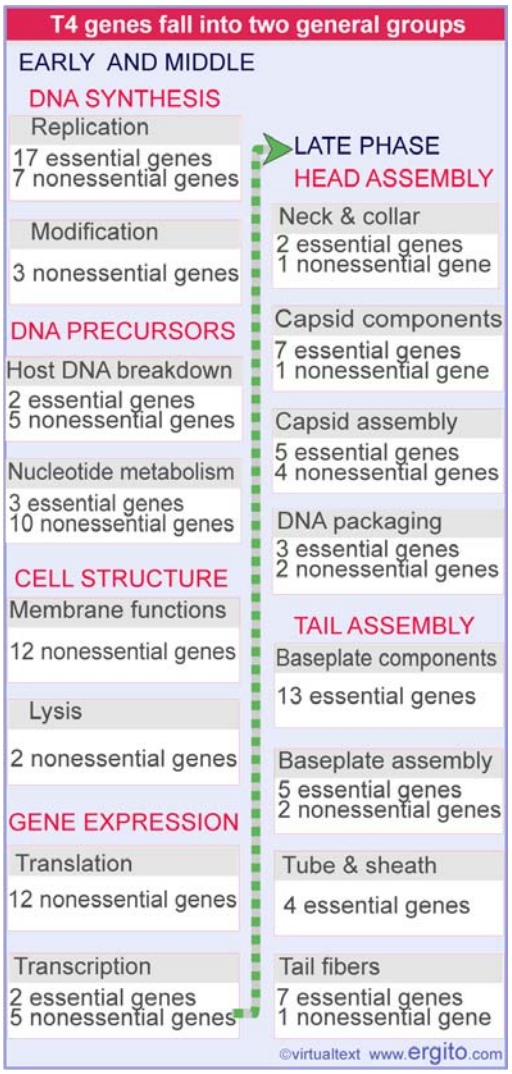

Figure 12.9 The phage T4 lytic cascade falls into two parts: early functions are concerned with DNA synthesis and gene expression; late functions are concerned with particle assembly.

The two essential genes in the "transcription" category fulfill a regulatory function: their products are necessary for late gene expression. Phage T4 infection depends on a mechanical link between replication and late gene expression. Only actively replicating DNA can be used as template for late gene transcription. The connection is generated by introducing a new sigma factor and also by making other modifications in the host RNA polymerase so that it is active only with a template of replicating DNA. This link establishes a correlation between the synthesis of phage protein components and the number of genomes available for packaging.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.3.12.5

PHAGE STRATEGIES

3.12.6 Lambda immediate early and delayed early genes are needed for both lysogeny and the lytic cycle

Key Concepts

- Lambda has two immediate early genes, *N* and *cro*, which are transcribed by host RNA polymerase.
- *N* is required to express the delayed early genes.
- Three of the delayed early genes are regulators.
- Lysogeny requires the delayed early genes *cII-cIII*.
- The lytic cycle requires the immediate early gene *cro* and the delayed early gene *Q*.

One of the most intricate cascade circuits is provided by phage lambda. Actually, the cascade for lytic development itself is straightforward, with two regulators controlling the successive stages of development. But the circuit for the lytic cycle is interlocked with the circuit for establishing lysogeny, as summarized in **Figure 12.10** (for review see 100).

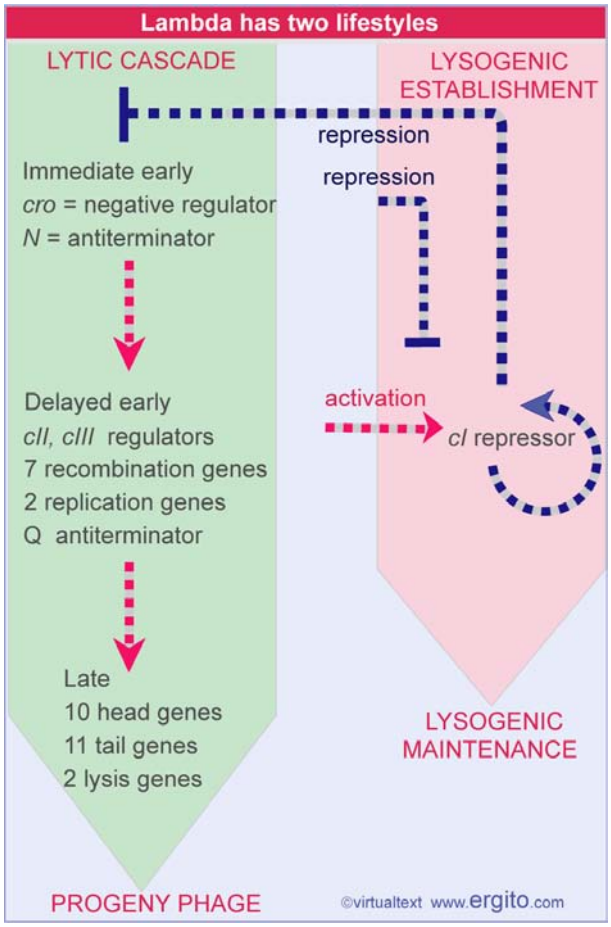

Figure 12.10 The lambda lytic cascade is interlocked with the circuitry for lysogeny.

When lambda DNA enters a new host cell, the lytic and lysogenic pathways start off the same way. Both require expression of the immediate early and delayed early genes. But then they diverge: lytic development follows if the late genes are expressed; lysogeny ensues if synthesis of the repressor is established.

Lambda has only two immediate early genes, transcribed independently by host RNA polymerase:

- *N* codes for an antitermination factor whose action at the *nut* sites allows transcription to proceed into the delayed early genes (see *Molecular Biology 3.9.24 Antitermination requires sites that are independent of the terminators*).
- *cro* has dual functions: it prevents synthesis of the repressor (a necessary action if the lytic cycle is to proceed); and it turns off expression of the immediate early genes (which are not needed later in the lytic cycle).

The delayed early genes include two replication genes (needed for lytic infection), seven recombination genes (some involved in recombination during lytic infection,

two necessary to integrate lambda DNA into the bacterial chromosome for lysogeny), and three regulators. The regulators have opposing functions:

- The *cII-cIII* pair of regulators is needed to establish the synthesis of repressor.
- The *Q* regulator is an antitermination factor that allows host RNA polymerase to transcribe the late genes.

So the delayed early genes serve two masters: some are needed for the phage to enter lysogeny, the others are concerned with controlling the order of the lytic cycle.

Reviews

100. Ptashne, M. (1992). *A genetic switch*. The Genetic Switch.

PHAGE STRATEGIES

3.12.7 The lytic cycle depends on antitermination

Key Concepts

- pN is an antitermination factor that allows RNA polymerase to continue transcription past the ends of the two immediate early genes.
- pQ is the product of a delayed early gene and is an antiterminator that allows RNA polymerase to transcribe the late genes.
- Because lambda DNA circularizes after infection, the late genes form a single transcription unit.

To disentangle the two pathways, let's first consider just the lytic cycle. **Figure 12.11** gives the map of lambda phage DNA. A group of genes concerned with regulation is surrounded by genes needed for recombination and replication. The genes coding for structural components of the phage are clustered. All of the genes necessary for the lytic cycle are expressed in polycistronic transcripts from three promoters.

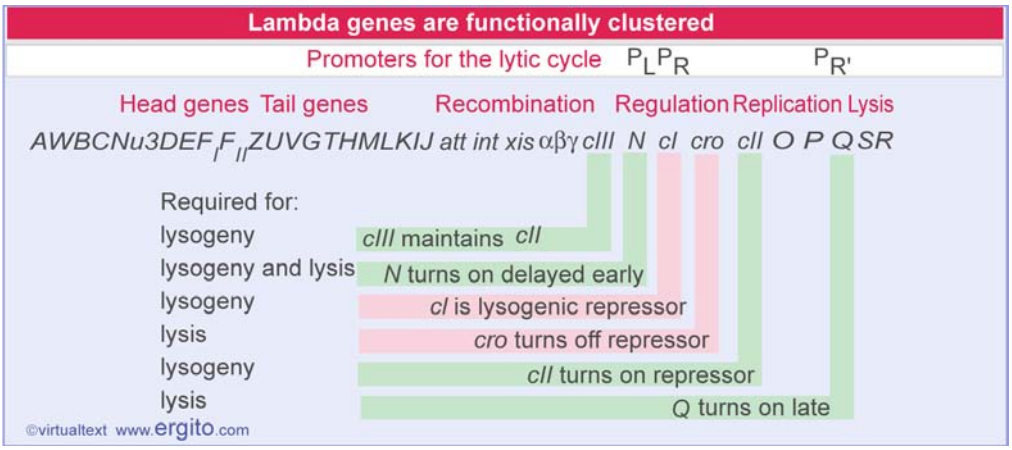

Figure 12.11 The lambda map shows clustering of related functions. The genome is 48,514 bp.

Figure 12.12 shows that the two immediate early genes, *N* and *cro*, are transcribed by host RNA polymerase. *N* is transcribed toward the left, and *cro* toward the right. Each transcript is terminated at the end of the gene. pN is the regulator that allows transcription to continue into the delayed early genes. It is an antitermination factor that suppresses use of the terminators t_L and t_R (see *Molecular Biology 3.9.25 Termination and anti-termination factors interact with RNA polymerase*). In the presence of pN, transcription continues to the left of *N* into the recombination genes, and to the right of *cro* into the replication genes.

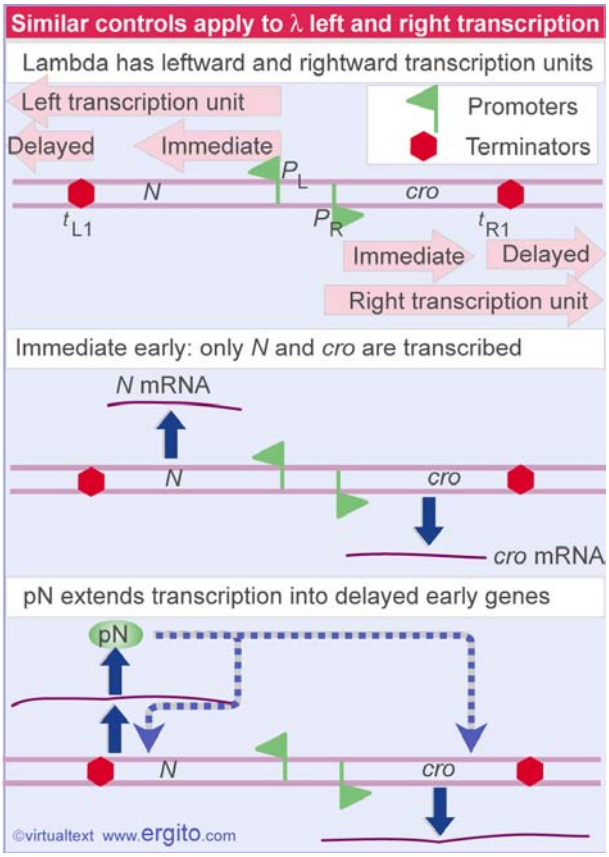

Figure 12.12 Phage lambda has two early transcription units; in the "leftward" unit, the "upper" strand is transcribed toward the left; in the "rightward" unit, the "lower" strand is transcribed toward the right. . Genes *N* and *cro* are the immediate early functions, and are separated from the delayed early genes by the terminators. Synthesis of N protein allows RNA polymerase to pass the terminators t_{L1} to the left and t_{R1} to the right.

The map in **Figure 12.11** gives the organization of the lambda DNA as it exists in the phage particle. But shortly after infection, the ends of the DNA join to form a circle. **Figure 12.13** shows the true state of lambda DNA during infection. The late genes are welded into a single group, containing the lysis genes *S-R* from the right end of the linear DNA, and the head and tail genes *A-J* from the left end.

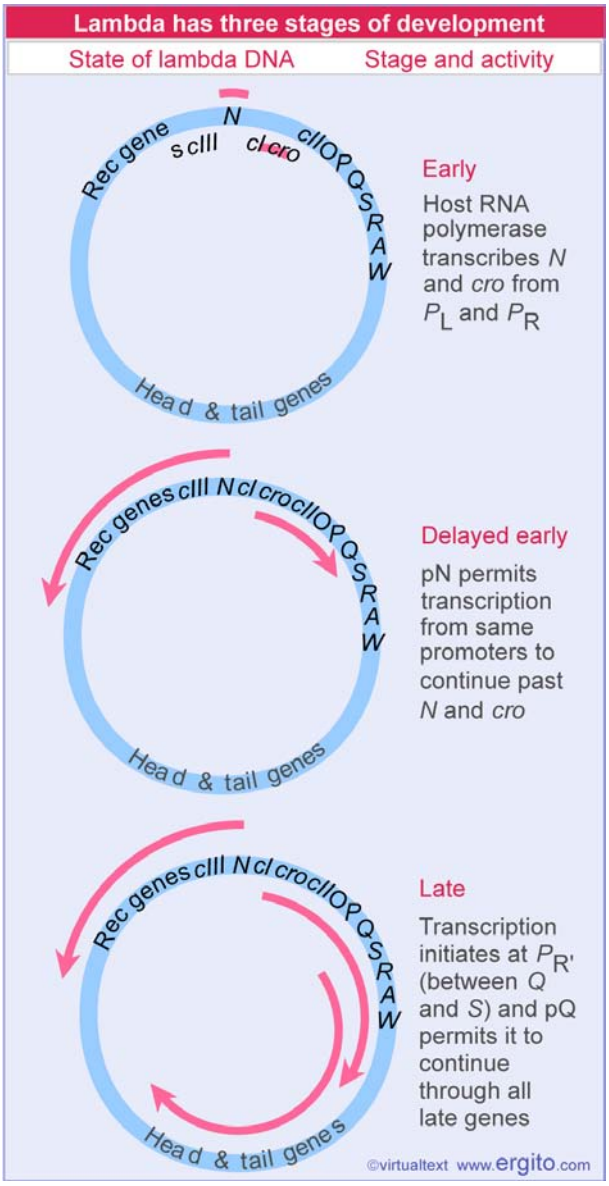

Figure 12.13 Lambda DNA circularizes during infection, so that the late gene cluster is intact in one transcription unit.

The late genes are expressed as a single transcription unit, starting from a promoter $P_{R'}$ that lies between *Q* and *S*. The late promoter is used constitutively. However, in the absence of the product of gene *Q* (which is the last gene in the rightward delayed early unit), late transcription terminates at a site t_{R3}. The transcript resulting from this termination event is 194 bases long; it is known as 6S RNA. When pQ becomes available, it suppresses termination at t_{R3} and the 6S RNA is extended, with the result that the late genes are expressed.

Late gene transcription does not seem to terminate at any specific point, but continues through all the late genes into the region beyond. A similar event happens with the leftward delayed early transcription, which continues past the recombination functions. Transcription in each direction is probably terminated before the

polymerases could crash into each other.

PHAGE STRATEGIES

3.12.8 Lysogeny is maintained by repressor protein

Key Terms

A **plaque** is an area of clearing in a bacterial lawn. It is created by a single phage particle that has undergone multiple rounds of lytic growth.

A **clear plaque** is a type of plaque that contains only lysed bacterial cells.

Key Concepts

- Mutants in the *cI* gene cannot maintain lysogeny.
- *cI* codes for a repressor protein that acts at the O_L and O_R operators to block transcription of the immediate early genes.
- Because the immediate early genes trigger a regulatory cascade, their repression prevents the lytic cycle from proceeding.

Looking at the lambda lytic cascade, we see that the entire program is set in train by initiating transcription at the two promoters P_L and P_R for the immediate early genes *N* and *cro*. Because lambda uses antitermination to proceed to the next stage of (delayed early) expression, the same two promoters continue to be used throughout the early period.

The expanded map of the regulatory region drawn in **Figure 12.14** shows that the promoters P_L and P_R lie on either side of the *cI* gene. Associated with each promoter is an operator (O_L, O_R) at which repressor protein binds to prevent RNA polymerase from initiating transcription. The sequence of each operator overlaps with the promoter that it controls; so often these are described as the P_L/O_L and P_R/O_R control regions.

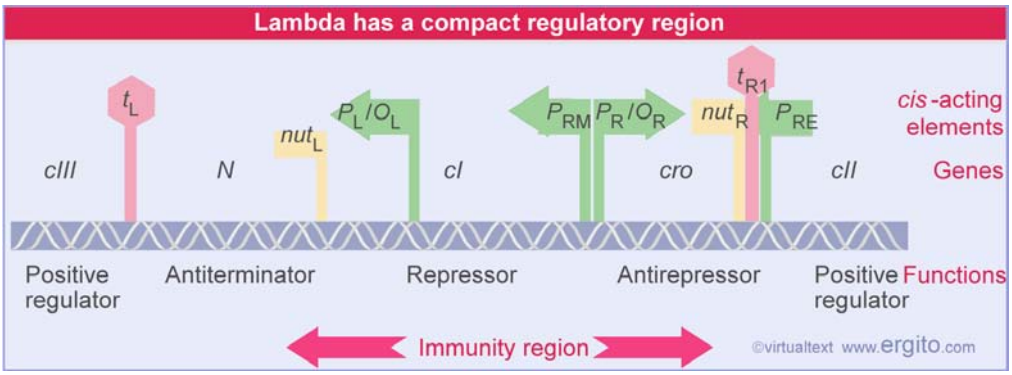

Figure 12.14 The lambda regulatory region contains a cluster of *trans*-acting functions and *cis*-acting elements.

Because of the sequential nature of the lytic cascade, the control regions provide a pressure point at which entry to the entire cycle can be controlled. *By denying RNA*

polymerase access to these promoters, a repressor protein prevents the phage genome from entering the lytic cycle. The repressor functions in the same way as repressors of bacterial operons: it binds to specific operators (525; 526; 527).

The repressor protein is coded by the *cI* gene. Mutants in this gene cannot maintain lysogeny, but always enter the lytic cycle. Since the original isolation of the repressor protein (525; see *Great Experiments 2.1 Isolation of repressor*), its characterization has shown how it both maintains the lysogenic state and provides immunity for a lysogen against superinfection by new phage lambda genomes.

When a bacterial culture is infected with a phage, the cells are lysed to generate regions that can be seen on a culture plate as small areas of clearing called **plaques**. With wild-type phages, the plaques are turbid or cloudy, because they contain some cells that have established lysogeny instead of being lysed. The effect of a *cI* mutation is to prevent lysogeny, so that the plaques contain only lysed cells. As a result, such an infection generates only **clear plaques**, and three genes (*cI*, *cII*, *cIII*) were named for their involvement in this phenotype. **Figure 12.15** compares wild-type and mutant plaques.

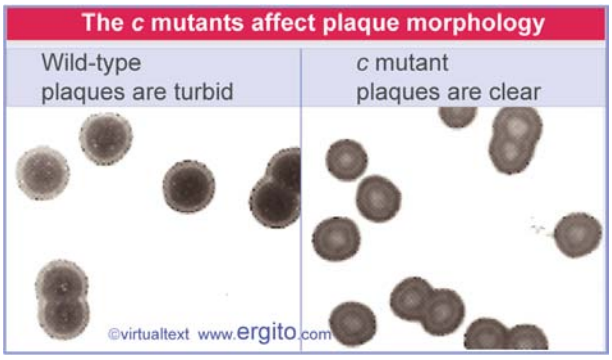

Figure 12.15 Wild-type and virulent lambda mutants can be distinguished by their plaque types. Photograph kindly provided by Dale Kaiser.

The *cI* gene is transcribed from a promoter P_{RM} that lies at its right end. (The subscript "RM" stands for repressor maintenance.) Transcription is terminated at the left end of the gene. The mRNA starts with the AUG initiation codon; because of the absence of the usual ribosome binding site, the mRNA is translated somewhat inefficiently, producing only a low level of repressor protein.

References

525. Ptashne, M. (1967). *Isolation of the λ phage repressor*. Proc. Natl. Acad. Sci. USA 57, 306-313.

526. Ptashne, M. (1967). *Specific binding of the λ phage repressor to DNA*. Nature 214, 232-234.

527. Pirrotta, V., Chadwick, P., and Ptashne, M. (1970). *Active form of two coliphage repressors*. Nature 227, 41-44.

PHAGE STRATEGIES

3.12.9 Repressor maintains an autogenous circuit

Key Concepts

- Repressor binding at O_L blocks transcription of gene *N* from P_L.
- Repressor binding at O_R blocks transcription of *cro* but also is required for transcription of *cI*.
- Repressor binding to the operators therefore simultaneously blocks entry to the lytic cycle and promotes its own synthesis.

The repressor binds independently to the two operators. It has a single function at O_L, but has dual functions at O_R. These are illustrated in the upper part of **Figure 12.16**.

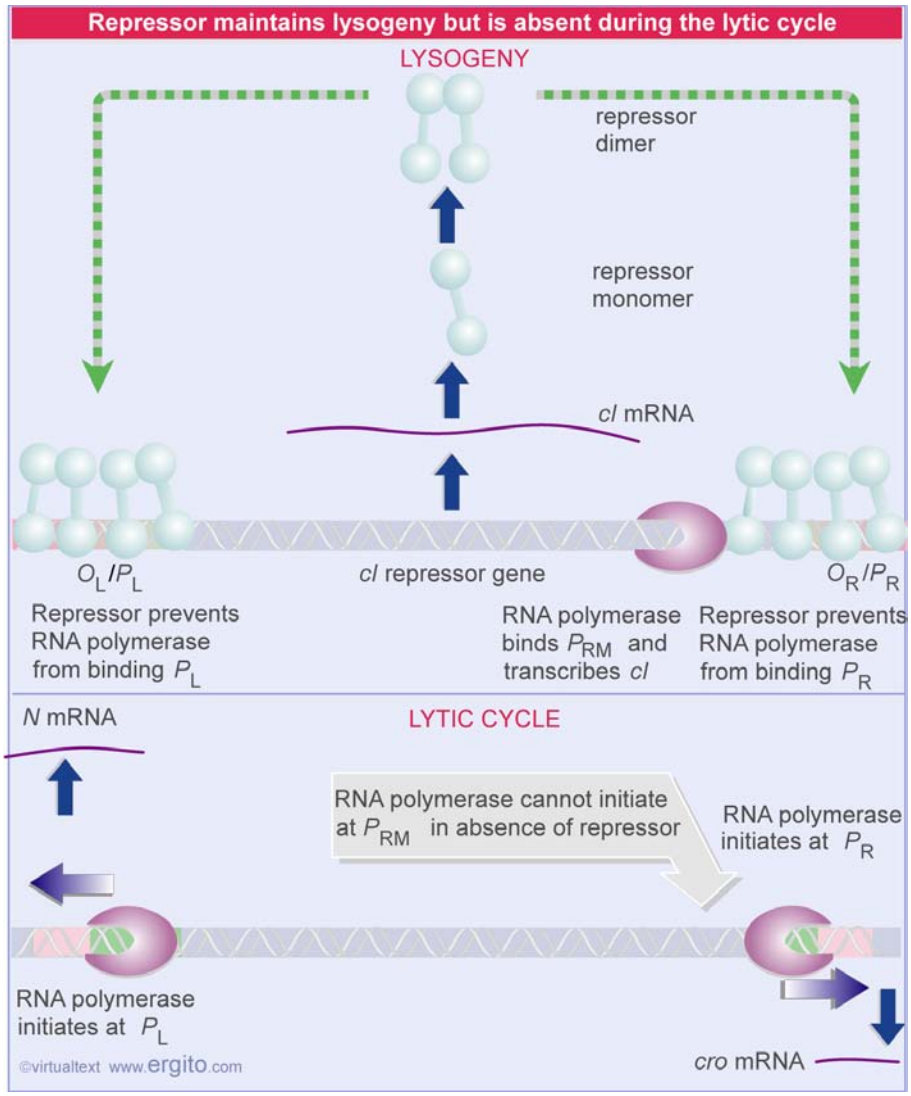

Figure 12.16 Lysogeny is maintained by an autogenous circuit (upper). If this circuit is interrupted, the lytic cycle starts (lower).

At O_L the repressor has the same sort of effect that we have already discussed for several other systems: it prevents RNA polymerase from initiating transcription at P_L. This stops the expression of gene *N*. Since P_L is used for all leftward early gene transcription, this action prevents expression of the entire leftward early transcription unit. *So the lytic cycle is blocked before it can proceed beyond the early stages.*

At O_R, repressor binding prevents the use of P_R. So *cro* and the other rightward early genes cannot be expressed. (We see later why it is important to prevent the expression of *cro* when lysogeny is being maintained.)

But the presence of repressor at O_R also has another effect. The promoter for repressor synthesis, P_{RM}, is adjacent to the rightward operator O_R. It turns out that *RNA polymerase can initiate efficiently at P_{RM} only when repressor is bound at O_R.* The repressor behaves as a positive regulator protein that is necessary for transcription of the *cI* gene (see *Molecular Biology 3.12.15 Repressor at O_R2*

interacts with RNA polymerase at P_{RM}). Since the repressor is the product of cI, *this interaction creates a positive autogenous circuit, in which the presence of repressor is necessary to support its own continued synthesis.*

The nature of this control circuit explains the biological features of lysogenic existence. Lysogeny is stable because the control circuit ensures that, so long as the level of repressor is adequate, there is continued expression of the *cI* gene. The result is that O_L and O_R remain occupied indefinitely. By repressing the entire lytic cascade, this action maintains the prophage in its inert form.

PHAGE STRATEGIES

3.12.10 The repressor and its operators define the immunity region

Key Terms

Immunity in phages refers to the ability of a prophage to prevent another phage of the same type from infecting a cell. It results from the synthesis of phage repressor by the prophage genome.

Virulent phage mutants are unable to establish lysogeny.

The **immunity region** is a segment of the phage genome that enables a prophage to inhibit additional phage of the same type from infecting the bacterium. This region has a gene that encodes for the repressor, as well as the sites to which the repressor binds.

Key Concepts

- Several lambdoid phages have different immunity regions.
- A lysogenic phage confers immunity to further infection by any other phage with the same immunity region.

The presence of repressor explains the phenomenon of **immunity**. If a second lambda phage DNA enters a lysogenic cell, repressor protein synthesized from the resident prophage genome will immediately bind to O_L and O_R in the new genome. This prevents the second phage from entering the lytic cycle.

The operators were originally identified as the targets for repressor action by **virulent** mutations (λ*vir*). These mutations prevent the repressor from binding at O_L or O_R, with the result that the phage inevitably proceeds into the lytic pathway when it infects a new host bacterium. And λ*vir* mutants can grow on lysogens because the virulent mutations in O_L and O_R allow the incoming phage to ignore the resident repressor and thus to enter the lytic cycle. Virulent mutations in phages are the equivalent of operator-constitutive mutations in bacterial operons.

Prophage is induced to enter the lytic cycle when the lysogenic circuit is broken. This happens when the repressor is inactivated (see next section). The absence of repressor allows RNA polymerase to bind at P_L and P_R, starting the lytic cycle as shown in the lower part of **Figure 12.16** (for review see 104).

The autogenous nature of the repressor-maintenance circuit creates a sensitive response. Because the presence of repressor is necessary for its own synthesis, expression of the *cI* gene stops as soon as the existing repressor is destroyed. So no repressor is synthesized to replace the molecules that have been damaged. This enables the lytic cycle to start without interference from the circuit that maintains lysogeny.

The region including the left and right operators, the *cI* gene, and the *cro* gene determines the immunity of the phage. Any phage that possesses this region has the same type of immunity, because *it specifies both the repressor protein and the sites on which the repressor acts.* Accordingly, this is called the **immunity region** (as marked in **Figure 12.14**). Each of the four lambdoid phages ϕ80, *21*, *434*, and λ has a unique immunity region. When we say that a lysogenic phage confers immunity to any other phage of the same type, we mean more precisely that the immunity is to any other phage that has the same immunity region (irrespective of differences in other regions).

Reviews

104. Friedman, D. I. and Gottesman, M. (1982). *Lytic mode of lambda development In Lambda*. Lambda II, 21-51.

PHAGE STRATEGIES

3.12.11 The DNA-binding form of repressor is a dimer

Key Concepts

- A repressor monomer has two distinct domains.
- The N-terminal domain contains the DNA-binding site.
- The C-terminal domain dimerizes.
- Binding to the operator requires the dimeric form so that two DNA-binding domains can contact the operator simultaneously.
- Cleavage of the repressor between the two domains reduces the affinity for the operator and induces a lytic cycle.

The repressor subunit is a polypeptide of 27 kD with the two distinct domains summarized in **Figure 12.17**.

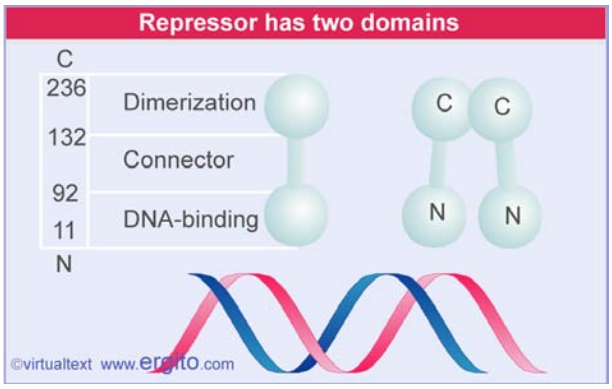

Figure 12.17 The N-terminal and C-terminal regions of repressor form separate domains. The C-terminal domains associate to form dimers; the N-terminal domains bind DNA.

- The N-terminal domain, residues 1-92, provides the operator-binding site.
- The C-terminal domain, residues 132-236, is responsible for dimerization.

The two domains are joined by a connector of 40 residues. When repressor is digested by a protease, each domain is released as a separate fragment.

Each domain can exercise its function independently of the other. The C-terminal fragment can form oligomers. The N-terminal fragment can bind the operators, although with a lower affinity than the intact repressor. So the information for

specifically contacting DNA is contained within the N-terminal domain, but the efficiency of the process is enhanced by the attachment of the C-terminal domain (529).

The dimeric structure of the repressor is crucial in maintaining lysogeny. The induction of a lysogenic prophage to enter the lytic cycle is caused by cleavage of the repressor subunit in the connector region, between residues 111 and 113. (This is a counterpart to the allosteric change in conformation that results when a small-molecule inducer inactivates the repressor of a bacterial operon, a capacity that the lysogenic repressor does not have.) Induction occurs under certain adverse conditions, such as exposure of lysogenic bacteria to UV irradiation, which leads to proteolytic inactivation of the repressor.

In the intact state, dimerization of the C-terminal domains ensures that when the repressor binds to DNA its two N-terminal domains each contact DNA simultaneously. But cleavage releases the C-terminal domains from the N-terminal domains. As illustrated in **Figure 12.18** this means that the N-terminal domains can no longer dimerize; this upsets the equilibrium between monomers and dimers, so that repressor dissociates from DNA, allowing lytic infection to start. (Another relevant parameter is the loss of cooperative effects between adjacent dimers.)

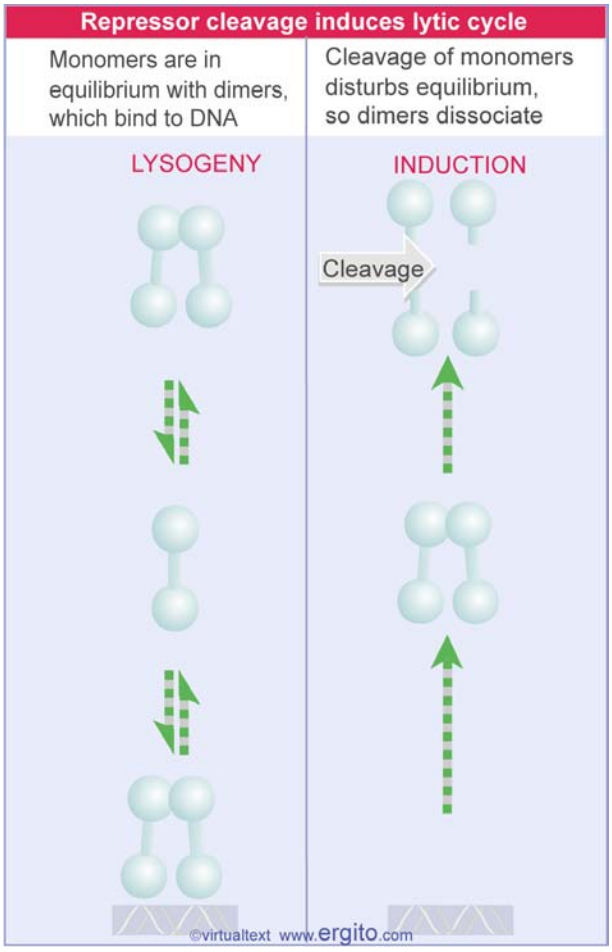

Figure 12.18 Repressor dimers bind to the operator. The affinity of the N-terminal domains for DNA is controlled by the dimerization of the C-terminal domains.

The balance between lysogeny and the lytic cycle depends on the concentration of repressor. Intact repressor is present in a lysogenic cell at a concentration sufficient to ensure that the operators are occupied. But if the repressor is cleaved, this concentration is inadequate, because of the lower affinity of the separate N-terminal domain for the operator. Too high a concentration of repressor would make it impossible to induce the lytic cycle in this way; too low a level, of course, would make it impossible to maintain lysogeny.

References

529. Pabo, C. O. and Lewis, M. (1982). *The operator-binding domain of λ repressor: structure and DNA recognition*. Nature 298, 443-447.

PHAGE STRATEGIES

3.12.12 Repressor uses a helix-turn-helix motif to bind DNA

Key Terms

The **helix-turn-helix** motif describes an arrangement of two α helices that form a site that binds to DNA, one fitting into the major groove of DNA and other lying across it.

Key Concepts

- Each DNA-binding region in the repressor contacts a half-site in the DNA.
- The DNA-binding site of repressor includes two short α-helical regions which fit into the successive turns of the major groove of DNA.
- A DNA-binding site is a (partially) palindromic sequence of 17 bp.

A repressor dimer is the unit that binds to DNA. It recognizes a sequence of 17 bp displaying partial symmetry about an axis through the central base pair. **Figure 12.19** shows an example of a binding site. The sequence on each side of the central base pair is sometimes called a "half-site". Each individual N-terminal region contacts a half-site. Several DNA-binding proteins that regulate bacterial transcription share a similar mode of holding DNA, in which the active domain contains two short regions of α-helix that contact DNA. (Some transcription factors in eukaryotic cells use a similar motif; see *Molecular Biology 5.22.14 Homeodomains bind related targets in DNA*.)

Figure 12.19 The operator is a 17 bp sequence with an axis of symmetry through the central base pair. Each half site is marked in green. Base pairs that are identical in each operator half are in red.

The N-terminal domain of lambda repressor contains several stretches of α-helix, arranged as illustrated diagrammatically in **Figure 12.20**. Two of the helical regions are responsible for binding DNA. The **helix-turn-helix** model for contact is illustrated in **Figure 12.21**. Looking at a single monomer, α-helix-3 consists of 9 amino acids, lying at an angle to the preceding region of 7 amino acids that forms α-helix-2. In the dimer, the two apposed helix-3 regions lie 34 Å apart, enabling them to fit into successive major grooves of DNA. The helix-2 regions lie at an angle that would place them across the groove (530). The symmetrical binding of dimer to the site means that each N-terminal domain of the dimer contacts a similar set of

bases in its half-site.

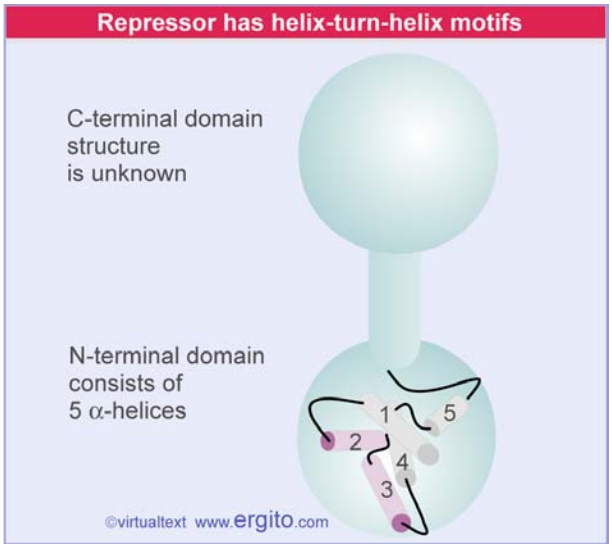

Figure 12.20 Lambda repressor"s N-terminal domain contains five stretches of α -helix; helices 2 and 3 are involved in binding DNA.

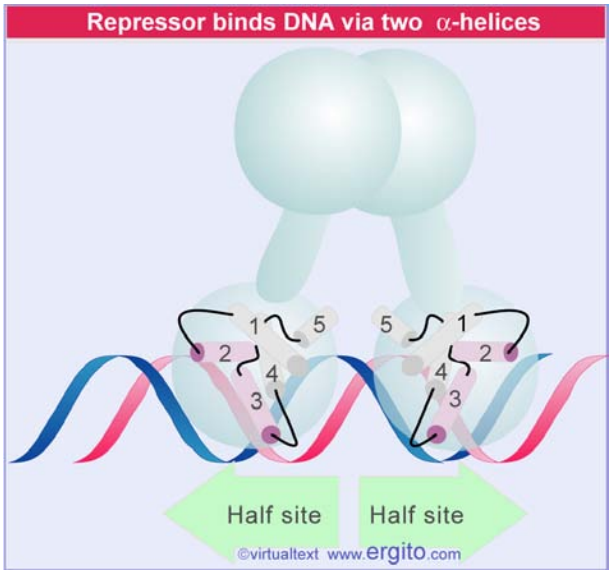

Figure 12.21 In the two-helix model for DNA binding, helix-3 of each monomer lies in the wide groove on the same face of DNA, and helix-2 lies across the groove.

Last updated on 2-28-2002

References

530. Sauer, R. T. et al. (1982). *Homology among DNA-binding proteins suggests use of a conserved super-secondary structure*. Nature 298, 447-451.

PHAGE STRATEGIES

3.12.13 The recognition helix determines specificity for DNA

Key Terms

The **recognition helix** is the one of the two helices of the helix-turn-helix motif that makes contacts with DNA that are specific for particular bases. This determines the specificity of the DNA sequence that is bound.

Key Concepts

- The amino acid sequence of the recognition helix makes contacts with particular bases in the operator sequence that it recognizes.

Related forms of the α-helical motifs employed in the helix-loop-helix of the lambda repressor are found in several DNA-binding proteins, including CRP, the *lac* repressor, and several other phage repressors. By comparing the abilities of these proteins to bind DNA, we can define the roles of each helix:

- Contacts between helix-3 and DNA rely on hydrogen bonds between the amino acid side chains and the exposed positions of the base pairs. This helix is responsible for recognizing the specific target DNA sequence, and is therefore also known as the **recognition helix**.
- Contacts from helix-2 to the DNA take the form of hydrogen bonds connecting with the phosphate backbone. These interactions are necessary for binding, but do not control the specificity of target recognition. In addition to these contacts, a large part of the overall energy of interaction with DNA is provided by ionic interactions with the phosphate backbone.

What happens if we manipulate the coding sequence to construct a new protein by substituting the recognition helix in one repressor with the corresponding sequence from a closely related repressor? The specificity of the hybrid protein is that of its new recognition helix. *The amino acid sequence of this short region determines the sequence specificities of the individual proteins, and is able to act in conjunction with the rest of the polypeptide chain* (531).

Figure 12.22 shows the details of the binding to DNA of two proteins that bind similar DNA sequences. Both lambda repressor and Cro protein have a similar organization of the helix-turn-helix motif, although their individual specificities for DNA are not identical:

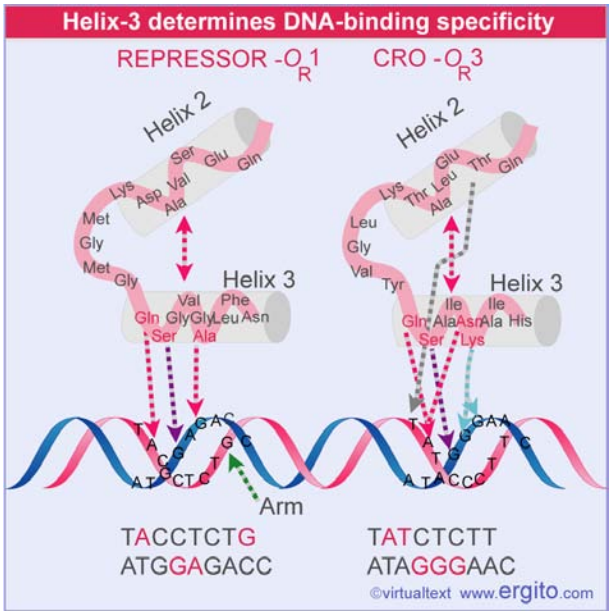

Figure 12.22 Two proteins that use the two-helix arrangement to contact DNA recognize lambda operators with affinities determined by the amino acid sequence of helix-3.

- Each protein uses similar interactions between hydrophobic amino acids to maintain the relationship between helix-2 and helix-3: repressor has an Ala-Val connection, while Cro has an Ala-Ile association.

- Amino acids in helix-3 of the repressor make contacts with specific bases in the operator. Three amino acids in repressor recognize three bases in DNA; the amino acids at these positions and also at additional positions in Cro recognize five (or possibly six) bases in DNA.

Two of the amino acids involved in specific recognition are identical in repressor and Cro (Gln and Ser at the N-terminal end of the helix), while the other contacts are different (Ala in repressor versus Lys and the additional Asn in Cro). Also, a Thr in helix-2 of Cro directly contacts DNA.

The interactions shown in the figure represent binding to the DNA sequence that each protein recognizes most tightly. The sequences shown at the bottom of the figure with the contact points in color differ at 3 of the 9 base pairs. The use of overlapping, but not identical contacts between amino acids and bases shows how related recognition helices confer recognition of related DNA sequences. This enables repressor and Cro to recognize the same set of sequences, but with different relative affinities for particular members of the group.

The bases contacted by helix-3 of repressor or Cro lie on one face of DNA, as can be seen from the positions indicated on the helical diagram in **Figure 12.22**. However, repressor makes an additional contact with the other face of DNA. Removing the last six N-terminal amino acids (which protrude from helix-1) eliminates some of the contacts. This observation provides the basis for the idea that the bulk of the N-terminal domain contacts one face of DNA, while the last six N-terminal amino

acids form an "arm" extending around the back. **Figure 12.23** shows the view from the back. Lysine residues in the arm make contacts with G residues in the major groove, and also with the phosphate backbone. The interaction between the arm and DNA contributes heavily to DNA binding; the affinity of the armless repressor for DNA is reduced by ~1000-fold (532).

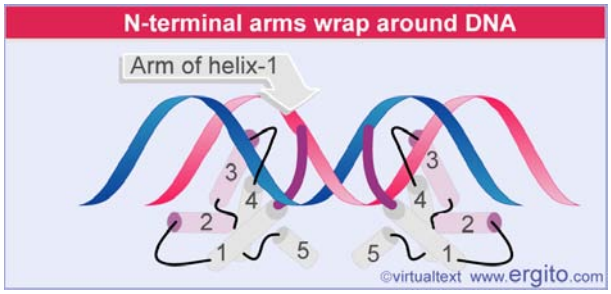

Figure 12.23 A view from the back shows that the bulk of the repressor contacts one face of DNA, but its N-terminal arms reach around to the other face.

Bases that are not contacted directly by repressor protein may have an important effect on binding. The related phage *434* repressor binds DNA via a helix-turn-helix motif, and the crystal structure shows that helix-3 is positioned at each half-site so that it contacts the 5 outermost base pairs but not the inner 2. However, operators with A·T base pairs at the inner positions bind *434* repressor more strongly than operators with G·C base pairs. The reason is that *434* repressor binding slightly twists DNA at the center of the operator, widening the angle between the two half-sites of DNA by ~3°. This is probably needed to allow each monomer of the repressor dimer to make optimal contacts with DNA. A·T base pairs allow this twist more readily than G·C pairs, thus affecting the affinity of the operator for repressor.

Last updated on 2-28-2002

References

531. Wharton, R. L., Brown, E. L., and Ptashne, M. (1984). *Substituting an α -helix switches the sequence specific DNA interactions of a repressor*. Cell 38, 361-369.

532. Brennan, R. G. et al. (1990). *Protein-DNA conformational changes in the crystal structure of a λCro-operator complex*. Proc. Natl. Acad. Sci. USA 87, 8165-8169.

PHAGE STRATEGIES

3.12.14 Repressor dimers bind cooperatively to the operator

Key Concepts

- Repressor binding to one operator increases the affinity for binding a second repressor dimer to the adjacent operator.
- The affinity is 10× greater for O_L1 and O_R1 than other operators, so they are bound first.
- Cooperativity allows repressor to bind the *O*1/*O*2 sites at lower concentrations.

Each operator contains three repressor-binding sites. As can be seen from **Figure 12.24**, no two of the six individual repressor-binding sites are identical, but they all conform with a consensus sequence. The binding sites within each operator are separated by spacers of 3-7 bp that are rich in A·T base pairs. The sites at each operator are numbered so that O_R consists of the series of binding sites O_R1-O_R2-O_R3, while O_L consists of the series O_L1-O_L2-O_L3. In each case, site 1 lies closest to the startpoint for transcription in the promoter, and sites 2 and 3 lie farther upstream.

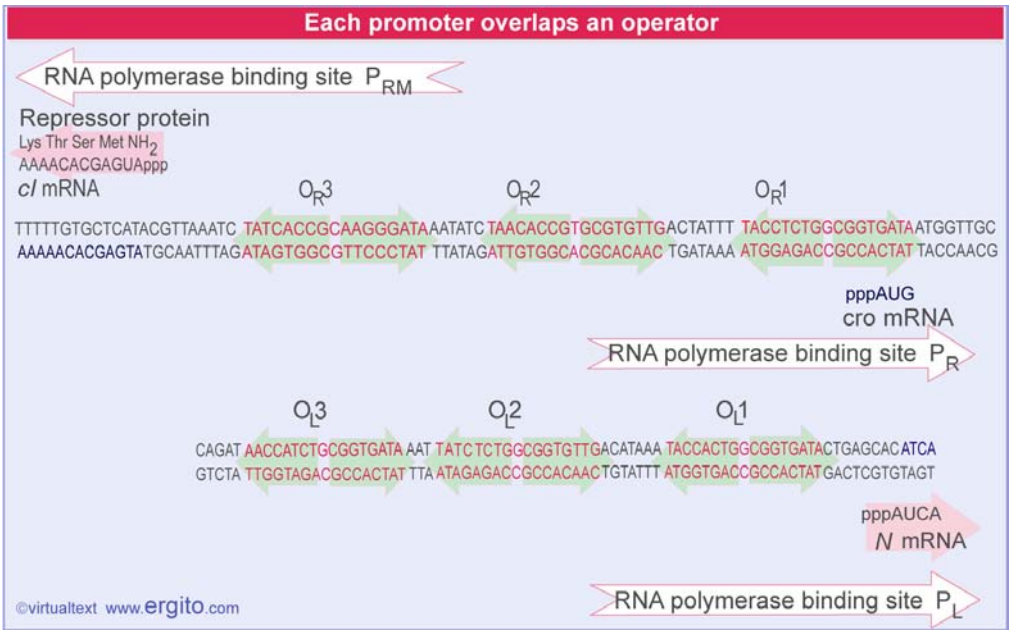

Figure 12.24 Each operator contains three repressor-binding sites, and overlaps with the promoter at which RNA polymerase binds. The orientation of O_L has been reversed from usual to facilitate comparison with O_R.

Faced with the triplication of binding sites at each operator, how does repressor

decide where to start binding? At each operator, site 1 has a greater affinity (roughly tenfold) than the other sites for the repressor. So the repressor always binds first to O_L1 and O_R1.

Lambda repressor binds to subsequent sites within each operator in a cooperative manner. The presence of a dimer at site 1 greatly increases the affinity with which a second dimer can bind to site 2. When both sites 1 and 2 are occupied, this interaction does *not* extend farther, to site 3. At the concentrations of repressor usually found in a lysogen, both sites 1 and 2 are filled at each operator, but site 3 is not occupied (528).

If site 1 is inactive (because of mutation), then repressor binds cooperatively to sites 2 and 3. That is, binding at site 2 assists another dimer to bind at site 3. This interaction occurs directly between repressor dimers and not via conformational change in DNA. The C-terminal domain is responsible for the cooperative interaction between dimers as well as for the dimer formation between subunits. **Figure 12.25** shows that it involves both subunits of each dimer, that is, each subunit contacts its counterpart in the other dimer, forming a tetrameric structure (1199).

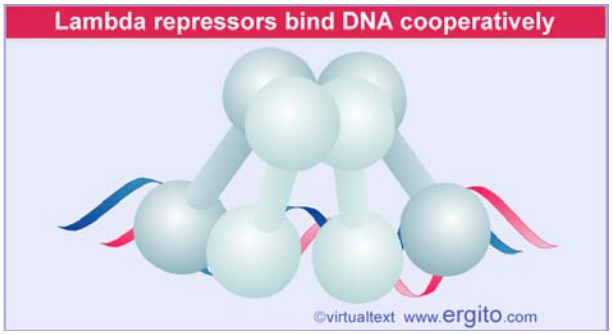

Figure 12.25 When two lambda repressor dimers bind cooperatively, each of the subunits of one dimer contacts a subunit in the other dimer.

A result of cooperative binding is to increase the effective affinity of repressor for the operator at physiological concentrations. This enables a lower concentration of repressor to achieve occupancy of the operator. This is an important consideration in a system in which release of repression has irreversible consequences. In an operon coding for metabolic enzymes, after all, failure of repression will merely allow unnecessary synthesis of enzymes. But failure to repress lambda prophage will lead to induction of phage and lysis of the cell.

From the sequences shown in **Figure 12.24**, we see that O_L1 and O_R1 lie more or less in the center of the RNA polymerase binding sites of P_L and P_R, respectively. Occupancy of O_L1-O_L2 and O_R1-O_R2 thus physically blocks access of RNA polymerase to the corresponding promoters.

References

528. Johnson, A. D., Meyer, B. J., and Ptashne, M. (1979). *Interactions between DNA-bound repressors govern regulation by the phage λrepressor*. Proc. Natl. Acad. Sci. USA 76, 5061-5065.

1199. Bell, C. E., Frescura, P., Hochschild, A., and Lewis, M. (2000). *Crystal structure of the lambda repressor C-terminal domain provides a model for cooperative operator binding.* Cell 101, 801-811.

PHAGE STRATEGIES

3.12.15 Repressor at O_R2 interacts with RNA polymerase at P_{RM}

Key Concepts

- The DNA-binding region of repressor at O_R2 contacts RNA polymerase and stabilizes its binding to P_{RM}.
- This is the basis for the autogenous control of repressor maintenance.

A different relationship is shown between O_R and the promoter P_{RM} for transcription of *cI*. The RNA polymerase binding site is adjacent to O_R2. This explains how repressor autogenously regulates its own synthesis. When two dimers are bound at O_R1-O_R2, the dimer at O_R2 interacts with RNA polymerase (see **Figure 12.16** in *Molecular Biology 3.12.9 Repressor maintains an autogenous circuit*). This effect resides in the amino terminal domain of repressor.

Mutations that abolish positive control map in the *cI* gene. One interesting class of mutants remain able to bind the operator to repress transcription, but cannot stimulate RNA polymerase to transcribe from P_{RM}. They map within a small group of amino acids, located on the outside of helix-2 or in the turn between helix-2 and helix-3. The mutations reduce the negative charge of the region; conversely, mutations that increase the negative charge enhance the activation of RNA polymerase. This suggests that the group of amino acids constitutes an "acidic patch" that functions by an electrostatic interaction with a basic region on RNA polymerase.

The location of these "positive control mutations" in the repressor is indicated on **Figure 12.26**. They lie at a site on repressor that is close to a phosphate group on DNA that is also close to RNA polymerase. So the group of amino acids on repressor that is involved in positive control is in a position to contact the polymerase. The interaction between repressor and polymerase is needed for the polymerase to make the transition from a closed complex to an open complex (see also **Figure 12.29**). The important principle is that *protein-protein interactions can release energy that is used to help to initiate transcription.*

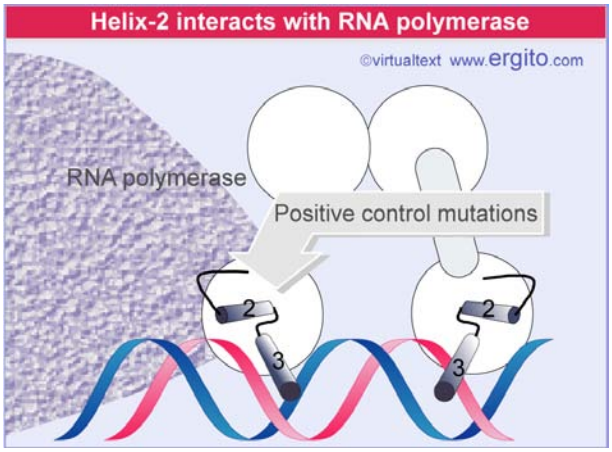

Figure 12.26 Positive control mutations identify a small region at helix-2 that interacts directly with RNA polymerase.

What happens if a repressor dimer binds to O_R3? This site overlaps with the RNA polymerase binding site at P_{RM}. So if the repressor concentration becomes great enough to cause occupancy of O_R3, the transcription of *cI* is prevented. This leads in due course to a reduction in repressor concentration; O_R3 then becomes empty, and the autogenous loop can start up again because O_R2 remains occupied.

This mechanism could prevent the concentration of repressor from becoming too great, although it would require repressor concentration in lysogens to reach unusually high levels. In the formal sense, the repressor is an autogenous regulator of its own expression that functions positively at low concentrations and negatively at high concentrations.

Virulent mutations occur in sites 1 and 2 of both O_L and O_R. The mutations vary in their degree of virulence, according to the extent to which they reduce the affinity of the binding site for repressor, and also depending on the relationship of the affected site to the promoter. Consistent with the conclusion that O_R3 and O_L3 usually are not occupied, virulent mutations are not found in these sites.

Last updated on 10-17-2000

PHAGE STRATEGIES

3.12.16 The cII and cIII genes are needed to establish lysogeny

Key Concepts

- The delayed early gene products cII and cIII are necessary for RNA polymerase to initiate transcription at the promoter P_{RE}.
- cII acts direct at the promoter and cIII protects cII from degradation.
- Transcription from P_{RE} leads to synthesis of repressor and also blocks the transcription of *cro*.

The control circuit for maintaining lysogeny presents a paradox. *The presence of repressor protein is necessary for its own synthesis.* This explains how the lysogenic condition is perpetuated. But how is the synthesis of repressor established in the first place?

When a lambda DNA enters a new host cell, RNA polymerase cannot transcribe *cI*, because there is no repressor present to aid its binding at P_{RM}. But this same absence of repressor means that P_R and P_L are available. So the first event when lambda DNA infects a bacterium is for genes *N* and *cro* to be transcribed. Then pN allows transcription to be extended farther. This allows *cIII* (and other genes) to be transcribed on the left, while *cII* (and other genes) are transcribed on the right (see **Figure 12.14**).

The *cII* and *cIII* genes share with *cI* the property that mutations in them cause clear plaques. But there is a difference. The *cI* mutants can neither establish nor maintain lysogeny. The *cII* or *cIII* mutants have some difficulty in establishing lysogeny, but once established, they are able to maintain it by the *cI* autogenous circuit.

This implicates the *cII* and *cIII* genes as positive regulators whose products are needed for an alternative system for repressor synthesis. The system is needed only to *initiate* the expression of *cI* in order to circumvent the inability of the autogenous circuit to engage in *de novo* synthesis. They are not needed for continued expression.

The cII protein acts directly on gene expression. Between the *cro* and *cII* genes is another promoter, called P_{RE}. (The subscript "RE" stands for repressor establishment.) This promoter can be recognized by RNA polymerase only in the presence of cII, whose action is illustrated in **Figure 12.27**.

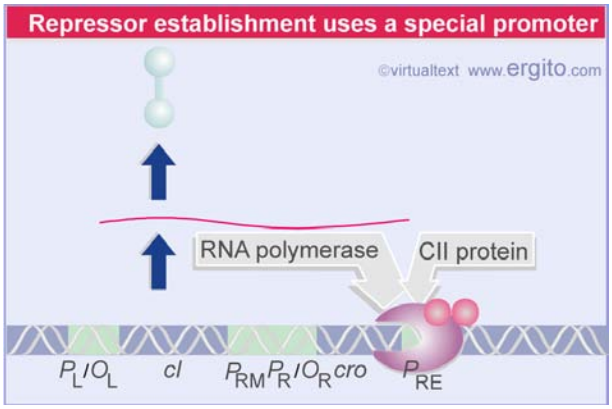

Figure 12.27 Repressor synthesis is established by the action of cII and RNA polymerase at P_{RE} to initiate transcription that extends from the antisense strand of *cro* through the *cI* gene.

The cII protein is extremely unstable *in vivo*, because it is degraded as the result of the activity of a host protein called HflA. The role of cIII is to protect cII against this degradation.

Transcription from P_{RE} promotes lysogeny in two ways. Its direct effect is that *cI* is translated into repressor protein. An indirect effect is that transcription proceeds through the *cro* gene in the "wrong" direction. So the 5′ part of the RNA corresponds to an antisense transcript of *cro;* in fact, it hybridizes to authentic *cro* mRNA, inhibiting its translation. This is important because *cro* expression is needed to enter the lytic cycle (see *Molecular Biology 3.12.19 The cro repressor is needed for lytic infection*).

The *cI* coding region on the P_{RE} transcript is very efficiently translated, in contrast with the weak translation of the P_{RM} transcript. In fact, repressor is synthesized ~7-8 times more effectively via expression from P_{RE} than from P_{RM}. This reflects the fact that the P_{RE} transcript has an efficient ribosome-binding site, whereas the P_{RM} transcript has no ribosome-binding site and actually starts with the AUG initiation codon.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.3.12.16

PHAGE STRATEGIES

3.12.17 A poor promoter requires cII protein

Key Concepts

- P_{RE} has atypical sequences at –10 and –35.
- RNA polymerase binds the promoter only in the presence of cII.
- cII binds to sequences close to the –35 region.

The P_{RE} promoter has a poor fit with the consensus at –10 and lacks a consensus sequence at –35. This deficiency explains its dependence on *cII*. The promoter cannot be transcribed by RNA polymerase alone *in vitro*, but can be transcribed when cII is added. The regulator binds to a region extending from about –25 to –45. When RNA polymerase is added, an additional region is protected, extending from –12 to +13. As summarized in **Figure 12.28**, the two proteins bind to overlapping sites.

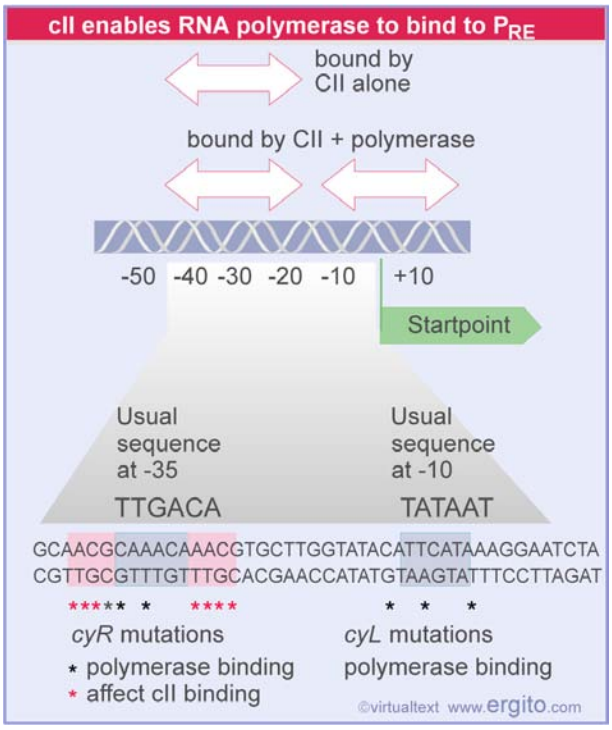

Figure 12.28 RNA polymerase binds to P_{RE} only in the presence of cII, which contacts the region around -35.

The importance of the –35 and –10 regions for promoter function, in spite of their lack of resemblance with the consensus, is indicated by the existence of *cy* mutations. These have effects similar to those of *cII* and *cIII* mutations in preventing the establishment of lysogeny; but they are *cis*-acting instead of *trans*-acting. They fall

into two groups, *cyL* and *cyR*, localized at the consensus operator positions of –10 and –35.

The *cyL* mutations are located around –10, and probably prevent RNA polymerase from recognizing the promoter.

The *cyR* mutations are located around –35, and fall into two types, affecting either RNA polymerase or cII binding. Mutations in the center of the region do not affect cII binding; presumably they prevent RNA polymerase binding. On either side of this region, mutations in short tetrameric repeats, TTGC, prevent cII from binding. Each base in the tetramer is 10 bp (one helical turn) separated from its homologue in the other tetramer, so that when cII recognizes the two tetramers, it lies on one face of the double helix.

Positive control of a promoter implies that an accessory protein has increased the efficiency with which RNA polymerase initiates transcription. **Figure 12.29** reports that either or both stages of the interaction between promoter and polymerase can be the target for regulation. Initial binding to form a closed complex or its conversion into an open complex can be enhanced.

Positive regulation influences initiation

Promoter	Regulator	Polymerase Binding (equilibrium constant, K_B)	Closed-Open Conversion (rate constant, k_2)
P_{RM}	repressor	no effect	11X
P_{RE}	CII	100X	100X

Figure 12.29 Positive regulation can influence RNA polymerase at either stage of initiating transcription.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.3.12.17

PHAGE STRATEGIES

3.12.18 Lysogeny requires several events

Key Concepts

- cII/cIII cause repressor synthesis to be established and also trigger inhibition of late gene transcription.
- Establishment of repressor turns off immediate and delayed early gene expression.
- Repressor turns on the maintenance circuit for its own synthesis.
- Lambda DNA is integrated into the bacterial genome at the final stage in establishing lysogeny.

Now we can see how lysogeny is established during an infection. **Figure 12.30** recapitulates the early stages and shows what happens as the result of expression of *cIII* and *cII*. The presence of cII allows P_{RE} to be used for transcription extending through *cI*. Repressor protein is synthesized in high amounts from this transcript. Immediately it binds to O_L and O_R.

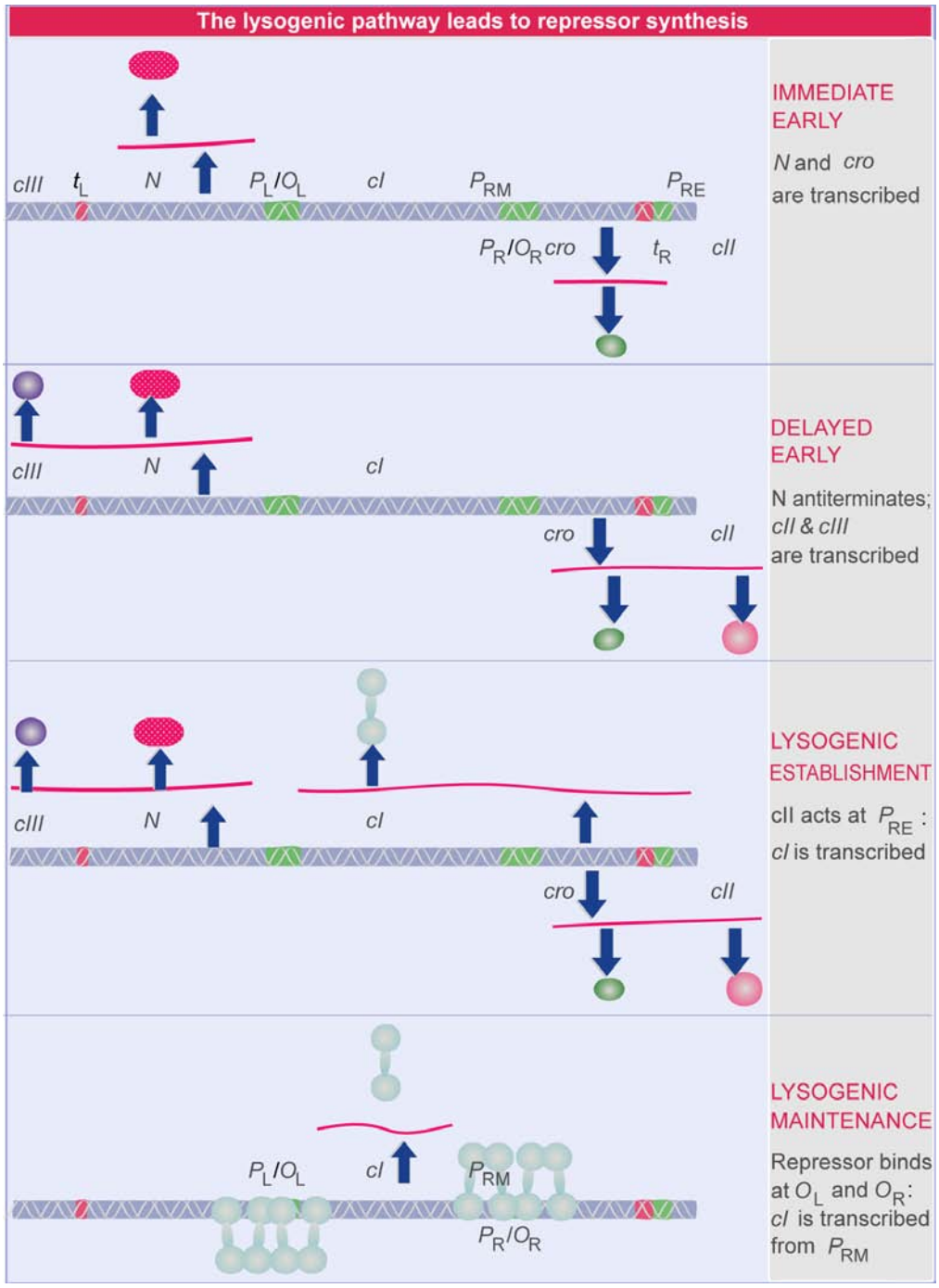

Figure 12.30 A cascade is needed to establish lysogeny, but then this circuit is switched off and replaced by the autogenous repressor-maintenance circuit.

By directly inhibiting any further transcription from P_L and P_R, repressor binding turns off the expression of all phage genes. This halts the synthesis of cII and cIII, which are unstable; they decay rapidly, with the result that P_{RE} can no longer be used. So the synthesis of repressor via the establishment circuit is brought to a halt.

But repressor now is present at O_R. It switches on the maintenance circuit for

expression from P_{RM}. Repressor continues to be synthesized, although at the lower level typical of P_{RM} function. So the establishment circuit starts off repressor synthesis at a high level; then repressor turns off all other functions, while at the same time turning on the maintenance circuit, which functions at the low level adequate to sustain lysogeny.

We shall not now deal in detail with the other functions needed to establish lysogeny, but we can just briefly remark that the infecting lambda DNA must be inserted into the bacterial genome (see *Molecular Biology 4.15.16 Specialized recombination involves specific sites*). The insertion requires the product of gene *int*, which is expressed from its own promoter P_I, at which cII also is necessary. The sequence of P_I shows homology with P_{RE} in the cII binding site (although not in the –10 region). The functions necessary for establishing the lysogenic control circuit are therefore under the same control as the function needed to integrate the phage DNA into the bacterial genome. So the establishment of lysogeny is under a control that ensures all the necessary events occur with the same timing.

Emphasizing the tricky quality of lambda's intricate cascade, we now know that cII promotes lysogeny in another, indirect manner. It sponsors transcription from a promoter called $P_{anti\text{-}Q}$, which is located within the *Q* gene. This transcript is an antisense version of the *Q* region, and it hybridizes with *Q* mRNA to prevent translation of Q protein, whose synthesis is essential for lytic development. So the same mechanisms that directly promote lysogeny by causing transcription of the *cI* repressor gene also indirectly help lysogeny by inhibiting the expression of *cro* (see above) and *Q*, the regulator genes needed for the antagonistic lytic pathway.

PHAGE STRATEGIES

3.12.19 The cro repressor is needed for lytic infection

Key Concepts

- Cro binds to the same operators as repressor but with different affinities.
- When Cro binds to O_R3, it prevents RNA polymerase from binding to P_{RM}, and blocks maintenance of repressor.
- When Cro binds to other operators at O_R or O_L, it prevents RNA polymerase from expressing immediate early genes, which (indirectly) blocks repressor establishment.

Lambda has the alternatives of entering lysogeny or starting a lytic infection. Lysogeny is initiated by establishing an autogenous maintenance circuit that inhibits the entire lytic cascade through applying pressure at two points. The program for establishing lysogeny proceeds through some of the same events that are required for the lytic cascade (expression of delayed early genes via expression of *N* is needed). We now face a problem. How does the phage enter the lytic cycle?

The key influence on the lytic cycle is the role of gene *cro,* which codes for another repressor. *Cro is responsible for preventing the synthesis of the repressor protein;* this action shuts off the possibility of establishing lysogeny. *cro* mutants usually establish lysogeny rather than entering the lytic pathway, because they lack the ability to switch events away from the expression of repressor.

Cro forms a small dimer (the subunit is 9 kD) that acts within the immunity region. It has two effects:

- It prevents the synthesis of repressor via the maintenance circuit; that is, it prevents transcription via P_{RM}.
- It also inhibits the expression of early genes from both P_L and P_R.

This means that, when a phage enters the lytic pathway, Cro has responsibility both for preventing the synthesis of repressor and (subsequently) for turning down the expression of the early genes.

Cro achieves its function by binding to the same operators as (*cI*) repressor protein. Cro includes a region with the same general structure as the repressor; a helix-2 is offset at an angle from recognition helix-3. (The remainder of the structure is different, demonstrating that the helix-turn-helix motif can operate within various contexts.) Like repressor, Cro binds symmetrically at the operators.

The sequences of Cro and repressor in the helix-turn-helix region are related, explaining their ability to contact the same DNA sequences (see **Figure 12.22**). Cro makes similar contacts to those made by repressor, but binds to only one face of DNA; it lacks the N-terminal arms by which repressor reaches around to the other side.

How can two proteins have the same sites of action, yet have such opposite effects? The answer lies in the different affinities that each protein has for the individual binding sites within the operators. Let us just consider O_R, where more is known, and where Cro exerts both its effects. The series of events is illustrated in **Figure 12.31**. (Note that the first two stages are identical to those of the lysogenic circuit shown in **Figure 12.30**.)

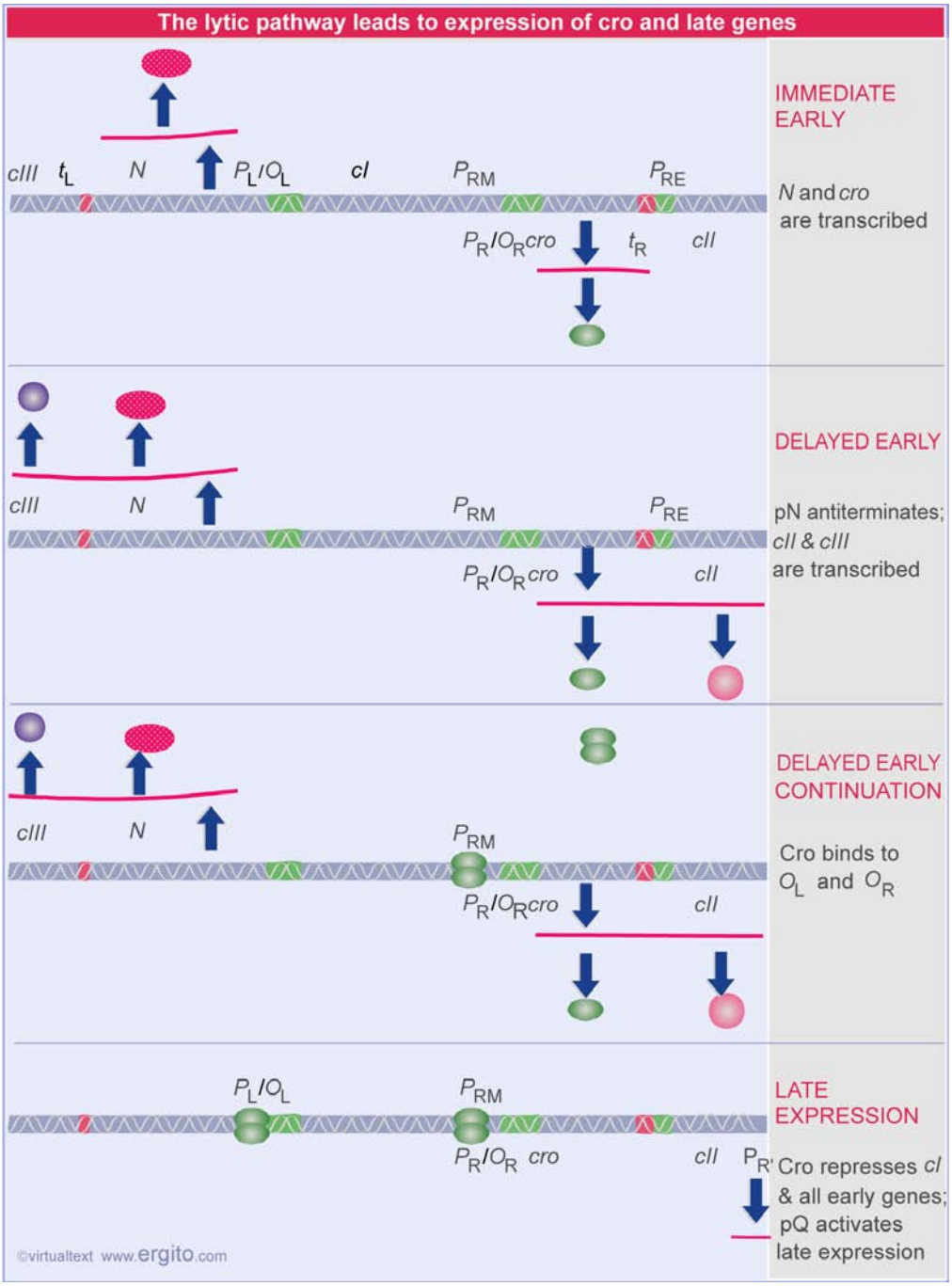

Figure 12.31 The lytic cascade requires Cro protein, which directly prevents repressor maintenance via P_{RM}, as well as turning off delayed early gene expression, indirectly preventing repressor establishment.

The affinity of Cro for O_R3 is greater than its affinity for O_R2 or O_R1. So it binds first to O_R3. This inhibits RNA polymerase from binding to P_{RM}. So Cro's first action is to prevent the maintenance circuit for lysogeny from coming into play.

Then Cro binds to O_R2 or O_R1. Its affinity for these sites is similar, and there is no cooperative effect. Its presence at either site is sufficient to prevent RNA polymerase

from using P_R. This in turn stops the production of the early functions (including Cro itself). Because cII is unstable, any use of P_{RE} is brought to a halt. So the two actions of Cro together block *all* production of repressor.

So far as the lytic cycle is concerned, Cro turns down (although it does not completely eliminate) the expression of the early genes. Its incomplete effect is explained by its affinity for O_R1 and O_R2, which is about eight times lower than that of repressor. This effect of Cro does not occur until the early genes have become more or less superfluous, because pQ is present; by this time, the phage has started late gene expression, and is concentrating on the production of progeny phage particles.

PHAGE STRATEGIES

3.12.20 What determines the balance between lysogeny and the lytic cycle?

Key Concepts

- The delayed early stage when both Cro and repressor are being expressed is common to lysogeny and the lytic cycle.
- The critical event is whether cII causes sufficient synthesis of repressor to overcome the action of Cro.

The programs for the lysogenic and lytic pathways are so intimately related that it is impossible to predict the fate of an individual phage genome when it enters a new host bacterium. Will the antagonism between repressor and Cro be resolved by establishing the autogenous maintenance circuit shown in **Figure 12.30**, or by turning off repressor synthesis and entering the late stage of development shown in **Figure 12.31**?

The same pathway is followed in both cases right up to the brink of decision. Both involve the expression of the immediate early genes and extension into the delayed early genes. The difference between them comes down to the question of whether repressor or Cro will obtain occupancy of the two operators.

The early phase during which the decision is taken is limited in duration in either case. No matter which pathway the phage follows, expression of all early genes will be prevented as P_L and P_R are repressed; and, as a consequence of the disappearance of cII and cIII, production of repressor via P_{RE} will cease.

The critical question comes down to whether the cessation of transcription from P_{RE} is followed by activation of P_{RM} and the establishment of lysogeny, or whether P_{RM} fails to become active and the pQ regulator commits the phage to lytic development. **Figure 12.32** shows the critical stage, at which both repressor and Cro are being synthesized.

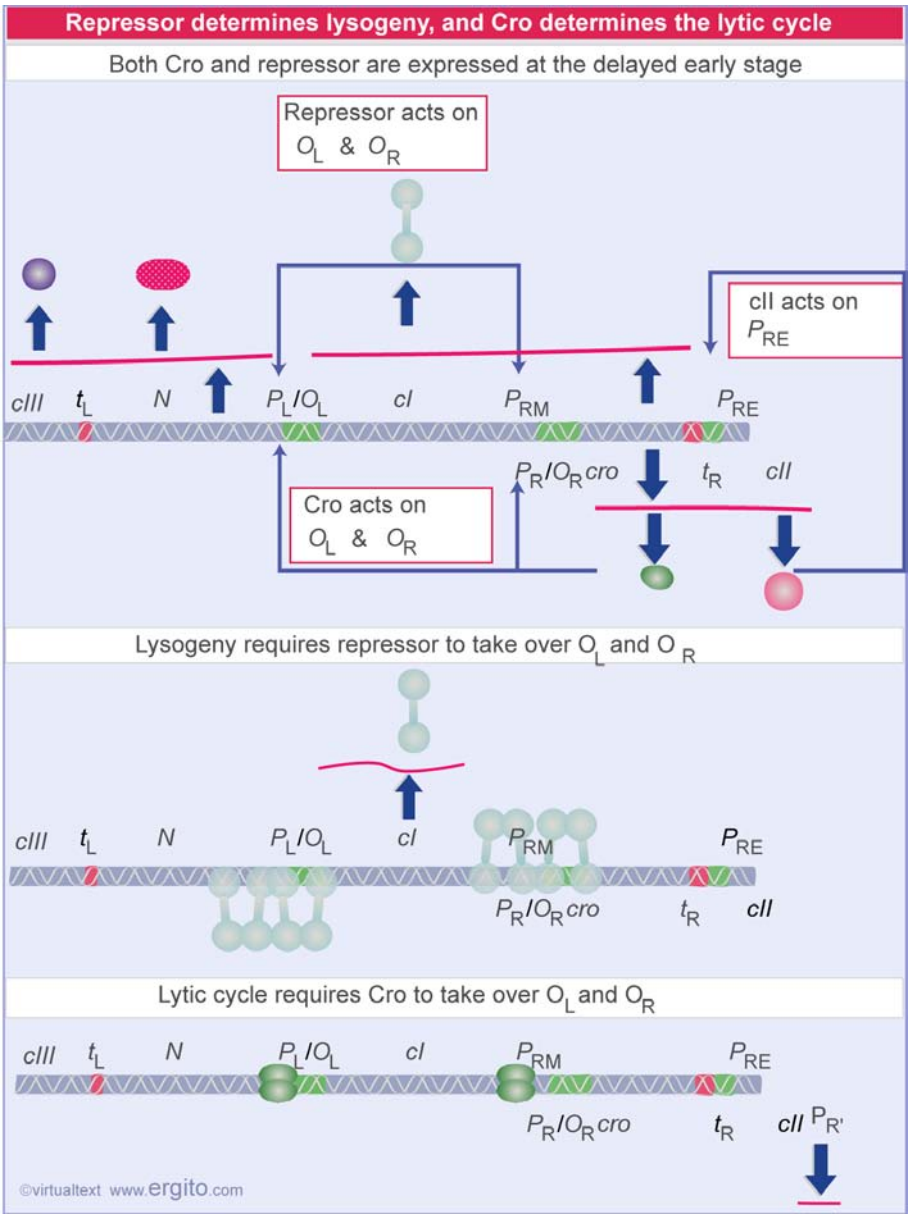

Figure 12.32 The critical stage in deciding between lysogeny and lysis is when delayed early genes are being expressed. If cII causes sufficient synthesis of repressor, lysogeny will result because repressor occupies the operators. Otherwise Cro occupies the operators, resulting in a lytic cycle.

The initial event in establishing lysogeny is the binding of repressor at O_L1 and O_R1. Binding at the first sites is rapidly succeeded by cooperative binding of further repressor dimers at O_L2 and O_R2. This shuts off the synthesis of Cro and starts up the synthesis of repressor via P_{RM}.

The initial event in entering the lytic cycle is the binding of Cro at O_R3. This stops the lysogenic-maintenance circuit from starting up at P_{RM}. Then Cro must bind to O_R1 or O_R2, and to O_L1 or O_L2, to turn down early gene expression. By halting production of cII and cIII, this action leads to the cessation of repressor synthesis via

P_{RE}. The shutoff of repressor establishment occurs when the unstable cII and cIII proteins decay.

The critical influence over the switch between lysogeny and lysis is cII. If cII is active, synthesis of repressor via the establishment promoter is effective; and, as a result, repressor gains occupancy of the operators. If cII is not active, repressor establishment fails, and Cro binds to the operators.

The level of cII protein under any particular set of circumstances determines the outcome of an infection. Mutations that increase the stability of cII increase the frequency of lysogenization. Such mutations occur in *cII* itself or in other genes. The cause of cII's instability is its susceptibility to degradation by host proteases. Its level in the cell is influenced by *cIII* as well as by host functions.

The effect of the lambda protein cIII is secondary: it helps to protect cII against degradation. Although the presence of cIII does not guarantee the survival of cII, in the absence of cIII, cII is virtually always inactivated.

Host gene products act on this pathway. Mutations in the host genes *hflA* and *hflB* increase lysogeny – *hfl* stands for *h*igh *f*requency *l*ysogenization. The mutations stabilize cII because they inactivate host protease(s) that degrade it.

The influence of the host cell on the level of cII provides a route for the bacterium to interfere with the decision-taking process. For example, host proteases that degrade cII are activated by growth on rich medium, so lambda tends to lyse cells that are growing well, but is more likely to enter lysogeny on cells that are starving (and which lack components necessary for efficient lytic growth).

PHAGE STRATEGIES

3.12.21 Summary

Phages have a lytic life cycle, in which infection of a host bacterium is followed by production of a large number of phage particles, lysis of the cell, and release of the viruses. Some phages also can exist in lysogenic form, in which the phage genome is integrated into the bacterial chromosome and is inherited in this inert, latent form like any other bacterial gene.

Lytic infection falls typically into three phases. In the first phase a small number of phage genes are transcribed by the host RNA polymerase. One or more of these genes is a regulator that controls expression of the group of genes expressed in the second phase. The pattern is repeated in the second phase, when one or more genes is a regulator needed for expression of the genes of the third phase. Genes of the first two phases code for enzymes needed to reproduce phage DNA; genes of the final phase code for structural components of the phage particle. It is common for the very early genes to be turned off during the later phases.

In phage lambda, the genes are organized into groups whose expression is controlled by individual regulatory events. The immediate early gene *N* codes for an antiterminator that allows transcription of the leftward and rightward groups of delayed early genes from the early promoters P_R and P_L. The delayed early gene *Q* has a similar antitermination function that allows transcription of all late genes from the promoter $P_{R'}$. The lytic cycle is repressed, and the lysogenic state maintained, by expression of the *cI* gene, whose product is a repressor protein that acts at the operators O_R and O_L to prevent use of the promoters P_R and P_L, respectively. A lysogenic phage genome expresses only the *cI* gene, from its promoter P_{RM}. Transcription from this promoter involves positive autogenous regulation, in which repressor bound at O_R activates RNA polymerase at P_{RM}.

Each operator consists of three binding sites for repressor. Each site is palindromic, consisting of symmetrical half-sites. Repressor functions as a dimer. Each half binding site is contacted by a repressor monomer. The N-terminal domain of repressor contains a helix-turn-helix motif that contacts DNA. Helix-3 is the recognition helix, responsible for making specific contacts with base pairs in the operator. Helix-2 is involved in positioning helix-3; it is also involved in contacting RNA polymerase at P_{RM}. The C-terminal domain is required for dimerization. Induction is caused by cleavage between the N- and C-terminal domains, which prevents the DNA-binding regions from functioning in dimeric form, thereby reducing their affinity for DNA and making it impossible to maintain lysogeny. Repressor-operator binding is cooperative, so that once one dimer has bound to the first site, a second dimer binds more readily to the adjacent site.

The helix-turn-helix motif is used by other DNA-binding proteins, including lambda Cro, which binds to the same operators, but has a different affinity for the individual operator sites, determined by the sequence of helix-3. Cro binds individually to operator sites, starting with O_R3, in a noncooperative manner. It is needed for progression through the lytic cycle. Its binding to O_R3 first prevents synthesis of repressor from P_{RM}; then its binding to O_R2 and O_R1 prevents continued expression

of early genes, an effect also seen in its binding to O_L1 and O_L2.

Establishment of repressor synthesis requires use of the promoter P_{RE}, which is activated by the product of the *cII* gene. The product of *cIII* is required to stabilize the *cII* product against degradation. By turning off *cII* and *cIII* expression, Cro acts to prevent lysogeny. By turning off all transcription except that of its own gene, repressor acts to prevent the lytic cycle. The choice between lysis and lysogeny depends on whether repressor or Cro gains occupancy of the operators in a particular infection. The stability of cII protein in the infected cell is a primary determinant of the outcome.