# Regulation of Gene Expression

# Regulation of Gene Expression

## *Molecular Mechanisms*

By

**Gary H. Perdew, PhD**
**John P. Vanden Heuvel, PhD**
**Jeffrey M. Peters, PhD**

*Center for Molecular Toxicology and Carcinogenesis*
*Department of Veterinary and Biomedical Sciences*
*The Pennsylvania State University*
*University Park, PA*

Humana Press Totowa, New Jersey

999 Riverview Drive, Suite 208
Totowa, New Jersey 07512

**www.humanapress.com**



This publication is printed on acid-free paper. ∞
ANSI Z39.48-1984 (American Standards Institute) Permanence of Paper for Printed Library Materials.

Cover design by Patricia F. Cleary

For additional copies, pricing for bulk purchases, and/or information about other Humana titles, contact Humana at the above address or at any of the following numbers: Tel.: 973-256-1699; Fax: 973-256-8341; E-mail: orders@humanapr.com; or visit our Website: www.humanapress.com



Printed in the United States of America. 10 9 8 7 6 5 4 3 2 1

eISBN 1-59745-228-9

**Library of Congress Cataloging-in-Publication Data**

Perdew, Gary H.
Regulation of gene expression : molecular mechanisms / by Gary H. Perdew, John P. Vanden Heuvel, Jeffrey M. Peters.
p. ; cm.
Includes bibliographical references and index.
ISBN 1-58829-265-7 (alk. paper)
1. Genetic regulation. I. Vanden Heuvel, John P. II. Peters, Jeffrey M. (Jeffrey Maurice) III. Title.
[DNLM: 1. Gene Expression Regulation--physiology. 2. Molecular Biology. QU 475 P433r 2006]
QH450.P45 2006
572.8'65--dc22

2006020090

# Preface

The use of molecular biology and biochemistry to study the regulation of gene expression has become a major feature of research in the biological sciences. Many excellent books and reviews exist that examine the experimental methodology employed in specific areas of molecular biology and regulation of gene expression. However, we have noticed a lack of books, especially textbooks, that provide an overview of the rationale and general experimental approaches used to examine chemically or disease-mediated alterations in gene expression in mammalian systems. For example, it has been difficult to find appropriate texts that examine specific experimental goals, such as proving that an increased level of mRNA for a given gene is attributable to an increase in transcription rates. *Regulation of Gene Expression: Molecular Mechanisms* is intended to serve as either a textbook for graduate students or as a basic reference for laboratory personnel. Indeed, we are using this book to teach a graduate-level class at The Pennsylvania State University. For more details about this class, please visit http://moltox.cas.psu.edu and select "Courses." The goal for our work is to provide an overview of the various methods and approaches to characterize possible mechanisms of gene regulation. Further, we have attempted to provide a framework for students to develop an understanding of how to determine the various mechanisms that lead to altered activity of a specific protein within a cell. We expect the reader will have a good working knowledge of basic biochemistry and cell biology, although detailed understanding of molecular biology techniques is not required.

Each of the three parts of *Regulation of Gene Expression: Molecular Mechanisms* is self-contained. Thus the order of reading does not need to follow the order of presentation, although the parts have been arranged in the way that investigators often approach the study of gene regulation. We have thoughtfully selected key references only and included their details in the page margins for ready reference, as this work is intended as a textbook, not a review of the literature. Key points as well have been placed in the margins in order to emphasize important issues.

Part I, written by John P. Vanden Heuvel, presents the experimental approaches that can be utilized to study control of mRNA expression and the determination of target genes for a given transcription factor. Part II, written by Gary H. Perdew, examines the experimental approaches utilized to determine how proteins can regulate each other by mediating synthesis, degradation, protein–protein interactions, and posttranslational modification. Finally, Part III, written by Jeffrey M. Peters, explores how gene targeting techniques in mice can provide insight into protein function. The point of view is that of a molecular toxicologist, but we have kept in mind a wider range of graduate students and professionals in the biological sciences. As toxicologists, however, we are primarily concerned with mammalian systems and with determining how chemicals can modify gene expression. This has clearly influenced

the biological systems utilized in the experimental approaches suggested throughout this text.

We thank those who contributed to the completion of this book, in particular, Marcia H. Perdew and Cheryl Brown for their excellent editorial assistance. Also, thanks go to the many students who have directly or indirectly contributed to the overall concept of this book, and who also read many segments of this book. We are indebted to Dr. C. Channa Reddy for his support and vision in providing an excellent research environment and establishing a molecular toxicology group at The Pennsylvania State University. Finally, we would like to acknowledge all of our mentors who have contributed to our careers and have inspired us to be the best scientists and mentors possible.

***Gary H. Perdew, PhD***
***John P. Vanden Heuvel, PhD***
***Jeffrey M. Peters, PhD***

# Contents

# COMPANION CD

Illustrations appearing in the book may be found on the Companion CD attached to the inside back cover. The image files are organized into folders by chapter number and are viewable in most Web browsers. The number following "f" at the end of the file name corresponds to the figure in the text. The CD is compatible with both Mac and PC operating systems.

# PART I

## GENE EXPRESSION CONTROL AT THE mRNA LEVEL

**John P. Vanden Heuvel, PhD**

### Contents

# 1

## Overview

### 1. Concepts

In the Part I, the basic approaches to understanding how a treatment or condition results in mRNA accumulation will be described. Although the focus and many of the examples will center on gene regulation by xenobiotics, the approaches are applicable to any treatment/condition that alters gene expression. The subsequent chapter will contain an overview of the molecular biology involved in each step of the process; however, details can be found elsewhere. For a good basic overview of transcriptional control of gene expression, the reader is directed to Molecular Biology of the Cell [1]. In addition, there have been several excellent review articles on eukaryotic transcriptional control [2–8]. Posttranscriptional gene regulation is, in general, more difficult to examine experimentally but is a very important determinant of cellular events. We will briefly discuss mRNA processing and stability, with emphasis on events altered by xenobiotics and the methods used to examine these events. Detailed laboratory procedures are available from several sources [9–13] and will not be emphasized here. Instead, we focus on the approaches to be used and the rationale behind these decisions. Translational and posttranslational regulation of gene control will not be examined in detail in this section, but they are discussed subsequently in Part II.

A xenobiotic or disease can affect the accumulation of mRNA for a particular gene in many ways, as shown in Fig. 1-1. The first level of control is at the level of chromatin packaging of DNA (chromatin control). It has been well established that

*The term xenobiotic refers to any chemical that is foreign to that organism and includes drugs, pollutants, and so forth*


1. Alberts, B. et al. In: Molecular Biology of the Cell. Garland Publishing, New York, NY, 1994, pp. 401–476.
2. Pugh, B.F. Gene 25 (2000) 1–14.
3. Dillon, N., Sabbattini, P. Bioessays 22 (2000) 657–665.
4. Huang, L. et al., Crit. Rev. Eukaryot. Gene Expr. 9 (1999) 175–182.
5. Kornberg, R.D. Trends Cell. Biol. 9 (1999) M46–9.
6. Chin, J.W. et al. Curr. Biol. 9, (1999) R929–932.
7. Roeder, R.G. Cold Spring Harb. Symp. Quant. Biol. 63 (1998) 201–218.
8. Franklin, G.C. Results Probl. Cell Differ. 25 (1999) 171–187.
9. Ausubel, F.M. et al. Current Protocols in Molecular Biology (John Wiley, New York, NY, 1994).
10. Sambrook, J. et al. (eds.), Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Press, Cold Spring Harbor, NY, 1989.
11. Davis, L. G. et al. Basic Methods in Molecular Biology. Appleton & Lange, Norwalk, CT, 2nd ed. 1994.
12. Kaufman, P.B. Handbook of Molecular and Cellular Methods in Biology and Medicine. CRC Press, Boca Raton, FL. 1995.
13. Vanden Heuvel, J.P. et al. In: PCR Protocols in Molecular Toxicology (J.P. Vanden Heuvel, ed). CRC Press, Boca Raton, FL. 1998.

*Messenger RNA accummulation may be caused by effects at the level of chromatin, transcription, or RNA processing, transport, or stability*

remodeling of DNA structure is required for gene transcription. In the fully packaged state ("closed"), the DNA is inaccessible to the transcriptional machinery, whereas once remodeled ("open"), transcription may proceed. Second, and perhaps most importantly from a xenobiotic standpoint, is controlling how and when a given gene is transcribed (transcriptional control). This is the predominant, but not the sole, manner in which ligands for nuclear receptors (NRs) and other soluble receptors affect gene expression. Many important signal transduction molecules including kinases and phosphatases will ultimately result in regulation of transcription factor activity and control mRNA levels in this manner. Also, certain oncogenes (including c-myc, fos and jun) encode for proteins that are transcriptionally active. Third, a chemical may control how the primary RNA transcript is spliced or otherwise processed (RNA processing control). Alternatively spliced transcripts may result in different protein products or may result in mRNA with different rates of turnover or translation. Although not studied in great detail, there are several examples of chemically-induced or disease-specific splice variants. Fourth, a chemical may select which completed mRNAs in the nucleus are exported to the cytoplasm (RNA transport control). To date, there are no examples of this means of regulation by xenobiotics, although the potential exists. Fifth, chemicals may affect gene expression by selecting which mRNAs in the cytoplasm are translated by ribosomes (translational control). Similar to RNA processing control, there are few concrete examples of chemicals specifically regulating the expression of a gene via this pathway. However, many chemicals affect the expression of ribosomal proteins that may in turn affect translation efficiency. It is also of note that the control of the rate of translation is coupled to RNA processing (i.e., capping) and transport. Sixth, certain mRNA molecules may be stabilized or destabilized in the cytoplasm (mRNA degradation control). This may be the most under-

*Receptors are macromolecules that contain functions for recognition (i.e., binding to a xenobiotic) and transduction (ability to regulate gene expression).*

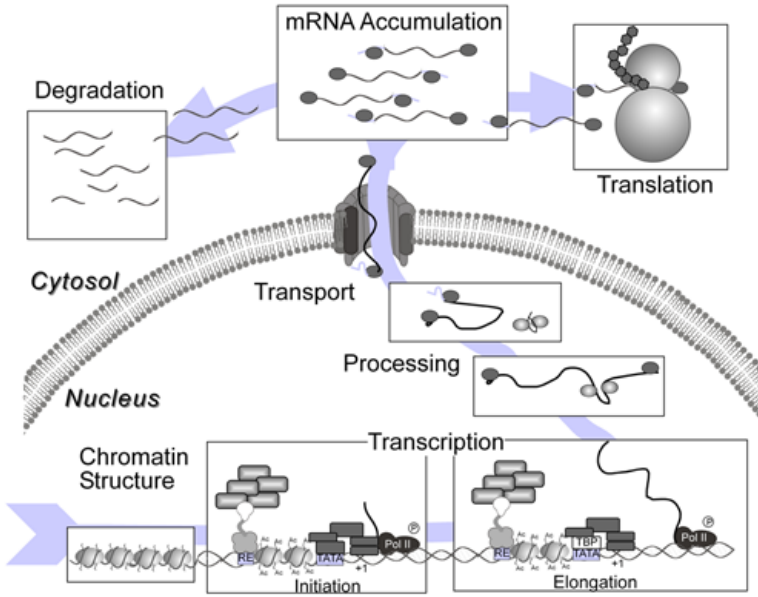


Fig. 1-1. Control of gene expression. In the Part I, the focus is on mRNA accumulation and hence transcription. mRNA processing and stability will be described. (Adapted from ref. [1]).

appreciated area in chemically-induced and disease-mediated alteration in mRNA levels. There are many examples of soluble receptor ligands that affect the stability of mRNA thereby resulting in alteration in corresponding protein levels.

Certainly, affecting protein activity is an important mechanism by which chemicals result in their biological and toxicological effects and how diseases are manifested. However, often it is the mRNA levels that are being examined and it is assumed (often falsely) that altered protein levels and activity will result. This is partially attributable to the fact that the examination of mRNA levels is at a much higher level of sophistication than are equivalent protein measurements. For example, microarray techniques make the concurrent examination of thousands of mRNA transcript levels feasible. Other high-throughput quantitation, such as real-time and competitive reverse transcriptase-polymerase chain reaction (RT-PCR), allow for detailed examination of when and how a transcript is affected by a xenobiotic. Equivalent methods for protein and enzyme activity are being developed but are not readily available at this time.

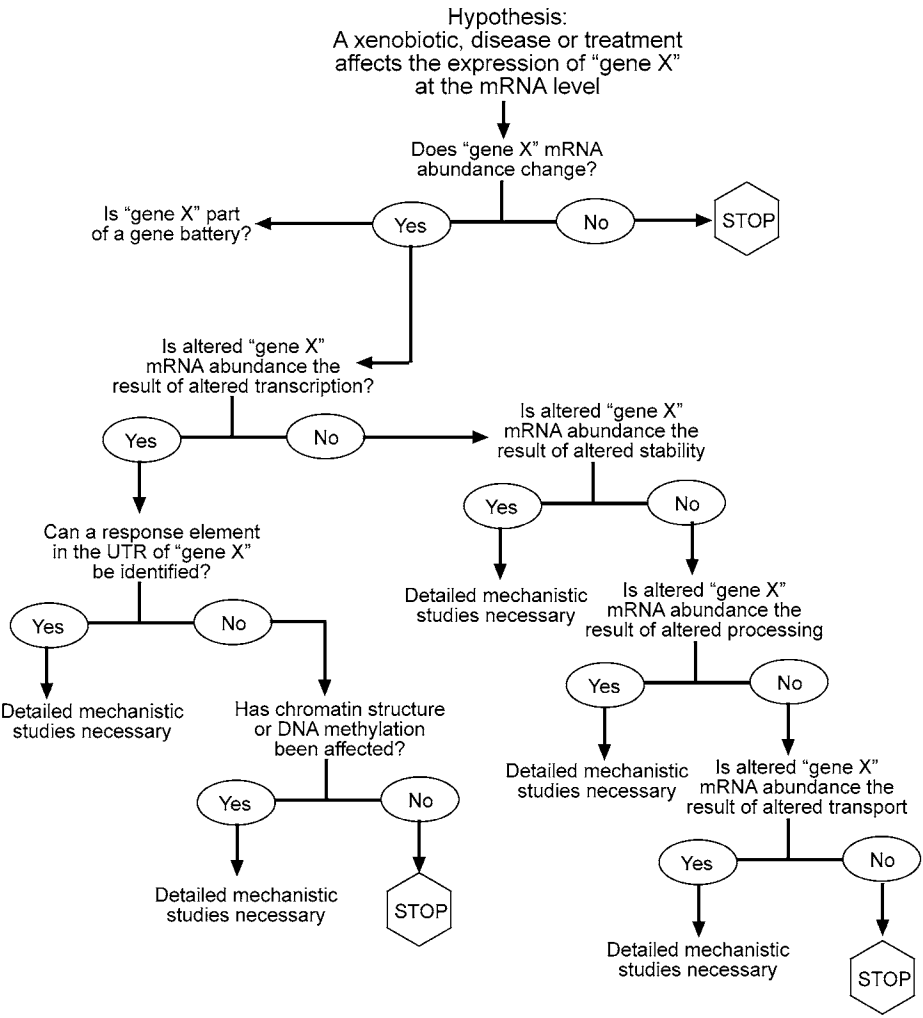


Fig. 1-2. Decision tree to determine how a xenobiotic treatment or disease affects gene expression at the mRNA level. *See* text for details.

Although the linear scheme depicted in Fig. 1-1 conveys the mechanism of mRNA accumulation, it is rarely utilized as an outline to determine the mechanism by which a gene is being regulated. More often, the scheme of events is more of a decision tree as shown in Fig. 1-2. Briefly, the hypothesis to be tested is generally "Does the particular treatment or condition affect mRNA accumulation?" If accumulation occurs, the goals of the subsequent experiments are to determine the mechanism of this regulation. The approaches to answering each

of the hypotheses listed in this figure will be discussed in detail in the subsequent chapters.

## 2. Example

The decision tree (Fig. 1-2) depicts a situation in which a xenobiotic results in altered gene expression via one mechanism only. By answering a series of yes/no questions (utilizing the approaches in the subsequent chapters), the mechanism by which a gene is being regulated can be found. However, this is rarely, if ever, the case and multiple mechanisms are likely for any particular gene and xenobiotic. A good example of this phenomena is altered expression of the proto-oncogene c-myc mRNA via the tumor promoter tetradecanoyl phorbol acetate (TPA).

The c-myc proto-oncogene is an important regulator of cellular differentiation, proliferation, growth, and apoptosis. Because of its central role in determining cell fate, the concentration of c-myc is under extremely tight and integrated control at both the mRNA and the protein level. For c-myc transcript production, control is exerted at the levels of transcription initiation and elongation, RNA processing, RNA stability, and translation. TPA and other phorbol esters are activators of protein kinase C (PKC) and regulate gene expression through AP-1 as well as other transcription factors. The effects of TPA on cell fate depend on the type of cells being examined and can include increased (hepatocytes) or decreased (hematopoietic cells) proliferation. Although not an exhaustive analysis of all the studies performed on TPA regulation of c-myc mRNA, Table 1-1 illustrates the complexity that is inherent in studies performed on mRNA levels for a growth regulatory gene.

**Table 1-1**
**Multifactorial Regulation of c-myc mRNA by Phorbol Esters**

| Question | Answer | Comment | References |
|---|---|---|---|
| Does c-myc mRNA abundance change? | Yes (either up or down) | Induction of c-myc mRNA by phorbol esters can occur through a protein kinase C (PKC)-dependent pathway in human astrocytoma cells. | Blackshear, P.J., et al. J. Biol. Chem. 262 (1987) 7774–7781. |
| | | In proliferating JURKAT cells, high levels of c-myc mRNA were found, which diminished rapidly following TPA-induced cessation of growth. | Makover D., et al. Oncogene 6 (1991) 455–460. |
| Is c-myc part of a gene battery? | Yes | Subtractive hybridization identified 16 sequences, including c-myc, that are down regulated by TPA-induced monocytic differentiation of HL60 cells. | Herblot, S., et al. FEBS Lett. 414 (1997) 146–152. |
| Is altered c-myc mRNA the result of altered transcription? | Yes | PMA decreases c-myc mRNA by blocking transcription elongation at sites near the first exon/intron border. | Chen, L., et al. J. Biol. Chem. 275 (2000) 32,227–32,233. |
| | | In THP-1 cells PMA altered c-myc expression dependent on the AP-1 enhancer activity. | Matikainen, S., Hurme, M. Int. J. Cancer 57 (1994) 98–103. |
| | | PMA stimulated transcriptional activity dependent on the phorbol 12-myristate 13-acetate-responsive element (TRE) of c-myc in MCF-7 cells. | Wosikowski, K., et al. Biochem. Biophys. Res. Commun. 188 (1992) 1067—1076. |
| Is altered mRNA abundance the result of altered stability? | Yes/No | Treatment of WEHI 231 cells, derived from a murine B-cell lymphoma, with PMA caused dramatic increase in the rate of c-myc gene transcription, as well as from partial stabilization of the mRNA in the cytoplasm. | Levine, R. A., et al. Mol. Cell. Biol. 6 (1986) 4112-4116. |

| | | |
|---|---|---|
| Yes | In human leukemia cell lines c-myc mRNA was regulated primarily by posttranscriptional mechanisms with altered mRNA half-life after TPA treatment. | Meinhardt, G., Hass, R. Leuk. Res. 19 (1995) 699-705. |
| | TPA-induced destabilization of c-myc mRNA requires post-ribosomal and polysome-associated factors. In dividing cells c-myc mRNA decays via a sequential pathway involving removal of the poly(A) tract followed by degradation of the mRNA body, TPA activates a deadenylation-independent pathway. | Brewer, G. J. Biol. Chem. 275 (2000) 33,336-33,345. |
| No | The treatment of transfected cells with phorbol ester (TPA) revealed that polyadenylated mRNAs from some genes were stabilized, while mRNAs bearing 3'-end structure of c-myc genes were not. | Sato, K. et al. Nucleic Acids Symp. Ser. (1989) 23–24. |

# 2

# Messenger RNA Accumulation

## 1. Concepts

The detection of mRNA levels of a particular gene is one of the cornerstones of molecular biology. There are many ways that mRNA can be detected, each with its strengths and weaknesses. Surprisingly, few investigators give much thought as to whether their methodologies and approaches are appropriate. Yet, the quantification and interpretation of results depends on understanding some key points. Three key factors contribute to the complexity and difficulty in examining differentially expressed genes. First, genes are not present at the same abundance and can vary from less than one up to thousands of copies per cell. This has implications for methods that must be used to accurately detect and quantify the expression of an mRNA. Second, the intensity of response varies greatly from gene to gene; that is, when comparing two treatments or conditions, an mRNA can be twofold or several orders of magnitude different between samples. Certain methods have a robust linear range and can handle both levels of response, while others are biased toward either the low or high responder. Last, there are many ways to alter gene expression, and some or all of the particular mechanisms may be at play. Approaches must be used that are capable of isolating, or at least accounting for, the competing possibilities so that hypotheses can be tested confidently. We will briefly discuss these parameters as they pertain to examining altered gene expression, and how these factors impinge on developing an optimal model system. Much of the following has been described in more detail in reference [14] as it pertained to cloning xenobiotically-induced genes.

### 1.1. mRNA Abundance

The mammalian genome of $3 \times 10^9$ base pairs (bp) has sufficient DNA to code for approx 300,000 genes, assuming a length of 10,000 bp per gene [15]. Obviously, not every gene is expressed by every cell; also, not every segment of DNA may be associated with a gene product. In fact, hybridization experiments in the mammalian cell have shown that approx 1 to 2% of the total sequences of nonrepetitive DNA are represented in mRNA [15]. Thus, if 70%

Gene Expression Control at the mRNA Level by J.P. Vanden Heuvel
From: *Regulation of Gene Expression*
By: G. H. Perdew et al. © Humana Press Inc., Totowa, NJ

of the total genome is nonrepetitive, 10,000–15,000 genes are expressed at a given time.

The average number of molecules of each mRNA per cell is called its representation or abundance. Of the 10,000 to 15,000 genes being expressed, the mass of RNA being produced per gene is highly variable. Usually, only a few sequences are providing a large proportion of the total mass of mRNA. Hybridization and kinetic experiments between excess mRNA and cDNA in solution identifies several components of mRNA complexity. Most of the mass of RNA (50%) is accounted for by a component with few mRNA species. In fact, approx 65% of the total mRNA may be accounted for in as few as ten mRNA species. The remaining 35% of the total RNA represents other genes being expressed in that tissue. Of course, the genes present in each category may be present in very different amounts and represent a continuum of expression levels. For means of this discussion, we will divide the three major components [16] into abundant, moderate, and scarce, representing approx 100,000 copies, 5000 copies and <10 copies per cell, respectively.

There are several reasons for discussing the components of mRNA. First, in a differential screen procedure (i.e., subtractive hybridization, differential display, microarray) what is actually being compared is two populations of mRNA, and you are examining the genes that overlap or form the intersection between these groups. When comparing two extremely divergent populations, such as liver and oviduct, as much as 75% of the sequences are the same [15] amounting to 10,000 genes that are identical and approx 3000 genes that are specific to the oviduct or liver. This suggests that there may be a common set of genes, representing required functions, that are expressed in all cell types. These are often referred to as housekeeping or constitutive genes. Second, there are overlaps between all com-

14. Vanden Heuvel, J.P. In: Toxicant-Receptor Interactions, Denison, M.S., Helferich, W.G., eds. Taylor and Francis, Philadelphia, PA. 1998, pp. 217–235

15. Lewin, B. Genes IV, Oxford University Press, Oxford UK.1990

16. Bishop, J. O. et al. Nature 270 (1974) 199–204.

15. Lewin, B. Genes IV, Oxford University Press, Oxford UK.1990

ponents of mRNA, regardless of the number of copies per cell. That is, differentially expressed genes may be abundant, moderate, or scarce. In fact, the scarce mRNA may overlap extensively from cell to cell, on the order of 90% for the liver-to-oviduct comparison. However, it worth noting that a small number of differentially expressed genes are required to denote a specialized function to that cell, and the level of expression does not always correlate with importance of the gene product.

As discussed in sections following, the key to developing an effective model for the study of differential gene expression may be to keep the differences in the abundant genes to a minimum. This is due to the fact that a small difference in expression of a housekeeping gene, say twofold, will result in a huge difference in the number of copies of that message from cell type to cell type (i.e., an increase of 10,000 copies per cell). Also, it is important to have a screening method that can detect differences in the scarce component. If the two populations to be compared have little difference in the abundant genes and you have optimized your screening technique to detect differences in the scarce population, the odds of cloning genes that are truly required for a specialized cellular function have increased dramatically.

### 1.2. Intensity of Response

A basic pharmacologic principle is that drugs and chemicals have different affinities for a receptor, and drug-receptor complexes will have different efficacies for producing a biological response, i.e., altering gene expression. A corollary of this principle states that not every gene being effected by the same drug-receptor complex will have identical dose-response curves. That is, when comparing two responsive genes, the affinity of the drug-receptor complex for the DNA response elements found in the two genes and the efficacy of the drug-receptor—DNA complex at effecting transcription could be quite different. In fact, similar DNA response elements may cause a repression or an induction of gene expression, depending on the context of the surrounding gene. Therefore, when comparing two populations of mRNAs (i.e., control versus treated), there may be dramatic differences in the levels of induction and repression regardless of the fact that all the genes are affected by the same drug-receptor complex.

Needless-to-say, the extent of change is important in the detection of these differences but not the importance of that deviation. For technical reasons, it is often difficult to detect small changes in gene expression (less than twofold). However, a twofold change in a gene product may have dramatic effects on the affected cell, especially if it encodes a protein with a very specialized or nonredundant function. Also, the detection of a difference between two cell popula-

tions is easier if the majority of the differences are in scarce mRNAs. Once again, this is owing to technical aspects of analyzing gene expression whereby the change from 500 to 1000 copies per cell is a dramatic effect compared to a change from $1 \times 10^5$ to $2 \times 10^5$, an effect that may be virtually unnoticed.

### 1.3. Specificity of Response

The last factor we will discuss regarding the complexity of mRNA species is that regulation of gene expression is multifaceted. The analysis of differential gene expression is most often performed by comparing steady-state levels of mRNA; that is, the amount of mRNA that accumulates in the cell is a function of the rate of formation (transcription) and removal (processing, stability, degradation). If differences in protein products are being compared, add translation efficiency and posttranslational processing to the scenario. With all the possible causes for altered gene expression, the specificity of response must be questioned. Is the difference in mRNA or protein observed an important effect on expression or is it secondary to a parameter in your model system you have not controlled or accounted for?

In the best-case situation, the key mechanism of gene regulation that results in the end point of interest should be known. At least, one should have criteria in mind for the type of response that is truly important. With most receptor systems, early transcriptional regulation may predominate as this key event. However, the other modes must be acknowledged, at least when novel genes are being characterized. By assuming that the key event is mRNA accumulation, the true initiating response such as protein phosphorylation or processing may be overlooked. Also, the extent and diversity of secondary events, i.e., those that require the initial changes in gene expression, may far exceed the primary events. The amplification of an initial signal (i.e., initial response [gene A] causes regulation of secondary response [gene B]) can confuse the interpretation of altered mRNA accumulation. Once again, one must have a clear understanding of whether a primary or secondary event is the key response and design your model accordingly.

## 2. Basic Methods and Approaches

For specific procedures on the analysis of gene expression the reader is directed to laboratory manuals such as Current Protocols in Molecular Biology [9] and Molecular Cloning [10]. Popular methods for examining mRNA accumulation are outlined in Fig. 2-1. As with any laboratory procedure, you must keep in mind the limitations of the procedure and your own technical expertise. In this section, the methods will be outlined broadly but with emphasis on the positive and negative aspects of each approach.

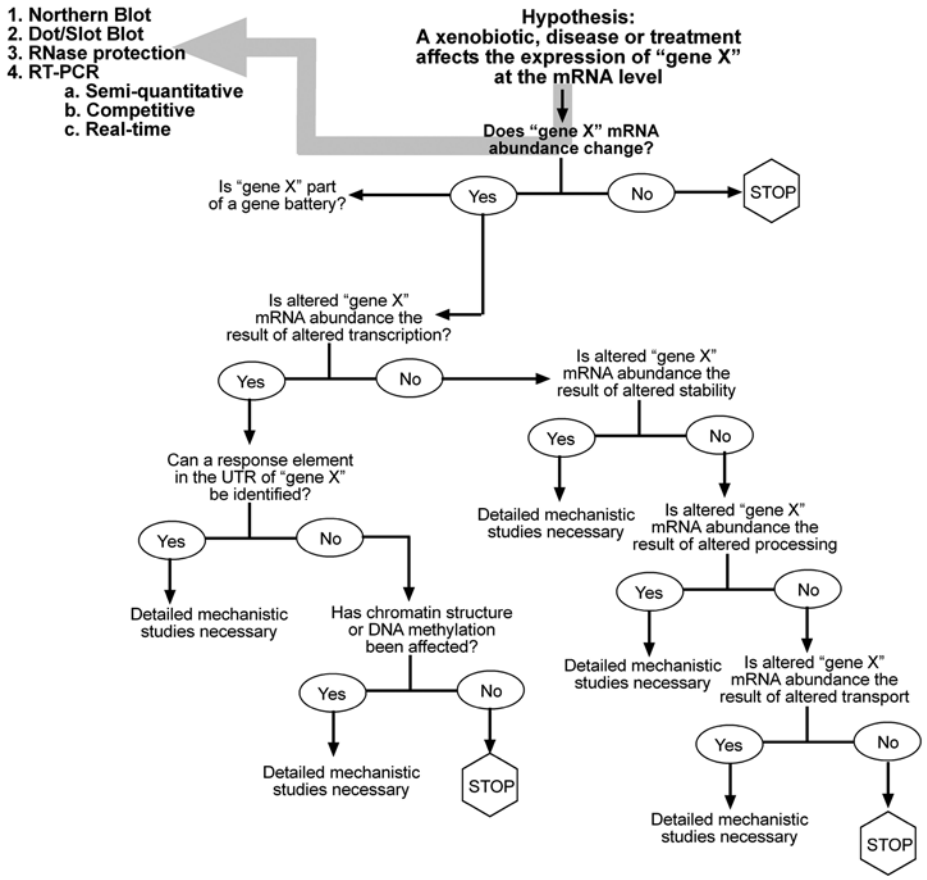


Fig. 2-1. Outline of approaches to examine mRNA accumulation.

Many are aware of the need for defining the model system when examining genome-wide mRNA accumulation, as when microarrays or serial analyses of gene expression (SAGE) are performed (Chapter 3). However, similar considerations should be made when using Northern blots or reverse transcriptase-polymerase chain reaction (RT-PCR) to examine a specific gene. Proper controls are necessary to minimize variability, and the studies must be designed in a way that eliminates extraneous factors (e.g., stress, contamination) that are not under investigation. In this section, approaches used to quantify the differences between two populations of mRNA for a particular gene will be examined. An assumption will be made that the model system is adequately designed to specifically test the hypothesis.

## 2.1. Northern Blots

*An excellent resource for RNA analysis methods is www.ambion.com.*

Despite the advent of more powerful techniques (RT-PCR, RNase protection), Northern blot analysis remains a standard method for detection and quantitation of mRNA levels. Northern blot analysis provides a relative comparison of message abundance between samples, is the preferred method for determining transcript size, and can be used for detecting alternatively spliced transcripts. The Northern blot procedure is straightforward and inexpensive, using common equipment and supplies present in most molecular biology laboratories.

*The Northern blot measures mRNA levels. Some other "directional" assays include Southern (DNA), Western (protein) and Southwestern (protein–DNA) blots.*

There are significant limitations associated with Northern blot analysis. First, despite the relative simplicity of the methodology, great pains must be taken to avoid RNase contamination of solutions and plasticware. If RNA samples are even slightly degraded, the quality of the data and the ability to quantitate expression are severely compromised. Second, Northern blotting is much less sensitive than nuclease protection assays and RT-PCR. Rare genes require highly specific probes and large amounts of RNA. Enriching the sample for polyadenylated mRNA and optimizing hybridization conditions can improve sensitivity to some extent. Third, Northern blots are not amenable to high-throughput analysis. For example, although it is possible to detect more than one gene per blot, often this requires stripping the nylon membrane and reprobing,which is time consuming and problematic, since harsh treatment is required to strip conventional probes from blots. Last, this procedure is most often used as a relative quantitation method, since the gene in question is being examined as a function of the expression of a housekeeping gene. Although internal standards can be synthesized and spiked into the sample (e.g., a synthetic gene with the hybridization site of the probe) in order to generate a standard curve and to obtain absolute quantification, this is rarely performed. Thus Northern blotting requires a large difference between samples (five- to tenfold) to be significant.

The general approach of Northern blot analysis is as follows (*see* Fig. 2-2). Extraction of high-quality intact RNA is a critical step in performing Northern analysis. This is generally performed by cell lysis with detergents or solvent, inhibition of ribonuclease, and ultimately separation of proteins and DNA from the RNA. This later event can be performed by liquid phase separation or via oligo(dT) chromatography. Once RNA samples are isolated, denaturing agarose gel electrophoresis is performed. Formaldehyde has traditionally been used as the denaturant, although the glyoxal system has several advantages over formaldehyde. All buffers and apparatus must be painstakingly treated with RNase inhibitors. Following separation by denaturing agarose gel electrophoresis, the RNA is transferred to a positively charged nylon membrane and then immobilized for subsequent hybridization. The transfer may be performed using a passive, slightly alkaline elution or via commercially available active transfer methods (electroblotter, pressure blotter). The membrane is crosslinked by ultraviolet light or by baking. Northern blots can be probed with radioactively or nonisotopically labeled RNA, DNA, or oligodeoxynucleotide probes. DNA probes may originate from plasmids or PCR. RNA probes can be produced by in vitro transcription reactions. Radioactivity may be incorporated during the PCR or in vitro transcription reactions or may be performed via end-labeling or random priming. Prehybridization, or blocking, is required prior to probe hybridization to prevent the probe from coating the membrane. Good blocking is necessary to minimize background problems. Although double-stranded DNA probes must be denatured prior to use, RNA probes and single-stranded DNA probes can be diluted and then added to the prehybridized blot. After hybridization, unhybridized probe is removed by washing in several changes of buffer. Low-stringency washes (e.g., with 2X SSC) remove the hybridization solution and unhybridized probe. High-stringency washes (e.g., with 0.1X SSC)

*There are many commercially available kits for the extraction of total RNA (mRNA, tRNA, rRNA) or poly(A) RNA (mRNA). Most methods for total RNA utilize chloroform:phenol extraction, whereas mRNA can be enriched via oligo(dT) chromatography.*

*SSC buffers are commonly employed in nucleic acid hybridizations. 20X SSC consists of 3 M NaCl, 0.3 M Na3 citrate.*

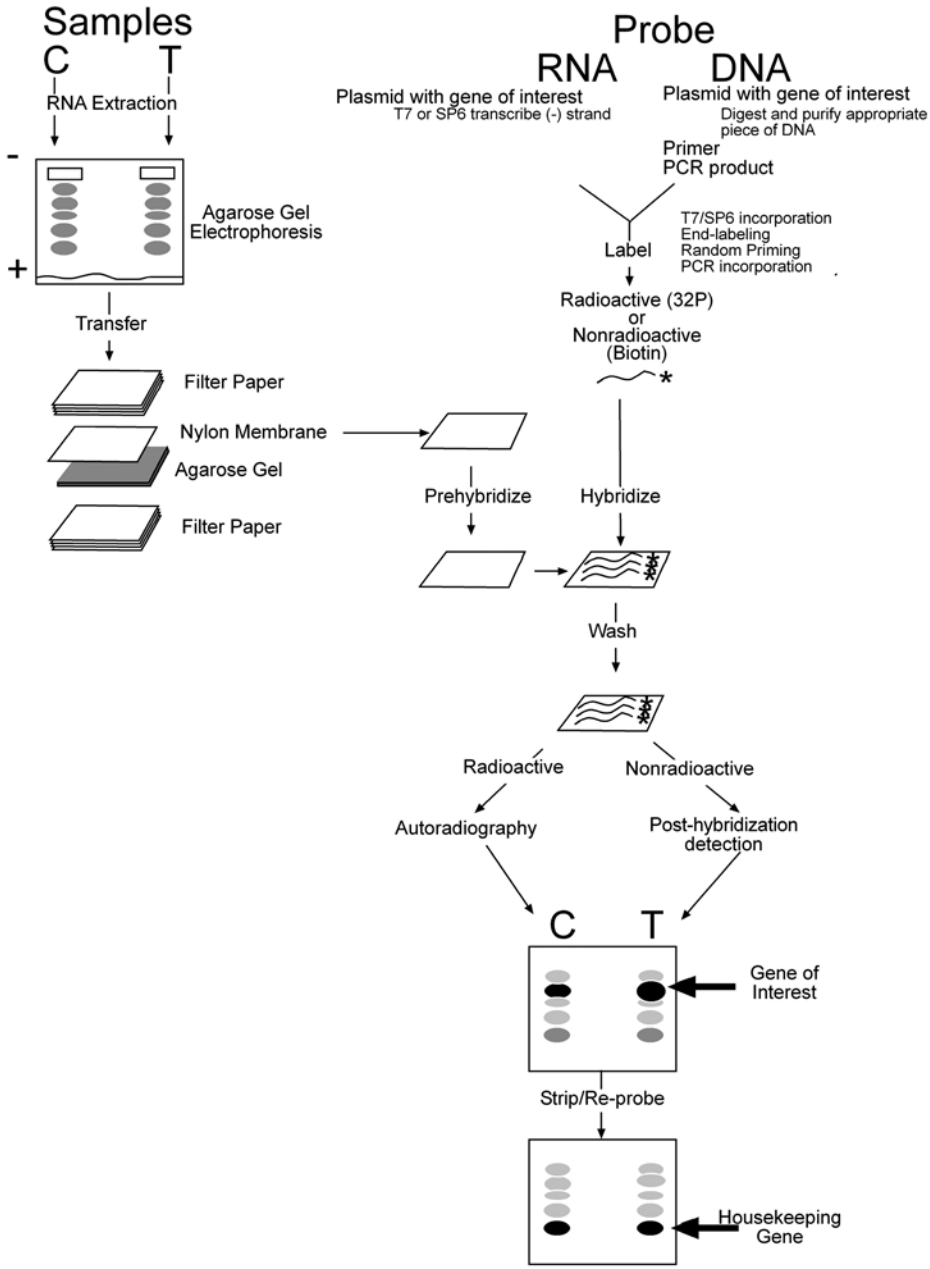


Fig. 2-2. Basic procedures in Northern blotting.

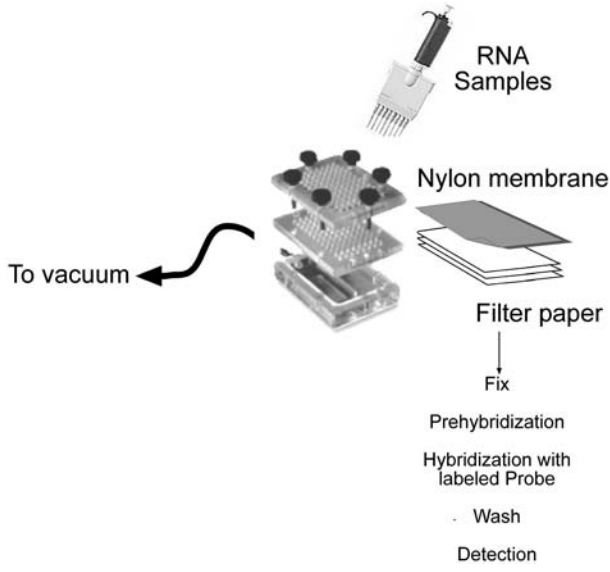


Fig. 2-3. Dot- or slot-bot analysis of RNA.

remove partially hybridized molecules. If a radiolabeled probe was used, the blot can be wrapped in plastic wrap to keep it from drying and then immediately exposed to film for autoradiography. If a nonisotopic probe was used, the blot must be treated with nonisotopic detection reagents prior to film exposure. Standard methods for removing probes from blots to allow subsequent hybridization with a different probe often include harsh treatments with boiling 0.1% SDS or autoclaving.

## 2.2. Dot/Slot Blots

Dot or slot blot analysis takes its name from the apparatus used to apply the sample to the nylon membrane, as shown in Fig. 2-3. This method is analogous to the Northern blot in most ways but has a much higher throughput, as dozens of samples can be examined simultaneously. RNA samples are applied to the membrane and fixed by UV or heat crosslinking. The subsequent application of probe, washing and detection is as described for Northern blots. The main advantage of the dot/slot blot is the number of samples that can be run simultaneously and the relative ease of quantitation resulting from uniformity in bands to be analyzed. There may be some increased sensitivity in the assay because more sample can be loaded onto the matrix, as long as the nitrocellulose or nylon does not become saturated. In addition, the RNA can be of slightly lower quality and still give a detectable signal. However, sensitivity is still signifi-

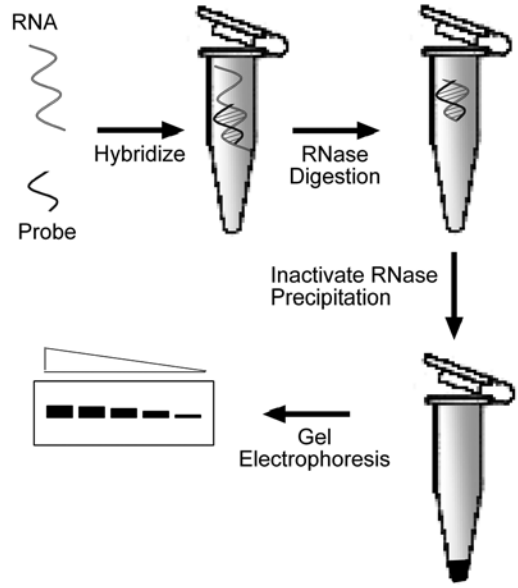


Fig. 2-4. Basics of RNase protection assays (RPAs).

cantly less than that for ribonuclease protection assay or RT-PCR. The dot/slot blot does not result in any information about transcript size, and conditions must be optimized to ensure that the probes are specific for the target gene.

It is also possible to perform a "reverse Northern" using the dot/slot blot format. In this procedure, the probe is spotted onto the nylon and the RNA sample is labeled (isotopic or nonisotopic) and added to the hybrizidation solution. This allows for the examination of multiple genes in one sample. The reverse Northern is predecessor to the microarray assay that will be discussed in detail in Chapter 3.

## 2.3. RNase Protection Assays (RPA)

Ribonuclease (or RNase) protection assays (RPAs) are an extremely sensitive method for the detection and quantitation of specific RNAs in a complex mixture of total cellular RNA. RPA is a solution hybridization of a single-stranded antisense probe to an RNA sample (*see* Fig. 2-4). The small volume solution hybridization is far more efficient than more common membrane-based hybridization and can accommodate much more RNA (hence increasing sensitivity over Northern blotting). After hybridization, any remaining unhybridized probe and sample RNA are removed by digestion with a mixture of nucleases (usually RNase A and T1; *see* Table 2-1). The nucleases are inacti-

vated, usually by extraction with chloroform/phenol, and the remaining RNA-probe hybrids are precipitated. These products are separated on a denaturing polyacrylamide gel and are visualized by autoradiography or secondary detection. To quantitate mRNA levels using RPAs, the intensities of probe fragments protected by the sample RNA are compared to the intensities generated from either an endogenous internal control (relative quantitation) or known amounts of sense strand RNA (absolute quantitation).

As just mentioned, one major advantage of RPAs over membrane-constrained hybridization is sensitivity. Low-abundance genes are detectable in RPAs, or a small sample size may be used for moderate or high expressing species. In addition, RPAs are the method of choice for the simultaneous detection of several RNA species. During solution hybridization and subsequent analysis, individual probe-target interactions are for all intents and purposes independent of one another. Thus, several RNA targets and housekeeping genes can be assayed simultaneously if the protected fragments are of different lengths. In addition to their use in quantitation of RNA, various RNases can be used to map the structure of transcripts, as will be discussed in more detail in Chapter 4 and 5.

The basic steps of the typical RNase protection assay are relatively straightforward and utilize common materials found in molecular biology laboratories (Fig. 2-4). RNA is extracted as described for the Northern blot analysis. RPA requires RNA probes, most often prepared using in vitro transcription assays and either radiolabeled or nonisotopically labeled. Oligonucleotides and other single-stranded DNA probes can only be used in assays containing S1 nuclease. Optimization of an RPA assay is very probe- and gene-specific and depends mostly on transcript concentration. Probe concentration must be in molar excess over the target mRNA to ensure rapid hybridization. Thus, highly abundant targets

**Table 2-1**
**RNase Properties**

| RNase | Specificity |
|---|---|
| **RPA** | |
| A | 3' of ss C and U |
| I | 3' of ssNTP |
| T1 | 3' of ss G |
| S1 nuclease | ssRNA |
| **RNA structure** | |
| H1 | RNA of RNA/DNA |
| V1 | dsRNA |

ss, single stranded; ds, double stranded

*Ambion (www.ambion.com) suggests the following dilutions of labeled to unlabeled probe:*

- *Low abundance, 1:0*
- *Moderate abundance,1:50*
- *Abundant, 1:10,000*

require the use of low specific-activity probes for abundant targets (dilute labeled probe with unlabeled) whereas probes for scarce RNA targets require no dilution. Generally, 10 µg total RNA is hybridized overnight with approx 50,000 cpm for each probe.

*Comparison of S1 nuclease and RNase A/T1 RPA can be found at: http:/www.promega.com/pnotes/38/38_01/38_01.htm. Accessed 03/24/06.*

Although S1 nuclease is able to cleave after every residue in both RNA and single-stranded DNA, some problems are associated with its use in RPA assays. Lower incubation temperatures and high salt concentrations (>200 m*M*) must be used to favor single-stranded over double-stranded cleavage, and S1 nuclease is prone to nonspecific cleavage in AU-rich regions. Thus, mixtures of RNase A and T1 typically have been used for digestion of unhybridized RNA in solution hybridization experiments. Nuclease P1 has been substituted for RNase A when using AU-rich probes which require lower reaction temperatures. The amount of enzyme(s), buffer constituents and incubation conditions vary widely from protocol to protocol and may require some optimization. Following digestion, the RNases need to be inactivated. Phenol/ chloroform extraction is required to inactivate Nuclease P1 and RNases A and T1. This extraction is not required for some of the commercially available RNase mixtures and addition of ethylenediaminetetraacetic acid (EDTA) and ethanol precipitation is sufficient.

The percentage of polyacrylamide in gel electrophoresis depends on the size of the products to be examined. Typically a 6% denaturing acrylamide gel is used to resolve fragments of 300–1000 nucleotides. The detection and analysis of the band intensities depends on whether radioactivity or nonisotopic methods were used.

## 2.4. RT-PCR

Polymerase chain reaction is an enzymatic assay which is capable of producing large amounts of a specific DNA fragment from a small amount of a complex mixture (Fig. 2-5; reviewed in [17]). In RT-PCR, the mRNA must first be converted to a

17. Mattes, W. B. In: PCR Protocols in Molecular Toxicology, Vanden Heuvel, J.P., ed. CRC Press, Boca Raton, FL. 1997, pp. 1–40.

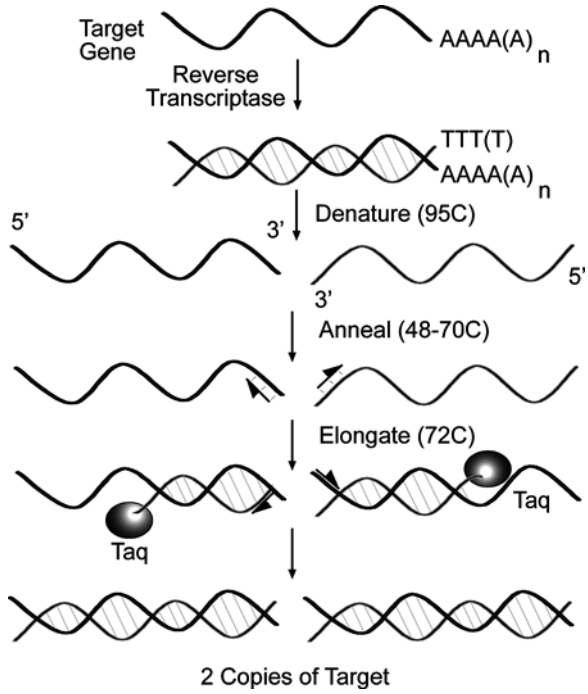


Fig. 2.5. Reverse transcription and one cycle of PCR.

double-stranded molecule by using the enzyme reverse transcriptase (RT). The thermostable DNA polymerase (i.e., Taq) and the use of specific "primers" are the key features of any PCR reaction. All known DNA polymerases require deoxyribonucleotide triphosphates (dNTPs), a divalent cation ($Mg^{2+}$ or $Mn^{2+}$), a DNA or cDNA template, and a region of that template that is double stranded adjacent to a single-stranded nick or gap. The double-stranded region is provided by the primer annealing to its complementary region of the DNA template. If the starting mixture includes not only a single-stranded polynucleotide template, but also (1) its complementary strand and (2) two oligonucleotide primers that hybridize to both strands, copies of both of these strands will be produced each cycle, and these copies can be used as templates for subsequent cycles. Short DNA fragments whose ends are defined by the position of the two oligonucleotide primers will accumulate in an exponential fashion, i.e., like a chain reaction. If 30 cycles of PCR are performed, theoretically one will achieve a $2^{30}$ amplification of the target gene's cDNA. The product which is formed is specific for a particular transcript as dictated by the design of the oligonucleotide primers.

**Common RT-PCR Definitions**

Internal standard (IS): A type of control molecule that can be used to minimize tube-to-tube variability in amplification efficiency. Normally an IS is a synthetic molecule that contains the same recognition sequences as the gene of interest. This type of amplification control is spiked into the PCR reaction.

External standard (ES): A type of control that can be used to minimize differences in template (mRNA, cDNA, or DNA) concentration from sample to sample. Most often an ES is a housekeeping gene that is used in a "coamplification" type of quantitation. This control is not added to each sample, as it is present in a finite amount in each tube. Care must be taken to ensure that the housekeeping gene is not affected by the treatment condition. Typical external standards include actin, tubulin, 18S rRNA, or glyceraldehyde 3-phosphate dehydrogenase (GAPDH)

Template: Any cDNA or DNA that contains primer recognition sites and can be PCR amplified.

Target gene: The gene of interest; to be differentiated from the internal standard, external standard, or artifact templates.

Linker gene: Specifically used to describe a template used to create a synthetic molecule in the synthesis of an internal standard

Forward primer (FP): Analogous to the 5' or "upstream" primer (usp).

Reverse primer (RP): Same as the 3' or "downstream" primer (dsp).

Cross-over point: In competitive RT-PCR, the concentration of internal standard at which the PCR products for the target and the internal standard are equivalent.

Therefore, RT-PCR is a tool to examine the messenger RNA expression of a target gene in that the amount of product formed is a function of the amount of starting template. Of course, the examination of mRNA accumulation can be determined in many cases by hybridization procedures such as Northern blots, dot or slot blots, and RNase protection assays. Nonetheless, in terms of amount of sample required, detection of small differences in expression and ability to examine many genes in a large number of samples, RT-PCR stands above the more conventional procedures.

### *2.4.1. Competitive RT-PCR*

The predominant negative of RT-PCR is related to its ability to amplify the products. That is, an internal standard (IS) is necessitated in these assays owing to the fact that there is a large amount of tube-to-tube variability in amplification efficiency. For example, if ten tubes of seemingly identical reagents are PCR amplified, there could be as much as a threefold difference in the amount of product formed. If an IS was coamplified with the target, the efficiency of amplification in each tube could be corrected and this threefold difference could easily be negated.