Cells having identical genomic content exhibit differences in their physiology and metabolism due to differential gene expression. Regulation of gene expression was initially discovered in the lac operon (has genes required for the uptake and degradation of lactose) of Escherichia coli\(^1\). Later it was found out that regulation of gene expression is a universal phenomenon in living organisms. Depending on the stage at which the expression of a gene is regulated, this has been categorized into four types: (1) Transcriptional regulation: determines whether the gene will be transcribed or not; (2) Post-transcriptional regulation: regulation of the processing, transport and longevity of mRNAs; (3) Translational regulation: determination of whether or not and to what extent an m-RNA will be translated; and (4) Post-translational regulation: regulation of protein modification and activity. For example, in the case of lac operon, gene expression is regulated at the transcriptional level. In the absence of lactose, the repressor protein binds to the operator region and blocks transcription initiation from the promoter. In the presence of lactose, the conformation of the repressor protein is changed by binding to allo-lactose (a derivative of lactose) so that it is no longer able to bind to the operator region. This leads to the transcription initiation from the promoter. Expression of heat shock sigma factor gene (\(\sigma^{32}\)) in E. coli is regulated at post-translational level. Transcription and translation of \(\text{rpoH}\) remain usual at both normal as well as heat shock conditions but \(\sigma^{32}\) level increases dramatically after the heat shock. This is because, at normal condition, \(\sigma^{32}\) is unstable in the cytosol due to its association with DnaK, another cytosolic protein. After the heat shock, \(\sigma^{32}\) is released from DnaK and becomes stable\(^2\). If the expression of \(\text{rpoH}\) gene is regulated at the transcriptional level instead of at the post-translational level then its expression will be delayed after the heat shock. This will delay the expression of the heat shock protein genes whose expression is dependent on the level of \(\sigma^{32}\) in the cell. This is likely to be disadvantageous for the cell in evoking an adaptive response to the heat shock. In bacteria, a single cell carries out all functions of a living organism. Though regulating the expression of a gene at a later stage after transcription has temporal advantage, there will be spatial as well as economical constraints on the cell. This is because the message will be stored in an amplified stage (in the form of RNA and protein) which might not be required during the lifetime of the cell and these molecules are not stable like DNA. These imply that different gene regulation systems have evolved in organisms striking a compromise between economy and survivability of the cell. In comparison to transcriptional regulation, translational regulation is difficult for genes that are sequentially present in an operon. In a translational regulation, the ribosomal binding site (RBS) in the m-RNA is blocked either due to binding of a protein to it or formation of a secondary structure in the RNA. In a poly-cistronic m-RNA, each open reading frame (ORF) is preceded by a RBS and each RBS has to be blocked for translational regulation. Therefore, an attenuation mode of gene regulation is observed in bacteria. Attenuation was initially discovered in the \(\text{trp}\) operon (involved in tryptophan biosynthesis) of E. coli\(^3\). In this mode of regulation, transcription from the promoter continues both in the ON and the OFF state of the gene expression.

During the OFF state, a premature transcription termination takes place and in the ON state full transcript is made. Attenuation mode of gene regulation, mediated by t-RNA, is limited to amino acid biosynthesis and aminoacyl-tRNA synthetase genes in bacteria. Like amino acids, nucleotides and vitamins are important metabolites of the cell and genes for their biosynthesis are arranged in operons. The expression of the genes for synthesis of thiamin and guanine is known to be regulated through a feedback mechanism in bacteria\(^4,5\). A search for a protein involved in this mechanism was not successful. A t-RNA-mediated attenuation process has been found to be irrelevant to regulate the expression of these genes. So the mechanism of regulation of these genes was not revealed.

Recent studies on the expression of genes for the biosynthesis of nucleotides (purines), and vitamins (thiamin, riboflavin, cobalamin) in Bacillus subtilis and E. coli have given insights into a new mechanism of gene regulation in these bacteria, which is called as riboswitch control\(^6\). The riboswitches are made up of the three-dimensional structure of RNA, in which RNA can undergo two mutually exclusive conformations in response to an environmental signal in the form of a metabolite (Figure 1). In case of B. sub-

**Figure 1.** Schematic representation of riboswitch-mediated gene regulation: The 5′ UTR of the nascent transcript can form two mutually exclusive conformations in relation to the availability of the metabolite. Interaction of the metabolite (FMN, TPP, guanine, Glic6NP, etc.; discussed in the text) with the nascent transcript induces conformational changes of the RNA which results in either transcription termination or prevention of translation initiation or the cleavage of the m-RNA by activating the ribozyme. In the absence of the metabolite gene expression continues.
The translation initiation from the thiC operon requires for thiamin biosynthesis have been shown to be regulated by riboswitches. In case of the rib-operon, the 5′ UTR (untranslated region; ~300 nts) of the nascent transcript can fold into two mutually exclusive conformations in relation to its binding to FMN (flavin mononucleotide). In the low concentration/absence of FMN, the nascent transcript comes out of the RNA polymerase and folds into a conformation that allows the transcription read through the proximal intrinsic termination site. As a result, the full transcript is formed and the genes in the operon are expressed. However, in the presence of FMN (1 µM), its binding to the nascent transcript folds the RNA in a way that results into transcription termination at the proximal intrinsic termination site. This premature transcription termination results into no expression of genes in the operon. This is a new mechanism of attenuation mediated by riboswitch. Similar mechanism is used to control the gene expression in the thi-operon and xpt-pbuX operon in response to thiamin pyrophosphate (TPP) and guanine respectively. In E. coli, thiM and thiC genes are required for thiamin biosynthesis. The translation initiation from thiM and thiC mRNAs is sensitive to the presence of thiamin. The 5′ UTR of these RNA folds into a conformation that can bind to thiamin. Thiamin binding changes the conformation that precludes the RBS from the ribosomal subunit access and inhibits translation. In the absence of thiamin, the RBS is free to access by the ribosomal subunit. A recent report about the regulation of glmS gene (encodes glutamine-fructose-6-phosphate amidotransferase) in B. subtilis provides further an interesting way to regulate gene expression mediated by riboswitch. In this case, a riboxygen function is associated with the riboswitch. The 5′ UTR of glmS m-RNA can fold into a conformation that can bind to GlcN6P (glucosamine-6-phosphate; product of GlmS activity). Upon binding of the GlcN6P, the riboswitch activity of the riboswitch gets activated due to conformational change. This results in the self-cleavage of the m-RNA by an internal phosphoester transfer, catalysed by the 2′-OH group of the ribose sugar.

Why this RNA-mediated mechanism of gene regulation? The advantages of the system are: (i) riboswitches are very specific to their substrates, suggesting a stringent control. For example, the riboswitch–ribozyme of glmS RNA is specifically activated by GlcN6P but not with related compounds GlcN (glucosamine) and Glc6P (glucose-6-phosphate; Figure 2); (ii) binding of the substrate to the nascent transcript is required for a short time. Once the transcript is terminated the ligand can be dislodged from the site and can be reused either for the same or for other purposes; and (iii) there is no involvement of another gene product, which reduces the effect of mutation that will deregulate the mechanism. This might be the reason why it took long time to identify this control system. In spite of these advantages, the lack of universality might be due to some limitations like: (i) the 5′ UTR may not form tertiary structures for all different kinds of metabolites; and (ii) riboswitch-mediated translational regulation is limited to mono-cistronic m-RNA.

The novelty in this mechanism of gene regulation lies in that there is no involvement of protein and it is RNA-mediated. Therefore, it has been speculated that this mechanism is a remnant of the RNA world. Till now discovered riboswitches have been shown to function in repressing gene expression (negative regulation) in response to metabolites. If riboswitch is a primitive mode of gene regulation then it suggests that in the RNA world negative regulation was predominant. There are also speculations that riboswitch-mediated control mechanism might also be playing an important role in eukaryotic gene expression. Unlike prokaryotes, eukaryotes have distinct compartments for transcription and translation and m-RNA is monocistronic. These favour post-transcriptional and translational gene regulation. A riboswitch-mediated gene regulation is expected to be acting either at post-transcriptional or translational level. Future studies in eukaryotic gene regulation might reveal the presence of this control system and the exact mechanism.


ACKNOWLEDGEMENTS. I thank Prof. B. K. Konwar, Head, Dept of MBBT, Tezpur University, for his support, Mr Eeshan Kalita for useful discussions and Dr Alak K. Buragohain for comments on the manuscript.

Suvendra Kumar Ray is in the Department of Molecular Biology and Biotechnology, Tezpur University, Tezpur 784 028, India. e-mail: suven@tezu.ernet.in

Figure 2. Chemical structures of GlcN6P, GlcN, and Glc6P.