

ARGENT™
Regulated Transcription Plasmid Kit

Version 2.0

ARIAD Pharmaceuticals, Inc.
26 Landsdowne St
Cambridge, MA, 02139

RegulationKits@ariad.com
Tel: (617) 494-0400

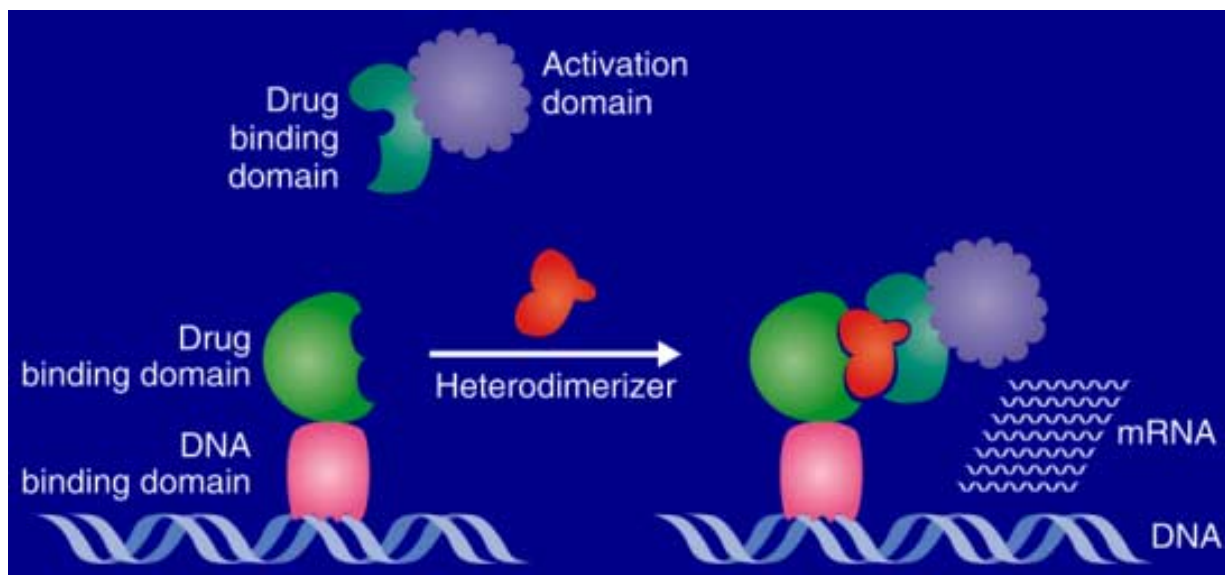
09/09/02

Table of Contents

ARGENT™ Regulated Transcription Plasmid Kit	1
Table of Contents	2
Overview	3
Applications of the ARGENT Regulated Transcription Plasmid Kit.....	4
Design of the kit components	5
Notes on the use of this kit	6
Kit contents	6
Transcription factor plasmids	7
pC ₄ N ₂ -R _H S/ZF3.....	7
pC ₄ N ₂ -R _H S3H/ZF3	8
Target gene plasmids	9
pZ ₁₂ l-PL-2	9
pZ ₁₂ l-hGH-2.....	9
General vector information.....	10
Transient transfection protocol	10
Addition of a selectable marker to transcription factor vectors	10
Procedure for making stable cell lines.....	10
Antibodies to detect fusion proteins.....	11
Viral vectors.....	11
hGH assays.....	11
Additional pC ₄ expression vector information	11
AP21967	12
Description.....	12
Reconstituting AP21967.....	12
Storage and handling of AP21967	12
Using AP21967 <i>in vitro</i>	12
Use of AP21967 in animals.....	13
Expected results	13
Conditions of use	13
References	14
Appendix.....	15
pC ₄ N ₂ -R _H S/ZF3 Annotated Sequence.....	16
pC ₄ N ₂ -R _H S3H/ZF3 Annotated Sequence	18
pZ ₁₂ l-PL-2 Annotated Sequence	20
pZ ₁₂ l-hGH-2 Annotated Sequence.....	20

ARGENT™ Regulated Transcription Plasmid Kit

ARGENT™ Regulated Transcription Kits contain reagents for placing the transcription of a target gene under the control of a small molecule “dimerizer”. The kits can be used to achieve tightly regulated conditional expression of genes of interest, allowing gene function to be investigated *in vitro* or *in vivo*. In the Plasmid Kit described here, the regulatory system has been incorporated into plasmid vectors. A [retroviral vector-based](#) version is also available.



[Controlling Gene Expression Using Regulated Transcription](#)

Overview

Activation of gene expression in eukaryotes is controlled by the induced binding of transcription factor proteins to target genes. Transcription factors are bifunctional proteins that recognize specific DNA sequences near target genes and then recruit the transcriptional machinery of the cell to activate transcription. The two domains responsible for these activities, the DNA binding domain and the transcriptional activation domain, are functionally separable and can reconstitute a sequence-specific transcriptional activator even when expressed as individual proteins and brought together via a noncovalent interaction.

This modular architecture has been exploited to develop a general method for controlling gene transcription using small molecules. The strategy is based on the use of chemical inducers of dimerization, or “dimerizers”, to induce the interaction of engineered proteins (1). A dimerizer is a cell-permeant organic molecule with two separate motifs that each bind with high affinity to a specific protein module. By fusing such modules to a DNA binding domain and an activation domain, the reconstitution of a functional transcription factor, and therefore the expression of a target gene, can be made absolutely dependent on the presence of dimerizer (see the figure).

In principle, the expression of any cloned gene can be brought under dimerizer control, by equipping the gene with upstream sequences that are recognized by the engineered DNA binding domain. Following introduction of the modified gene into cells that also express the engineered transcription factor proteins, addition of dimerizer will lead to dose-dependent activation of target gene expression.

Because the transcription factor fusion proteins have no affinity for one another in the absence of dimerizer, regulation is characterized by an extremely low, usually undetectable, basal level of gene expression. In addition, the highly potent activation domains incorporated into the system typically lead to high maximal levels of induced gene expression, often in excess of levels obtained with strong constitutive promoters/enhancers.

There are two classes of dimerizers. **Homodimerizers** incorporate two identical binding motifs, whereas **heterodimerizers** have two different binding motifs, allowing the specific dimerization of two different proteins when they are fused to two appropriate ligand binding domains. Regulated transcription is usually accomplished using a heterodimerizer, as shown in the figure, since this leads to the most efficient reconstitution of an active transcription factor. The ARGENT™ Regulated Transcription Plasmid Kit described here provides a heterodimerizer, vectors encoding the engineered transcription factor fusion proteins, and target gene vectors into which genes of interest can be inserted. The vectors are provided as plasmids— a separate kit is available in which these constructs are instead provided as retroviral vectors (see [ARGENT Regulated Transcription Retrovirus Kit](#)).

Applications of the ARGENT Regulated Transcription Plasmid Kit

The ability to control the transcription of specific genes using small molecules has broad utility in biological research. Varying the expression level of a gene is a powerful way to study its function, allowing the creation of inducible alleles in cell culture and in transgenic animals. By precisely varying the expression level using the dose of dimerizer, detailed questions can be asked about the physiological role of the gene, and the protein it encodes.

The plasmid reagents provided in this kit have been used to achieve regulated gene expression in a variety of contexts. The reagents have been used to transiently or stably transfect many different cell types *in vitro* (2-5). The kit components are also suitable for *in vivo* studies: the plasmid vectors have been used as starting points to create adenovirus and adeno-associated virus (AAV) vectors for delivering the system to mice and primates (6, 7), and to create constructs for generating transgenic mice expressing the engineered transcription factors and target genes. In addition, the dimerizer provided with the kit, AP21967, is suitable for *in vivo* use and has been used successfully to achieve regulated gene expression in mice. Under all conditions tested, the system has allowed tight, dose-dependent control of gene expression. A complete list of publications describing the use of ARGENT regulated transcription reagents can be found in the [Regulation Kit Bibliography](#).

A key feature of regulation of transcription using the dimerizer system is the extremely low level of basal expression in the absence of inducer. This is critical for applications where even minimal “leakiness” (expression in the absence of inducer) is unacceptable. We recently demonstrated the isolation of stable cell lines inducibly expressing the highly toxic diphtheria toxin A chain gene (5), suggesting that the dimerizer system is particularly suitable for analysis of the many interesting genes that promote cell death, block the cell cycle, or are otherwise toxic.

Tight, ligand-inducible control of gene expression also has many applications in functional genomics research and drug discovery. A particularly promising application is the generation of conditional gene knockouts in mice through dimerizer-inducible expression of recombinases. Inducible animal models can be established of disease states associated with overexpression of a particular gene (such as an oncogene). In addition, cell lines in which expression of a single gene can be chemically induced may be useful in the configuration of targeted cell-based assays for small molecule drugs.

Design of the kit components

Rapamycin and its analogs

The reagents in the ARGENT™ Regulated Transcription Plasmid Kit, like those of the other ARGENT kits, are based on the human protein FKBP12 (FKBP, for FK506 binding protein) and its small molecule ligands. FKBP is an abundant cytosolic protein that serves as the initial intracellular receptor for the natural product immunosuppressive drugs FK506 and rapamycin. Both these drugs act naturally as heterodimerizers, and both have been used to control transcription (2, 3), as has FK-CsA, a cyclosporin-FK506 hybrid molecule (8). We have focused on the use of rapamycin, because it has well understood chemistry and has favorable pharmacokinetic properties. Rapamycin functions by binding with high affinity to FKBP, and then to the large PI3K homolog FRAP (RAFT, mTOR), thereby acting as a heterodimerizer to join the two proteins together (9). To control transcription of a target gene, a DNA binding domain is fused to one or more FKBP domains, and a transcriptional activation domain is fused to a 93 amino acid portion of FRAP, termed FRB, that is sufficient for binding the FKBP-rapamycin complex (10). Only in the presence of rapamycin are the two fusion proteins dimerized and therefore capable of activating the transcription of a gene equipped with binding sites for the DBD (2).

In some cases, the use of rapamycin may be compromised by its cell cycle inhibitory effects (the result of inhibiting FRAP kinase activity, which in T cells leads to immunosuppression). To overcome this limitation, we have engineered the system to function with non-immunosuppressive analogs of rapamycin, which we call rapalogs. These compounds have been chemically modified so that they no longer can bind to wild-type endogenous FRAP, greatly reducing immunosuppressive activity. The compounds can however bind to a modified FRAP that contains a single designed amino acid change (T2098L). Incorporation of this mutation into the FRB domain fused to the activation domain allows a rapalog to be used to specifically heterodimerize the engineered transcription factor fusion proteins without interfering with the activity of endogenous FRAP.

The redesigned rapamycin system forms the basis of this kit, which includes the mutant FRB sequence, and a non-immunosuppressive rapalog, AP21967. It is important to note that the redesigned system retains the ability to respond to rapamycin itself, as well as to AP21967. Therefore experiments can be carried out with either dimerizer, as appropriate.

Transcription factor components

Since we are developing regulatory systems for use in human gene therapy, we have built our transcription system using only human proteins to minimize the potential for immunogenicity in clinical applications. The DNA binding domain we use is called ZFHD1, a composite human DBD with novel DNA recognition specificity (11). ZFHD1 is composed of two zinc finger domains from the human transcription factor Zif268, joined to a homeodomain derived from the human transcription factor Oct-1. ZFHD1 binds with high affinity and specificity to a unique composite DNA binding sequence, but not to Zif268 or Oct-1 binding sites. Typically, multiple copies of the ZFHD1 binding site are included upstream of target genes to obtain robust gene activation.

The activation domain used in the first version of our system consisted of the carboxy terminal 191 amino acids from the p65 subunit of human NF- κ B (12). In our system, this p65-derived domain substantially outperforms the commonly used activation domain from the herpesvirus VP16 protein. Since the level of activation of a target gene is directly proportional to the potency of the transcriptional activation domain, we have invested significant effort in trying to identify even more potent domains. We recently described a new activation domain consisting of the carboxy terminal 271 amino acids of p65 fused to the activation domain from human Heat Shock Factor 1 (5). Use of this stronger composite activation domain, which we call S3H, typically leads to significantly higher levels of gene activation, especially under conditions in which the transcription factors are expressed at low levels. Both activation domains are provided in this kit.

Notes on the use of this kit

Use of the previous kit based on AP1510

This kit, along with the companion retrovirus-based kit, replaces the original Regulated Transcription kit that was based on the homodimerizer AP1510 and FKBP fusion proteins (13). We have found that the rapamycin-based reagents significantly out-perform the original reagents in all applications tested. In particular, the pharmacological properties of AP1510 preclude its use for *in vivo* studies, whereas rapamycin and AP21967 are well suited to these applications. If you are already using the AP1510-based kit, we will continue to provide dimerizer for *in vitro* experiments upon request. Please note that rapamycin and AP21967 will **not** activate transcription from constructs generated using the previous kit.

Regulating heterodimerization of proteins other than transcription factors

Regulated heterodimerization is generally applicable to the study of signaling pathway components and other proteins that function through protein-protein interactions. However the kit described here has been designed specifically for use in regulating target genes, and the genes encoding the fusion proteins cannot readily be reconfigured for other uses. For applications other than regulated transcription of target genes, please request the [ARGENT Regulated Heterodimerization Kit](#).

Rapamycin analog AP22565

Please note that the rapalog provided in this kit, AP21967, is different from AP22565, the analog used in our recent publication (5), although the two molecules are from the same chemical class and are highly related. AP21967 can be used equivalently in all the applications described in the paper.

Kit contents

The ARGENT™ Regulated Transcription Plasmid Kit contains plasmid-based reagents that can be used to make transcription of a target gene dependent on the presence of a rapalog. The kit includes

- Two versions of the transcription factor expression vector
- An empty target gene vector, for inserting the gene of interest
- A control target gene vector
- An aliquot of the rapalog AP21967

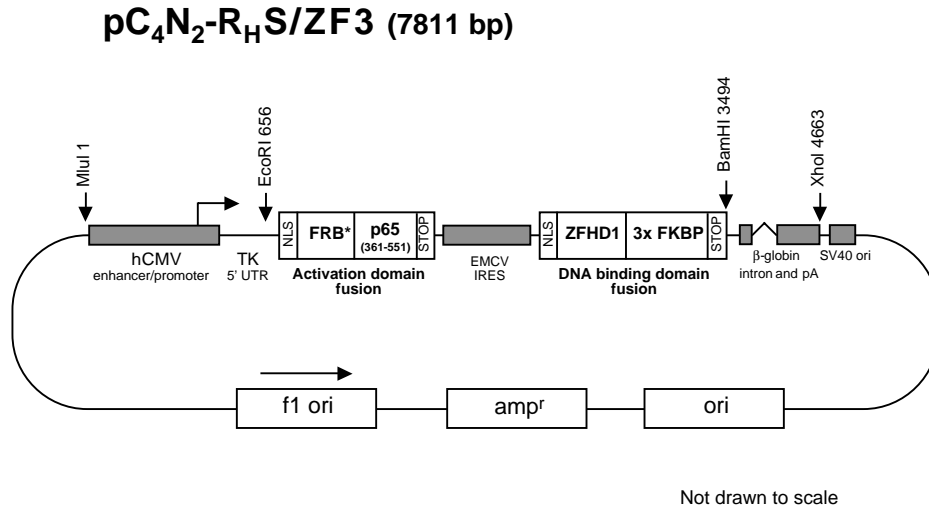
The transcription factor vectors express both the activation domain and the DNA-binding domain fusions from a single bicistronic transcript. In these vectors the activation domain is fused to the mutant FRB domain and the DNA binding domain is fused to 3 copies of FKBP. The two transcription factor vectors differ in the activation domains they carry (see “Design of the kit components”, above). pC₄N₂-R_HS/ZF3 contains the original p65 activation domain (called “S”), while pC₄N₂-R_HS3H/ZF3 contains the more potent composite p65-HSF1 domain, called S3H. Both vectors use ZFHD1 as the DNA binding domain.

An empty target vector plasmid, pZ₁₂I-PL-2, is provided which contains a minimal promoter carrying binding sites for the DBD followed by a polylinker into which the gene of interest can be cloned. A control vector, pZ₁₂I-hGH-2, is also provided in which the expression of human growth hormone (hGH) is placed under rapalog control.

Transcription factor plasmids

pC₄N₂-R_HS/ZF3

Description



- The transcription factor vector pC₄N₂-R_HS/ZF3 contains a CMV enhancer/promoter (C) driving expression of the following fusion proteins from a bicistronic transcript:
 - an activation domain fusion (R_HS) which consists of the FRB fragment of human FRAP (R_H), fused to an activation domain derived from the p65 subunit of human NFκB (S; amino acids 361 to 551). The FRB domain consists of amino acids 2021-2113 of FRAP, in which the threonine at amino acid 2098 is mutated to leucine. This mutation allows the protein to bind to rapamycin analogs (e.g. AP21967) which no longer bind appreciably to endogenous FRAP.
 - a DNA binding domain fusion (ZF3) which consists of the ZFHD1 DNA binding domain (Z) and three tandemly repeated copies of human FKBP12 (F3).
- Both fusion proteins contain an amino-terminal nuclear localization signal (N₂, from the human *c-myc* gene).
- The two coding regions are separated by an internal ribosome entry sequence (IRES) derived from the encephalomyocarditis virus to allow translation of the second cistron (ZF3).
- We have found the optimal configuration of the transcription factor fusions to be that in which ZFHD1 is fused to three tandemly-reiterated FKBP12 domains, and the p65 activation domain to a single FRB domain. This configuration theoretically allows recruitment of up to 3 activation domains per DNA binding site.

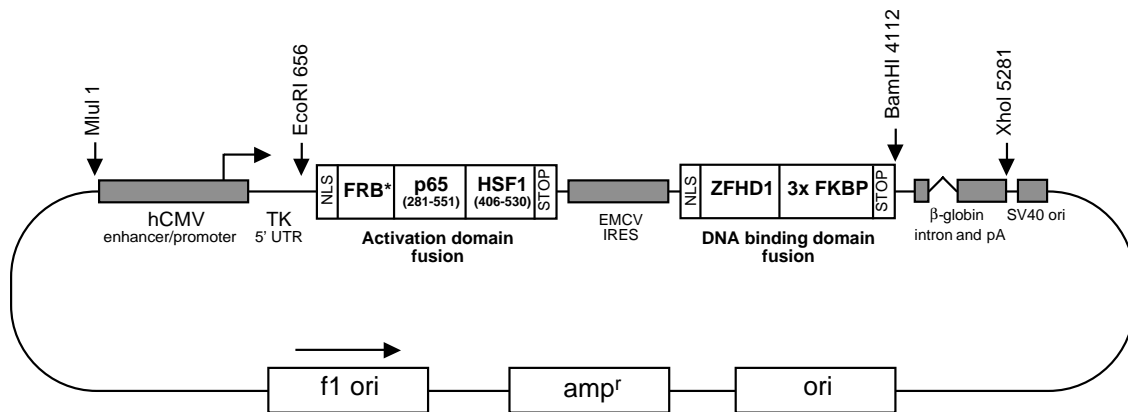
[Annotated Sequence](#)

[Raw sequence](#)

pC₄N₂-R_HS3H/ZF3

Description

pC₄N₂-R_HS3H/ZF3 (8429 bp)



Not drawn to scale

- The vector pC₄N₂-R_HS3H/ZF3 is identical to pC₄N₂-R_HS/ZF3 except that the activation domain present in pC₄N₂-R_HS/ZF3 (S; amino acids 361 to 551 of the p65 subunit of human NFκB) has been replaced by a more potent version, S3H, which consists of a larger portion of the activation domain from p65 (S3; amino acids 281 to 551) fused to the activation domain from heat shock factor 1 (H, amino acids 406-530).

[Annotated Sequence](#)

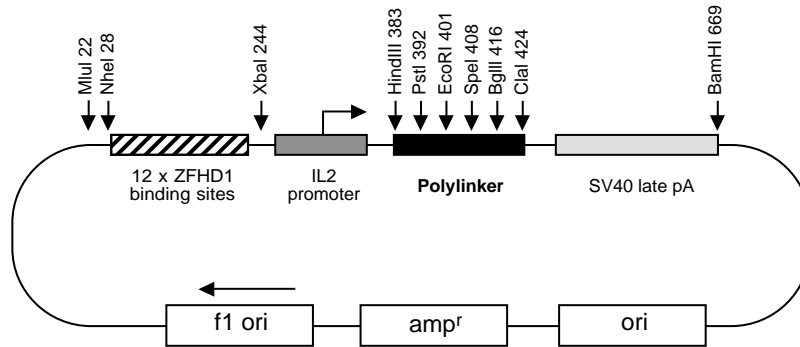
[Raw sequence](#)

Target gene plasmids

pZ₁₂I-PL-2

Description

pZ₁₂I-PL-2 (3528 bp)



Not drawn to scale

- pZ₁₂I-PL-2 contains 12 ZFHD1 binding sites and a minimal human interleukin-2 gene promoter (Z₁₂I), upstream of a polylinker (PL) followed by a 3' UTR containing a polyadenylation signal from the SV40 late gene (-2). Insertion of the gene of interest into the polylinker puts its expression under control of the dimerizer-regulated transcription factors.
- Alternatively, the Z₁₂I control region can be removed from this vector (using 5' MluI or NheI sites and 3' HindIII, PstI, EcoRI, SpeI, BglII or ClaI sites) and inserted upstream of the gene of interest located in a different vector.
- This vector has very low basal expression (2). In our experiments, the use of the minimal IL-2 promoter has been found to be important as the substitution of a minimal SV40 or CMV promoter results in significantly higher levels of basal expression.

[Annotated Sequence](#)

[Raw sequence](#)

pZ₁₂I-hGH-2

Description

- This control vector contains the human growth hormone cDNA inserted into pZ₁₂I-PL-2 as a 759 bp HindIII-EcoRI fragment.

[Annotated Sequence](#)

[Raw sequence](#)

General vector information

Transient transfection protocol

To introduce the components necessary to regulate expression of a target gene by transient transfection of plasmids, as a starting point, we recommend that transcription factor and target gene plasmids be introduced in a 2:1 ratio. Note that if the expression level of the transcription factors is too high, then transcription of a transiently transfected target gene can be reduced or “squashed” (14). Therefore it may be necessary to further optimize plasmid ratios for a given situation and to limit the amount of transcription factor plasmid introduced by adding carrier DNA.

Following transfection, replace the transfection mix with medium \pm 50 nM AP21967 (or try a range of concentrations).

After overnight incubation (or longer), assay for target gene expression.

Addition of a selectable marker to transcription factor vectors

Stable integration of the transcription factor vectors into cells is greatly facilitated by co-expression of the transcription factors and the selectable marker gene on the same mRNA transcript. Such multicistronic mRNAs can be created by expressing the transcription factors downstream of an enhancer/promoter and upstream of an internal ribosome entry sequence (IRES) which drives expression of a selectable marker gene.

Using “pIRES-” vectors available from Clontech (e.g. pIRESneo2, pIRESHyg2, pIRESpuro2, pIRESbleo), two alternative approaches can be used to generate such a vector.

- The BamHI-XhoI fragment from the transcription factor vectors can be replaced by a BamHI-XhoI fragment from a pIRES vector. This replaces the intron and polyA sequence from the vector in this kit with an intron, IRES element, selectable marker gene and a polyA sequence.
- Alternatively, an EcoRI-BamHI fragment containing the transcription factor cassette can be cloned into the MCS of a pIRES vector.

The advantage of using this approach to introduce a selectable marker is that essentially all drug resistant cells will express the transcription factors since any mRNA that expresses the selectable marker as its third cistron should also express the transcription factors since they are the first two cistrons in the mRNA.

Procedure for making stable cell lines

We have found that when making stable cell lines (a situation in which squashing generally does not apply - 14) the most important consideration is that the level of induction of target gene expression correlates with the level of expression of the activation domain fusion. It is therefore critical that the enhancer that drives expression of the transcription factor fusion be optimal for the cells being transduced. The CMV enhancer is generally suitable for this purpose, but in certain cell types alternate enhancers might be more appropriate.

To stably introduce the transcription factor vector into cells we recommend first cloning an IRES-selectable marker cassette downstream of the transcription factor fusions as described above. This will result in the creation of a tricistronic mRNA in which the last cistron encodes the selectable marker. In this configuration essentially all drug resistant cells also express both the activation and DNA binding domain fusions. It is also generally true that increasing the stringency of selection selects for clones that express higher levels of the transcription factor fusions.

The target gene vector provided in this kit does not contain a selectable marker and should therefore either be co-transfected with a selectable marker plasmid (not provided) or subcloned into a plasmid that contains a separate selectable marker (e.g. pTK-Hyg, Clontech).

To generate stable cell lines containing both the transcription factor and target gene plasmids we recommend transfecting them into cells sequentially. For example,

1. Stably integrate the transcription factor vector plasmid.
2. Screen individual clones by transiently transfecting the target gene of interest or an easily assayed target gene (e.g. pZ₁₂-hGH-2).
3. Select the clone with lowest background and highest AP21967-dependent induction.
4. Stably integrate the target gene plasmid by co-transfection with a selectable marker plasmid.
5. Screen individual clones for the lowest background and highest levels of AP21967-dependent target gene expression.

Antibodies to detect fusion proteins

Anti-p65 antibodies (Santa Cruz Biotechnology #sc-372) can be used to detect R_HS (~48 kDa) and R_HS3H (~72 kDa) activation domain fusion proteins.

Viral vectors

We have successfully incorporated the system into AAV (7), adenoviral (6) and retroviral vectors (5; see ARGENT™ Regulated Transcription Retrovirus Kit). To do so, simply clone the appropriate portion of the transcription factor and target gene vectors into an appropriate viral vector.

hGH assays

Several kits are available for the quantitation of hGH levels including an RIA from Nichols Diagnostic (# 40-2155) and an ELISA from Roche (# 1 585 878).

Additional pC₄ expression vector information

Origin of vector

pC₄ expression vectors are derived from the vector pCGNN (15). To create pC₄ several existing restriction sites were eliminated and several others added in order to have all functional regions of the plasmid be flanked by unique restriction sites (i.e. MluI, EcoRI, XbaI, SpeI, BamHI and XhoI).

Configuration of vectors

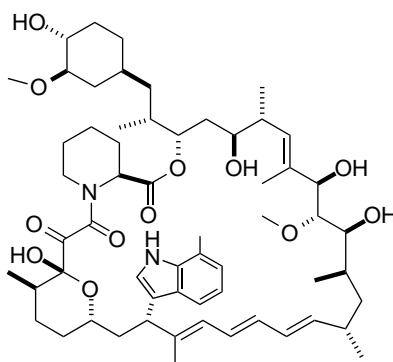
- **MluI-EcoRI:** Contains the enhancer/promoter from human CMV and the 5' UTR from the Herpes simplex virus TK gene. This fragment can be replaced with an alternate enhancer/promoter of choice (i.e. to promote tissue-specific expression)
- **EcoRI-XbaI-SpeI-BamHI:** Contains the coding region broken down as follows:
 - *EcoRI-XbaI:* leader sequence (e.g. a nuclear localization sequence)
 - *XbaI-SpeI:* coding sequence for regulated transcription factors
 - *SpeI-BamHI:* Carboxy-terminal sequence (e.g. stop codon).
- **BamHI to XhoI:** contains an 830-bp portion of the 3'UTR of the rabbit β-globin gene that includes its final intron and poly(A) signal.

Production of single stranded DNA for mutagenesis

All pC₄ plasmids contain f1 origins for rescue of single stranded DNA. The strand generated upon rescue is indicated by the arrow in plasmid maps. For example, in pC₄N₂-R_HS/ZF3 the antisense strand is generated, therefore oligonucleotides used for mutagenesis should correspond to the sense stand of the vector (the strand shown in the vector sequences).

AP21967

Description



AP21967

AP21967 is a chemically modified derivative of rapamycin that can be used to induce heterodimerization of FKBP and FRB_{T2098L}-containing fusion proteins. AP21967 is greater than 1000-fold less immunosuppressive than rapamycin as measured in an *in vitro* splenocyte proliferation assay. In all studies to date, AP21967 is non-toxic to cells at up to 1 μ M concentrations, or mice at up to 30 mg/kg doses.

To date, AP21967 has only been tested *in vitro* and in mice. We do not yet know whether it crosses the blood-brain barrier in mice or whether it works in yeast or any other model organisms.

AP21967 cannot be used to heterodimerize proteins containing a wild type FRB domain. If you have already made constructs using the wild type FRB domain, you must use rapamycin as the heterodimerizer.

Note, however, that the presence of the T2098L mutation in FRB has little or no detrimental effect on the binding of rapamycin. Therefore, as noted earlier, rapamycin can also be used to dimerize fusion proteins made using the reagents in this kit. Rapamycin is available commercially from Sigma (cat # R0395) or Affinity BioReagents (cat # IR-022).

Reconstituting AP21967

AP21967 (molecular mass 1017.4 Da) is provided in lyophilized form which should be reconstituted as a concentrated stock in an organic solvent. We recommend dissolving the lyophilized material in absolute ethanol to make a 1 mM solution (e.g. dissolve 250 μ g AP21967 in 246 μ l ethanol). After adding the appropriate volume of ice-cold ethanol, seal and vortex periodically over a period of a few minutes to dissolve the compound. Keep on ice during dissolution to minimize evaporation.

Storage and handling of AP21967

Once dissolved, the stock solution can be kept at -20°C indefinitely, in a glass vial or a microfuge tube. Further dilutions in ethanol can be similarly stored. At the bench, solutions in ethanol should always be kept on ice, and opened for as short a time as possible, to prevent evaporation and consequent changes in concentration.

Using AP21967 *in vitro*

Working concentrations of AP21967 can be obtained by adding compound directly from ethanol stocks, or by diluting serially in culture medium just before use. In the latter case we recommend that the highest concentration does not exceed 5 μ M, to ensure complete solubility in the (aqueous)

medium. In either case, the final concentration of ethanol in the medium added to mammalian cells should be kept below 0.5% (a 200-fold dilution of a 100% ethanol solution) to prevent detrimental effects of the solvent on the cells.

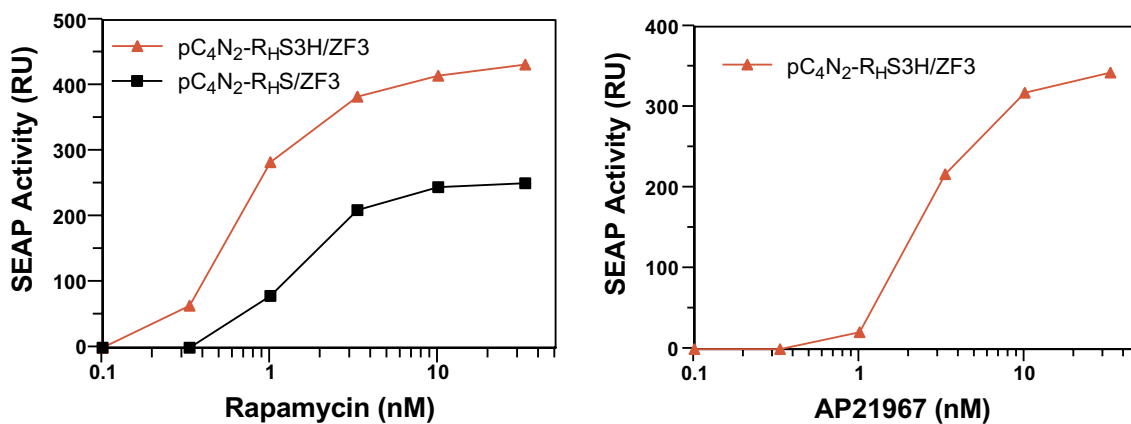
Use of AP21967 in animals

Once preliminary *in vitro* experiments have been carried out successfully we will be happy to provide quantities of AP21967 necessary for use in animals.

Expected results

The figures below show the effects upon reporter gene expression of adding increasing amounts of rapamycin or AP21967 to HT1080 cells stably transfected with a secreted alkaline phosphatase (SEAP) target gene and the pC₄N₂-R_HS/ZF3 or pC₄N₂-R_HS3H/ZF3 transcription factor vectors.

In the absence of dimerizer, target gene expression is undetectable. In the cell line transfected with pC₄N₂-R_HS/ZF3 half-maximal induction occurs at 2 nM rapamycin or 6 nM AP21967 (data not shown). In the cells transfected with pC₄N₂-R_HS3H/ZF3 half-maximal induction occurs at approximately 3-fold lower doses of rapamycin or AP21967 and the absolute levels of peak reporter gene expression are approximately two-fold higher. The difference in potency between these two vectors is even greater when the expression level of the transcription factors is lower, as occurs in cells in which the CMV enhancer is not as active as it is in HT1080 cells. In initial experiments we recommend that AP21967 be tested across a broad range of concentrations (e.g. .01 to 1000 nM) to provide a complete dose-response profile.



Conditions of use

Please bear in mind that these materials will be provided to you pursuant to a Material Transfer Agreement (MTA). Our MTA contains, among other provisions, certain restrictions on the transfer to others of our materials and any derivatives you create using or incorporating our materials. If you wish to share the materials or derivatives with colleagues or collaborators, they must first complete our MTA. Please also be aware that our Kits are not to be used in research funded by, or conducted on behalf of, a commercial or for-profit entity. Those situations require a [commercial agreement](#).

We certainly hope that you obtain interesting results and that they are presented and published without delay. But please note that under the terms of the MTA, you need to give us advance notice of any such presentations or publications, including talks, posters, and submissions of abstracts or manuscripts for publication. Also, in the event of a patent filing, a copy of the patent application must be provided to ARIAD. Advance notice is usually 4 weeks prior to submission, but please check your MTA for specific details.

Please also be aware that the use of intellectual property or materials of others, in conjunction with the Regulation Kit, may have additional ramifications. For example, if you plan to use a Regulation Kit together with human embryonic stem cells from WiCell (WARF), we and you are required to execute an additional MTA which will be provided to you.

We appreciate your cooperation in this regard.

References

References cited here are listed below. A complete list of articles that have used dimerizers to regulate transcription can be found in the [Regulation Kits Bibliography](#).

1. Spencer, D. M., Wandless, T. J., Schreiber, S. L. & Crabtree, G. R. (1993) [Controlling signal transduction with synthetic ligands](#). Science 262: 1019-24.
2. Rivera, V. M., Clackson, T., Natesan, S., Pollock, R., Amara, J. F., Keenan, T., Magari, S. R., Phillips, T., Courage, N. L., Cerasoli, F., Jr., Holt, D. A. & Gilman, M. (1996) [A humanized system for pharmacologic control of gene expression](#). Nat Med 2: 1028-32.
3. Ho, S. N., Biggar, S. R., Spencer, D. M., Schreiber, S. L. & Crabtree, G. R. (1996) [Dimeric ligands define a role for transcriptional activation domains in reinitiation](#). Nature 382: 822-6.
4. Liberles, S. D., Diver, S. T., Austin, D. J. & Schreiber, S. L. (1997) [Inducible gene expression and protein translocation using nontoxic ligands identified by a mammalian three-hybrid screen](#). Proc Natl Acad Sci U S A 94: 7825-30.
5. Pollock, R., Issner, R., Zoller, K., Natesan, S., Rivera, V. M. & Clackson, T. (2000) [Delivery of a stringent dimerizer-regulated gene expression system in a single retroviral vector](#). Proc Natl Acad Sci U S A 97: 13221-6.
6. Rivera, V. M., Ye, X., Courage, N. L., Sachar, J., Cerasoli, F., Jr., Wilson, J. M. & Gilman, M. (1999) [Long-term regulated expression of growth hormone in mice after intramuscular gene transfer](#). Proc Natl Acad Sci U S A 96: 8657-62.
7. Ye, X., Rivera, V. M., Zoltick, P., Cerasoli, F., Jr., Schnell, M. A., Gao, G., Hughes, J. V., Gilman, M. & Wilson, J. M. (1999) [Regulated delivery of therapeutic proteins after in vivo somatic cell gene transfer](#). Science 283: 88-91.
8. Belshaw, P. J., Ho, S. N., Crabtree, G. R. & Schreiber, S. L. (1996) [Controlling protein association and subcellular localization with a synthetic ligand that induces heterodimerization of proteins](#). Proc Natl Acad Sci U S A 93: 4604-7.
9. Choi, J., Chen, J., Schreiber, S. L. & Clardy, J. (1996) [Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP](#). Science 273: 239-42.

10. Chen, J., Zheng, X. F., Brown, E. J. & Schreiber, S. L. (1995) [*Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue.*](#) Proc Natl Acad Sci U S A 92: 4947-51.
 11. Pomerantz, J. L., Sharp, P. A. & Pabo, C. O. (1995) [*Structure-based design of transcription factors.*](#) Science 267: 93-6.
 12. Schmitz, M. L. & Baeuerle, P. A. (1991) [*The p65 subunit is responsible for the strong transcription activating potential of NF-kappa B.*](#) Embo J 10: 3805-17.
 13. Amara, J. F., Clackson, T., Rivera, V. M., Guo, T., Keenan, T., Natesan, S., Pollock, R., Yang, W., Courage, N. L., Holt, D. A. & Gilman, M. (1997) [*A versatile synthetic dimerizer for the regulation of protein-protein interactions.*](#) Proc Natl Acad Sci U S A 94: 10618-23.
 14. Natesan, S., Rivera, V. M., Molinari, E. & Gilman, M. (1997) [*Transcriptional squelching re-examined.*](#) Nature 390: 349-50.
 15. Attar, R. M. & Gilman, M. Z. (1992) [*Expression cloning of a novel zinc-finger protein that binds to the c-fos serum response element.*](#) Mol. Cell. Biol. 12: 2432-2443.
-

Appendix

pC₄N₂-R_HS/ZF3 Annotated Sequence

MluI <--
1 acgccttcgagctcgccccgttacataactacggtaaatggccccctggctgaccgccaacgacccccgccattgacgtcaataatgacgtatggt 100

101 cccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaacgcccacttgccagtcacatcaagtgatcatatgcca 200

201 gtacgccccctattgacgtcaatgacggtaaatggccccctggcattatgcccagtcacatgaccttatgggactttctacttggcagtcacatctacgt 300

301 attagtcacgtcattaccatggatgagcgggtttggcagtcacatcaatggcgtagcgggtttgactcacggggattccaagtctccacccatt 400

401 gacgtcaatgggagttgtttttggcaccaaaatcaacgggactttccaaaatgctgtaaacactccgccccattgacgcaaatggcggtaggcgtgtac 500

501 ggtgggaggtctatataagcagagctcgttttagtgaaccgctcagatcgctggagacgccatccacgctgttttgacctccatagaagacaccgggaccg 600

601 atccagcctccgggggatcttgggtggcgtgaaactcccgcagatcttggccagcgaaattcagaagccacc ATG GAC TAT CCT GCT GCC AAG 693
1 M D Y P A A K 7

694 AGG GTC AAG TTG GAC TCT AGA ATC CTC TGG CAT GAG ATG TGG CAT GAA GGC CTG GAA GAG GCA TCT CGT TTG TAC 768
8 R V K L D S R I L W H E M W H E G L E E A S R L Y 32

769 TTT GGG GAA AGG AAC GTG AAA GGC ATG TTT GAG GTG CTG GAG CCC TTG CAT GCT ATG ATG GAA CGG GGC CCC CAG 843
33 F G E R N V K L G M F E V L E P L H A M M E R G P A Q 57

844 ACT CTG AAG GAA ACA TCC TTT AAT CAG GCC TAT GGT CGA GAT TTA ATG GAG GCC CAA GAG TGG TGC AGG AAG TAC 918
58 T L K E T S F N Q A Y G R D L M E A Q E W C R K Y 82

919 ATG AAA TCA GGG AAT GTC AAG GAC CTC CTC CAA GCC TGG GAC CTC TAT TAT CAT GTG TTC CGA CGA ATC TCA AAG 993
83 M K S G N V K D L L Q A W D L Y Y H V F R R I S K 107

994 ACT AGA GAT GAG TTT CCC ACC ATG GTG TTT CCT TCT GGG CAG ATC AGC CAG GCC TCG GCC TTG GCC CCG GCC CCT 1068
108 T R D E F T M V F P S Q I A S Q A L A P 132

1069 CCC CAA GTC CTG CCC CAG GCT CCA GCC CCT GCC CCT GCT CCA GCC ATG GTA TCA GCT CTG GCC CAG GCC CCA GCC 1143
133 P Q V L P Q A P A P A P A P A M V S A L A Q A P A 157

1144 CCT GTC CCA GTC CTA GCC CCA GGC CCT CCT CAG GCT GTG GCC CCA CCT GCC CCC AAG CCC ACC CAG GCT GGG GAA 1218
158 P V P V L A P G P P Q A V A P P A P K P T Q A G E 182

1219 GGA ACG CTG TCA GAG GCC CTG CTG CAG CTG CAG TTT GAT GAT GAA GAC CTG GGG GCC TTG CTT GGC AAC AGC ACA 1293
183 G T L S E A L L Q L Q F D D E D L G A L L L G A L S T 207

1294 GAC CCA GCT GTG TTC ACA GAC CTG GCA TCC GTC GAC AAC TCC GAG TTT CAG CAG CTG CTG AAC CAG GGC ATA CCT 1368
208 D P A V F T D L A S V D N S E F Q Q L L N Q G I P 232

1369 GTG GCC CCC CAC ACA ACT GAG CCC ATG CTG ATG GAG TAC CCT GAG GCT ATA ACT CGC CTA GTG ACA GGG GCC CAG 1443
233 V A P H T T E P M L M E Y P E A I T R L V T G A Q 257

1444 AGG CCC CCC GAC CCA GCT CCT GCT CCA CTG GGG GCC CCG GGG CTC CCC AAT GGC CTC CTT TCA GGA GAT GAA GAC 1518
258 R P P D P A P L G A P G L P N G L L S G D E D 282

1519 TTC TCC TCC ATT GCG GAC ATG GAC TTC TCA GCC CTG CTG AGT CAG ATC AGC TCC ACT AGT TAT TAA ggatctccgggt 1596
283 F S S I A D M D F S A L L S Q I S S T S Y * 304

1597 attttccacatattgcccgtcttttggcaatgtgagggccggaaacctggccctgtctcttgcagagcattcctaggggtctttccctctcgcacaaa 1696

1697 ggaatgcaaggtctgtgaaatgctggaaggaagcagttcctctggaagcttcttgaagacaaacaacgctgttagcagaccctttgcaggcagcggaaacc 1796

1797 cccacactggcgacaggtgctctcggccaaaagccacgtgtataagatacactgcgaaagggcgcacacccagtgccacggtgtgagttggatagt 1896

1897 tgtggaagagtgcaaatggctctcctcaagcgtattcaacaaggggctgaaggatgccagaaggtaccccattgtatgggatctgatctggggcctcgg 1996

1997 tgcacatgctttacatgtgttttagtgcaggttaaaaaacgtctagggcccccgaaaccggggcagctggttttctcttgaaaaaacacgatgataatacc 2095

2096 ATG GAC TAT CCT GCT GCC AAG AGG GTC AAG TTG GAC TCT AGA GAA CGC CCA TAT GCT TGC CCT GTC GAG TCC TGC 2170
1 M D Y P A A K R V K L D S R P E R P Y A C P V E S C 25

2171 GAT CGC CGC TTT TCT TCT CGC TCG GAT GAG CTT ACC CGC CAT ATC CGC ATC CAC ACA GGC CAG AAG CCC TTC CAG TGT 2245
26 D R R F S R S D E L T R H I R I H T G Q K P F Q C 50

2246 CGA ATC TGC ATG CGT AAC TTC AGT CGT AGT GAC CAC CTT ACC ACC CAC ATC CGC ACC CAC ACA GGC GGC GGC CGC 2320
51 R I C M R N F S R S D H L T T H I R T H T G G G R 75

2321 AGG AGG AAG AAA CGC ACC AGC ATA GAG ACC AAC ATC CGT GTG GCC TTA GAG AAG AGT TTC TTG GAG AAT CAA AAG 2395
76 R R K K R T S I E T N I R V A L E K S F L E N Q K 100

2396 CCT ACC TCG GAA GAG ATC ACT ATG ATT GCT GAT CAG CTC AAT ATG GAA AAA GAG GTG ATT CGT GTT TGG TTC TGT 2470
101 P T S E E I T M I A D G Q L N M E K E V I R V W F C 125

2471 AAC CGC CGC CAG AAA GAA AAA AGA ATC AAC ACT AGA GGA GTG CAG GTG GAA ACC ATC TCC CCG GGA GAC GGG CGC 2545
126 N R R Q K E K R I N T R G V Q V E T I S P G D G R 150

2546 ACC TTC CCC AAG CGC GGC CAG ACC TGC GTG GTG CAC TAC ACC GGG ATG CTT GAA GAT GGA AAG AAA TTT GAT TCC 2620
151 T F P K R G Q T C V V H Y T G M L E D G K K F D S 175

2621 TCC CGG GAC AGA AAC AAG CCC TTT AAG TTT ATG CTA GGC AAG CAG GAG GTG ATC CGA GGC TGG GAA GAA GGG GTT 2695
176 S R D R N K P F K M L G K G Q E V I R G W E E G V 200

2696 GCC CAG ATG AGT GTG GGT CAG AGA GCC AAA CTG ACT ATA TCT CCA GAT TAT GCC TAT GGT GCC ACT GGG CAC CCA 2770
201 A Q M S V G Q R A K L T I S P D Y A Y G A T G H P 225

2771 GGC ATC ATC CCA CCA CAT GCC ACT CTC GTC TTC GAT GTG GAG CTT CTA AAA CTG GAA GTC GAG GGC GTG CAG GTG 2845
226 G I I P P H A T L V F D V E L L K L E V E G V Q V 250

2846 GAA ACC ATC TCC CCA GGA GAC GGG CGC ACC TTC CCC AAG CGC GGC CAG ACC TGC GTG GTG CAC TAC ACC GGG ATG 2920
251 E T I S P G D G R T F P K R Q T C V V H Y T G M 275

2921 CTT GAA GAT GGA AAG AAA TTT GAT TCC TCC CGG GAC AGA AAC AAG CCC TTT AAG TTT ATG CTA GGC AAG CAG GAG 2995
276 L E D G K K F D S S R D R N K P F K F M L G K Q E 300
2996 GTG ATC CGA GGC TGG GAA GAA GGG GTT GCC CAG ATG AGT GTG GGT CAG AGA GCC AAA CTG ACT ATA TCT CCA GAT 3070
301 V I R G W E E G V A Q M S V G Q R A K L T I S P D 325
3071 TAT GCC TAT GGT GCC ACT GGG CAC CCA GGC ATC ATC CCA CCA CAT GCC ACT CTC GTC TTC GAT GTG GAG CTT CTA 3145
326 Y A Y G A T G H P G I I P P H A T L V F D V E L L 350
3146 AAA CTG GAA ACT AGA GGA GTG CAG GTG GAA ACC ATC TCC CCA GGA GAC GGG CGC ACC TTC CCC AAG CGC GGC CAG 3220
351 K L E T R G G V Q V E T I S P G D T G R T F K R G G Q 375
3221 ACC TGC GTG GTG CAC TAC ACC GGG ATG CTT GAA GAT GGA AAG AAA TTT GAT TCC TCC CGG GAC AGA AAC AAG CCC 3295
376 T C V V H Y T G M L E D G K K F T D S S R D R N K P 400
3296 TTT AAG TTT ATG CTA GGC AAG CAG GAG GTG ATC CGA GGC TGG GAA GAA GGG GTT GCC CAG ATG AGT GTG GGT CAG 3370
401 F K F M L G K Q E V I R G W E E G V A Q M S V G Q 425
3371 AGA GCC AAA CTG ACT ATA TCT CCA GAT TAT GCC TAT GGT GCC ACT GGG CAC CCA GGC ATC ATC CCA CCA CAT GCC 3445
426 R A K L T I S P D Y A Y G A T G H P G I I P P H A 450
3446 ACT CTC GTC TTC GAT GTG GAG CTT CTA AAA CTG GAA ACT AGT TAT TAA ggatcctgagaacttcagggtgagtttggggaccct 3529
451 T L V F D V E L L K L E T S Y *
3530 tgattgttctttctttctttctgctattgtaaaattcatgttatatggagggggcaaagttttcagggtgtgtttagaatgggaagatgtcccttgatcac 3629
3630 catggaccctcatgataattttgtttctttcactttctactctgttgacaaccattgtctctcttattttcttttcattttctgtaactttttcgtaa 3729
rabbit β-globin intron
3730 acttttagcttgcaattgttaacgaatttttaaaattcacttttgtttatttgcagattgtaagtaactttctctaatacactttttttcaaggcaatcaggg 3829
3830 tatattatattgtacttcagcacagtttttagagaacaattgttataaataatgataaggtagaatattttctgcataataaattctggctggcgtggaat 3929
3930 attcttattggtagaacaactacatcctggctcatcatcctgctttctctttatgggttacaatgatatacactgtttgagatgaggataaaataactctg 4029
-->
4030 agtccaaaccgggcccctctgctaaccatgttcatgccttcttcttttctacagctcctgggcaacgtgctggtgtgtgtgctctcatcattttgg 4129
and 3' UTR
4130 caaaggattcactcctcaggtgcaggctgacctatcagaaggtggctggtgtgtggccaatgcctggctcacaataaccactgagatctttttccctct 4229
Poly A signal
4230 gccaaaaattatggggacatcatgaagccccttgagcatctgacttctggctaaataaaggaaattttttcattgcaatagtggttgaattttttgt 4329
4330 gtctctcactcgggaaggacatatgggagggcaaatactttaaaacatcagaatgagatatttggttttagagtttgcaacatatgccatagctggctgcc 4429
4430 atgaacaaaaggtggctataaagaggtcatcagtatatgaaacagccccctgctgtccattcttattccatagaaaagccttgacttgaggttagatatt 4529
4530 ttttatattttgtttgtgttattttttttaaatacctaaaattttctcttacatgttttactagccagatttttctcctctcctgactactccca 4629
XhoI
4630 gtcatagctgtccctcttctcttatgaagatccgtcgaggagctttttgcaaaaagccttaggcctcAAAAAagcctcttctactacttctggaatagctc 4729

pC₄N₂-R_HS3H/ZF3 Annotated Sequence

```

MluI <--
1  accgcttcgagctcgccttacataacttacggttaaatggcccgctggctgaccgccaacgacccccgccattgacgtcaataatgacgtatggt 100
101 cccatagtaacgccaatagggactttccattgacgtcaaatgggtggagtatttacggtaaacgcccacttggcagtacatcaagtgtatcatatgcca 200
201 gtacgccccctattgacgtcaatgacggtaaatggcccgctggcattatgccagtacatgaccttatgggactttctacttggcagtacatctacgt 300
301 attagtcatcgtattaccatggatgacggttttggcagtacatcaatggcggtgatagcggtttgactcacggggattccaagtctccacccatt 400
401 gacgtcaatgggagttgtttttggcaccaaaatcaacgggactttccaaaatgctgtaaacactccgccccattgacgcaaatggcggtgagcggtgac 500
501 ggtgggaggtctatataagcagagctcgttttagtgaaccgacagatcgctggagagcccatccacgctgttttgacctccatagaagacacggggaccg 600
601 atccagcctccggggatcttgggtggcgtgaaactcccgcagatcttggccagcgaatttcagaagccacc ATG GAC TAT CCT GCT GCC AAG 693
1  M D Y P A A K 7
694 AGG GTC AAG TTG GAC TCT AGA ATC CTC TGG CAT GAG ATG TGG CAT GAA GGC CTG GAA GAG GCA TCT CGT TTG TAC 768
8 R V K L D S R I L W H E M W H E G L E E A S R L Y 32
769 TTT GGG GAA AGG AAC GTG AAA GGC ATG TTT GAG GTG CTG GAG CCC TTG CAT GCT ATG ATG GAA CGG GGC CCC CAG 843
33 F G E R N V K G M F E V L E P L H A M M E R G P Q 57
844 ACT CTG AAG GAA ACA TCC TTT AAT CAG GCC TAT GGT CGA GAT TTA ATG GAG GCC CAA GAG TGG TGC AGG AAG TAC 918
58 T L K E T S F N Q A Y G R D L M E A Q E W C R K Y 82
919 ATG AAA TCA GGG AAT GTC AAG GAC CTC CTC CAA GCC TGG GAC CTC TAT TAT CAT GTG TTC CGA CGA ATC TCA AAG 993
83 M K S G N V K D L L Q A W D L Y Y H V F R R I S K 107
994 ACT AGA AGT GAG CCC ATG GAA TTT CAG TAC CTG CCA GAT ACA GAC GAT CGT CAC CGG ATT GAG GAG AAA CGT AAA 1068
108 T R S E P M E F Q Y L P D T D GAT R H R I E E K R K 132
1069 AGG ACA TAT GAG ACC TTC AAG AGC ATC ATG AAG AAG AGT CCT TTC AGC GGA CCC ACC GAC CCC CGG CCT CCA CCT 1143
133 R T Y E T F K S I M K K S P F S G P T D P R P P P 157
1144 CGA CGC ATT GCT GTG CCT TCC CGC AGC TCA GCT TCT GTC CCC AAG CCA GCA CCC CAG CCC TAT CCC TTT ACG TCA 1218
158 R R I A V P S R S S A S V P K P A P Q P Y P F T S 182
1219 TCC CTG AGC ACC ATC AAC TAT GAT GAG TTT CCC ACC ATG GTG TTT CCT TCT GGG CAG ATC AGC CAG GCC TCG GCC 1293
183 S L S T I N Y D E F P T M V F P S G Q I S Q A S A 207
1294 TTG GCC CCG GCC CCT CCC CAA GTC CTG CCC CAG GCT CCA GCC CCT GCC CCT GCT CCA GCC ATG GTA TCA GCT CTG 1368
208 L A P A P Q V L P Q A P A P A P A P A P A P A L 232
1369 GCC CAG GCC CCA GCC CCT GTC CCA GTC CTA GCC CCA GGC CCT CCT CAG GCT GTG GCC CCA CCT GCC CCC AAG CCC 1443
233 A Q A P A P V P V L A P G P P Q A V A P P A P K P 257
1444 ACC CAG GCT GGG GAA GGA ACG CTG TCA GAG GCC CTG CTG CAG CTG CAG TTT GAT GAT GAA GAC CTG GGG GCC TTG 1518
258 T Q A G E G T L S E A L L Q L Q F D D E D L G A L 282
1519 CTT GGC AAC AGC ACA GAC CCA GCT GTG TTC ACA GAC CTG GCA TCC GTC GAC AAC TCC GAG TTT CAG CAG CTG CTG 1593
283 L G N S T G D P A V F T D L A S V D N S E F Q Q L L 307
1594 AAC CAG GGC ATA CCT GTG GCC CCC CAC ACA ACT GAG CCC ATG CTG ATG GAG TAC CCT GAG GCT ATA ACT CGC CTA 1668
308 N Q G I P V A P H T T E P M L M E Y P E A I T R L 332
1669 GTG ACA GGG GCC CAG AGG CCC CCC GAC CCA GCT CCT GCT CCA CTG GGG GCC CCG GGG CTC CCC AAT GGC CTC CTT 1743
333 V T G A Q R P P D P A P A P L G A P G L P N G L L 357
1744 TCA GGA GAT GAA GAC TTC TCC TCC ATT GCG GAC ATG GAC TTC TCA GCC CTG CTG AGT CAG ATC AGC TCC ACT AGA 1818
358 S G D E T G F T C S I A D M D F S A L L L S Q I A S T R 382
1819 GGC TTC AGC GTG GAC ACC AGT GCC CTG CTG GAC CTG TTC AGC CCC TCG GTG ACC GTG CCC GAC ATG AGC CTG CCT 1893
383 G F S V D T S A L L D L F S P S V T V P D M S L P 407
1894 GAC CTT GAC AGC AGC CTG GCC AGT ATC CAA GAG CTC CTG TCT CCC CAG GAG CCC CCC AGG CCT CCC GAG GCA GAG 1968
408 D L D S S L A S I Q E L L S P Q E P P R P P E A E 432
1969 AAC AGC AGC CCG GAT TCA GGG AAG CAG CTG GTG CAC TAC ACA CGC CAG CCG CTG TTC CTG CTG GAC CCC GGC TCC 2043
433 N S S P D T S G K Q L V H Y T A A Q P L F L L D P G S 457
2044 GTG GAC ACC GGG AGC AAC GAC CTG CCG GTG CTG TTT GAG CTG GGA GAG GGC TCC TAC TTC TCC GAA GGG GAC GGC 2118
458 V D T G S N D L P V L F E L G E G S Y F S E G D G 482
2119 TTC GCC GAG GAC CCC ACC ATC TCC CTG CTG ACA GGC TCG GAG CCT CCC AAA GCC AAG GAC CCC ACT GTC TCC ACT 2193
483 F A E D P T I S L L T G S E P P K A K D P T V S T 507
2194 AGA TAG tgaagatctccgggtattttccaccatattgcccgtcttttggcaatgtgagggcccgaaacctggccctgtcttcttgacgagcattccta 2291
508 R * 509
2292 ggggtctttcccctctcgcacaaaggaatgcaaggtctgttgaatgctgtaaggaagcagttcctctggaagcttcttgaagacaaacaacgtctgtagc 2391
2392 gaccctttgcaggcagcggaacccccccactggcgacaggtgocctctcggccaaaagccacgtgtataagatacacctgcaaaggcgccacaaccccgag 2491
2492 tgccacgttgtgagttggatggttggaaagagtcaaatggctctcccaagcgtattcaacaaggggctgaaggatgccagaaggtaccccaattgta 2591
2592 tgggatctgatctggggcctcgtgacatgctttacatgctttagtgcaggttaaaaaacgtctagggcccccaaacacggggacgtggttttctt 2691
2692 tgaaaaacacgatgataatacc ATG GAC TAT CCT GCT GCC AAG AGG GTC AAG TTG GAC TCT AGA GAA CGC CCA TAT GCT 2770
1 M D Y P A A K R V K L D S R E R P Y A 19
2771 TGC CCT GTC GAG TCC TGC GAT CGC CGC TTT TCT CGC TCG GAT GAG CTT ACC CGC CAT ATC CGC ATC CAC ACA GGC 2845
20 C P V E S C D R R F S D E L T R H I R I H T G 44
2846 CAG AAG CCC TTC CAG TGT CGA ATC TGC ATG CGT AAC TTC AGT CGT AGT GAC CAC CTT ACC ACC CAC ATC CGC ACC 2920
45 Q K P F Q C R I C M R N F S R S D H L T T H I R T 69

```


pZ₁₂I-PL-2 Annotated Sequence

```

MluI  NheI
1  ccggggaggtaccgagctcttaccgctgctagctaatgatggcgctcgagtaatgatggcgctcgactaatgatggcgctcgagtaatgatggcgct 100
                                     12x ZFHD1 sites
101 ctagctaatgatggcgctcgagtaatgatggcgctcgactaatgatggcgctcgagtaatgatggcgctctagctaatgatggcgctcgagtaatg 200
                                     XbaI <--
201 atggcgctcgactaatgatggcgctcgagtaatgatggcgctctagaaacgcgaattttaacaacattttgacacccccataatattttccagaatta 300
      IL2 promoter  <--> HindIII PstI
301 acagtataaattgcatctctgttcaagagttccctatcactctctttaatcactactcacagtaaacctcaactcctgccacaagcttgccctgacagcg 400
      EcoRI SpeI BglIII ClaI <-- SV40 late 3'UTR
401 gaattccactagctcgagatctccatcgatggcgcttcgagcagacatgataagatacattgatgagtttgacaaaaccacaactagaatgcagtgaaaa 500
                                     Poly A signal
501 aaatgctttatttgtgaaatttgtgatgctattgctttatttgaaccattataagctgcaataaacagtttaacaacaacaattgcattcattttatgt 600
                                     BamHI
601 ttcaggttcagggggaggtgtgggaggttttttaagcaagtaaacctctacaatgtggtaaaatcgatccgctcgaccgatgccttgagagccttc 700

```

pZ₁₂I-hGH-2 Annotated Sequence

```

MluI  NheI
1  ccggggaggtaccgagctcttaccgctgctagctaatgatggcgctcgagtaatgatggcgctcgactaatgatggcgctcgagtaatgatggcgct 100
                                     12x ZFHD1 sites
101 ctagctaatgatggcgctcgagtaatgatggcgctcgactaatgatggcgctcgagtaatgatggcgctctagctaatgatggcgctcgagtaatg 200
                                     XbaI <--
201 atggcgctcgactaatgatggcgctcgagtaatgatggcgctctagaaacgcgaattttaacaacattttgacacccccataatattttccagaatta 300
      IL2 promoter  <--> HindIII
301 acagtataaattgcatctctgttcaagagttccctatcactctctttaatcactactcacagtaaacctcaactcctgccacaagcttaccactcaggg 400
                                     <--
401 cctgtggacagctcacctagctgca ATG GCT ACA GGC TCC CGG ACG TCC CTG CTC CTG GCT TTT GGC CTG CTC TGC CTG 479
      1 M A T G S R T S L L L A F G L L C L 18
480 CCC TGG CTT CAA GAG GGC AGT GCC TTC CCA ACC ATT CCC TTA TCC AGG CCT TTT GAC AAC GCT ATG CTC CGC GCC 554
      19 P W L Q E G S A F P T I P L S R P F D N A M L R A 43
555 CAT CGT CTG CAC CAG CTG GCC TTT GAC ACC TAC CAG GAG TTT GAA GAA GCC TAT ATC CCA AAG GAA CAG AAG TAT 629
      44 H R L H Q L A F D T Y Q E F E A Y I P K E Q K Y 68
630 TCA TTC CTG CAG AAC CCC CAG ACC TCC CTC TGT TTC TCA GAG TCT ATT CCG ACA CCC TCC AAC AGG GAG GAA ACA 704
      69 S F L Q N P Q T S L C F S E S I P T P S N R E E T 93
705 CAA CAG AAA TCC AAC CTA GAG CTG CTC CGC ATC TCC CTG CTG CTC ATC CAG TCG TGG CTG GAG CCC GTG CAG TTC 779
      94 Q Q K S N L E L R I S L L L I Q S W L E P V Q F 118
780 CTC AGG AGT GTC TTC GCC AAC AGC CTG GTG TAC GGC GCC TCT GAC AGC AAC GTC TAT GAC CTC CTA AAG GAC CTA 854
      119 L R S V F A N S L V Y G A S D S N V Y D L L K D L 143
855 GAG GAA GGC ATC CAA ACG CTG ATG GGG AGG CTG GAA GAT GGC AGC CCC CGG ACT GGG CAG ATC TTC AAG CAG ACC 929
      144 E E G I Q T L M G R L E D G S P R T G Q I F K Q T 168
930 TAC AGC AAG TTC GAC ACA AAC TCA CAC AAC GAT GAC GCA CTA CTC AAG AAC TAC GGG CTG CTC TAC TGC TTC AGG 1004
      169 Y S K F D T N S H N D A T G A C L L K N Y G L L Y C F R 193
1005 AAG GAC ATG GAC AAG GTC GAG ACA TTC CTG CGC ATC GTG CAG TGC GGC TCT GTG GAG GGC AGC TGT GGC TTC TAG 1079
      194 K D M D K V E T F L R I V Q C R S V E G S C G F * 218
1080 ctgcccgggtggcatccctgtgaccctccccagtgctctcctggccctggaagttgccacgaattccactagctcgagatctccatcgatggcgcttc 1179
                                     <-- SV40 late 3'UTR
1180 gagcagacatgataagatacattgatgagtttgacaaaaccacaactagaatgcagtgaaaaaatgctttatttgtgaaatttgtgatgctattgcttt 1279
      Poly A signal
1280 atttgtaaccattataagctgcaataaacagtttaacaacaacaattgcattcattttatgcttcagggttcagggggaggtgtgggaggttttttaagc 1379
                                     BamHI
1380 aagtaaacctctacaatgtggtaaaatcgatccgctcgaccgatgccttgagagccttcaaccagtcagctccttcgggtggcgggggcagtcag 1479

```