

3

Gene Expression and Its Regulation

JOHN D. HELMANN

Department of Microbiology, Cornell University, Ithaca, New York 14853-8101

I. Introduction	47
II. RNA Polymerase and the Process of Transcription ..	48
A. Structure of RNAP	48
B. The Bacterial Transcription Cycle—Overview	48
C. Promoter Structure	50
III. Ribosomes and the Process of Translation	53
A. Structure of Ribosomes	53
B. The Bacterial Translation Cycle	55
IV. Transcriptional Regulation—Repressors and Activators	56
A. Regulation by Repressors	57
B. Regulation by Activators	59
V. Transcriptional Regulation—Other Mechanisms	61
A. Alternative Sigma Factors	61
B. Direct Regulation of RNAP Activity by NTPs ...	63
C. RNAP Substitution or Modification during Phage Infection	66
D. Regulation of Transcription Termination	
1. Overview	66
2. Attenuation	67
VI. Translational Regulation	70
A. Regulation of Translation Initiation	70
B. Regulation of Translation Elongation and Termination	72
VII. Regulation by DNA Modifications	74
VIII. Conclusions	78

I. INTRODUCTION

Bacteria have a remarkable ability to adapt to a rapidly changing environment. In most cases adaptation requires that new proteins be synthesized to adjust the metabolic capacity of the organism to the available nutrients or to defend against chemical or physical toxins. In this chapter we will survey the diverse ways that bacteria have evolved

to coordinate gene expression with environmental signals.

Gene expression begins with the copying of discrete segments of the DNA into RNA, a process known as *transcription*. The products of transcription include four classes of RNA molecules: messenger RNA (mRNA),

ribosomal RNA (rRNA), transfer RNA (tRNA), and regulatory RNA. For protein-coding genes, the corresponding mRNA molecule binds to ribosomes and directs the synthesis of one or more specific proteins in a process called *translation*. These processes are so central to all living things that the informational transfer of "DNA makes RNA makes protein" has been referred to as the "central dogma" of molecular biology.

In bacteria, gene expression is most frequently regulated at the level of transcription. That is, bacteria only transcribe the subset of their genes that are necessary for growth and survival under the existing environmental conditions. The remaining regions of the genome are silent. In some cases, however, genes are transcribed into mRNA even when their protein products may not be needed. In these cases the process of translation is likely to be the focus of regulation. To appreciate the diverse mechanisms that allow bacteria to regulate transcription and translation, we will first review these processes and the enzymes that catalyze them.

II. RNA POLYMERASE AND THE PROCESS OF TRANSCRIPTION

A. Structure of RNAP

The first step in gene expression is the transcription of an RNA molecule complementary to the DNA template catalyzed by DNA-dependent RNA polymerase (RNAP). As befits its central role in the cell, RNAP is highly conserved and very complex (McClure, 1985; Young, 1991). All cells contain a multisubunit RNAP with two large subunits and a variable number of smaller subunits. In bacteria, a catalytically active RNAP core enzyme has minimal subunit composition $\beta\beta'\alpha_2$ (indicated as "E") (Burgess et al., 1969). Additional small proteins, including the omega (ω) polypeptide (Gentry and Burgess, 1993; Mukherjee and Chatterji, 1997) and, in some, gram-positive bacteria, the delta (δ) subunit (Juang and Helmann, 1994; Lopez de Saro et al., 1995; Lopez de

Saro et al., 1999), are also often present. The core enzyme can faithfully copy DNA into RNA over many thousands of base pairs, but by itself is incapable of recognizing promoter elements.

Promoter recognition requires a separate, dissociable specificity protein known as σ (Gross et al., 1992; Helmann, 1994). The complex formed by binding of σ to the core enzyme is called *holoenzyme* (Fig. 1) and is often identified by the associated σ factor. In *Escherichia coli*, for example, the primary σ factor is 70 kDa in size and is referred to as σ^{70} . The complex formed by the binding of σ^{70} to the core RNAP is the σ^{70} holoenzyme ($\beta\beta'\alpha_2\sigma^{70}$ or $E\sigma^{70}$). As we will see, substitution of one σ factor by an alternative specificity subunit is a powerful mechanism for activating the transcription of new sets of genes (Table 1).

Eukaryotes have three nuclear RNAP forms that have between 10 and 12 protein subunits each, including two similar in sequence to the large β and β' subunits that make up the bulk of the bacterial core enzyme. Recently the three-dimensional structures of the RNAP from the thermophilic bacterium *Thermus aquaticus* (Zhang et al., 1999) and from the yeast *Saccharomyces cerevisiae* (Cramer et al., 2000; Fu et al., 1999) have been determined at atomic resolution. The resulting structures reveal that those regions that are highly similar in sequence between the bacterial and eukaryotic RNAP subunits are closely clustered around the active site for RNA synthesis.

B. The Bacterial Transcription Cycle—Overview

In general, processes of macromolecular synthesis can be divided into three major phases: initiation, elongation, and termination. In the case of RNA synthesis (Fig. 2A), the initiation phase involves the interaction of RNAP with specific *promoter* sites that identify the start point of an RNA molecule (deHaseth et al., 1998). Once bound to the promoter, initially as a *closed complex*, RNAP locally separates the two DNA