

Lecture 15 - Gene Expression in Cultured Cells (AMG text pp. 175-191)
October 11, 2001

Molecular genetic methods used to characterize cloned genes fall into four basic categories;

- **gene expression surveys:** Where and When is the gene expressed?
- **transcriptional regulation studies:** What is the mechanism of its regulation?
- **ectopic expression experiments:** What happens if the gene is expressed aberrantly?
- **protein characterization assays:** What does the gene encode?

Expression Vectors

Similar to endogenous genes, expression vectors can contain DNA sequences that direct low level basal transcription, or more commonly, can have promoter elements linked to upstream regulatory sequences that serve as recognition sites for enhancer binding proteins (EBPs).

There are four basic types of expression vectors used for modulate gene expression:

1. **Minimal promoters** used to study gene regulatory elements such as enhancer elements.
2. **Constitutive promoters** used to direct expression of gene products.
3. **Cell-specific promoters** used to specify expression to target cells.
4. **Regulated promoters** used to control the on/off expression of cloned genes.

Expression vectors contain DNA sequences flanking the multiple cloning site (MCS) that function as sequence-specific binding sites for transcription factors. Common DNA sequence motifs found in many types of promoters are represented by SP1, CAAT and AP1 binding sites. HREs are hormone response elements.

What explains the cell-specific expression of genes, isn't the DNA sequence within all cells the same?

Why would you want to use a generic "minimal" promoter to analyze gene regulatory sequences of a recently cloned gene, why not just use the promoter of the gene you are analyzing?

What is an "HRE" and why would 4 be better than 1, what does that tell you about the mechanism of regulated gene expression?

Reporter Genes

One way to accelerate the rate at which promoter mapping studies can be done is to create an artificial "reporter gene" by fusing the regulatory region of a the test gene, to a

heterologous gene coding sequence that directs the synthesis of a readily detectable protein product.

Assuming that the reporter gene assays will be performed under conditions where the steady-state levels of mRNA and protein are directly correlated, reporter genes must also meet several other criteria if they are going to be reliable and useful for promoter mapping studies.

1. The reporter gene must not encode a protein activity that is similar to one already present in the target cell.
2. The protein assay needs to be sensitive over several orders of magnitude, reproducible, and easy to perform. In this regard, it is also helpful if the assay is cost-effective.
3. The reporter protein function must not interfere with host cellular processes in a way that will alter intracellular signaling pathways or metabolic rates.

In this hypothetical example, the AMG gene 5' regulatory region was cloned into a reporter gene vector containing the bacterial chloramphenicol acetyl transferase (CAT) coding sequence. Transfection of two different cell types (A or B) with either the promoterless CAT vector (pBasic), the AMG-CAT test plasmid (pAMG), or a constitutive CMV-CAT reporter plasmid (pCMV), would be used to determine if this cloned portion of the AMG regulatory region encodes a cell-specific promoter.

One way to measure the level of CAT enzyme activity in cell extracts is to quantitate the production of mono- and diacetylated ¹⁴C-chloramphenicol using thin layer chromatography (TLC) followed by phosphorimaging. As depicted in this example, it would be concluded that the cloned fragment of the AMG promoter selectively directs transcription of the CAT reporter gene in cell type B. The constitutive cytomegalovirus (CMV) promoter functions in both cell types.

Commonly Used Reporter Genes

chloramphenicol acetyl transferase (CAT) - This is a bacterial enzyme that catalyzes the transfer of acetyl groups from acetyl-coenzyme A to the antibiotic chloramphenicol. This enzyme provides protection to the bacterium by inactivating chloramphenicol through acetylation. A disadvantage of CAT reporter genes is that the CAT assay is laborious and expensive.

luciferase (luc) - Protein expressed from the luc gene of the firefly species *Photinus pyralis* can be easily detected in cell extracts using a specially designed luminometer that measures fluorescence emitted from the luc-catalyzed ATP-dependent oxidation of compounds called luciferans. Reporter genes containing the luc gene from the sea pansy *Renilla reniformis* can be used in combination with the firefly luc reporter gene as an internal control for DNA transfection efficiencies.

b-galactosidase (b-gal or lacZ) - This bacterial enzyme is one of the most versatile reporter genes available because b-gal enzyme activity can both be measured in cell extracts using a spectroscopic assay that detects cleavage of ONPG (o-nitrophenyl-b-D-galactopyranoside), and by histochemical methods that measure the appearance of a blue precipitate produced from the cleavage of X-gal.

Green fluorescent protein (GFP) - The gene for this autofluorescent protein has rapidly become one of the most useful reporter genes for marking transfected cells *in vivo*. GFP is expressed in the outer dermal layer of the Pacific Northwest jellyfish *Aequorea victoria*.

GFP autofluorescence provides a green bioluminescent light to the jellyfish in response to a calcium-dependent energy transfer step involving a GFP-associated protein called aequorin. It was found that exposure to ultraviolet light causes GFP to autofluoresce a bright green color within living cells, even in the absence of aequorin, calcium, or any other cofactor or substrate.

The molecular structure of purified *A. victoria* GFP reveals a striking barrel-like arrangement of eleven β -sheets surrounding the internal cyclic tripeptide chromophore. The structure below is of a GFP dimer showing the orientation of the two proteins relative to each other. A GFP variant was generated that emits "blue" light by mutating Tyrosine 66 to Histidine.

The molecular basis for this UV-induced autofluorescence is a cyclization process involving the tripeptide Ser65-Tyr66-Gly67 which functions as the chromophore in the intact protein. Variants of GFP have been developed that fluoresce at other wavelengths, either by mutating the GFP gene, or by isolating other fluorescent proteins. Here are some examples of GFP applications in molecular genetics:

Green Transgenic Mice.

Osaka University spermatologist Masaru Okabe and colleagues at Osaka's Research Institute for Microbial Diseases added a form of a gene for green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* into mice, producing animals that are green through and through when exposed to blue light. "It's very beautiful," Okabe says. The gene had previously been transferred into fruit flies and zebrafish, but not mammals.

Bacterial Expression of GFP Variants.

Double-labelling experiments using multiple autofluorescent protein genes.

Fluorescence Resonance Energy Transfer (FRET) is another application of autofluorescent proteins. In this strategy, excitation of GFP will result in emission from a nearby protein such as blue fluorescent protein (BFP) if it is physically close enough. The best FRET pairs are actually the cyan and yellow mutants of GFP, referred to as CFP and YFP.

Schematic depiction of the cameleon system for measuring protein-protein interactions by fluorescence resonance energy transfer (FRET). Increasing local $[Ca^{2+}]$ leads to association of the intervening calmodulin protein and the M13 myosin light chain

kinase peptide with accompanying increase in FRET between CFP and YFP. (b) FRET response of individual HEK/293 stable transfectant cells, expressing yellow cameleon-3 protein, to addition of a Ca^{2+} ionophore and additional extracellular Ca^{2+} .

Ligand-dependent FRET illustrated for full-length CFP-RAR bound to DNA as a heterodimer with retinoid X receptors and YFP-tagged nuclear receptor interaction domain of SRC-1. Imaging the interaction between CFP-RAR full-length and SRC-1-YFP by FRET in live HeLa cells. The ratio of FRET channel emission and donor CFP emission in response to the RAR ligand TTNPB is shown.

What two properties of GFP make it so useful as "molecular tag" in molecular genetic applications?

What is required for FRET to work, i.e., what is the underlying assumption about the spike in 530 nm emission?

The original Aequora victoria GFP protein gene has been "humanized" by altering the nucleotide sequence without changing the amino acid sequence. How and why was this done?

Inhibitory "Anti-Gene" Expression Strategies

Several gene expression-based methods have been developed that use a form of "antisense genetics". These methods are based on introducing sequence-specific nucleotide binding molecules into cells that cause a transient inhibition of gene expression at the transcriptional or post-transcriptional level. The broader implications of developing antisense reagents for gene-specific inhibition, are their potential as **therapeutic agents**.

There are basically four types of anti-gene strategies:

1. Antisense oligodeoxynucleotides that **block translation** of mRNA by targeting the AUG codon.
2. Antisense oligonucleotides derivatives that **block gene transcription** at the promoter.
3. RNA molecules called **Hammerhead Ribozymes** that degrade RNA targets catalytically.
4. Double strand RNA (**dsRNA**) **gene silencing** that specifically degrades mRNA targets in some cells.

Ribozymes are single strand RNA molecules that contain a catalytic site, called a "hammerhead structure". Ribozymes can be targeted to an RNA substrate by complementary base pairs engineered into regions flanking the hammerhead structure. A number of proof-of-principle ribozyme experiments have been described using cultured cells and several clinical trials have been undertaken to determine if ribozyme-based gene therapy approaches are effective. [Click here to download a review article on ribozymes.](#)

Double-stranded RNA (dsRNA) has recently been shown to trigger sequence-specific gene silencing in a wide variety of organisms, including nematodes, plants, trypanosomes, fruit flies, plants and mice. The mechanism of **RNA interference (RNAi)** is thought to involve cellular proteins that recognize double strand RNA as a "**foreign invader**", e.g. a viral genome, and triggers a cycle of sequence specific RNA degradation.

This form of gene silencing was first discovered in *C. elegans* and has since been demonstrated in numerous organisms, including mammals (mouse embryos). By adding gene-specific exogenous dsRNA to cells, it has been found that this cellular process will **degrade the entire pool of target mRNA** in the cell. One recent model describing how RNAi functions in gene silencing invokes the formation of a cycling complex.

Injection of mouse embryos with dsRNA directed against GFP mRNA has been shown to abolish green fluorescence, similar experiments have been done with dsRNA directed against tissue plasminogen activator.

Methods to express genes in cell lines

Much of what we know about gene regulation in higher eukaryotes, especially mammalian systems, has come from studies using **immortalized cell lines**. Many types of cancer cells have an unlimited capacity to proliferate in serum-containing medium. The term "immortalizing functions" to refer to unlimited proliferative capacity in cell culture, whereas, "transforming functions" refer to the ability of immortalized cells to also form tumors in animals.

Table 7.2 lists some of the mammalian cell lines that are most commonly used in molecular genetic applications. Most cell lines in use today can be obtained for a nominal fee from the American Type Culture Collection (ATCC), a nonprofit organization in Rockville, Maryland, that has collected an archive of over 4000 immortalized cell lines representing at least 150 different species.

Gene Transfer Methods

Three basic DNA transfection strategies have been developed to deliver cloned DNA to cells that I have chosen to call the 1) **Trojan Horse**, 2) **SWAT Team** and 3) **Biological Warfare**, strategies of DNA transfection.

The **Trojan Horse** strategy is based on the use of a positively-charged carrier molecules that are mixed with the experimental DNA in vitro and then applied directly to the cell culture media. These carrier-DNA complexes associate with cell membranes and thus **facilitate DNA uptake**. Three most commonly used transfection methods are **calcium phosphate precipitation**, **DEAE dextran-mediated gene transfer** and **liposome-mediated gene transfer**.

Three **SWAT Team** DNA transfection methods have been developed; **electroporation**, **biolistics (gene gun)**, and **microinjection**. The overall efficiency of these brute force transfection methods are relatively cell type independent.

The **Biological Warfare** strategy uses recombinant eukaryotic viruses to deliver DNA to host cells. The two most efficient mammalian cell viral vectors are derived from retroviruses and adenoviruses.

Liposome-mediated transfection

A mixture of a **polycationic lipid** and a **neutral lipid** will result in the formation of unilamellar liposome vesicles that have a net positive charge due to the **highly-positive amine** head groups on these molecules. Liposomes have become the method of choice for carrier molecules in routine cell line DNA transfections because liposome-mediated gene delivery is technically easy, highly reproducible, and very efficient.

Electroporation

Electroporation is the most versatile method of DNA transfection because it has been shown to work for such a wide variety of cell types, which includes primary cells from tissue isolates, plant protoplasts and bacterial cells. By varying the **electric field strength**, and the length of time the cells are exposed to the electric field, it is possible to optimize electroporation parameters for essentially any cell type. The use of square wave pulses is an improvement.

Retroviral infection

Retroviral expression vectors are **plasmid-based shuttle vectors** that contain deleted versions of the viral genome. The recombinant viral shuttle vector needs to be stably-transfected into a special cell line that constitutively expresses the viral reverse transcriptase and capsid proteins. Several improvements in the use of **retroviral vectors** have made them commercially available, however, NIH **Biosafety** precautions for viruses must be used in the laboratory.

Retroviral expression vectors provide an efficient gene transfer method for expressing cloned genes in appropriate host cells. Non-replicating recombinant retroviral stocks can be made by stably transfecting a **retrovirus packaging cell line** with recombinant retroviral vector using liposome-mediated transfection. Retroviral expression systems can be used for both transient and stable transfection strategies.

Transient Transfection Assays

For many types of gene expression assays, it is possible to transfect cell lines with an appropriate reporter gene, and then collect the data 24-48 hours later. Relative to the generation time of most cell lines (~16-24 hours), this represents a **relatively short time period** between DNA transfection and cell harvesting, and hence the term "transient" transfection.

Transient transfection protocols **need to be optimized** for each cell type being analyzed due to inherent differences in DNA uptake efficiencies. In the example shown below,

the electroporation conditions of 250V and 1180 μ F would be chosen as optimal for efficient transient transfection of this cell type.

Stable DNA Transfection

Stable DNA transfection of cell lines is basically done the same way that transient transfections are, with the exception that stable integration of the transfected DNA needs to be positively selected for by including a marker gene on the expression vector. Most often the marker gene encodes an enzyme that inhibits the function of a toxic compound. A description of the most commonly used **dominant selectable marker genes** for stable transfection of eukaryotic cell lines is given below.

Regulated Expression Systems

An important molecular genetic strategy for studying the cellular function of cloned genes, is to use an expression vector that contains a **transcriptionally-regulated promoter**. One of such regulated expression system was developed by Manfred Gossen and Hermann Bujard. Their strategy uses the **tetracycline-resistance operon** of the E. coli Tn10 transposon to control the expression of genes that have been cloned downstream of a promoter containing tetracycline receptor (TetR) binding sites.

The "**Tet-off**" system is repressed in the presence of the tetracycline analogue doxycycline. The VP16 transcriptional activation domain (VP) is derived from the herpes simplex virus VP16 protein.

In the "**Tet-on**" system, the DNA binding domain of the Tet-on regulator (rTetR) contains mutations that convert it from a repressor that only binds DNA in the absence of ligand, to a ligand-dependent DNA binding protein.

The use of RNAi for gene silencing requires the production of double strand RNA molecules of ~500 bp in length. It has been found that some regions of the target transcript work better as RNAi molecules than others. How is this dsRNA prepared and what might account for the variability in efficacy across a region of mRNA?

What are the advantages and disadvantages of using transient transfection versus stable transfection strategies to study gene function?

How are most immortalized cell lines established and what characteristic phenotype of these biological reagents limits their usefulness for cell-specific studies?

What two steps are required to generate a useful Tet-regulated stable cell line? What parameter of tetracycline regulation in any given stable cell line is the most desirable? How is this optimized?