Introduction

Bacteriophages (phages) are vital for biogeochemical processes, gene transformation, and prokaryotic diversity. Phages have been valuable research tools since the time they were involved in establishing DNA as the hereditary molecule. Today, phages serve as excellent model systems in a wide range of applications, ranging from genetics, molecular biology, drug discovery, genomics, and proteomics. However, the first step toward manipulating phages for research purpose is to understand how phages operate in nature.

Phages can go through two forms of life cycles: (1) lysogenic and (2) lytic life cycle. Lysogenic phage DNA (e.g., phage lambda-λ) integrates into host DNA after the initial infection and exists as a prophage until it encounters unfavorable environment, upon which it induces lytic cycle. In the lytic cycle (e.g., phage T4), phage encoded DNA captures the host transcriptional machinery to synthesize phage proteins, subsequently followed by phage assembly, host lysis, and phage release. A study observed reduction in glucose intake, accompanying lyses of recombinant Escherichia coli (E. coli) cells, during 1-phenylalanine production in the induction phase.

Transcription is the first step in gene regulation, and a set of proteins involving DNA-dependent RNA polymerase (RNAP) and its associated factors interact with DNA to transcribe mRNA that translates into functional proteins. RNAP holoenzyme consisting of four catalytic units and one regulatory sigma unit is recruited to appropriate DNA promoter sequences to initiate transcription. Therefore, it is not surprising that the phages attack the host RNAP and other transcriptional activators and/or regulators. However, the intricacies in the mechanism of confiscating the host transcription apparatus varies with different types of phages.

For example, P23–45 phage encoded Gp39 protein binds to host RNAP and inhibits transcriptional initiation, while λ phage-encoded N protein acts via processive anti-termination.

In this review, we have summarized a wide range of mechanisms regulating transcription in model organisms such as E. coli phage and Thermus thermophilus (T. thermophilus) phage. We have also discussed how the different transcriptional strategies correlate with phage diversity and similarity.

Gene Transcription Regulation of Phage without Bacteriophage RNAP

Classification of phage transcriptional regulatory mechanisms is primarily based on the presence or absence of phage RNAP. Because of the absence of endogenous RNAP, phages such as T4 and λ take over host RNAP to direct expression of their own genes. The next few sections will discuss how these phages have generated these highly effective and sophisticated regulatory mechanisms.

Gene Transcription Regulation of Phage T4

Temporal regulation of phage T4 phage primarily occurs at the transcriptional level, and involves early, middle, and late specific promoters. Transcription activator proteins bind
to the regulatory sites to stimulate transcription from these specific promoters. Following T4 phage infection, host RNAP immediately recognizes the early promoters, and transcription of phage early genes begins. Studies indicate that the protein Alt, which is injected into the host cell simultaneously with the phage DNA, enhances the activity of the phage early promoter.

One minute after infection, T4 middle promoters are active, and initiate transcription of the middle phase genes. There are two mechanisms that underlie the transcription of middle phase genes: (1) Extension of early gene transcripts to downstream middle phase genes and (2) Sigma (σ)-dependent activation of the middle promoter. Genomic analysis has revealed over 45 middle promoters in the T4 phage genome. While RNAP modification is indispensable, various regulatory proteins induce the activation of the middle promoters. Recent studies have shown that a T4 activator MotA, which engages the C-terminus of σ70 activates the middle promoter. Furthermore, Truncaye et al. identified 12 new MotA-dependent middle promoters by using MotA box (a/t)(a/t)(a/t)TGCTTtA, which was centered at the -30 position. The MotA-DNA molecular interaction within the sigma complex is mapped by isolating σ-appropriated complex using iron bromoacetamidobenzyl-EDTA. Single amino acid substitutions based on the C-terminus of σ70 determined the interaction between σ70 residues; it was also found that MotA was negatively affected by the substitution of several residues (e.g., Leucine 607) with alanine. Activator MotA binds more tightly to the promoter in the presence of the co-activator AsiA and RNAP. The T4 co-activator (AsiA) is a small protein consisting of 90 amino acid residues that firmly bind to the σ70 subunit of RNAP, leading to a conformational change that facilitates tighter binding of MotA to RNAP. In thermophilic bacteria, a hydrophobic pocket separating the RNA from the DNA-RNA hybrid in RNAP may affect the interactions among RNAP, MotA, and AsiA.

Finally, T4 late gene transcription acts in parallel with the phage DNA replication. The late promoter simply has a conserved sequence TATAAATA, which is similar to the -10 region of the bacterial chromosome. The phage activator proteins gp45 and gp55 together interact with the co-activator gp33 and promote host RNAP recognition of the late promoter to initiate late gene transcription. Conformational change in RNAP that leads to a termination-resistant complex. It must be noted here that evolutionary changes in the N-protein contributes to the diversity of phages such as HK022, in which Nun protein acts an anti-termination factor for early transcripts.

Genetic analysis of N-protein-dependent anti-termination indicates a possible role for the α subunit of RNAP in facilitating specific functions of NusA and NusE proteins. Structure-function studies of the ternary NusB-NusE-BoxA-RNA complex revealed that protein S10 is at the core of the NusB-S10 transcription anti-termination complex, thus offering better insight into the assembly of the anti-termination complex. The expression of late genes is mediated by the protein Q, which is an anti-termination protein converting the host RNAP into a termination-resistant form.

**Gene Transcription Regulation of Other Phages**

The genome of the extremely thermophilic T. thermophilus phage P23–45 does not possess RNAP coding sequences. The host RNAP cannot identify the conserved 11-bp motif in the early gene promoters. Nevertheless, host RNAP in combination with phage gp76 and gp39 proteins regulate the transcription of middle and late genes. T. thermophilus phage P23–45 cannot correctly complete the temporal transcription of middle-late phage genes (e.g., gp39) required to regulate the function studies of the ternary NusB-NusE-BoxA-RNA complex revealed that protein S10 is at the core of the NusB-S10 transcription anti-termination complex, thus offering better insight into the assembly of the anti-termination complex. The expression of late genes is mediated by the protein Q, which is an anti-termination protein converting the host RNAP into a termination-resistant form.

**Gene Transcription Regulation of Phage λ**

Phage λ utilizes two separate mechanisms to regulate early and late transcripts. Expression of early genes is regulated by the phage encoded N-protein that forms a termination-resistant RNAP complex. Here, the N-terminal arginine-rich motif of the N-protein binds to the NUT region consisting of RNA sequences boxA, boxB, and boxC on the λ chromosome. The N protein further binds with E. coli proteins (NusA, NusB, NusG, and NusE), inducing a conformational change in RNAP that leads to a termination-resistant complex. It must be noted here that evolutionary changes in the N-protein contributes to the diversity of phages such as HK022, in which Nun protein acts an anti-termination factor for early transcripts.

Unlike phage T4, these phages encode their own RNAPs and phage-specific promoters. Hence, phages T7 and T3 exert mechanisms distinct from those of phage T4 to regulate temporal expression of the middle and late gene transcripts. The phages utilize the host E. coli RNAP to transcribe early genes, but then continue to use their own RNAP to activate middle and late promoters. In general, phage protein factors modify the function of host RNAP to direct their own gene expression. The next few sections will discuss regulation of gene transcription in such phages.
Gene Transcription Regulation of Phage T7

Phage T7 RNAP is one of the simplest and the most well characterized RNA polymerase. Similar to T3, K11, and SP6 RNA polymerases, it is composed of a single subunit. T7 RNAP has stringent transcriptional regulation, and transcribes late genes in the absence of auxiliary protein factors. Overall, phage T7 controls transcription at two levels: (1) the early and (2) the middle-late stages; both mechanisms induce inhibition of RNA synthesis by the host RNAP. Early gene transcription of phage T7 is coupled to the host RNAP, and is followed by synthesis of middle gene products. Phage T7 RNAP identifies the middle and late stage-specific promoters to initiate transcription. The phage protein gp2, a middle T7 gene product, binds to the host RNAP to inhibit the transcriptional activity of the RNAP, and consequently, repress the host transcriptional processes.

Gene Transcription Regulation of Other Phages

Transcriptional termination of T3-infected Escherichia, Yersinia, and Pseudomonas species is dependent on the T7 protein gp2 homolog. Current research investigates the transcriptional strategies of novel phages such as φKMV. Phage φKMV infects Pseudomonas aeruginosa and encodes its own RNAP. It induces transcriptional inhibition via gp2-like proteins, suggesting an evolutionary conservation in the regulatory mechanisms of the phage φKMV and phage T7.

Gene Transcription Regulation of Phage Xp10

In contrast to the regulatory mechanisms mentioned above, Xanthomonas oryzae phage Xp10 (a siphovirus) has a distinct and a unique transcriptional regulation strategy. Despite encoding for its own RNAP, Xp10 requires the host RNAP to initiate partial gene transcription. Soon afterward, the host RNAP activity quickly decreases. Several factors, including phage genome size, gene composition, gene structure, hosts, and the environment, determine the transcriptional regulation of Xp10.

Considering the elaborate phage diversity, phage Xp10 is a sophisticated combination of two widely different phage genomes-T7 and phage λ. The genome structure and transcriptional activity of Xp10 reflects this characteristic mosaic. While it encodes its own RNAP (similar to phage T7), the genomic structure of Xp10 is analogous to that of phage λ. Approximately half of Xp10 genome codes for structural and lytic proteins, while the other half, regulates host RNAP inhibition, virus replication, and enzyme function. These two groups of genes diverge at the transcription level, and are separated by a regulatory region consisting of diverse promoters that are recognized by RNAP. Xp10 genomic analysis determined that 10% of its genes originated in NHN family endonucleases and DNA polymerase. Interestingly, the DNA polymerase of Xp10 was similar to that of Leishmania species. Yuzenkova et al. used a combination of genetic and biochemical methods to map the Xp10-encoded p7 transcription factor interaction site and explain the role of p7 in transcription initiation and termination.

Other Transcription Regulation Mechanisms

It was reported that Pseudomonas aeruginosa YuA phage includes two different types of promoters. The second promoter involving the regulator σ54 induces late gene transcription. Subunit σ54 varies from the σ70 class, and requires additional activation to initiate transcription. Further investigation is required to elucidate the mechanisms underlying transcriptional regulation of phage YuA. Yet another E. coli phage φEcoM-GJ1 encodes its own single subunit RNAP to initiate viral gene transcription. Phage φEcoM-GJ1 is similar to the phage φKMV mentioned earlier in the review; however, if the phage encodes the host RNAPs regulators or not is uncertain.

Conclusions

The phages use several different mechanisms to appropriate bacterial processes to benefit their development. The phages can either encode a specialized class of protein factors (e.g., P23–45 phage Gp39 protein, λ phage N protein) to regulate its own gene transcription. However, some transcriptional regulatory mechanisms described to date are still not very clear. Development of genomics, transcriptomics, and proteomics offer a better platform to understand the phage-host interactions via transcriptional regulation. Furthermore, understanding the multilevel (molecular and structural) operation of phage-encoded RNAP inhibitors will pave the path for the development of anti-bacterial drugs that specifically target bacterial RNAPs.

Disclosure of Potential Conflicts of Interest

No potential conflict of interest was disclosed.

Acknowledgments

This project was financially supported by the Open Project Program of the Key Laboratory of Industrial Biotechnology, Ministry of Education, China (KLIB-KF201301), the Fundamental Research Funds for the Central Universities (JUSRP11429), the 111 Project (111-2-06), Natural Science Foundation of Jiangsu Province (BK20140152), the Priority Academic Program Development of Jiangsu Higher Education Institutions, and the Jiangsu province “Collaborative Innovation Center for Advanced Industrial Fermentation” industry development program.
Twelve new MotA-dependent middle promoters of RNA polymerase. Proc Natl Acad Sci U S A 2011; 108:19690-6; PMID:22155590; http://dx.doi.org/10.1073/pnas.1107169108

Bacteriophage transcription regulator inhibits displacement. Nucleic Acids Res 2014; 42:4294-305; PMID:24482445; http://dx.doi.org/10.1093/nar/gku080

H. Bacteriophage prehistory: Is or is not Hankin, 1896, a phage reference? Bacteriophage 2011; 1:174-8; PMID:22164351; http://dx.doi.org/10.4161/bact.1.3.16591

Liu B, Shadrin A, Sheppard C, Meeker V, Xu Y, Severinov K, Matthews S, Wigneshweraraj S. A bacteriophage transcription regulator inhibits bacterial transcription initiation by σ-factor degradation. Nucleic Acids Res 2014; 42:4294-305; PMID:24482445; http://dx.doi.org/10.1093/nar/gku080

Zhou H, Liao X, Liu L, Wang T, Du G, Chen J. Enhanced L-phenylalanine production by bacteriophage T4 activator MotA/promoter DNA resistant to bacteriophage BP-1 via a two-stage feeding approach. J Ind Microbiol Biotechnol 2011; 38:1219-27; PMID:21104195; http://dx.doi.org/10.1007/s10295-010-0980-0

Park S, Court C, Lublowska L, Jin DJ, Kashlev M, Court DL. Bacteriophage A protein inhibits transcription slippage by Escherichia coli RNA polymerase. Nucleic Acids Res 2014; 42:5823-9; PMID:24711367; http://dx.doi.org/10.1093/nar/gku203

Hsieh ML, James TD, Knipping L, Waddell MB, White S, Hinton DM. Architecture of the bacteriophage T4 activator MotA/promoter DNA interaction during sigma factor degradation. J Biol Chem 2013; 288:27607-18; PMID:23902794; http://dx.doi.org/10.1074/jbc.M113.475434

Pavlova O, Lavish D, Klukim E, Djordjevic M, Ravcheev DA, Gelfand MS, Severinov K, Akulenko N. Temporal regulation of viral transcription of the Escherichia coli bacteriophage phiX174. J Mol Biol 2012; 416:389-99; PMID:22622652; http://dx.doi.org/10.1016/j.jmb.2012.01.002

Hinton DM. Transcriptional control in the prereplicative phase of T4 development. Virol J 2010; 7:289; PMID:20192433; http://dx.doi.org/10.1186/1743-422X-7-289

Truncarte I, Piesiniene I, Kolesinskien G, Zajankauskait E, Driukas A, Klaua V, Ninivaks R. Twelve new MotA-dependent middle promoters of bacteriophage T4: consensus sequence revised. J Mol Biol 2003; 327:335-46; PMID:12628241; http://dx.doi.org/10.1006/jmbi.2003.0801

Bonocora RP, Decker PK, Glass S, Knipping L, Hinton DM. Bacteriophage T4 MotA activator and the β-flip tip of RNA polymerase target the same set of σ70 carboxyl-terminal residues. J Biol Chem 2011; 286:39299-6; PMID:21911499; http://dx.doi.org/10.1074/jbc.M111.278762
39. Yuzenkova Y, Zenkin N, Severinov K. Mapping of RNA polymerase residues that interact with bacteriophage Xp10 transcription antitermination factor p7. J Mol Biol 2008; 375:29-35; PMID:18021805; http://dx.doi.org/10.1016/j.jmb.2007.10.054

40. Ceyssens PJ, Mesyanzhinov V, Sykilinda N, Briërs Y, Roucourt B, Lavigne R, Robben J, Domashin A, Miroshnikov K, Volckaert G, et al. The genome and structural proteome of YuA, a new Pseudomonas aeruginosa phage resembling M6. J Bacteriol 2008; 190:1429-35; PMID:18065532; http://dx.doi.org/10.1128/JB.01441-07

41. Nechaev S, Severinov K. The elusive object of desire - interactions of bacteriophages and their hosts. Curr Opin Microbiol 2008; 11:186-93; PMID:18400552; http://dx.doi.org/10.1016/j.mib.2008.02.009

42. Jamalludeen N, Kropinski AM, Johnson RP, Lingohr E, Hazel J, Gyles CL. Complete genomic sequence of bacteriophage phiEcoM-GJ1, a novel phage that has myovirus morphology and a podovirus-like RNA polymerase. Appl Environ Microbiol 2008; 74:516-25; PMID:18039824; http://dx.doi.org/10.1128/AEM.00990-07