A REGULATORY SYSTEM FOR TARGET GENE EXPRESSION

Mark M. Burcin, Bert W. O’Malley and Sophia Y. Tsai

Department of Cell Biology, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA.

Received 1/17/98 Accepted 2/20/98

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Regulatory Systems for target gene expression
   3.1. ITPG based regulatory systems
   3.2. Tetracyclin based regulatory systems
   3.3. CID based regulatory system
   3.4. Ecdysone based regulatory system
   3.5. Estrogen-based regulatory system
   3.6. A novel regulatory system based on a mutated progesterone receptor ligand binding domain.
      3.6.1. Modifications of GLVP
      3.6.2. Creation of ligand inducible target genes in transgenic mice
      3.6.3. Inducible gene repression
      3.6.4. Future applications for the regulatory system
4. Acknowledgments
5. References

1. ABSTRACT

Temporally-regulated expression of endogenous genes is a desirable goal in stable cell line and transgenic animal systems, as well as in clinical gene therapy. Protocols for introducing genes into stable cell lines and experimental animals are often unsatisfactory due to the constitutive expression of such transgenes. To circumvent this problem we have demonstrated specific and temporally regulable expression of a target gene in vivo effected by a chimeric regulator in response to an orally-administered, non-toxic chemical. This regulatory system utilizes a chimeric regulator GLVP, consisting of a mutated human progesterone receptor ligand binding domain (PRLBD-delta) fused to the yeast GAL4 DNA binding domain (DBD) and the HSV VP16 transcriptional activation domain and whose activity is solely regulable by non-physiological doses of RU486 but not by progesterone or other endogenous progestins. Replacing the activation domain of the chimeric regulator with a transcriptional repression domain results in inducible repression of target gene expression in vitro. Our regulatory system functions in transient and stable transfections as well as in transgenic animals, and will have a wide variety of potential applications.

2. INTRODUCTION

The level of intracellular proteins in an eukaryotic cell is regulated at multiple levels including transcription of mRNA, mRNA stability, splicing and protein translation, modifications and stability. In this cascade, transcriptional regulation plays a key role in production of intracellular secreted proteins. The initiation of transcription is thought to be mediated by the sequential formation of the preinitiation complex. As a first step, TFIIID recognizes the TATA box followed by the binding of TFIIB and TFIIB. Subsequently, RNA polymerase and other TFII proteins enter to form the preinitiation complex thereby allowing transcription to be initiated. Binding of transcription factors to specific DNA elements influences the formation and stability of this preinitiation complex (1). Steroid/thyroid hormone receptors, the largest known family of transcription factors, have been shown to bind specific DNA sequences and activate or repress gene transcription upon binding their respective hormonal ligands. These steroid hormones such as progesterone, androgens, estrogen, glucocorticoids and mineralocorticoids, secreted by endocrine cells travel via the blood stream to their target cells and enter these cells by passive diffusion. Upon binding to hormone, steroid receptors undergo a “transformation”, bind specific DNA sequences and activate or repress gene transcription. Thus, the function of steroid receptors can be regulated by exogenously added ligands. The structure of nuclear receptors can be subdivided into different modular functional domains which are exchangeable between different receptors to form chimeric constructs (2). The modular structure of these receptors makes them readily applicable for designing novel transactivators.

Based on this knowledge we designed a chimeric regulatory system (3-5) containing three important modules that together constitute a regulable transcription factor: (I) A domain able to bind a specific DNA-sequence, (II) a domain that regulates the ability of the receptor to bind DNA upon administration of an exogenous chemical (ligand) and (III) a domain with the potential to activate or repress transcription. In order to regulate a specific target gene in vivo, a desirable chimeric regulator should contain the following features: The DNA-binding domain should not bind to DNA elements in the nucleus and alter endogenous gene expression. The ligand binding domain (LBD) must be stimulated by an easily administerable, nontoxic drug which retains biological activity at low concentrations without the interference of endogenous...
Inducible gene expression

compounds. The regulation mediated by the LBD should only occur in the presence of the ligand to guarantee a finely controlled regulation without “leaky” expression of the target gene in absence of the drug. The activation- or repressor domain must be sufficiently potent to generate high expression or tight repression of the coupled target genes.

This review will give a short summary of different regulatory systems that have been developed over the last few years and will then focus on the RU486-inducible regulatory system created by our group.

3. REGULATORY SYSTEMS FOR TARGET GENE EXPRESSION

Over the last few years, a variety of different regulatory systems have been developed (6,7). The inducing agents include heavy metal ions, heat shock, isopropyl beta-D-thiogalactoside (beta-gal), antibiotics and steroid hormones (8). Even if these systems are able to regulate gene expression, a variety of inducers are toxic to mammalian cells, can influence the expression of endogenous genes and in the absence of the inducer show high background expression of their target genes. Recently, however some approaches have been used successfully to induce regulable target gene expression in mammalian systems.

3.1. IPTG based regulatory systems

This inducible system is based on the bacterial lactose repressor protein (lacR) (9). The lacR binds in the presence of isopropyl beta-D-thiogalactoside (IPTG) to a specific DNA element, the lac operator (lacO) (10-11). To be useful for the regulation of eukaryotic genes the promoter had to be modified by inserting the lac operon upstream to a TATA box. The lac repressor protein was then modified by excision of its repressor function and fusion to the activation domain of the herpes simplex virus protein (VP16) to the IPTG- and DNA-binding site of lacR. Thus creating a lacI activator protein (LAP) that can drive the expression of lacO-bearing genes. Administration of IPTG causes the chimeric regulator to bind lacO and activate the TATA box coupled reporter gene. Mouse cell lines containing both the LAP gene and a LAP-inducible chloramphenicol acetyltransferase (CAT) reporter gene showed a 1200-fold induction of CAT activity within 24 h after addition of IPTG (11). Using this IPTG based regulatory system it is possible to achieve high induction of target genes, however, a disadvantage is that IPTG is toxic to mammalian cells, thereby limiting the scope of its application.

3.2. Tetracycllin based regulatory systems

The tetracycllin (tet)-inducible system comprises of the tet transactivator protein (tTA), a chimeric protein containing the VP16 activation domain fused to the tet repressor of Escherichia coli (12). In the absence of tetracycllin, tTA exhibits a high binding affinity to the tet resistance operator (tetO). In the presence of tet, a conformational change in the tet repressor prevents tTA from binding to tetO. By fusing multiple copies of tetO in front of a minimal promotor linked to a reporter gene, constitutive expression of tTA causes gene activation in the absence of tet. By addition of tet, only a very low expression level of the target gene can be observed. This system has been used to investigate many different stably expressed proteins under the control of the tet-regulator. One of these proteins is ICP47, expressed by the herpes simplex virus (HSV) in the cytoplasm, which has been postulated to sequester with peptide transporters (TAP) that are responsible for the translocation of antigenic peptides into the lumen of the endoplasmic reticulum (ER). By expressing ICP47 under control of the tet-regulatory system, peptide transport across the ER membrane was inhibited, indicating a potential mechanism for viral immune evasion (13). In addition the tTA system also has been used to create transgenic mice which express beta-galactosidase in different tissues. This model could be used to investigate temporal expression of transgenes in vivo. These and various other examples display the broad range of applications in which the tet regulatory system has been applied. However, a major problem of this system is that the continuous presence of the tet is required to suppress the target gene expression and that the expression of the tTA protein is toxic to cells (9).

A modified system contains a reverse transactivator (rtTA) that only binds tetO in the presence of doxycyclin, a derivative of tetracycllin. In this case administration of doxycyclin induces expression of target genes (12,14). Due to problems with constitutively expressed tTA causing toxic effects in eukaryotic cells a autoregulatory tetracycllin-responsive gene expression system was developed by placing the tetO sequence in the tTA gene promoter. In this case the absence of tet causes autoactivation of tTA and suppression of the chimeric regulator expression in the presence of tet. Transfection assays using the autoregulatory system demonstrated inducible expression of a larger fraction of cells than that achieved using a constitutive tet regulatory system.

Creation of transgenic mice haboring a luciferase transgene under control of the autoregulatory tet system showed inducible expression levels of the transgene which seem to be substantially higher than luciferase expression in transgenic mice created with the constitutive system (15). Hence this autoregulatory system is a considerable improvement over the constitutive expressed tet regulatory system.

The tetracycllin inducible system has been used in a variety of different studies such as stable mammalian cell
Inducible gene expression

3.3. CID based regulatory system

This regulatory system is based on the dimerization activity caused by “chemical inducers of dimerization” (CID)(7). The system is based on the ability of rapamycin to induce heterodimerization of two cellular proteins FKBP12 and FRAP. A DNA binding domain (ZFHD1) that binds to a novel DNA response element and is not recognized by any endogenous transcription factors was fused to FKBP12. To be able to activate target gene expression, the activation domain of p65 was added to the second partner, FRAP. Dimerization of these two chimeric proteins, induced by rapamycin administration, caused gene expression of ZFHD1 binding site coupled reporter genes in transiently and stably transfected cells. In this case cells were stably transfected with three different plasmid constructs: one for each of the two chimeric rapamycin-binding proteins and the third one using the human growth hormone (hGH) coupled to multiple ZFHD1 binding sites as a target gene. The stably transfected cells were then implanted into muscles of nude mice and hGH was detectable 17 hours after a single i.v. rapamycin injection. Subsequent expression of hGH was dose dependent (16) and gives this inducible system the potency for various applications. However, besides these advantages, rapamycin is involved in many signaling transduction pathways which will induce other possible side effects in eukaryotic cells.

3.4. Ecdysone based regulatory system

Using the insect steroid hormone ecdysone as an inducer to activate target gene expression this regulatory system consists of a truncated ecdysone receptor (EcR) attached to the VP16 activation domain (17). To bind its response element, the chimeric EcR must heterodimerise with its partner ultra spirical (USP), the insect homologue of the mammalian retinoid X receptor (RXR). To prevent the binding of the regulator to possible endogenous response elements in mammalian cells, a mutation in the DNA binding domain of the EcR was introduced that altered its binding specificity to mimic that of the glucocorticoid receptor.

To enable this modified regulator to bind DNA, a hybrid response element containing a glucocorticoid response element and a type II nuclear receptor element (like RXR) was created (17). This artificial element called E/GRE is not able to bind any receptors other than the modified ecdysone inducible regulator which, by itself, does not bind any other response elements. Using this inducible system, ecdysone dependent gene expression was demonstrated in stable cell lines and transgenic mice (17). Administration of ecdysone induced upwards of 20,000 fold increased expression of target genes, while having a very low basal level of reporter gene expression in the absence of hormone. Due to the fact that the activation potency of this regulatory system depends on interaction with RXR, pleiotropic effects in mammalian cells are very likely.

3.6. Estrogen-based regulatory system

This estrogen-regulable chimeric transcription factor consists of a yeast GAL4 DNA-binding domain (GAL) fused to the ligand binding domain of the estrogen receptor (ER) and a VP16 activation domain (VP16) (18). In the presence of estrogen the chimeric regulator dimerizes, binds to GAL4-responsive promotors and causes activation of target gene expression. In transient transfections, GAL-ER-VP16 showed a 100-fold activation of reporter genes upon induction with estrogen. However, this regulatory system has several drawbacks. First, endogenous estrogen will activate target gene expression. Second, since it is known, that ER interferes with a variety of other endogenous genes the use of this regulator in transgenic mice and human gene therapy will cause unintentional target gene activation. Third, additional administration of estrogen is likely to cause deregulation of endogenous genes and will probably result in pleiotropic effects.

3.6. A novel regulatory system based on a mutated progesterone receptor ligand binding domain

We recently described a novel regulable system which can be used to reversibly and specifically express a target gene in vivo at a selected site and at a desired time under the control of an oral, nontoxic chemical (3). This system utilizes a chimeric regulator, GLVP “figure 1”, which consists of a mutated human progesterone receptor ligand binding domain (PRLBD-delta) fused to the yeast
**Inducible gene expression**

![Figure 2](image)

**Figure 2.** DNA transfection experiments in HeLa cells using a 17x4 tk luciferase reporter as a target gene. The transfection was done in the absence (-) and in the presence (+) of RU486. **A.** Inducible reporter gene activation by the regulator GLVP. Relative luciferase units (RLU) are indicated. **B.** Directed gene repression with gene switch construct fused to a KRAB repression domain (GLKRAB). The fold repression is indicated.

GAL4 DNA binding domain (DBD) and the HSV VP16 transcriptional activation domain. In the presence of the anti-progestin mifepristone (RU486), but not any endogenous molecules present in mammalian tissues and organs, this chimeric regulator binds to a target gene containing the 17-mer GAL4 upstream activation sequence (UAS) and results in efficient ligand-inducible transactivation of the target gene “figure 1”. Most importantly the gene regulator responds to RU486 at a concentration that has no endogenous anti-progesterone or anti-glucocorticoid activity. Using this regulatory system in transient transfections we have demonstrated RU486-dependent activation of the luciferase reporter gene. In the presence of RU486 a 30-50-fold activation of the luciferase reporter gene was observed in cultured cells, whereas in the absence of RU486, no significant activation was observed “figure 2A”.

### 3.6.1. Modifications of GLVP

To evaluate the potential of the chimeric regulator GLVP, four different modifications have been tested to achieve a higher induction of target gene expression at lower ligand concentrations (4). The first modification was to replace the VP16 activation domain. Transcriptional activators can be characterized by different functional activation domains, which presumably facilitate different interactions with the basal transcriptional machinery. In the case of VP16, the amino acid content is highly acidic in nature, whereas other activation domains are glutamine-rich, proline-rich, or serine- and threonine-rich (19). Fusion of a poly-glutamine (poly-Q) or poly-proline (poly-P) stretch to the VP16 activation domain previously has been shown to result in enhanced transcriptional activation (20). The addition of 10-34 Q residues increased transactivation by the regulator, whereas inclusion of 66-132 stretches of Q residues decreased activation of target genes. A second approach to increase the activational potential of GLVP was to add multiple VP16 activation domains to the chimeric regulator. In this case, multiple VP16 transactivation domains did not enhance inducible activation potential of GLVP. Third, considerable enhancement of the potency of the regulator was achieved by increasing the length of the truncated PR-LBD. GLVP has a C-terminal truncation of the PR-LBD at amino acid 891. Increasing the length from 891 AA to 914 AA resulted in a gradual increase in RU486-dependent activation of target gene expression. Further extension resulted in decreased response to RU486.

Finally, we examined the effect of location of the VP16 activation domain within the regulator on potency of activation of target gene expression. Transferring the VP16 from the N-terminus to the C-terminus of the regulator (GL914VPc’) resulted in increased activation.

Taken together, these approaches have resulted in considerable improvement on the chimeric regulator. In particular, the revised version has a greater potential to activate target gene expression at lower doses of ligand, which results in a more efficient induction of coupled reporter genes. Second GL914VPc’ activates a target gene at a 10-fold lower concentration of RU486 (0.01 mM) (4). This allows us to use RU486 at a concentration with little anti-progestin or anti-glucocorticoid activity, an important improvement with regard to the potential use of this regulatory system in transgenic animals or in human gene therapy.

### 3.6.2. Creation of ligand inducible target genes in transgenic mice

The next step was to introduce the inducible system into transgenic animals (5). Introduction of transgenes frequently results in prenatal or postnatal death due to deleterious effects of constitutive expression of these genes. The fine control afforded by our regulatory system in terms of the transgene inducibility makes it a powerful system for transgenic research in mice. In this system, transcription of the target transgene is kept silent until turned on by the administration of the exogenous compound RU486. In order to obtain a tissue-specific expression of the transgene we incorporated the enhancer/promoter region of the trans-thyretin (TTR) gene which has been shown to target transgene expression specifically to the liver (21-22). Towards this goal, two mouse lines were created; one with liver-specific expression of the transgene harboring both transgenes (hereafter referred to as TTR-GLVP/hGH).

hGH expression is minimal in these bigenic mice in the absence of RU486, demonstrating that activation of this inducible binary system is independent of endogenous hormones. However, after a single intraperitoneal injection of RU486, strong induction (5800- to 33,000-fold) of hGH expression was detected in the serum. Over time the serum
Inducible gene expression

level of hGH returned back to its baseline after RU486 was metabolized by the animal, and expression of the hGH transgene could be reinduced with another dose of RU486. These results demonstrate that the target gene can be turned on and off repeatedly by pulsed administration of an exogenous compound. This RU486-inducible binary regulatory system represented the first study of inducible and dose-dependent transgene activation in vivo (30).

3.6.3. Inducible gene repression

In contrast to regulatory gene activation, another possible use for the inducible system would be the regulable repression of target gene expression. For this reason we replaced the VP16 activation domain of the chimeric regulator with a heterologous repression domain. We chose the krüppel-associated box (KRAB), a potent repressor domain identified in the kidney specific transcription factor Kid-1 (23-25). By replacing VP16 with the KRAB repression domain, we achieved RU486-dependent regulable repression of target genes using the reporter plasmids 17x4-tk-LUC (containing the thymidine kinase promoter) and 17x4-SV-CAT (containing the SV40 enhancer). Using these regulatory elements in transient transfection assays, we were able to downregulate both basal promoter “figure 2B” as well as enhancer-driven gene expression (4).

Our novel negative regulatory system could conceivably be used to downregulate genes involved in apoptosis, an underlying process in a variety of neurodegenerative diseases (26-28). Another desirable goal would be the repression of oncogenes responsible for tumorigenesis. By exchanging the GAL4 DNA binding domain for the DNA binding domain of another protein which participates in the regulation of a gene involved in tumorigenesis, RU486-dependent repression of this target gene may be possible. This strategy also could be used to achieve ligand-inducible gene knockouts in transgenic animals. Since a variety of gene knockouts in transgenic mice are lethal in early embryonic stages, it would be helpful to selectively repress the function of a gene at different developmental stages. By introducing a GAL4-DNA binding element into the promoter region of a given gene, RU486-inducible binding of the negative regulator would cause down regulation or even ablation of transcription enabling one to study the physiological consequences of the decreased gene product levels.

3.6.4. Future applications for the regulatory system

Since human gene therapy is a potential future application for the regulatory system, the replacement of the viral VP16 activation domain is essential for its success in this regard. This is particularly necessary, because it is assumed that the VP16 will evoke an immune response in the human body and prolonged high expression levels of the VP16 activation domain may induce an inappropriate immune response in mammalian tissues. For this reason, the replacement of the viral activation domain with a less immunogenic human activation domain is desirable. Recently discovered transcriptional coactivators such as SRC-1 (29) and CBP (30), factors that coactivate transcription by interacting with nuclear hormone receptors, harbor activation domains that could be used to replace VP16. The mechanism of this coactivation may involve the stabilization and the modification of the general transcription complex, hence resulting into more efficient transcription. The use of the independent activation domains of the human coactivators in the context of the regulatory system should lead to the stabilization of the transcription initiation complex and to transcriptional activation of a linked gene. Another possible source of a potent activation domain is derived from p65 (31), a component of the NF kappa B complex, which is known to possess a strong transactivation potential. Using the p65 activation domain in our regulatory system we achieved high inducibility of target genes. The use of this human activation domain to replace the viral VP16 of the regulator could circumvent the possible problems for the use in human gene therapy due to immunogenic responses.

Further potential problems, however, are still possible in the design of a non-immunogenic system, due to the presence of the yeast GAL4-DNA binding domain, or because the chimeric receptor itself, with its novel combination of functional domains, might generate an immune response.

As previously discussed an important potential use for our system could be the creation of transgenic mice to generate disease models of tumor induction by regulable expression of oncogenes. Oncogenes such as int-2, fgf-3 (32) or oncogene products such as large T antigen (33) known to be involved in tumor development could be linked to a minimal promoter containing a GAL4-DNA binding site and then used to generate transgenic animals. Using tissue specific promoters to express the regulator would restrict oncogene expression to selected tissues like the lung, mammary gland or pancreas. By crossing mice containing a tissue specific regulator with mice haboring an oncogene as a target for the regulator, RU486 inducible expression of the oncogene in a specific tissue could cause tumor formation. Such models in which tumor formation can be induced at a specific tissue and time point, would be very important for the investigation of the initiation and progression of tumor formation thereby allowing the development of potential therapeutic approaches to oncogenesis.

Using the regulatory system in stable cell lines or transgenic animals carries with it the problem of random nature of gene integration. The expression of an integrated gene is contingent upon: (A), the influence of proximal positive regulatory elements in terms of constitutive activity of the gene, (B), the influence of a negative locally repressive chromatin milieu with resultant suboptimal expression of the target gene. These problems could be offset by linking target genes to matrix attachment regions (34) or boundary elements like insulators (35), to insulate integrated transgenes from the effects of neighboring regulatory elements.
Inducible gene expression

Another problem using our regulatory system in different cell lines and tissues is the heterogeneity of expression of transgenes found in different tissues. Therefore it will be important to use cell specific promoters for transactivator expression in order to guarantee successful regulator function in a specific tissue.

In order to facilitate initial delivery of the inducible system in vivo, adenoviral vector-mediated gene delivery is one possible approach. Previously, such a means of delivery had inherent problems: (i) expression of viral proteins in infected cells is believed to trigger a cellular immune response that precludes long-term expression of the transferred gene; and (ii) the insert capacity of adenoviral vectors has been previously limited to 8 kb of transgenic sequence. Recently an adenoviral vector has been constructed (36) which contains no viral coding sequences and possesses a very large insert capacity (28 kb). Using this adenoviral vector in combination with our inducible regulatory system, could provide a regulable gene delivery system in vivo for use in human gene therapy.

Another application for the regulatory system would be the use of the Cre-loxP recombination system (37-39) to create conditional knockout mice at different developmental stages in order to overcome the problem of knockouts being lethal in early embryonic stages. In this case, two transgenic mice strains have to be generated. The first one will carry two constructs: A. Our regulator under the control of a tissue specific promoter and B. a cre transgene linked to a tk-promoter and a GAL4 DNA binding site. The second mouse strain carries the target gene flanked by two loxP sites. After crossing these mice, administration of RU486 will result in tissue specific expression of Cre, leading to a Cre-loxP site-dependent recombination, culminating in tissue-specific deletion of the target gene. Such conditional gene targeting would permit the knockout of a transgene in a specific tissue at a desired developmental stage.

The regulatory system described in this review shows the tremendous progress that has been made in its development over the past few years. It represents a powerful tool capable of addressing a variety of scientific questions. In the coming years, we envisage a great variety of applications of this system to study the effect that expression of specific proteins might have on cell differentiation and development. Since RU486 can be widely distributed in an organism and even penetrates the blood/brain barrier, this regulatory system is applicable even to the study of the central nervous system. By combining the RU486-inducible regulator with novel vectors, it is likely that this system will contribute to future approaches in clinical gene therapy.

4. ACKNOWLEDGMENTS

We thank M.-J. Tsai for helpful discussion, and N. McKenna, S. Chua and V. Krishnan for critical reading of the manuscript. This work was partially supported by a Fritz Thyssen Foundation- and DFG- Postdoctoral fellowship from Germany (M.M.B); NIH grants to S.Y.T. and B.W.

5. REFERENCES


15. Shockett P., V. Difilippantonio, N. Hellman & D.G. Schatz: A modified tetracycline-regulated system provides autoregulatory, inducible gene expression in...
Inducible gene expression


**Key words:** Inducible system, Gene therapy, Transgenic mouse, Gene activation, Repression, RU486

Send correspondence to: Sophia Y. Tsai, Department of Cell Biology, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA. Tel: (713)-798-6251, Fax: (713)-798-8227, Email: stsai@bcm.tmc.edu