Cell Programmers Technical Report

Use of C_2H_2 Zn-finger Proteins in Yeast-based circuits

Summary -

The intention is to build simple toggle switches in the yeast *Saccharomyces Cereviscae* utilizing recombinant zinc finger proteins that recognize novel DNA-motifs. This report describes the first stage of the project, namely the analysis of a series of C_2H_2 Zinc fingers to determine their potential and the creation of a simple regulatable switch. The work described is intended to take one year.

Introduction -

The aim of the project is to create engineered circuits which can be regulated at will *in vivo*. The simplest of these circuits is the 'toggle-switch' (**Figure 1**) which has been shown to work in *E.coli* and this is the first that will be created in the yeast *Saccharomyces Cereviscae.* The toggle-switch in *E.coli* was regulated by the Lac-operon / λC_1 repressor system, but the switch from prokaryotes to eukaryotes will be accompanied by a more appropriate method of transcriptional regulation.

Transcription factors fall into many classes depending on their structure and exact mode of action but their function can be simply stated – the regulation of transcription. A transcription factor regulates gene expression by recognizing and binding to a specific DNA sequence (recognition motif) associated with a particular gene and causing that gene to be

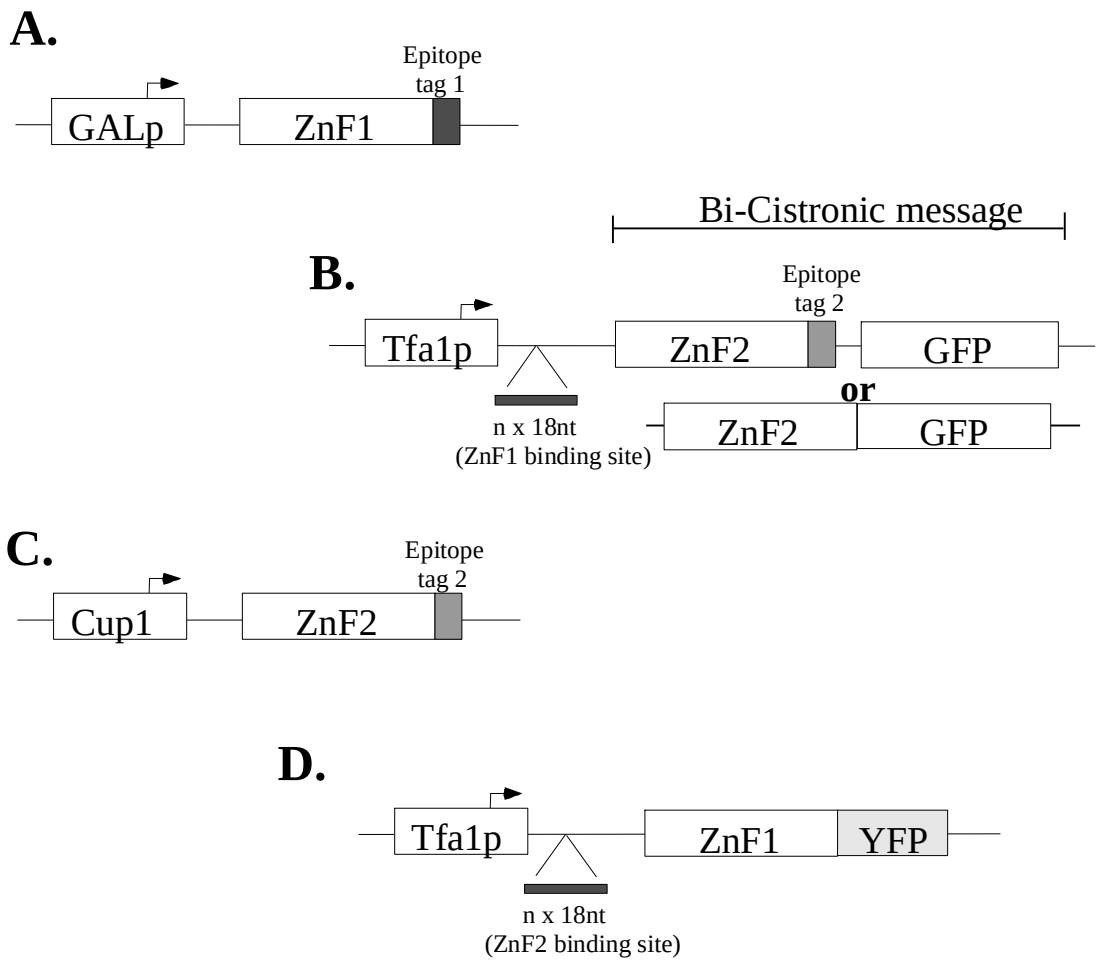


Figure 1 - Toggle Switch

Comprised of four independent constructs (A - D). B and D are under the control of the constitutive promoter Tfa1 while A and C are regulated by the inducible promoters; Galactose and Copper respectively. Two six finger C_2H_2 factors are utilised in this circuit (ZnF1 and ZnF2) and both are differentially epitope tagged for ease of identification biochemically.
Construct B can encode a bi-cistronic message, namely two individual proteins encoded by the same mRNA), This is achieved by encoding a strong ribozyme binding site between the two ORFs. An alternative is the expression of a ZnF2 / GFP fusion protein. Thus the circuit encodes two fluorescent readouts of expression which can be easily measured on a FACS machine.

diverse DNA motifs, with each zinc finger being highly specific in its recognition sequence (16).

Structural studies show that each zinc finger module, which contains about 30 amino acids, forms a compact ββα structure. Four conserved residues in each finger, two cysteine and two histidine, are ligands for a central zinc ion that stabilizes this small globular domain. On association with a specific DNA motif, the α-helix fits into the major groove. Zinc-finger specificity can be explained with a zinc finger-DNA recognition code that correlates specific amino acid at key positions in the α-helix with specific bases in each position of the corresponding DNA motif (**Figure 2**). Indeed the crystal structure of Zinc finger – DNA complexes show a semi-conserved pattern of interactions in which three residues from the α-helix contact three adjacent bases in DNA. In its simplest terms, the mode of DNA recognition is a one-to-one interaction between amino acids in the Zinc fingers and nucleotides in the DNA strand. This is where the most interesting facet of C_2H_2 Zinc fingers comes into play; because the fingers function as independent modules, fingers with different triplet specificities can be combined to give specific recognition of longer DNA sequences.

Naturally occurring transcription factors using these C_2H_2 fingers typically contain sets of two, three or four fingers that bind to neighbouring subsites on the DNA although proteins with many more fingers have been described. *Xenopus* Xfin, for example contains 37 such fingers (19) although there are questions about how many of these actually contact DNA and how this effects affinity and specificity of recognition. However many groups have successfully exploited individual finger adaptability to randomly mutagenise and isolate novel fingers with altered DNA binding specificities (10, 12, 20, 21). The majority of the derived factors contain three fingers and recognise a DNA motif of 9 – 11 nucleotides. These recombinant factors have affinities and specificities for their sites comparable to

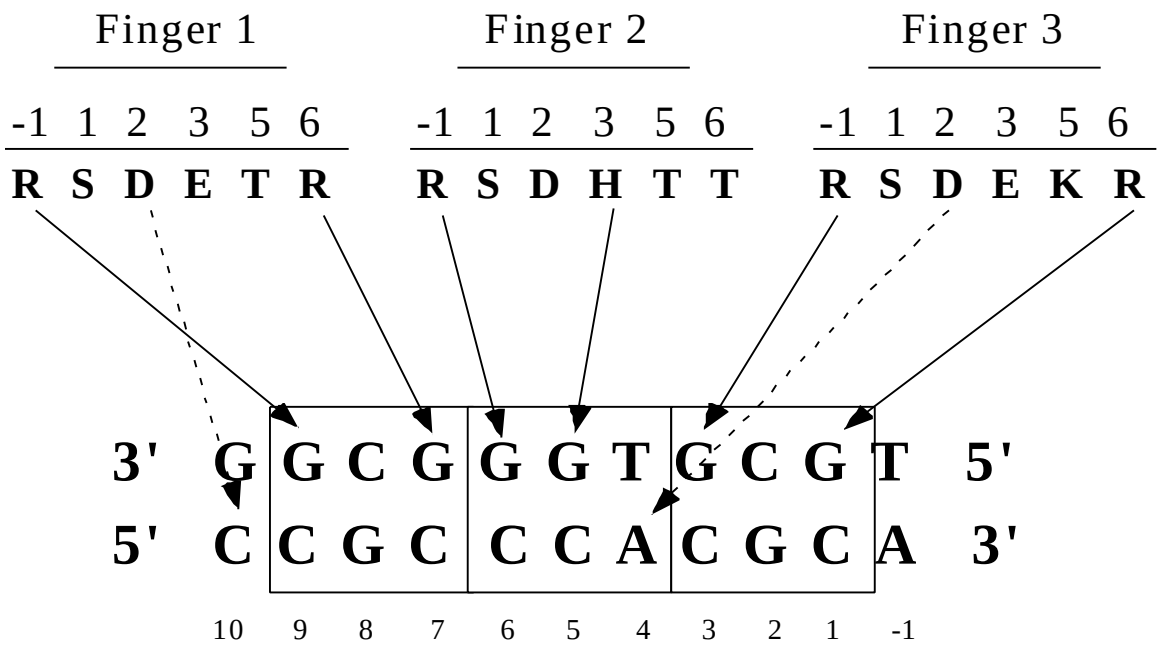


Figure 2 - Side chain-base contacts in Zif286-DNA complex

Amino acids in the α-helix of each C2H2 Zinc finger are indicated. Arrows represent observed side chain-base interactions and DNA involves residues - 1, 3 and 6 of the helix. Each finger also uses residue 2 to contact a flanking base in the secondary strand of the DNA, although the hydrogen binding geometry is not ideal for those involving fingers one and three (represented by broken lines). The c-terminal finger binds near the 5' end of the primary strand, and thus the fingers (proceeding in the conventional N - > C order) bind ' anti-parallel' with respect to the conventional (5' ->3') direction of the primary DNA strand. (Reproduced from (8)).

Optimisation of C_2H_2 Zinc fingers –

Numerous studies have used phage display methods or design methods to explore

been performed to optimise the linker length between the two three-finger proteins and the ideal consensus DNA motif. The use of short linkers (TGEKP) to connect fingers gives only modest increases in affinity, probably due to unfavourable docking of the resulting six finger protein on DNA, i.e. the helical periodicity of the zinc fingers does not quite match the helical periodicity of slightly relaxed B-DNA. With the utilization of longer, more flexible linkers (such as the 11aa RQKDGGGSERPY), two three zinc finger domains have been linked resulting in a six finger factor which binds the appropriate 18-bp motif >6000 fold more tightly than the individual parent three finger factors (12).

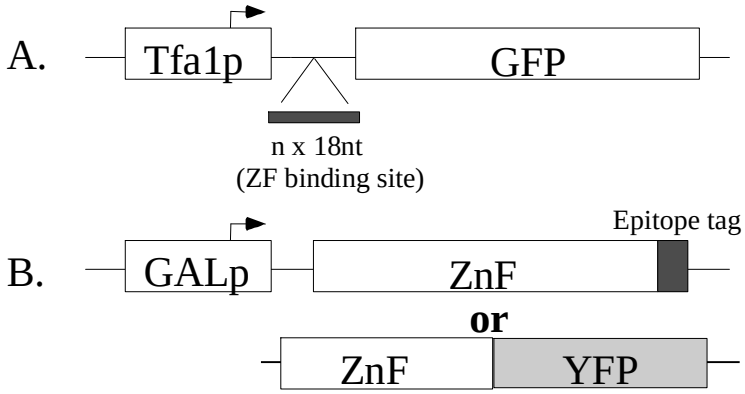
Experimental design -

The construction of a simple repression switch (**Figure 3**) in the budding yeast *Saccharomyces cereviscae* will be undertaken to identify appropriate Zinc-fingers for the toggle swtch. *S. cereviscae* is chosen as a model organism for its ease of use and representation of the eukaryotic system. A number of factors will be addressed in these initial experiments.

1. Of the C_2H_2 Zinc fingers available which are the most potent ?
2. Do they regulate the expression of any endogenous genes (non-specific binding ?)
3. Do the reporter constructs demonstrate dose-dependence ?
4. How stable are the proteins ? If too stable they will be unsuitable for a switch.
5. Characterization of the transfer curve

Choice of C_2H_2 Zinc finger -

As stated, many transcription factors in eukaryotes from yeast to humans contain



A & B tested as plasmids or integrated into genome

Glucose
A. GFP constitutively on
B. Zn expression repressed

Switch

Galactose
A. GFP expression ?
B. Zn expression induced

Switch ? Level of GFP repression ?

Glucose
A. GFP expression recovery ?
B. Zn expression repressed

Recovery of GFP expression dependent on Zn finger stability

Figure 3 - Simple Repression Switch

The complexity of the yeast genome (≈1.8Mb) makes the design of gene-specific transcription factors a formidable challenge. Ideally the chosen factor(s) will specifically regulate the expression of the few target genes among the thousands present in the genome. The standard DNA motif recognised by a C_2H_2 Zinc finger covers 9 - 11 nucleotides and this can be bound with a Kd of 2 – 120 pM with a ratio of specific to non-specific binding approaching 300,000 in the best case (21) (**Table 1**). A combinatorial approach to improve specificity still further is possible. This can take the form of targeting the gene of interest with two or more recombinant factors or directly fusing two zinc finger proteins, each recognizing adjacent but non-overlapping 9 base pair motifs to yield highly specific DNA-binding proteins (9, 12, 14). Designer proteins containing six zinc finger domains have been described in which two three-finger proteins are fused via improved linker segments and recognise 18 – 20bp motifs with femtomolar dissociation constants (10, 12).

An alternative method of increasing the affinity of zinc fingers has also been described – the addition of a Gal4 domain (18) or peptide sequences (20) which mediate dimerization to a characterized ZnF significantly increases its' affinity. The increase in affinity can be of the order of 1000-fold in the best case. This is of interest for a number of reasons – the most relevant being that these proteins might lend themselves more readily to destabilisation and the creation of a useful repression switch.

The repression switch -

The design of the first toggle-switch is shown in **Figure 1**. The steps to achieve this are outlined in **Figure 4**. A simple repression switch will be created to assay the ability of each finger to act as a repressor (**Figure 3**).

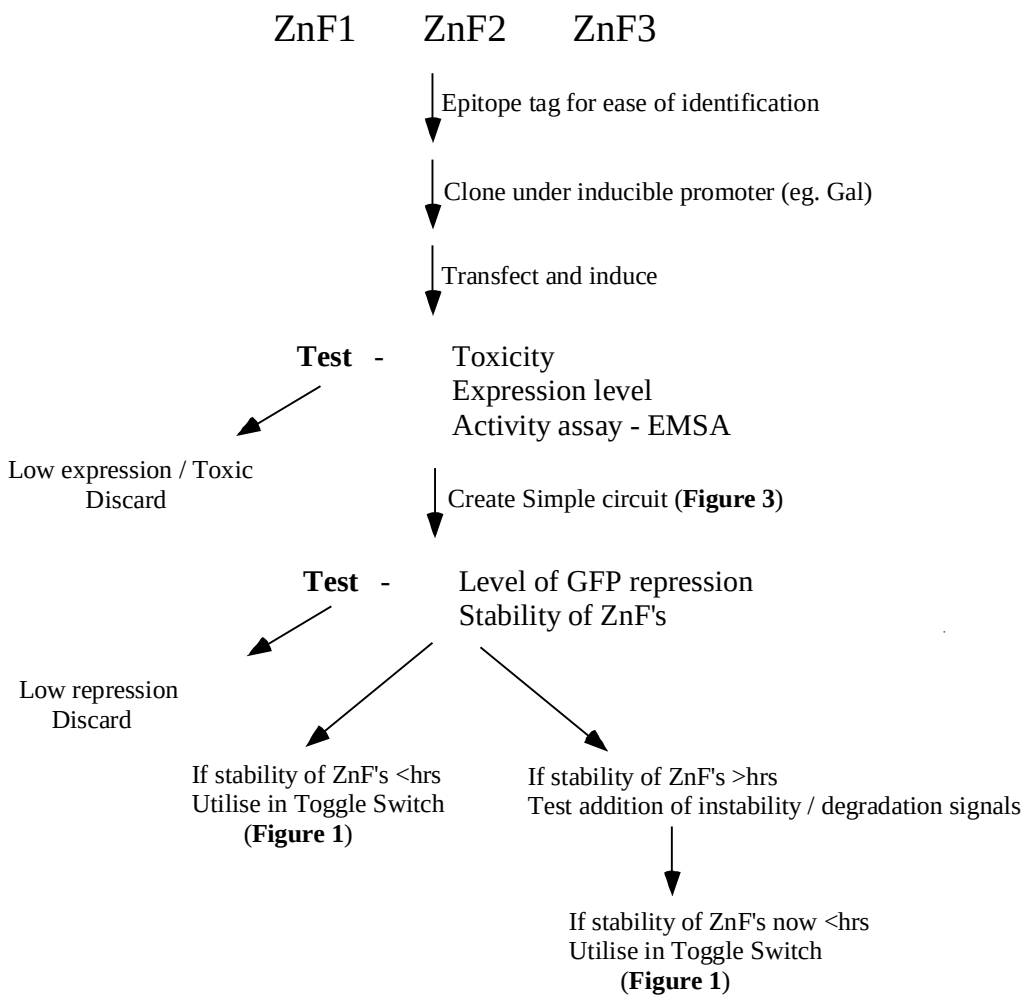


Figure 4 - Experimental design

Experimental design for testing the activity of each Zinc finger (ZnF) is depicted. The relation of each step to the creation of the simple repression circuit (**Figure 3**) and toggle switch (**Figure 1**) is shown. The utilisation of a zinc finger in the toggle switch is dependent on it satisfying the indicated criteria. Thus a desirable zinc finger will show;

(i) minimal background binding to the yeast genome (orders of magnitude below binding to its specific recognition sequence),

b. Whole cell extracts will be prepared from strains containing the Zinc fingers and checked for expression level and binding to the appropriate DNA motif in Gel shift assays. The creation and expression of a ZnF/YFP fusion will allow us to determine the expression level on a cell by cell basis to determine if there are any cell cycle effects. Also examined at this stage will be the dose dependence of the Gal promoter.
c. The creation of the simple repression circuit will allow us to guage the binding characteristics of each finger *in vivo*. Also as the Zinc fingers will be under the control of the Galactose promoter it is a simple matter of switching the sugar source to allow expression or repession of the proteins. Expression of a ZnF/YFP fusion will allow us to determine the expression level on a cell by cell basis to directly relate gene expression with expression level. Flow measurement will be used for quantification of expression levels and identity of input-output relationship (YFP:GFP). Simultaneous measurement of all by FACS as described (22).

The stability of the zinc-fingers is expected to be a problem in the first instance. Ecdysone mediated inducible expression has been performed in human cells with either the three finger 293 or six finger 293/NRE, binding of which leads to transcriptional repression of the marker gene Luciferase (10). Following withdrawal of the inducer, 293/NRE mediated repression was found to be long lived (>48hrs) due to the extreme stability of the protein. This makes 293/NRE unsuitable for the creation of a switch in its unmodified form. It is possible, however, to reduce the half-life of proteins significantly by the addition of specific peptide sequences which target them for degradation (15). This approach will be attempted to modify the stability of the zinc fingers to increase their usefulness.

which can give effective repression when bound to a site far downstream of the transcription start site (3). In general however, it appears that isolated C_2H_2 Zinc fingers can block transcriptional initiation but not elongation (10).

It is possible to increase the blocking ability of C_2H_2 Zinc fingers by linking them directly to transcriptional repression domains such as the Kruppel-associated box (KRAB) domain (1, 14) or the engrailed repressor domain (8). Care needs to be taken with this approach however: zinc finger-KRAB fusions are likely to affect the expression of many genes other than that intended as the KRAB domain has the ability to act over large distances or somehow sequester essential transcription factors (10). C_2H_2 Zinc finger – activation domain fusions have also been described (2).

The toggle-switch -

A functional toggle switch in *Saccharomyces cereviscae* is the aim of this project. The steps leading to this are outlined above and a schematic of the switch is outlined in **Figure 1**. The circuit described is envisioned to employ two six-finger C_2H_2 proteins with 18bp recognition motifs.

Four constructs are utilised in the circuit -

A. HA-tagged ZnF1 under the control of the inducible GAL promoter

B. Constitutive Tfa1 promoter regulating expression of a di-cistronic mRNA encoding two ORFs; ZnF2 and GFP (or more likely a ZnF2:GFP fusion protein). The expression of di-cistronic messages and active internal ribosome entry sites (IRES) has been previously described in *Saccharomyces cereviscae* (7, 17). There is a ZnF1 binding site between the Tfa1 promoter and the ZnF2/GFP mRNA to allow

Thus the circuit encodes two fluorescent proteins, GFP and YFP, both yeast-optimised forms. Simultaneous measurement by FACS will be used for quantification of expression levels (22) and identity of input-output relationship (YFP:GFP).

If it works, what then ? -

A demonstration that the toggle switch is functional in yeast is a large step forward and one that will be relatively easy to transfer to other eukaryotic systems - as mentioned C_2H_2 Zinc-finger proteins are used by all eukaryotes. We have a number of possible applications in mind.

References -

1. **Beerli, R. R., D. J. Segal, B. Dreier, and C. F. Barbas III.** 1998. Proc NAtl Acad Sci USA **95:**14628-14633.
2. **Choo, Y., A. Castellanos, B. Garcia-Hernandez, I. Sanchez-Garcia, and A. Klug.** 1997. Promoter-specific activation of gene expression directed by bacteriophage-selected zinc fingers. J Mol Biol **273:**525-32.
3. **Choo, Y., I. Sanchez-Garcia, and A. Klug.** 1994. Nature **372:**642-645.
4. **Corbi, N., V. Libri, M. Fanciulli, J. M. Tinsley, K. E. Davies, and C. Passananti.** 2000. The artificial zinc finger coding gene 'Jazz' binds the utrophin promoter and activates transcription. Gene Ther **7:**1076-1083.
5. **Desjarlais, J. R., and J. M. Berg.** 1992. Toward rules relating zinc finger protein sequences and DNA binding site preferences. Proc Natl Acad Sci U S A **89:**7345-9.
6. **Desjarlais, J. R., and J. M. Berg.** 1993. Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins. Proc Natl Acad Sci U S A **90:**2256-60.
7. **Greger, I. H., and N. J. Proudfoot.** 1998. Poly(A) signals control both transcriptional termination and initiation between the tandem GAL10 and GAL7 genes of Saccharomyces cerevisiae. Embo J **17:**4771-9.
8. **Jaynes, J. B., and P. H. O'Farrell.** 1991. Active repression of transcription by the Engrailed homeodomain protein. EMBO J **10:**1427-1433.
9. **Kamiuchi, T., M. Imanishi, E. Abe, and Y. Sugiura.** 1997. New multi zinc finger protein: biosynthetic design and characteristics of DNA recognition. Nucleic Acids Symp Ser **:**153-4.
10. **Kang, J. S., and J. S. Kim.** 2000. Zinc finger proteins as designer transcription factors. J Biol Chem **275:**8742-8.
11. **Kim, J. S., J. Kim, K. L. Cepek, P. A. Sharp, and C. O. Pabo.** 1997. Design of TATA box-binding protein/zinc finger fusions for targeted regulation of gene expression. Proc Natl Acad Sci U S A **94:**3616-20.
12. **Kim, J. S., and C. O. Pabo.** 1998. Getting a handhold on DNA: design of poly-zinc finger proteins with femtomolar dissociation constants. Proc Natl Acad Sci U S A **95:**2812-7.
13. **Kim, J. S., and C. O. Pabo.** 1997. Transcriptional repression by zinc finger

16. **Pabo, C., and R. T. Sauer.** 1992. Transcription factors: structural families and principles of DNA recognition. Annu Rev Biuochem **61:**1053-1095.
17. **Paz, I., L. Abramovitz, and M. Choder.** 1999. Starved Saccharomyces cerevisiae cells have the capacity to support internal initiation of translation. J Biol Chem **274:**21741-5.
18. **Pomerantz, J. L., S. A. Wolfe, and C. O. Pabo.** 1998. Structure-based design of a dimeric zinc finger protein. Biochemistry **37:**965-70.
19. **Ruiz i Altaba, A., H. Perry O' Keefe, and D. A. Melton.** 1987. EMBO J **6:**3065-3070.
20. **Wang, B. S., and C. O. Pabo.** 1999. Dimerization of zinc fingers mediated by peptides evolved in vitro from random sequences. Proc Natl Acad Sci U S A **96:**9568-73.
21. **Wolfe, S. A., H. A. Greisman, E. I. Ramm, and C. O. Pabo.** 1999. Analysis of zinc fingers optimized via phage display: evaluating the utility of a recognition code. J Mol Biol **285:**1917-34.
22. **Zhu, J., M. L. Musco, and M. J. Grace.** 1999. Three-color flow cytometry analysis of tricistronic expression of eBFP, eGFP, and eYFP using EMCV-IRES linkages. Cytometry **37:**51-9.