BMC Biotechnology

Methodology article

A family of *E. coli* expression vectors for laboratory scale and high throughput soluble protein production

Lisa D Cabrita, Weiwen Dai and Stephen P Bottomley*

Address: Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, 3800, Australia

Email: Lisa D Cabrita - lisa.cabrita@med.monash.edu.au; Weiwen Dai - weiwen.dai@med.monash.edu.au; Stephen P Bottomley* - steve.bottomley@med.monash.edu.au

* Corresponding author

Published: 01 March 2006

*BMC Biotechnology*2006, **6**:12 doi:10.1186/1472-6750-6-12

Received: 15 September 2005
Accepted: 01 March 2006

This article is available from: http://www.biomedcentral.com/1472-6750/6/12

Abstract

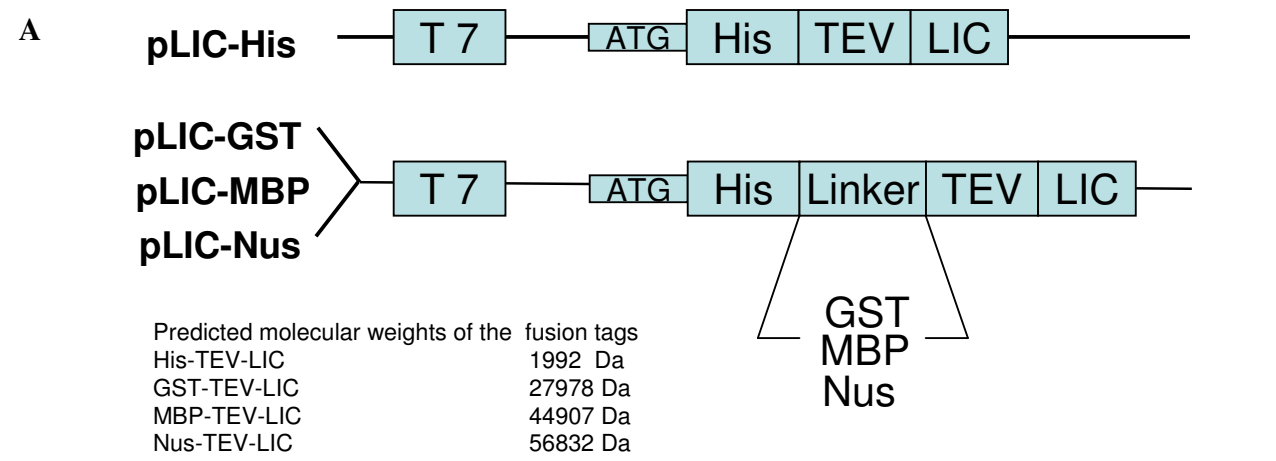
Background: In the past few years, both automated and manual high-throughput protein expression and purification has become an accessible means to rapidly screen and produce soluble proteins for structural and functional studies. However, many of the commercial vectors encoding different solubility tags require different cloning and purification steps for each vector, considerably slowing down expression screening. We have developed a set of *E. coli* expression vectors with different solubility tags that allow for parallel cloning from a single PCR product and can be purified using the same protocol.

Results: The set of *E. coli* expression vectors, encode for either a hexa-histidine tag or the three most commonly used solubility tags (GST, MBP, NusA) and all with an N-terminal hexa-histidine sequence. The result is two-fold: the His-tag facilitates purification by immobilised metal affinity chromatography, whilst the fusion domains act primarily as solubility aids during expression, in addition to providing an optional purification step. We have also incorporated a TEV recognition sequence following the solubility tag domain, which allows for highly specific cleavage (using TEV protease) of the fusion protein to yield native protein. These vectors are also designed for ligation-independent cloning and they possess a high-level expressing T7 promoter, which is suitable for auto-induction. To validate our vector system, we have cloned four different genes and also one gene into all four vectors and used small-scale expression and purification techniques. We demonstrate that the vectors are capable of high levels of expression and that efficient screening of new proteins can be readily achieved at the laboratory level.

Conclusion: The result is a set of four rationally designed vectors, which can be used for streamlined cloning, expression and purification of target proteins in the laboratory and have the potential for being adaptable to a high-throughput screening.

Background

The establishment of rapid expression and purification procedures for recombinant proteins has become a major challenge for biotechnology and indeed any laboratory which studies proteins. The genomic and structural genomic communities have driven the development of high-throughput cloning, expression and purification technologies to a large extent. Recent developments including filter-plate based assays for cloning, expression and purification, ligation-independent cloning (LIC) [1],



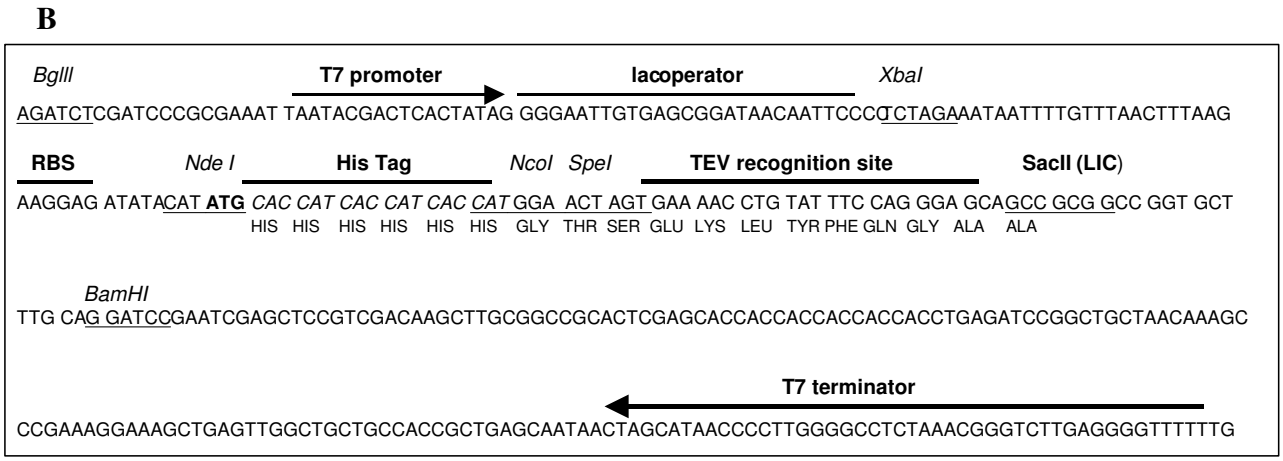


Figure 1
Basic expression vector design. (A) The basic vector design for the pLIC vectors. (B) Shows the nucleotide sequence for the features common to all the vectors. The pLIC-His vector is identical in sequence except it does not contain the NcoI-SpeI restriction enzyme sites between the His tag and the TEV recognition sequence.

Gateway technology [2] and auto-induction for protein expression [3] are now readily coupled to robotic pipelines that have made the parallel production of proteins a relatively simple, cost-effective approach. *E. coli* remains the system of first-choice for expressing proteins, as it is cheap and easy to handle, however many mammalian proteins cannot be successfully expressed in *E. coli* [4]. This leaves the researcher to either explore expression space using a range of alternative *E. coli* strains, different temperatures, solubility tags or choose an alternative expression host [5,6].

There are numerous commercial and non-commercial *E. coli* expression vectors available that incorporate fusion tags (for both purification and enhanced solubility), Gateway or LIC technologies, protease cleavage sequences and regulable expression levels [7]. These vectors are generally designed to be used in diverse expression/purification applications and thus often contain extraneous sequences at either the N- or C-termini of the protein. In addition, the use of different fusion tags mean that alternative purification protocols are required for each protein expressed from a different vector, increasing the time it takes to find

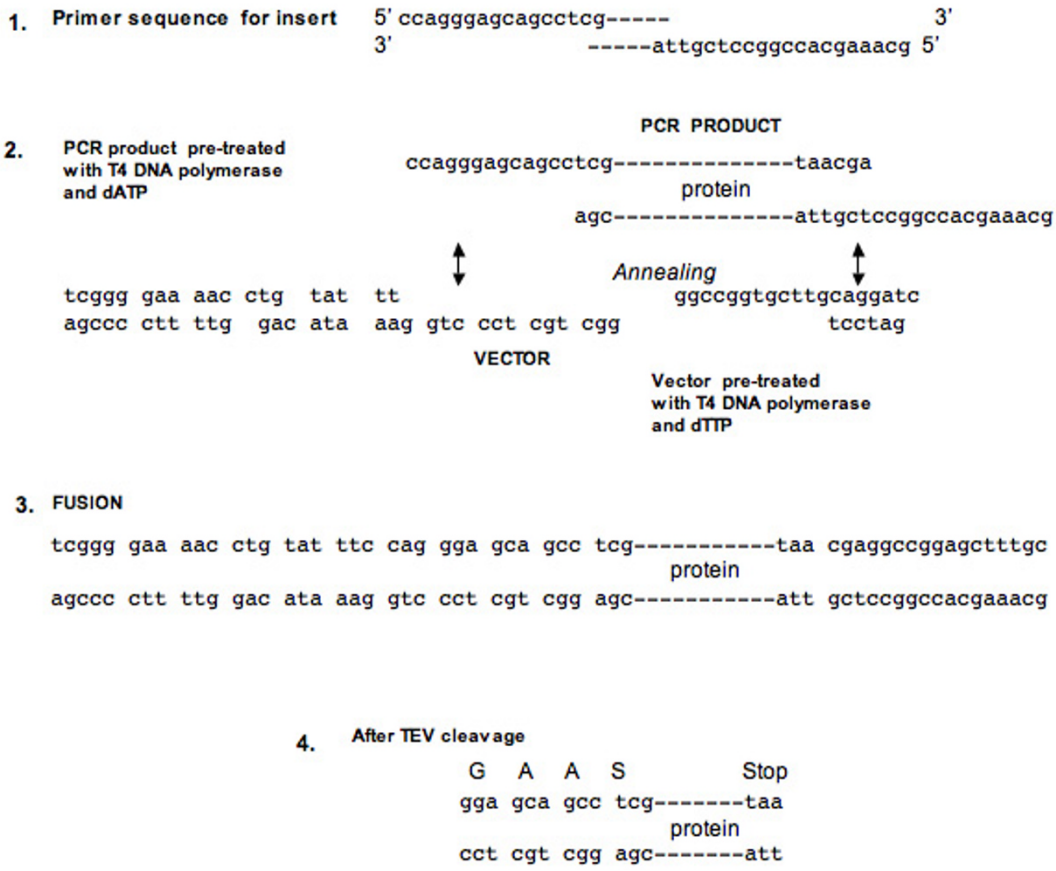


Figure 2
Ligation-Independent Cloning (LIC) Methodology. Shows the general procedure for LIC (1) The primer sequences required for amplification of the gene of interest. (2) Both the vector and insert are subject to treatment with T4 DNA polymerase, in the presence of either dTTP (vector) or dATP (insert). (3) The result after successful annealing. (4) After cleavage with TEV protease, the tags are removed and the protein retains 4 amino acids (GAAS).

successful expression and purification conditions. As a result of these limitations, we set out to create a set of *E. coli* expression vectors that incorporate these essential features: i) protein expressed from each vector is initially purified in the same manner; ii) the cloning procedure is the same for all vectors and iii) a range of solubility tags are available. Moreover these vectors have been developed for use at the general laboratory level so that a researcher can rapidly screen a range of expression conditions.

Here, we describe a set of four rationally designed expression vectors for *E. coli*. The first, (pLIC-His) encodes for a hexa-His tag followed by a TEV recognition sequence. The other three encode for three common solubility tags [MBP [8], NusA [9], GST [10] plus a TEV recognition sequence. TEV protease was selected due its superior specificity and adaptability to a range of buffering conditions [11]. The vectors are LIC compatible, and contain an N-terminal hexa-His tag enabling parallel cloning and purification, expression is via a T7 promoter allowing for auto-induction and they all possess a TEV recognition site between the fusion partner and the gene of interest. These features make the vectors ideally suitable for high-throughput cloning and expression screening.

We demonstrate that soluble expression was observed for a range of proteins and that they could be readily digested with TEV to yield the native protein. In addition, we illustrate the relevance of parallel cloning in deciding on a solubility tag that is suitable for a target protein. This demonstrates that the vectors can be used successfully at the laboratory level to rapidly screen target proteins (or

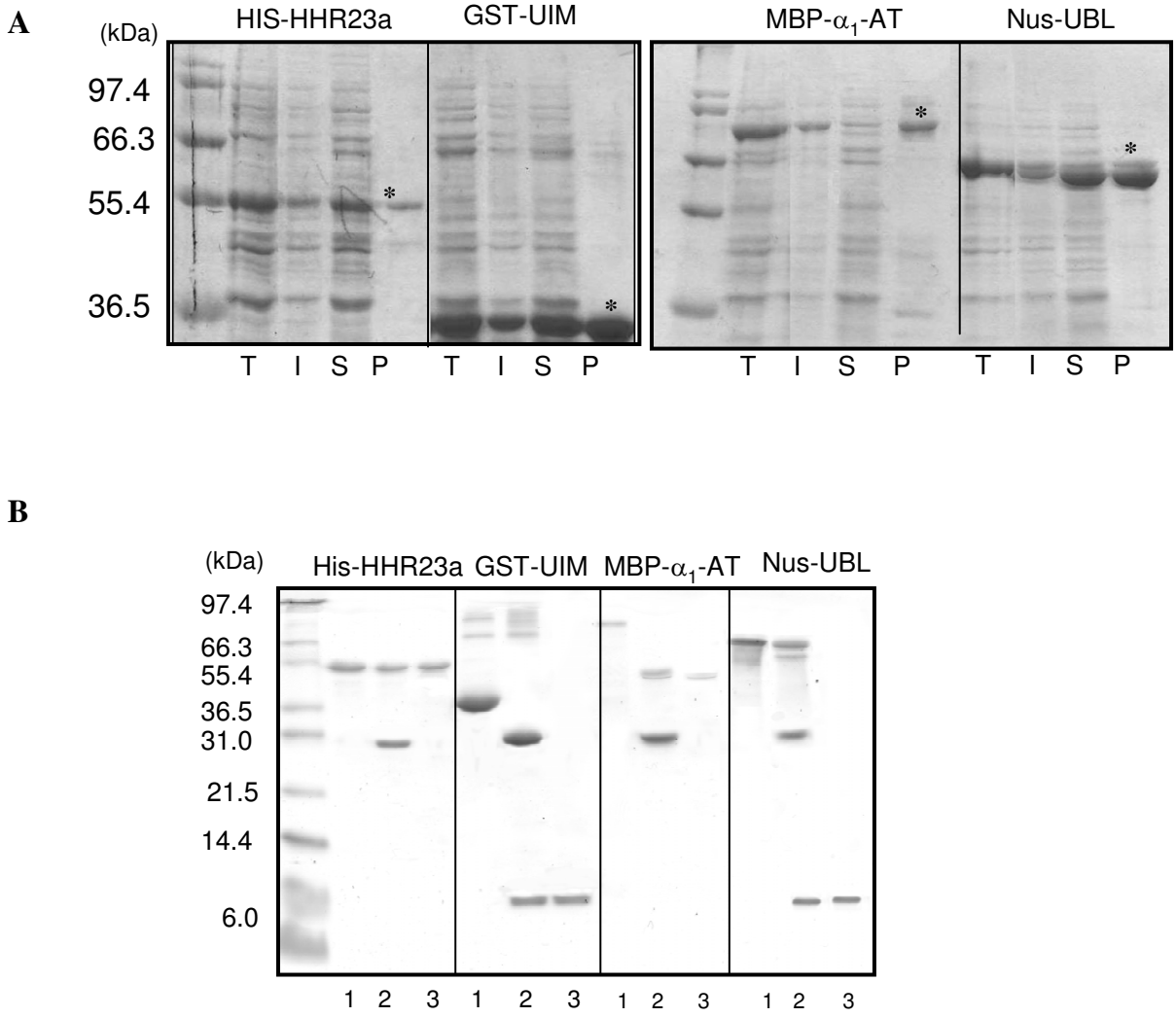


Figure 3

Small scale expression, purification and cleavage of recombinant proteins. (A) SDS-PAGE analysis of the total cell lysate (T), insoluble (I), soluble (S), and the purified (P) fractions for each of the fusions: His-HHR23a, GST-UIM, MBP-α_1-AT and Nus-UBL (the protein is marked with an asterisk). (B) SDS-PAGE analysis of each fusion during TEV protease digestion, (1) intact fusion (2) TEV digestion (3) purified protein after digestion. The band around 30 kDa observed in lane 2 of each experiment corresponds to the TEV protein.

their mutants) and can be readily adapted to the high-throughput process.

Results and discussion

Construction of the pLIC vectors and His-HHR23A, GST-UIM, MBP-α_1-AT and Nus-UBL

Several vector systems have been described that incorporate a variety of backbones, promoters, tag positions and cloning procedures. One drawback from several of these studies is the lack of consistency within the vector systems which does not allow a systematic study of protein expression [12-15]. Thus the aim of this study was to generate a set of vectors which enable parallel cloning of a given target(s) and enables a comparative analysis of the expression yields using different solubility tags. As a result, the pLIC set of prokaryotic vectors were created using pET21b(+) as their host. An oligonucleotide cassette was used to introduce the His tag, TEV site and LIC site as well as two unique restriction sites (NcoI-SpeI). These unique sites were used to introduce either the genes encoding for the GST, MBP or Nus tags and could in principle, be used to introduce alternative tags. The four vectors, termed pLIC-His, pLIC-GST, pLIC-MBP and pLIC-Nus were created as shown in Figure 1. It is to our knowledge, the first

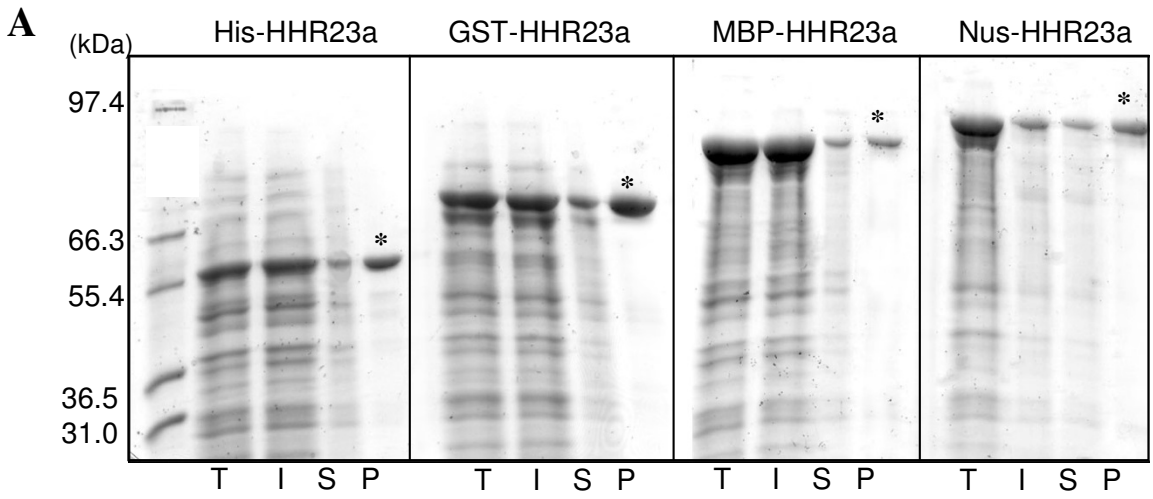


Figure 4
Parallel expression, purification and cleavage of HHR23A. HHR23A was cloned into all 4 vectors and the constructs were expressed and purified as described in the materials and methods. SDS-PAGE analysis of the total cell lysate (T), insoluble (I), soluble (S), and the purified (P) fractions for each of the fusions are shown.

set of *E. coli* expression vectors derived from the same backbone, with the only variation being in the solubility tag present. The genes HHR23A, UIM, α_1-AT and UBL were subsequently amplified and successfully introduced into each of the vectors using LIC. In addition, HHR23A was cloned into all vectors simultaneously to monitor the effect that different tags had on the solubility of this particular protein.

Two recently developed ligation methods (Gateway technology (Invitrogen) and LIC) have been used extensively by the structural genomics community. Both techniques offer the opportunity for parallel cloning, such that multiple targets can be easily cloned into several different vectors. With the LIC strategy in particular, target genes are amplified with 12–15 bp end sequences and the 5'-3' endonuclease activity of T4 DNA polymerase is then exploited [1,16] (Figure 2). This digestion event generates overhangs in the PCR product that allows subsequent insertion into a complementarily digested vector using a simple annealing step. As there is no need for restriction enzymes or ligation steps, LIC is a highly efficient and cost effective means of cloning. It also has the advantage that it does not require expensive reagents or multiple manipulations.

The LIC technique was found to be an efficient process in our hands, whereby up to 80% of the colonies screened were found to be positive. We observed, however, that success in LIC lies in the quality of the linearized vector, which was dependent, in part, on using a vast excess of SacII enzyme. Gel purification was used in the preparation of the vector and insert and it serves three purposes: by removing unwanted (and possibly active) DNA polymerase, removing excess dNTPs and selecting the correct PCR product over other misprimed events or linearized vector over undigested material. In a high-throughput format, a sizeable quantity of PCR product would be generated (and can be verified by sequencing) prior to parallel cloning into the vectors.

Expression and purification of the constructs

Each of the constructs was transformed into the *E. coli* host BL21(DE3) and expressed in 2 ml of auto-induction expression media. Following expression, the cell density of each of the cultures were normalised to an $OD_{600\ nm}$ of 5.0 prior to lysis, which allowed for a uniform number of cells between each sample. The cultures were chemically lysed using PopCulture (Novagen), which is a chemical method for lysis and particularly useful for multiple samples. After lysis, the cell lysates were subjected to a solubil-

ity assay (as described in the Materials and Methods) to assess the relative proportions of soluble and insoluble expression. As shown in Figure 3A, all of the constructs were found to have soluble expression, as judged by presence of a respective bands at the predicted molecular weights. Both His-HHR23A and MBP-α_1-AT have marginally greater proportions of insoluble material, whilst Nus-UBL and GST-UIM have slightly elevated proportions of soluble material. Optimization of soluble yields was not performed in this study, as this can vary greatly from protein to protein. It has been recently been shown however, that a lower temperature used during auto-induction may improve the yield of soluble material [3]. The soluble expression and purification of MBP-α_1-AT was particular interesting, as α_1-AT is generally produced in very high yields in an insoluble form and purified using a refolding method [17,18].

In addition to screening different targets in each of the vectors, we also cloned a single gene (HHR23A) into all four vectors. As can be seen in Figure 4, HHR23A expressed largely as inclusion bodies, however soluble material was obtained in all four vectors. More importantly however, is that the tags had differing effects on the expression of HHR23A. In this instance, GST was found to be the most efficient in improving the expression whereas NusA had little additional benefit (Figure 4). This illustrates therefore, the usefulness of parallel cloning and expression, as a protein may express more favourably with a different tag.

An additional observation of the small-scale expression was the importance for a purification step (Ni-NTA). In some cases, soluble expression of the protein showed a very faint band in the soluble fraction, but was purified at more detectable quantities (see α_1-AT band, Figure 3A). This however, appears to be dependent on the degree of soluble expression, as the UBL and UIM constructs for instance, showed very clear results in the solubility assay (Figure 3A). The use of the solubility assay in isolation should thus be used with caution, particularly for a poorly expressing protein that may produce unclear results. In some cases, it may be useful to complement the assay with a Western blot.

TEV cleavage

Generally, affinity tags are removed after purification as they can interfere with either the protein's function or with downstream processes. The most common example is if the protein is to be used for crystallographic or NMR studies. While there have been reports of proteins being crystallized with the assistance of an affinity tag [19-21] the general view is that they should be removed [22]. Each of the constructs described above has a TEV cleavage site which can be used to remove the affinity tags, leaving the native protein with 4 vector-derived residues (GAAS) (Figure 1). This is comparable to another LIC vector available which describes 3 vector-derived residues (S-N-A) [14]. Indeed the introduction of non-native amino acids at the N- terminus is not ideal and it is a common problem that generally exists both for the LIC and Gateway technologies. As a result, the effect of such residues (both in number and composition) on the expression, purification and integrity of a given protein must be verified experimentally [12].

To validate the TEV cleavage site, each of the proteins were digested with his-tagged TEV protease and purified. As seen in Figure 3B, after cleavage three species are present: the solubility tag (26 kDa GST, 42 kDa MBP, 54 kDa NusA), the his-tagged TEV protease (30 kDa) and the protein of interest. This mixture can be purified further to remove the his-tagged TEV protease and the solubility tag by selectively binding them to a Ni-NTA column. The protein of interest, on the other hand remains unbound to the Ni resin as it lacks a His tag (Figure 3B).

It is should be noted that although a given protein target may express solubly, its behaviour in the absence of the solubility tag may be very different, as this is a function of other factors (structure, amino acid sequence, aggregation potential, buffering conditions). Thus the integrity of the protein after TEV cleavage should be verified using other methods (dynamic light scattering, size exclusion chromatography, circular dichroism etc.).

Conclusion

Here, we have presented a set of four rationally designed T7-based *E. coli* expression vectors whose features include the incorporation of solubility tags to assist in expression/purification, a TEV cleavable sequence and a LIC sequence. We have shown soluble expression and purification of four different proteins as well the ability to remove the tags using TEV protease. These vectors allow for rapid parallel cloning and expression and are suitable for either laboratory-scale screening of target proteins, or alternatively for high-throughput screening.

Methods

Construction of pLIC-HIS and pLIC-MGN

The vector pET21b(+) (Novagen) was used as the basis for the pLIC vectors. pET21b(+) was manipulated to produce two different constructs, pLIC-His and pLIC-MGN. For the pLIC-HIS construct, the vector's unique NdeI-BamHI restriction sites were used. An oligonucleotide cassette containing the His-tag, TEV and LIC sites was created and ligated into the vector which had been previously digested with NdeI-BamHI (Figure 1). The cassette was comprised of the following oligonucleotides: 5'-tatgcaccatcaccatcaccatgaaaacctgtatttccagggagcagccgcggccg-

gtgctttgcag-3' and 3'-acgtggtagtggtagtggtacttttggacataaaggtc-cctcgtcggcgccggccacgaaacgtcctag-5'. pLIC-MGN was created by utilising the same NdeI-BamHI sites, but incorporating a cassette of the following oligonucleotides: 5'-tatgcaccatcaccatcaccatggaactagtgaaaacctgtatttccagggag cagccgcggccggtgctttgcag-3' and 3'-acgtggtagtggtagtggtaccttt-gatcacttttggacataaagg tccctcgtcggcgccggccacgaaacgtcctag-3'. This cassette encodes the His tag, a unique NcoI site, a unique SpeI site, TEV recognition sequence and the LIC site (Figure 1). It was also ligated into pET21b(+).

Construction of pLIC-GST, pLIC-MBP and pLIC-Nus

The introduction of the unique NcoI and SpeI site in pLIC-MGN allowed for the introduction of the solubility tags: GST, MBP and NusA. The GST gene was cloned from the vector pET41 (Novagen), the MBP gene from pMAL-2cx (New England Biolabs) and NusA from pET43.1 (Novagen). The oligonucleotides used for the amplification are presented in Additional file 1. As the NusA gene already contained a NcoI site, the QuikChange site-directed mutagenesis technique (Stratagene) was used to remove this site, using the vendor's instructions. Briefly, the plasmid was amplified using a pair of complementary primers (See Additional file 1) and using the polymerase *Pfu*. The PCR product was then digested with 10 u of DpnI, followed by transformation into competent DH5α *E. coli*.

The genes encoding GST, MBP and NusA were amplified and each of the PCR products were subjected to sequential digestion with NcoI and SpeI followed by gel purification and ligation into the pLIC-MGN vector. This yielded three different constructs termed: pLIC-GST, pLIC-MBP and pLIC-Nus.

Cloning of UIM, α_1-AT, HHR23A and UBL into the vectors

To clone into the expression vectors, PCR products of the gene of interest need to be generated with specific overhangs that are complementary to the sequence of the vector (Figure 2). After cleavage with TEV, the native protein will have four vector-derived residues (G-A-A-S). To validate the vectors, four genes were chosen to be expressed: HHR23A was cloned into pLIC-His, ubiquitin interacting motif (UIM) into pLIC-GST, α_1-antitrypsin (α_1-AT) in pLIC-MBP and ubiquitin-like domain from HHR23A (UBL) into pLIC-NUS.

The LIC site consists of a unique SacII site, which can be used to linearize the vector and thus permit T4 DNA polymerase treatment to take place (Figure 2). 30 units of SacII (New England Biolabs) was used to digest 1 μg of the target vector in a total reaction volume of 50 μl for a total of 3 hours. The linearized vector was then separated and gel purified from an agarose gel and the DNA concentration determined using UV-spectroscopy. 0.2 pmol of DNA was then treated with T4 DNA Polymerase (1 u/0.1 pmol DNA) and 2.5 mM dTTP at 22°C for 30 minutes followed by heat inactivation for 20 minutes.

Each of the four genes was amplified using *Pfu* polymerase (Promega) using the oligonucleotides detailed in Additional file 1. The amplification was for 30 cycles with the following conditions: 95°C 1 min, X°C 1 min, 72°C for Ymins. The annealing temperatures (X) and extension times (Y) for each of the constructs are shown in Additional file 1. After amplification, the PCR products were gel purified prior to being treated with T4 DNA Polymerase (1 u/0.1 pmol) and 2.5 mM dATP at 22°C for 30 mins, followed by a 20 min heat inactivation step.

For the annealing process, 1 μl of T4 treated vector was mixed with 2 μl of T4 treated PCR product and allowed to incubate at 22°C for 1 hour. 1 μl of 25 mM EDTA pH 8.0 was then added and the reaction incubated for 5 mins. The annealed mix was then transformed into competent JM107 *E. coli*. The colonies were assessed by colony PCR and the clones were verified using DNA sequencing.

Expression and purification of His-HHR23A, GST-UIM, MBP-α_1AT and Nus-UBL

Each of the clones were transformed into competent BL21(DE3) *E. coli* cells and streaked onto plates. A single colony was selected and expressed in 2 ml of Overnight Express Media (Novagen) for 18 hours at 30°C, 300 rpm. The $OD_{600\,nm}$ was recorded for cells after growth (typically ranging between 8–11) and the absorbance was normalised to $OD_{600\,nm}$ = 5.0 for each, which allowed for a uniform number of cells between samples. The cells were then lysed with 0.2 ml of PopCulture supplemented with 2 μl of lysozyme and benzoase, as per manufacturer's recommendation (Novagen). 200 μl of each cell lysate was retained for the Solubility Assay (see below). The remaining material was bound batch wise, to 200 μl of Ni-NTA resin which had been previously equilibrated in buffer (25 mM $NaPO_4$, 500 mM NaCl pH 8.0). The slurry was then collected in a 2 ml filter plate (supplied with the solubility assay kit, Novagen) and washed with 4 ml of buffer supplemented with 25 mM imidazole. The proteins were then eluted with 250 mM imidazole in 25 mM $NaPO_4$, 50 mM NaCl pH 8.0 in a 200 μl volume.

Solubility assay

The solubility assay (Novagen) was used to determine the relative proportions of soluble and insoluble material that were expressed for each of the constructs. After lysis with PopCulture, 200 μl samples of each lysate was passed through the wells of a 0.2 μm filter plate (supplied by the manufacturer) and collected. Here, the soluble protein passes through the plate whilst the insoluble inclusion bodies are retained. 200 μl of 4%(w/v) SDS was then

added to each well and following a 10 min incubation the insoluble proteins eluted. A sample of both the soluble and insoluble fractions were retained for SDS-PAGE analysis.

TEV protease cleavage

The concentration of protein was estimated using the BCA Assay (Pierce). Protein at a concentration of 0.5 mg/ml was incubated with 0.05 mg/ml of TEV and allowed to incubate at 4°C for 16 hours. The sample was then run on a SDS gel to check for the completion of cleavage

Authors' contributions

LC carried out the cloning and expression studies, participated in the design of the study and drafted the manuscript. WD participated in the molecular biology. SB conceived of the study, and participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Additional material

Additional File 1

Oligonucleotide sequences used to isolate the genes for GST, MBP and Nus, as well as the annealing and extension times used for amplification.

[http://www.biomedcentral.com/content/supplementary/1472-6750-6-12-S1.doc]

Acknowledgements

The authors would like to thank members of the lab for useful discussions and also to Peter Howley for providing the plasmid containing the HHR23A gene. S.P.B is a NH&MRC R.D. Wright Fellow and a Monash University Logan Senior Research Fellow. L.D.C is a NH&MRC CJ Martin Fellow.

References

1. Aslanidis C, de Jong PJ: **Ligation-independent cloning of PCR products (LIC-PCR).** *Nucleic Acids Res* 1990, **18(20):**6069-6074.
2. Hartley JL, Temple GF, Brasch MA: **DNA cloning using in vitro site-specific recombination.** *Genome Res* 2000, **10(11):**1788-1795.
3. Studier FW: **Protein production by auto-induction in high-density shaking cultures.** *Protein Expr Purif* 2005, **41(1):**207-234.
4. Stevens RC: **Design of high-throughput methods of protein production for structural biology.** *Structure Fold Des* 2000, **8(9):**R177-85.
5. Buckle AM, Devlin GL, Jodun RA, Fulton KF, Faux N, Whisstock JC, Bottomley SP: **The matrix refolded.** *Nat Methods* 2005, **2(1):**3.
6. Cabrita LD, Bottomley SP: **Protein expression and refolding - A practical guide to getting the most out of inclusion bodies.** *Biotechnol Annu Rev* 2004, **10:**31-50.
7. Hunt I: **From gene to protein: a review of new and enabling technologies for multi-parallel protein expression.** *Protein Expr Purif* 2005, **40(1):**1-22.
8. di Guan C, Li P, Riggs PD, Inouye H: **Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein.** *Gene* 1988, **67(1):**21-30.
9. De Marco V, Stier G, Blandin S, de Marco A: **The solubility and stability of recombinant proteins are increased by their fusion to NusA.** *Biochem Biophys Res Commun* 2004, **322(3):**766-771.
10. Smith DB, Johnson KS: **Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase.** *Gene* 1988, **67(1):**31-40.
11. Phan J, Zdanov A, Evdokimov AG, Tropea JE, Peters HK, Kapust RB, Li M, Wlodawer A, Waugh DS: **Structural basis for the substrate specificity of tobacco etch virus protease.** *J Biol Chem* 2002, **277(52):**50564-50572.
12. Dummler A, Lawrence AM, de Marco A: **Simplified screening for the detection of soluble fusion constructs expressed in E. coli using a modular set of vectors.** *Microb Cell Fact* 2005, **4:**34.
13. Korf U, Kohl T, van der Zandt H, Zahn R, Schleeger S, Ueberle B, Wandschneider S, Bechtel S, Schnolzer M, Ottleben H, Wiemann S, Poustka A: **Large-scale protein expression for proteome research.** *Proteomics* 2005, **5(14):**3571-3580.
14. Stols L, Gu M, Dieckman L, Raffen R, Collart FR, Donnelly MI: **A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site.** *Protein Expr Purif* 2002, **25(1):**8-15.
15. Woestenenk EA, Hammarstrom M, van den Berg S, Hard T, Berglund H: **His tag effect on solubility of human proteins produced in Escherichia coli: a comparison between four expression vectors.** *J Struct Funct Genomics* 2004, **5(3):**217-229.
16. Aslanidis C, de Jong PJ, Schmitz G: **Minimal length requirement of the single-stranded tails for ligation-independent cloning (LIC) of PCR products.** *PCR Methods Appl* 1994, **4(3):**172-177.
17. Bottomley SP, Stone SR: **Protein engineering of chimeric Serpins: an investigation into effects of the serpin scaffold and reactive centre loop length.** *Protein Eng* 1998, **11(12):**1243-1247.
18. Bird PI, Pak SC, Worrall DM, Bottomley SP: **Production of recombinant serpins in Escherichia coli.** *Methods* 2004, **32(2):**169-176.
19. Chao JA, Williamson JR: **Joint X-ray and NMR refinement of the yeast L30e-mRNA complex.** *Structure* 2004, **12(7):**1165-1176.
20. Ke A, Wolberger C: **Insights into binding cooperativity of MATa1/MATalpha2 from the crystal structure of a MATa1 homeodomain-maltose binding protein chimera.** *Protein Sci* 2003, **12(2):**306-312.
21. Smyth DR, Mrozkiewicz MK, McGrath WJ, Listwan P, Kobe B: **Crystal structures of fusion proteins with large-affinity tags.** *Protein Sci* 2003, **12(7):**1313-1322.
22. Bucher MH, Evdokimov AG, Waugh DS: **Differential effects of short affinity tags on the crystallization of Pyrococcus furiosus maltodextrin-binding protein.** *Acta Crystallogr D Biol Crystallogr* 2002, **58(Pt 3):**392-397.

Biotechnol Lett (2009) 31:543–549
DOI 10.1007/s10529-008-9909-9

ORIGINAL RESEARCH PAPER

Construction of a dual-tag system for gene expression, protein affinity purification and fusion protein processing

Hassan Motejadded · Josef Altenbuchner

Received: 31 October 2008 / Accepted: 5 December 2008 / Published online: 7 January 2009

Abstract An *E. coli* vector system was constructed which allows the expression of fusion genes via a L-rhamnose-inducible promotor. The corresponding fusion proteins consist of the maltose-binding protein and a His-tag sequence for affinity purification, the *Saccharomyces cerevisiae* Smt3 protein for protein processing by proteolytic cleavage and the protein of interest. The Smt3 gene was codon-optimized for expression in *E. coli*. In a second rhamnose-inducible vector, the *S. cerevisiae* Ulp1 protease gene for processing Smt3 fusion proteins was fused in the same way to maltose-binding protein and His-tag sequence but without the Smt3 gene. The enhanced green fluorescent protein (eGFP) was used as reporter and protein of interest. Both fusion proteins (MalE-6xHis-Smt3-eGFP and MalE-6xHis-Ulp1) were efficiently produced in *E. coli* and separately purified by amylose resin. After proteolytic cleavage the products were applied to a Ni-NTA column to remove protease and tags. Pure eGFP protein was obtained in the flow-through of the column in a yield of around 35% of the crude cell extract.

Keywords Gene expression · Protein affinity purification · Proteolytic cleavage · Ulp1 protease · SUMO protein

H. Motejadded · J. Altenbuchner (✉)
Institut für Industrielle Genetik, Universität Stuttgart, Allmandring 31, 70569 Stuttgart, Germany
e-mail: josef.altenbuchner@iig.uni-stuttgart.de

Introduction

The purification of proteins for enzyme characterization, crystallization, therapeutic applications etc. is a tedious and time-consuming process. The fusion of tags to recombinant genes for affinity purification facilitates the purification and leads to higher yield and purity of proteins. One of the most commonly used tags are six consecutive histidine residues (His-tag) which bind to divalent metal cations immobilized on a matrix by a metal chelator (IMAG). The His-tag may be fused to the *N*- or *C*-terminal end of the protein. There are numerous other tags like Strep-tag, S-tag, calmodulin-binding tag, chitin or maltose binding protein or glutathione-*S*-transferase available (for reviews see: Uhlen and Moks 1990; Terpe 2003). Sometimes proteins are even fused to two tags (dual tags or tandem affinity tags) to improve the purification of a protein (Nallamsetty and Waugh 2007). Some of the tags like His-tag are very small, others like maltose binding protein very bulky. Depending on the tag, fusion to the *N*- or *C*-terminal end of a protein or distance to the protein the tags may influence the properties of the proteins like enzyme activity or crystallization and for proteins destined for human therapeutic use they are not acceptable at all because of undesirable immune response. This means that for various purposes tags can not be used or have to be removed after purification. Several proteases are used for cleavage of the fusion proteins. These are mainly factor Xa, enterokinase or TEV

protease (Sassenfeld 2008). For cleavage, the recognition sequence for the proteases has to be inserted between the tag and the protein of interest. There are three main problems reported: (i) the distance between tag and protein of interest is too small and inefficiently cleaved; (ii) there are other related recognition sequences within the protein of interest which are recognized by the protease and (iii) some proteases do not remove their recognition sequence completely leading to artificial ends in the protein of interest.

An alternative to conventional proteases are SUMO proteases. SUMO proteins (small ubiquitin-related modifier) modulate protein functions in the cell through covalent modifications (Vertegaal 2007). An example is the *Saccharomyces cerevisiae* Smt3 protein, an essential protein which is involved in septin formation, chromosomal segregation and progression of the cell cycle (Li and Hochstrasser 1999; Tanaka et al. 1999; Johnson and Blobel 1999). Smt3 is first proteolytically processed at a conserved Gly-Gly motif near the *C*-terminal end of the protein by the Ulp1 cysteinyl protease, another essential enzyme of *S. cerevisiae* (Li and Hochstrasser 1999). The mature protein is covalently bound via a high-energy thioester linkage to an E1 enzyme. From here it is transferred to the E2 enzyme and finally to a lysine ε-amino group of the target protein. The sumoylation of target proteins modifies their activities in a reversible way since the Ulp1 protease is able to cleave the peptide bond between the two glycin residues of Smt3 and the lysine ε-amino group.

The great advantage of the Smt3-Ulp1 proteolytic system is the high specificity of the Ulp1 protease which recognizes the three dimensional structure of Smt3 and not just a short amino acid sequence. A commercial product was developed consisting of a His-tagged Smt3 and Ulp1 protease gene expressed under a T7 promotor system (LifeSensors Inc.). In the following work, we describe an improvement to this system by constructing dual-tagged Smt3 and Ulp1 genes for improved recovery of the protein of interest. Further modifications concern an *E. coli* codon-optimized version of the Smt3 gene to ameliorate fusion gene expression and the use of the tightly regulated *E. coli* rhamnose-inducible promotor to improve the stability of the production strains when genes with toxic effects are inserted in the expression vectors.

Material and methods

Bacterial strain and growth conditions

E. coli JM109 (Yanisch-Perron et al. 1985) was used throughout this study. The strain was grown on Luria-Bertani (LB) agar and in LB broth (Luria et al. 1960) at 37°C. Plasmids were selected by adding ampicillin (100 mg/l). For L-rhamnose induction an overnight culture was diluted in 50 ml LB broth with 100 mg/l ampicillin to 0.05 OD_{650} and grown at 37°C to 0.4 OD_{650} on a rotary shaker at 200 rpm. Then 0.2% (w/v) L-rhamnose was added and the culture incubated overnight (about 16 h) at 30°C until the cells were harvested by centrifugation.

Protein purification

L-Rhamnose-induced cells from a 50 ml culture were harvested by centrifugation (Sorvall SS34, 10 min, 7,000 rpm, 4°C), washed with cold 0.1 M $NaPO_4$, pH 7.0 and resuspended in the same buffer (20 ml). The cells were disrupted in a cooled high-pressure homogenizer (EmulsiFlex-C5, Avestin, Canada) and cell debris removed by centrifugation (15,000*g*, 30 min, 4°C). All further steps were done at room temperature. The supernatant was added to 4 ml amylose resin (New England Biolabs GmbH) in a column by using gravity flow. The column was washed 6 times with 4 ml $NaPO_4$ buffer each and finally the MalE fusion protein eluted with 4 × 1 ml $NaPO_4$ buffer supplemented with 20 mM maltose. The three elution fractions containing the majority of the MalE-6xHis-Smt3-eGFP-fusion protein were combined (3 ml) and treated for 1 h at 30°C with 60 µl of MalE-6xHis-Ulp1 protease purified in the same way on amylose resin. The protease cleavage reaction was added to 2 ml Ni-nitrilotriacetic acid (NTA)-agarose resin column (Quiagen GmbH) equilibrated with $NaPO_4$ buffer and the flow-through collected in fractions of 1 ml each. The fractions 2 and 3 containing high amounts of pure eGFP were combined.

Protein electrophoresis and protein determination

SDS-PAGE was done according to the method of Lämmli. Native protein was subjected to electrophoresis on a 9% polyacrylamide gel with native Tris/acetate/EDTA buffer. Protein concentrations were

determined by the method of Bradford using the Bio-Rad protein assay dye reagent and bovine serum albumin as standard.

Recombinant DNA techniques

Standard procedures were used for restriction enzyme analysis, ligation and transformation of *E. coli.* DNA sequencing was done by GATC Biotech, Germany. Plasmids were prepared by the QIAprep spin miniprep kit (Qiagen GmbH). Chromosomal DNA of the *Saccharomyces cerevisae* strain EGY48 (Gyuris et al. 1993) for amplification of the Smt3 and Ulp1 gene by PCR was prepared by boiling cells of an overnight culture for 5 min in water and removal of cell debris by centrifugation (10,000*g*, 5 min, room temperature). Standard PCR was done in 100 µl reactions using High Fidelity PCR Enzyme Mix (Fermentas). The Smt3 codon usage was optimized for *E. coli* with the software OPTIMIZER (Puigbo et al. 2007) on a web server. For synthesis of the 197 bp *C*-terminal *smt3* gene in an *E. coli* optimized for codon usage, 3 pmol each of the 10 oligonucleotides s3822–s3831 (see Table 1) were combined in 100 µl containing 10 µl 10-fold *Taq* buffer, 3 mM $MgCl_2$, 0.2 mM dNTP and 0.5 U *Taq* polymerase (New England BioLabs Inc.) and amplified in 15 cycles (45 s 92°C, 45 s 51°C and 45 s 72°C). One µl of this reaction mix was used again in a second PCR amplification containing 10 µl 10-fold *Taq* buffer, 3 mM $MgCl_2$, 0.2 mM dNTP and 1 pmol each of the two outmost primers s3822 and s3831. PCR was done under above conditions in 27 cycles. The fragment was separated on agarose gel and purified with GFX PCR DNA and band purification kit (GE Healthcare).

Measurement of fluorescence intensity

To determine the amount of eGFP in crude extract and purification fractions the protein solutions were diluted in 0.1 M $NaPO_4$, pH 7.0 and three aliquots of 100 µl were added to microplates (Greiner, Germany) and the fluorescence was measured with a GENios fluorometer (Tecan, Austria), at an excitation wavelength of 485 nm and an emission wavelength of 535 nm. Background fluorescence was obtained from $NaPO_4$ buffer in the microtiter plates.

Results

Construction of the plasmids pJOE4905.1 and pJOE4847.2

The plasmid pJOE4905.1 is a rhamnose-inducible expression vector derived from pTST101 (Stumpp

Table 1 Primers used in this work

S3684	AAA AAG GAT CCC TTG TTC CTG AAT TAA ATG AAA AAG	PCR amplification of *ulp1*
S3689	AAA AAA TGT ACA CTA TTT TAA AGC GTC GGT TAA AA	PCR amplification of *ulp1*
S3764	AAA AAA GGC GCC AGC AGG GCG AGG AGC TGT T	PCR amplification of *eGFP*
S3771	GGC TGA AAA TCT TCT CTC ATC	PCR amplification of *eGFP*
S3800	AAA AAA GGA TCC GAG GTC AAG CCA GAA GTC A	PCR amplification of *smt3*
S3801	AAA AAA AGC TTG GTA CCG GCG CCA CCA ATC TGT T	PCR amplification of *smt3*
s3822	AAG CTC TGA GAT CTT CTT CAA GAT CAA GAA AAC CA	Codon-optimized *smt3*
s3823	TTC CAT CAG CCG ACG CAG TGG GGT GGT TTT CTT GAT CTT GAA G	Codon-optimized *smt3*
S3824	TGC GTC GGC TGA TGG AAG CGT TCG CTA AAC GTC AGG GTA	Codon-optimized *smt3*
S3825	AGC GCA GGC TGT CCA TCT CTT TAC CCT GAC GTT TAG CGA	Codon-optimized *smt3*
S3826	GAT GGA CAG CCT GCG CTT TCT GTA TGA TG CAT TCG CAT T	Codon-optimized *smt3*
S3827	GGG GTC TGA TCC GCC TGA ATG CGA ATG CCA TCA TAC	Codon-optimized *smt3*
S3828	AGG CGG ATC AGA CCC CGG AAG ACC TGG ACA TGG AA	Codon-optimized *smt3*
S3829	GGC CTC AAT GAT ATC GTT ATC TTC CAT GTC CAG GTC TTC	Codon-optimized *smt3*
S3830	TAA CGA TAT CAT TGA GGC CCA TCG TGA ACA GAT TGG TG	Codon-optimized *smt3*
S3831	TTT TCT AGA AGC TTG GCG CCA CCA ATC TGT TCA CGA TG	Codon-optimized *smt3*
S3893	GAT CAC ACC ACC ATC ACC ATC ACC	His-Tag sequence
S3894	GAT CGG TGA TGG TGA TGG TGG TGT	His-Tag sequence

et al. 2000) and contains the *malE* gene fused with a His-tag sequence, the yeast *smt3* gene and eGFP. The *smt3* gene was amplified from *Saccharomyces cerevisae* with the primers s3800/s3801 (Table 1). The first non-essential 12 codons of the gene were deleted (Mossessova and Lima 2000) as well as the codons behind the two glycin codons representing the mature form of Smt3 (amino acids 13–101). Instead, a *Bam*HI and *Hin*dIII restriction site was added through the primers used for the PCR amplification. The restriction sites permitted the gene to be cloned into pJOE4056.1 (Stumpp et al. 2000) cleaved with the same enzymes. The *C*-terminal part of the cloned *smt3* gene showed a series of codons which are very rare in *E. coli* such as five AGA and one AGG Arg codons which might later affect translation efficiency of recombinant fusion genes (Altenbuchner and Mattes 2004). A 197 bp region from the *Bgl*II site down to the two glycine codons was synthesised in a codon usage adapted to *E. coli* (Fig. 1). In addition, the third base of last glycine codon was changed from T to C which gives a *Sfo*I restriction site and allows a precise fusion of recombinant genes to *smt3*. The fragment was synthesized with 10 overlapping and complementary oligonucleotides, which were mixed and amplified in a PCR reaction in three steps. The amplified fragment was cleaved with endoR *Bgl*II and *Hin*dIII and used to replace the corresponding fragment in the pJOE4056.1 plasmid containing the wild type sequence. As a reporter for gene expression and protein processing the eGFP gene was amplified with the primers s3764/s3771 and fused to the *C*-terminal end of the codon-optimized *smt3*. The primer s3764 added a *Sfo*I site to the *N*-terminal end of the eGFP and allowed a precise fusion of *smt3* and *egfp*. A fragment containing the *smt3-eGFP* fusion was isolated as a *Bam*HI-*Hin*dIII fragment and inserted into pTST101 cut with the same enzymes. This fused the *smt3-egfp* gene in frame to the *C*-terminal end of *malE*. In addition, the two complementary oligonucleotides S3893/S3894 were inserted into the *Bam*HI site between *malE* and *smt3* containing in-frame six histidine codons. This destroyed the *Bam*HI site. Since there was already a *Sfo*I site in the vector backbone, the plasmid was cut with *Nae*I and religated to the final plasmid pJOE4905.1. Hereby three small *Nae*I fragments were deleted together with the second *Sfo*I site.

The *S. cerevisiae* Ulp1 protease is a 621 amino acid protein with two domains. The proteolytic domain is

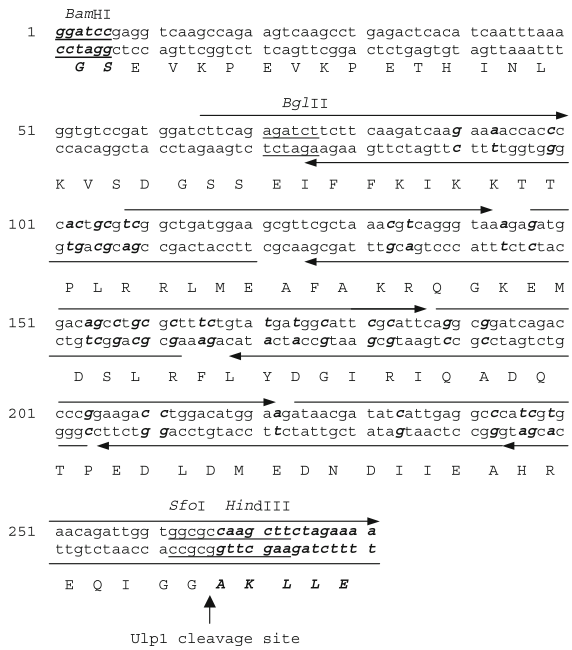


Fig. 1 Nucleotide and amino acid sequence of the codon-optimized *smt3* sequence used in vector pJOE4905.1. Nucleotides and amino acids in bold and italics indicate changes and additional sequences in comparison to the original sequence, *arrows* indicate the oligonucleotides used for the synthesis of the gene fragment, recognition sites for the restriction enzymes *Bam*HI, *Bgl*II, *Hin*dIII and *Sfo*I (isoschizomers *Nar*I, *Bbe*I, *Kas*I) are *underlined*

located in the *C*-terminal domain. Full proteolytic activity was obtained with Ulp1 fragments comprising residues 403–621 (Li and Hochstrasser 1999). Only this region was amplified from *S. cerevisae* with the primer s3684 and s3689. A *Bam*HI site in the upstream oligonucleotide and a *Bsr*GI site in the downstream oligonucleotide enabled the insertion of the PCR fragment into the vector pTST101 and in frame fusion with the *malE* gene. Again, the two complementary oligonucleotides s3893/s3894 with six his codons were inserted into the *Bam*HI site between MalE and Ulp1 encoding fragment. This new plasmid encoding a MalE-6xHis-Ulp1 fusion protein was called pJOE4847.2. The plasmids pJOE4905.1 and pJOE4847.2 are shown in Fig. 2.

Expression of MalE-His6-Smt3-eGFP, purification and protein cleavage

E. coli JM109 with the plasmid pJOE4905.1 was cultivated in 50 ml LB medium and after 2 h growth

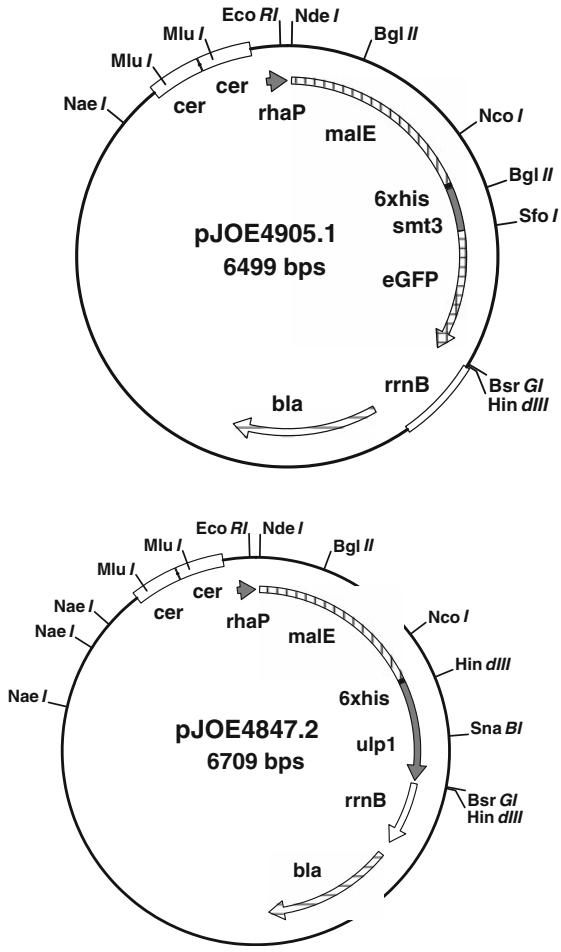


Fig. 2 Restriction maps of the plasmids pJOE4905.1 and pJOE4847.2

induced with L-rhamnose and incubated overnight at 30°C. The cells were harvested and crude extract prepared by high pressure homogenization. The crude extract was passed over amylose resin and the bound MalE fusion protein eluted with maltose. From each fraction protein content and fluorescence was determined. The crude extract of JM109 pJOE4905.1 contained 51.6 mg protein and showed a relative fluorescence (rF) of 1.48×10^8 rF. In the pooled elution fractions there was 9.81 mg protein producing 5.7×10^7 rF. JM109 pJOE4847.2 expressing the MalE-His-Ulp1 fusion protein was cultivated in the same way and the protein purified. The crude extract of JM109 pJOE4847.2 contained 18.6 mg protein and after the amylose resin there were 2.54 mg protein left.

To find out how much protease was needed for complete cleavage of MalE-6xHis-Smt3-eGFP 5 µl

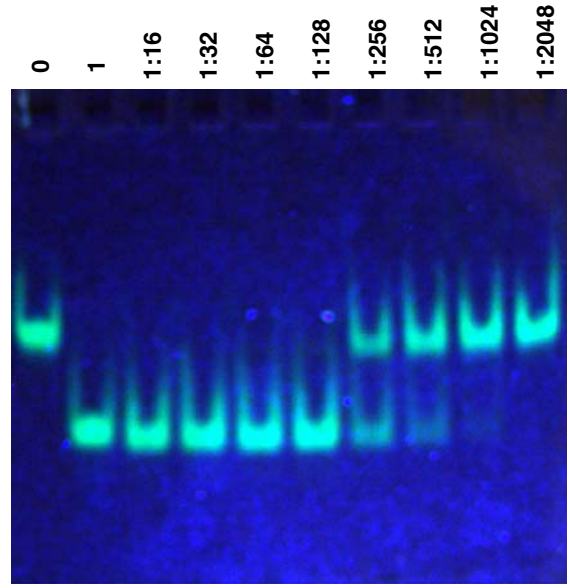


Fig. 3 Processing of the MalE-6xHis-Smt3-eGFP fusion protein by addition of different dilutions of the MalE-6xHis-Ulp1 protease. The cleavage reaction was done for 1 h, the products separated by a 9% non-denaturing PAGE and illuminated on a UV transilluminator (300 nm)

of the pooled eluate (~16 µg) was mixed with the same volume of the purified protease. The protease was added undiluted (6.4 µg) and in serial dilutions down to 1:2048. The reaction was done at 30°C for 1 h and analysed on a 9% nondenaturing polyacrylamide gel. The cleavage reaction could be easily identified on an UV transilluminator due to a mobility shift of the much smaller unfused eGFP. As can be seen on Fig. 3, a 124-fold dilution of the protease is sufficient for complete cleavage of MalE-6xHis-Smt3-eGFP.

For further purification of the 9.81 mg MalE-6xHis-Smt3-eGFP, which was in 3 ml buffer, 60 µl protease solution (corresponds to a 50-fold dilution and about a twofold excess of what would be necessary) was added and incubated at 30°C for 1 h. The cleaved eGFP protein was finally separated from the MalE-6xHis-Smt3 protein and MalE-6xHis-Ulp1 by passing the reaction mixture through a Ni-NTA matrix column. The processed eGFP was found in the flow-through fraction as expected. Altogether, 1.9 mg purified eGFP protein was obtained with a fluorescence activity of 5.1×10^7 rF. This corresponds to a yield of 35% of processed and pure protein and demonstrates the high efficiency of this

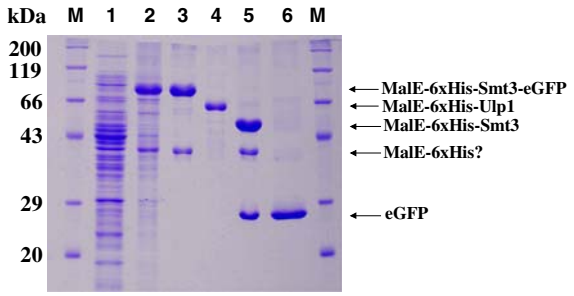


Fig. 4 SDS-PAGE analysis of the purification of eGFP from MalE-6xHis-Smt3-eGFP. Lanes: (M) standard protein marker, (1) crude cell extract of uninduced JM109 pJOE4905.1, (2) crude extract from rhamnose-induced cells, (3) elution fraction from amylose resin chromatography, (4) MalE-6xHis-Ulp1 protein purified by amylose resin, (5) cleavage reaction of MalE-6xHis-Smt3-eGFP and MalE-6xHis-Ulp1 mixed at a dilution of 1:50 and incubated for 1 h at 30°C, (6) flow-through fraction from IMAC

Table 2 Purification and processing of eGFP from the MalE-6xHis-Smt3-eGFP fusion protein

Purification	Total protein (mg)	Total volume (ml)	Total rFL ($\times 10^7$)	Yield (%)
Crude extract	51.6	20	14.5	–
Amylose resin	9.8	3	5.7	38.5
IMAC	1.9	3	5.1	34.5

system. A SDS-PAGE analysis of the purification steps is shown in Fig. 4, the protein concentrations and fluorescence determined during purification are summarized in Table 2.

Discussion

The *E. coli* vector system reported in this paper enables a stable and high level expression of recombinant genes in shake flasks as well as by high-cell-density fermentation and an efficient purification of the desired protein. Several aspects of the expression system positively contribute to this. There is a tight regulation of the rhamnose promotor with very low background expression in the uninduced state and the high transcriptional activity in the induced state. The *cer*-site in the vector for multimer resolution is especially important in Rec-proficient strains. Both properties contribute to the stability of the vectors in the host strains. In principle, a single His-tag would be sufficient to purify first the 6xHis-Smt3-eGFP protein, to cleave the fusion protein by a His-tagged Ulp1 and to remove the protease and tag by a second IMAC purification. But this would need an extensive and time-consuming removal of imidazole and divalent cations, for example, by dialysis. In addition, the *malE* gene in the vector provides an efficient translation initiation region. In comparison with other *N*-terminal tags like Strep-tag, His-tag or glutathione-*S*-transferase-tag with eGFP as reporter the *malE*-fusion proved to be superior (data not shown). Furthermore, MalE increases folding and solubility of fusion partners which might help to overcome inclusion body formation (Nallamsetty and Waugh 2006).

The proteolytic cleavage of MalE-6xHis-Smt3-eGFP is highly specific and efficient giving about 35% yield of pure eGFP despite having to use two chromatographic steps and a protein cleavage reaction. According to the distribution of the fluorescence most of the fusion protein was lost during separation on amylose resin. Some part was found in the flow-through and the other in the washing fractions. This cannot simply be further improved by using a larger volume of amylose resin. After binding of the fusion protein all the washing fractions showed a significant amount of fluorescence indicating a rather weak binding of the fusion protein to the amylose as a reason for product losses. Weak binding of MalE to amylose resin was also observed by others (Lichty et al. 2005) and was overcome by adding a second tag at the *N*-terminal end of MalE (Nallamsetty and Waugh 2007; Pryor and Leiting 1997). In addition, in the crude extract of the rhamnose-induced cells as well as after the amylose resin purification step a relative intense protein band was observed on SDS-PAGE (Fig. 4). The molecular mass of this protein of about 40 kDa is in the range of the maltose binding protein. Since this protein is removed by the IMAC purification, it is likely to be MalE-6xHis protein proteolytically processed by *E. coli* endogenous proteases. The Smt3-eGFP does not bind to the amylose resin column and therefore is lost in the flow-through fraction. Maybe one could avoid this by reducing the distance between His-tag and Smt3. On the other hand, this might affect binding of the His-tag to the IMAC matrix. Nevertheless, even without further improvements this new gene expression and protein processing system should be a valuable tool for production of pure proteins.

Acknowledgements I am grateful to Prof. Wolfgang Seufert and Prof. Ralf Mattes for helpful ideas and discussion and to Annette Schneck and Silke Weber for excellent technical assistance. This work was part of the project "Post-genomic datamining of enzymes for the synthesis of chiral pharmaceutical intermediates" supported by the European Community under contract No LSHB-CT-2003-503017.

References

Altenbuchner J, Mattes R (2004) *Escherichia coli*. Production of recombinant proteins. In: Gellisson G (ed) Production of recombinant proteins. Wiley-VCH, Washington, DC

Gyuris J, Golemis E, Chertkov H, Brent R (1993) Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. Cell 75:791–803

Johnson ES, Blobel G (1999) Cell cycle-regulated attachment of the ubiquitin-related protein SUMO to the yeast septins. J Cell Biol 147:981–994

Li SJ, Hochstrasser M (1999) A new protease required for cell-cycle progression in yeast. Nature 398:246–251

Lichty JJ, Malecki JL, Agnew HD, Michelson-Horowitz DJ, Tan S (2005) Comparison of affinity tags for protein purification. Protein Expr Purif 41:98–105

Luria SE, Adams JN, Ting RC (1960) Transduction of lactose-utilizing ability among strains of *E. coli* and *S. dysenteriae* and the properties of the transducing phage particles. Virology 12:348–390

Mossessova E, Lima CD (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. Mol Cell 5:865–876

Nallamsetty S, Waugh DS (2006) Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. Protein Expr Purif 45:175–182

Nallamsetty S, Waugh DS (2007) A generic protocol for the expression and purification of recombinant proteins in *Escherichia coli* using a combinatorial His6-maltose binding protein fusion tag. Nat Protoc 2:383–391

Pryor KD, Leiting B (1997) High-level expression of soluble protein in *Escherichia coli* using a His6-tag and maltose-binding-protein double-affinity fusion system. Protein Expr Purif 10:309–319

Puigbo P, Guzman E, Romeu A, Garcia-Vallve S (2007) OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res 35:126–131

Sassenfeld HM (2008) Engineering proteins for purification. TibTech 8:88–93

Stumpp T, Wilms B, Altenbuchner J (2000) Ein neues L-Rhamnose-induzierbares Expressionssystem für *Escherichia coli*. BIOspektrum 6:33–36

Tanaka K, Nishide J, Okazaki K, Kato H, Niwa O, Nakagawa T, Matsuda H, Kawamukai M, Murakami Y (1999) Characterization of a fission yeast SUMO-1 homologue, *pmt3p*, required for multiple nuclear events, including the control of telomere length and chromosome segregation. Mol Cell Biol 19:8660–8672

Terpe K (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 60:523–533

Uhlen M, Moks T (1990) Gene fusions for purpose of expression: an introduction. Methods Enzymol 185:129–143

Vertegaal AC (2007) Small ubiquitin-related modifiers in chains. Biochem Soc Trans 35:1422–1423

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119

US005877018A

United States Patent [19]

Filmus et al.

[11] **Patent Number: 5,877,018**

[45] **Date of Patent: *Mar. 2, 1999**

[54] **SYNTHETIC EUKARYOTIC PROMOTERS CONTAINING TWO INDUCIBLE ELEMENTS**

[75] Inventors: **Jorge Filmus**, Toronto; **Michel Klein**, Willowdale, both of Canada

[73] Assignee: **Connaught Laboratories Limited**, North York, Canada

[*] Notice: The term of this patent shall not extend beyond the expiration date of Pat. No. 5,559,027.

[21] Appl. No.: **633,289**

[22] Filed: **Apr. 16, 1996**

Related U.S. Application Data

[63] Continuation of Ser. No. 256,720, Oct. 20, 1994, Pat. No. 5,559,027.

[51] **Int. Cl.**[6] **C12N 15/85**; C12N 15/63; C07H 71/04; C12Q 1/68

[52] **U.S. Cl.** **435/325**; 435/6; 435/172.3; 435/320.1; 536/23.1; 536/24.1

[58] **Field of Search** 435/172.3, 240.2, 435/320.1, 6, 325; 536/23.1, 24.1

[56] **References Cited**

FOREIGN PATENT DOCUMENTS

WO83/01783	5/1983	WIPO .
WO91/10453	7/1991	WIPO .
WO91/12258	8/1991	WIPO .

OTHER PUBLICATIONS

Filmus, J. et al. "Synergistic induction of promoters containing metal–and glucocorticoid–responsive elements" Nucleic Acids Research. vol. 20, No. 11, Jun. 11, 1992, pp. 2755–2760.

Thompson, et al. "Trans–activation by thyroid hormone receptors: functional paralleles with steroid hormone receptors" –Proceedings of the National Academy of Sciences of USA, vol. 86, No. 10, May, 1989, pp. 3494–3498.

Hu et al; Molecular and Cellular Biology, vol. 10, No. 12; pp. 6141–6151, Dec., 1990.

Wright et al; Proceedings of the National Academy of Sciences USA, vol. 88, pp. 8283–8287; Oct., 1991.

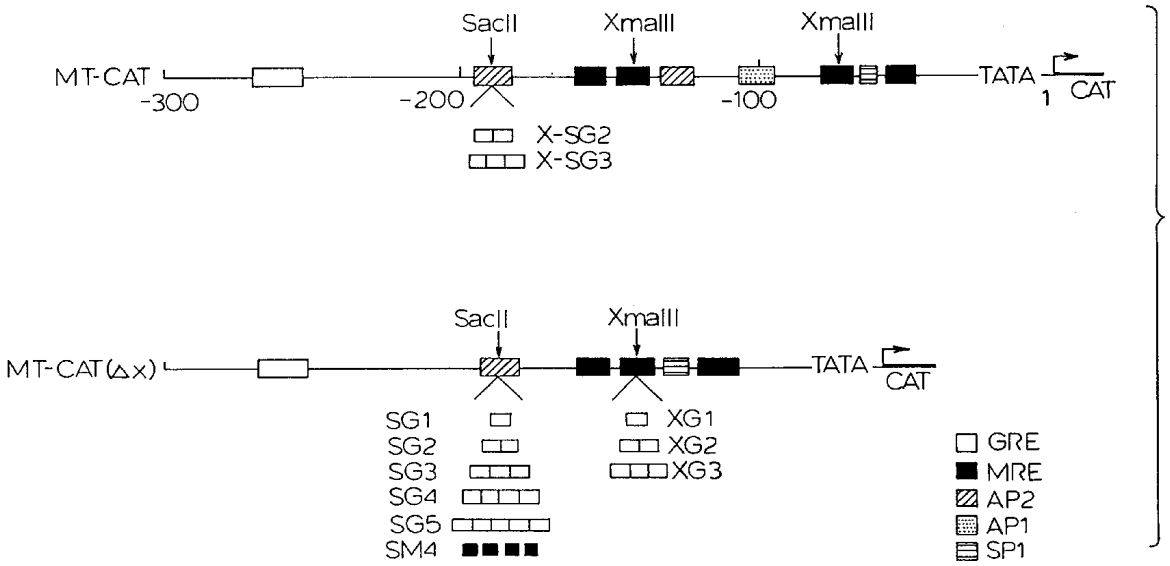
Allan et al; The Journal of Biologi cal Chemistry, vol. 266, No. 9; pp. 5905–5910; Mar. 25, 1991.

Primary Examiner—George C. Elliott
Assistant Examiner—Andrew Wang
Attorney, Agent, or Firm—Sim & McBurney

[57] **ABSTRACT**

Synthetic inducible eukaryotic promoters for the regulation of transcription of a gene achieve improved levels of protein expression and lower basal levels of gene expression. Such promoters contain at least two different classes of inducible elements, usually by modification of a native promoter containing one of the inducible elements by inserting the other of the inducible elements. In embodiments, additional metal responsive elements IR:Es) and/or glucocorticoid responsive elements (GREs) are provided to native promoters, particularly the hMT-IIA and MMTV-LTR promoters. One or more constitutive elements may be functionally disabled to provide the lower basal levels of gene expression.

28 Claims, 1 Drawing Sheet



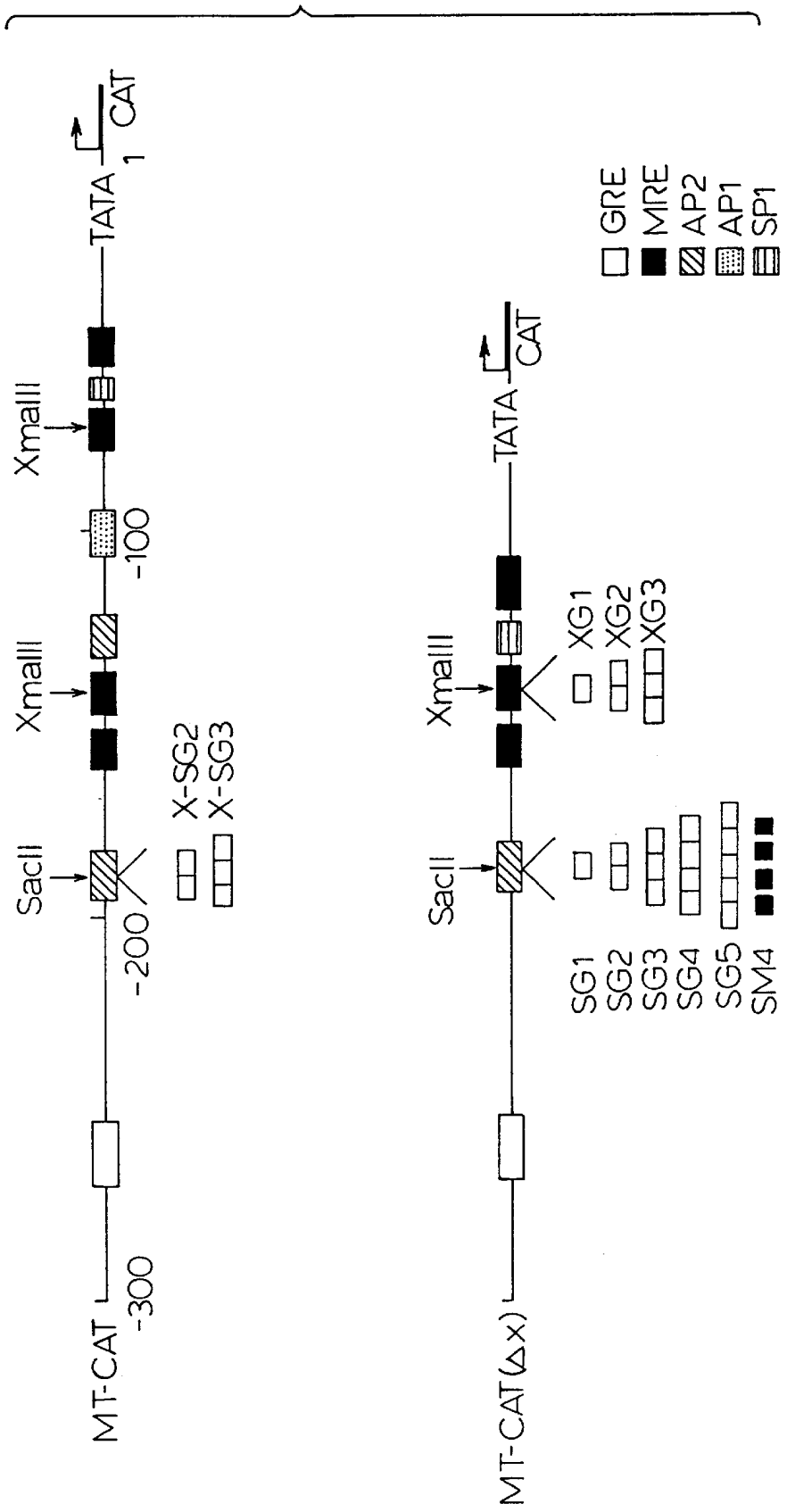

GRE
MRE
AP2
AP1
SP1
CAT
TATA
1
XmaIII
-100
SacII
-200
-300
X-SG2
X-SG3
MT-CAT
XG1
XG2
XG3
SG1
SG2
SG3
SG4
SG5
SM4
MT-CAT(Δx)

FIG.1.

SYNTHETIC EUKARYOTIC PROMOTERS CONTAINING TWO INDUCIBLE ELEMENTS

This application is a continuation of Ser. No. 08/256,720, filed Oct. 20, 1994, now U.S. Pat. No. 5,559,027.

FIELD OF INVENTION

The present invention relates to the generation of improved inducible mammalian expression systems.

BACKGROUND TO THE INVENTION

Mammalian expression systems are being widely used in the production, by recombinant techniques, of proteins that are extensively modified after translation. These systems can be either constitutive or inducible. It is advisable to use inducible systems for the expression of potentially cytotoxic proteins.

A key element in determining whether an expression system is constitutive or inducible is the promoter. Several mammalian promoters that can be induced in experimental systems have been characterized and promoters present in the metallothionein (MT) genes and in the mouse mammary tumour virus/long terminal repeat (MMTV-LTR) have been used extensively.

The best inducers for the MT promoter are heavy metal ions, such as cadmium (Cd) and zinc (Zn). The induction of the promoter is mediated by transcription factors which, after activation by metals, bind to the inducible metal responsive elements (MREs) that are present in the MT promoter. This promoter also contains several constitutive (non-inducible) elements that bind transcription factors which do not need to be activated and that are responsible for a basal level of gene expression. As a result of the presence of these constitutive elements, the non-induced level of expression of the MT promoter is significant and the induction ratio (the ratio between the inducible expression and the basal level of expression) is usually no greater than 5- to 10-fold. Attempts have been made to reduce the basal level of expression by removing some of the constitutive elements of the MT promoter. The removal of these elements, however, also reduces the inducible level of expression.

The native human MT-IIA promoter, besides having the MREs and the constitutive elements, contains a single inducible glucocorticoid responsive element (GRE) and glucocorticoids, such as dexamethasone (dex), induce low levels of expression from the MT-IIA promoter in its native context.

The native MMTV-LTR promoter contains four inducible GREs and can be strongly induced by glucocorticoids. The basal level of expression is lower than that obtained with the human MT-IIA promoter but the absolute level of inducible expression is not as high.

Nucleic acid sequences, such as inducible elements, involved in the regulation of gene expression, may be located 5' to, 3' to, or within the regulated gene.

SUMMARY OF INVENTION

In accordance with the present invention, there is provided a synthetic inducible eukaryotic promoter for the regulation of transcription of a gene, comprising at least two different classes of inducible elements. Classes of inducible elements with which the invention is concerned include hormone-responsive elements (including GREs), metal-responsive elements (MREs), heat shock-responsive elements, interferon-responsive elements and cytokine responsive elements.

In one embodiment, the synthetic promoter provided herein is derived from a native promoter and one of the different classes of inducible elements is a native inducible element while another of the different classes of inducible elements is provided, such as by insertion into the native promoter or by activation of a normally-inactive element in the native promoter. While, in general two different classes of inducible elements are present in the novel synthetic promoter of the invention, combinations of three or more may be present, if desired.

The utilization of different classes of inducible elements in the synthetic promoters enables synergistic induction of a expression of a gene product in a eukaryotic expression system, particularly a mammalian expression system. That is, the level of gene expression obtained by induction of multiple classes of inducible element is greater than the sum of the individual gene expressions achieved by separate induction of the individual classes of inducible elements. In addition, overall levels of gene expression may be enhanced.

The synthetic promoters provided herein generally are derived from natural promoters by modification, as described in more detail herein, although such promoters also may be produced synthetically.

As mentioned above, inducible promoters may contain at least one constitutive element, which provides a basal level of gene expression in the absence of induction. In one embodiment of the invention, at least one constitutive element is functionally disabled, which generally results in a decreased level of basal gene, expression and an increased ratio of induced gene expression to basal gene expression, when compared to the unmodified promoter. Such functional disablement of the at least one constitutive element may be effected by deletion from the native promoter and/or by insertion, for example, of an inducible element therein.

The present invention, therefore, provides, in preferred embodiments, improved inducible eukaryotic promoters containing not only native GREs and/or MREs but also additional GREs and/or MREs. Constitutive elements of native promoters may or may not be deleted in the improved promoters. The improved promoters may be synergistically induced when both a heavy metal ion and a glucocorticoid (such as dexamethasone) are used at the same time and both at least one MRE and at least one GRE are present. Synergistic induction results in levels of gene expression that are much higher than those observed with unmodified promoters, such as the human MT-IIA or MMTV-LTR promoters. The new promoters also may contain fewer constitutive elements than unmodified promoters, which allows for a lower basal level of gene expression.

Conveniently the unmodified promoter may be the human MT-IIA or MMTV-LTR promoter. The responsive elements may conveniently contain the consensus sequence for such elements, for example,

5'-GATCTTGCGCCCGGCCCG-3' (SEQ ID NO: 2) contains-the MRE consensus sequence, and

5'-GATCTGGTACAGGATGTTCTAGCTACG-3' (SEQ ID NO: 1) contains the GRE. consensus sequence used in the embodiments of this invention.

Advantages of the present invention include:

a) high overall levels of gene expression,

b) decreased levels of basal gene expression,

c) synergistic induction of expression of a gene,

d) promoters-customized with regard to induction ratio and/or responsiveness to convenient inducers.

BRIEF DESCRIPTION OF DRAWING

FIG. **1** is a genetic map of the hMT-IIA promoter and of a modified promoter with various modifications effected to the hMT-IIA promoter in accordance with one embodiment of the present invention.

GENERAL DESCRIPTION OF INVENTION

As noted above, the novel promoter provided herein may be derived from a native promoter. In one preferred embodiment of the invention, the promoter contains at least one native inducible element which is an MRE and at least one different inducible element which is a hormone responsive element, particularly a glucocorticoid responsive element (GRE) provided in the native promoter by insertion.

Such an inserted GRE may be a synthetic molecule comprising a pair of complementary oligonucleotides containing the GRE consensus sequence. A plurality of GREs may be inserted into the native promoter in the form of a multimeric head-to-tail self-ligated element.

A particularly preferred embodiment of the invention provides a human metallothionein gene (hMT-IIA) promoter modified to contain at least one inducible GRE, so as to obtain a synergy of gene expression upon induction of the inducible MREs and GREs in a eukaryotic expression system, particularly a mammalian expression system, and preferably combined with an enhanced overall level of gene product expression. In this particularly preferred embodiment, multimeric head-to-tail GREs may be inserted into the native hMT-IIA promoter.

It is preferred also to disable at least one constitutive element of the native hMT-IIA promoter, such as by deletion of such element and/or by insertion of at least one GRE therein. In one illustrative Example, both deletion of constitutive elements and insertion of single or multiple GREs are employed to disable constitutive elements.

In another preferred embodiment of the invention, the promoter contains at least one native inducible element which is an HRE, particularly a glucocorticoid responsive element (GRE), and at least one different inducible element which is a MRE provided by insertion.

Such inserted MRE may be a synthetic molecule comprising a pair of complementary oligonucleotides containing the MRE consensus sequence. A plurality of MREs may be inserted into the native promoter in the form of a multimeric head-to-tail self-ligated element.

A particularly preferred embodiment of the invention provides a mouse mammary tumor virus/long terminal repeat (MMTV-LTR) promoter, modified to contain at least one inducible MRE, so as to obtain a synergy of gene expression upon induction of the inducible GREs and MREs in a eukaryotic expression system,. and preferably combined with an enhanced overall level of gene expression. In this particularly preferred embodiment, multimeric head-to-tail MREs may be inserted into the native MMTV-LTR promoter.

The novel synthetic inducible eukaryotic promoter provided herein may be incorporated into a vector for eukaryotic expression of a gene product, particularly when operatively connected to a gene to be expressed by the expression system. Such expression system may comprise eukaryotic cells containing the vector, particularly mammalian cells, such as Vero, CHO, HeLa, RatII fibroblasts and intestinal epithelial cells.

DESCRIPTION OF PREFERRED EMBODIMENT

In FIG. **1**, there are shown different versions of a new promoter incorporating various modifications in accordance with embodiments of the present invention. The new series of promoters are generated using the following methodology. A KspI DNA fragment containing 800 bp of the 5' promoter region of the human MT-IIA gene (bases −740 to +60) was isolated from a plasmid containing the human MT-IIA gene (see Karin et al, (1982) Nature, 299, 797–802). After generating blunt ends, HindIII linkers were added and the fragment was inserted into pSVOATCAT, a plasmid containing the chloramphenicol acetyl transferase (CAT) gene used as a reporter gene, at the HindIII site 5' to the CAT gene. Two constitutive elements (AP1 and AP2 —see upper map, FIG. **1**) of the original MT-IIA promoter were deleted by removing an XmaIII fragment (bases −79 to −129).

A pair of complementary oligonucleotides containing the GRE consensus sequence, a 5' BamHI site and a 3' BglII site was synthesized. The positive strand oligonucleotide sequence was:

5'-GATCTGGTACAGGATGTTCTAGCTACG-3' (SEQ ID NO: 1) Multimeric head-to-tail GREs were prepared by self-ligating the synthetic GRE oligonucleotide in the presence of BamHI and BglII. Single and multimeric GREs were inserted into the SacII site of the promoter (at base −175) or the XmaIII site of the promoter (at base −129) (see lower map in FIG. **1**). The insertion at the SacII element destroys a second AP2 site.

A pair of complementary oligonucleotides containing the MRE consensus sequence, a 5' BamHI site and a 3' BglII site was synthesized. The positive strand nucleotide sequence was:

5'-GATCTTGCGCCCGGCCCG-3' (SEQ ID NO: 2) Such oligonucleotides may be used to synthesize multimeric head-to-tail elements and single or multiple MREs may be inserted into the hMT-IIA promoter in an analogous manner to the GREs.

The MMTV-CAT vector for effecting similar GRE and/or MRE insertions to and optionally constitutive element deletions from the MMTV-LTR promoter was removed from plasmid p201 (Majors et al, (1981), Nature, 283, 253–258) using PstI and, after generation of blunt ends, inserted into the HindIII site of pSVOATCAT.

The new promoters were tested in transient CAT expression assays using RAT II fibroblasts, CHO (chinese hamster ovarian cells), VERO (monkey fibroblasts) and Hela (human cervical tumour cells) cells, expressing the glucocorticoid receptor. The results, reproduced in the Examples below, indicated that these new promoters generate very high levels of expression when cells normally expressing the glucocorticoid receptor or transfected with the glucocorticoid receptor gene are simultaneously induced with heavy metal ions and dexamethasone. The induced levels of expression obtained with these promoters are significantly higher than those observed with the wild-type human MT-IIA or MMTV-LTR promoters. At the same time the basal level of expression was significantly lower than that observed with the wild-type human MT-IIA promoter.

EXAMPLES

The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples. These examples are described solely for purposes of illustration and are not intended to limit the scope of the invention. Changes in form and substitution of equivalents are contemplated as circumstances may suggest or render expedient. Although specific terms have been employed herein,

such terms are intended in a descriptive sense and not for purposes of limitations.

EXAMPLE 1

This Example illustrates the construction of modified hMT-IIA promoters containing additional GREs.

All MT expression vectors were derived from pSVOATCAT, a plasmid containing the chloramphenicol acetyl transferase (CAT) gene without any regulatory sequences (Gorman et al., Mol.Cell.Biol., 2, 1044, [1982]). MT-CAT, a control plasmid in which the CAT gene is under the regulation of the wild-type human MT-IIA promoter (hMT-IIA), was generated as described below. An 800 bp KspI fragment of the promoter region of the hMT-IIA (bases −740 to +60) (FIG. **1**) was isolated. After generating blunt ends, HindIII linkers were added and the fragment was inserted into the HindIII site of pSVOATCAT, 5' to the CAT gene. Plasmid MT-CAT-ΔX was generated by removing the XmaIII fragment (base −79 to −129) from the MT promoter of MT-CAT which contains the constitutive AP1–AP2 elements. To insert additional GREs, a pair of complementary oligonucleotides containing the GRE consensus sequence, a 5' BamHI site and a 3' BglII site were synthesized and multimeric head-to-tail elements were generated by self-ligating these synthetic sequences in the presence of BamHI and BglII. The positive strand nucleotide sequence was SEQ ID NO: 1, as specified above. Monomeric or multimeric GREs then were inserted at either the SacII or the XmaIII site of the MT-CAT-**66** X vector after generation of blunt ends (FIG. **1**). The number of GREs inserted was confirmed by DNA sequencing.

EXAMPLE 2

This Example illustrates the use of an expression vector containing additional GREs.

The expression vector used in this example was SG2, which is a pSVOATCAT-derived CAT expression vector containing a modified MT-IIA promoter in which two additional GREs were inserted at the SacII site of MT-CAT-ΔX (FIG. **1**). Fifteen μg of plasmid DNA were transfected into CHO cells using the calcium phosphate procedure (Graham et al (1973) Virology, 52, 456–467). After incubation for 5 hours at 37° C., the cells were shocked for 3 minutes with 15% glycerol in PBS. The monolayers then were incubated with the different inducers ($CdCl_2$ and/or dexamethasone) for 16 hours and cell extracts were prepared. The CAT activity then was measured using ^{14}C-Chloramphenicol as substrate and the radioactive acetylated product was extracted with xylene. Radioactive counts were determined in a scintillation counter.

In addition, the SG2 vector was compared with two other vectors that were constructed by inserting a wild-type MT-IIA promoter and the MMTV-LTR promoter into the HindIII site of the pSVOATCAT plasmid. Since CHO cells do not have glucocorticoid receptors, the cells were co-transfected with 10 μg of a glucocorticoid receptor expression vector (Giguere et al, (1986) Cell, 46, 645–652). CAT expression assays were performed in quadruplicate and the standard deviation did not exceed 10%. Protein concentration was measured in each cell lysate and CAT activity was calculated for equivalent amounts of protein. The results from these experiments are summarised in Table I below. (The Tables appear at the end of the descriptive text).

The results appearing in Table I show that the synergistic induction of the SG2 promoter with metals and dexamethasone generated a higher level of CAT gene expression than the wild-type MT-IIA and the MMTV-LTR promoters. At the same time, the induction ratio also was significantly improved.

EXAMPLE 3

This Example further illustrates the use of a vector containing additional GREs.

Using a procedure similar to that of Example 1, the activity of the SG2 promoter was compared with that of the native MT-IIA promoter in VERO cells engineered to express glucocorticoid receptors (Giguere et al, (1986) Cell, 46, 645–652). In this Example, the cells also were co-transfected with an expression vector in which the β-galactosidase gene was driven by a promoter, whose activity was not affected under the experimental conditions by heavy metals or glucocorticoids. After transfection and induction, an aliquot of the cell extract was used to measure the β-galactosidase (β-Gal) activity. This activity was used to standardize CAT activity measurements by taking into account the efficiency of transfection.

The results obtained are shown in Table II below, and it can be seen that they are very similar to those obtained with CHO cells (Table I) and demonstrate that dexamethasone acts synergistically with metal ions on the modified MT-IIA (SG2) promoter.

EXAMPLE 4

This Example illustrates further modification to the expression vector and the results obtained.

Additional modifications were effected to the hMT-IIA promoter to introduce additional numbers of GREs and multiple MREs at the SacII site and to introduce numbers of GREs at the XmaIII site, as detailed in FIG. **1**.

The resulting modified plasmid DNA was introduced into Vero cells as described in Example 3 and CAT gene expression was determined as described above. The results obtained are set forth in Table III below.

EXAMPLE 5

This Example illustrates the construction and use of a modified MMTV-LTR promoter containing additional GREs.

Two MREs were inserted, using a similar procedure to previous examples, at the BfrI site of the MMTV-LTR promoter, which contains four GREs but has no MREs (Majors and Varmus, Nature 283: 253–258). Table IV shows that while the unmodified MMTV-LTR promoter was not inducible by Zn plus Cd, the modified promoter (BM2-MMTV) displayed a ten-fold induction. When BM2-MMTV was induced by dexamethasone plus Zn plus Cd a two-fold synergy in CAT expression was observed.

The results of the experiments represented in Examples 1 to 5 and Tables I to IV show that it is possible to achieve synergistic activation of transcription in the context of a modified hMT-IIA promoter by inserting additional inducible elements in the form of GREs and in the context of a modified MMTV promoter by inserting additional inducible elements in the form of MREs. Addition of the GREs to the hMT-IIA promoter and MREs to the MMTV promoter did not increase the basal level of reporter gene expression and the inducibility and transcriptional strength of the modified promoters were significantly improved over those of their wild-type counterpart. In contrast the exclusive insertion of four extra MREs (vector SM4) to the hMT-IIA promoter resulted only in a moderate improvement in MT promoter

transcriptional strength and this improvement was accompanied by a significant increase in basal expression.

The unmodified hMT-IIA promoter in the MT-CAT vector could not be induced by dexamethasone in Vero cells transfected with the glucocorticoid receptor gene. However, the insertion of at least one additional GRE to the promoter was enough to confer glucocorticoid responsiveness and gene expression.

To analyze the impact of the number of additional GREs inserted and the site of insertion, two series of modified promoters were generated in the Examples by adding one or more GREs at either SacII site (SG series) or the XmaIII site (XG series) of MT-CAT-**66** X. All vectors were inducible by $CdCl_2$ and glucocorticoids. However, a minimum of two adjacent GREs was necessary to generate synergistic inducibility by simultaneous treatment of transfected Vero cells with $CdCl_2$ and dexamethasone, regardless of the site of insertion.

The induction ratio calculated for the modified hMT-IIA promoters was increased up to 6-fold as compared to the wild-type promoter. The fact that the insertion of additional GREs did not increase the basal level of gene expression in, for example, SG3 is an important factor in the improvement of this ratio. This observation emphasizes one of the advantages of generating synergistic transcription activation by adding different classes of inducible elements rather than constitutive ones, in accordance with the present invention.

SUMMARY OF DISCLOSURE

In summary of this disclosure, the inventors provide for the engineering and use of novel and improved inducible mammalian expression systems, in particular, the preparation and use of modified human MT-IIA promoters containing one or several additional glucocorticoid-responsive elements which can be synergistically induced by glucocorticoids and metal ions while maintaining a low level of basal gene expression. The induction ratio may be increased further by deleting constitutive elements. A similar strategy may be used to generate improved mouse mammary tumour virus (MMTV) promoter by inserting additional metal-responsive elements. Modifications are possible within the scope of this invention.

TABLE I

Promoter	Inducer	CAT Activity (cpm)
MT-IIA	Control	5932
MT-IIA	100 μM $ZnCl_2$ + 2 μM $CdCl_2$	70235
MT-IIA	1 μM Dexamethasone	3935
MT-IIA	100 μM $ZnCl_2$ + 1 μM Dexamethasone	70119 (12x)*
SG 2	Control	2893
SG 2	100 μM $ZnCl_2$ + 2 μM CCl_3	22901
SG 2	1 μM Dexamethasone	97068
SG 2	100 μM $ZnCl_2$ + 2 μM $CdCl_2$ + 1 μM Dexamethasone	147713 (57x)*
MMTV-LTR	Control	751
MMTV-LTR	1 μM Dexamethasone	20310 (27x)*

*Induction Ratio

TABLE II

Promoter	Inducer	Standardised CAT Activity (U CAT/β-GAL)
MT-IIA	Control	19
MT-IIA	5 μM $CdCl_2$	574
MT-IIA	1 μM Dexamethasone	40
MT-IIA	5 μM $CdCl_2$ + μM Dexamethasone	526 (27x)*
SG 2	Control	8
SG 2	5 μM $CdCl_2$	114
SG 2	1 μM Dexamethasone	230
SG 2	5 μM $CdCl_2$ + 1 μM Dexamethasone	1072 (134x)*

*Induction Ratio

TABLE III

Promoter	Inducer	Relative CAT activity (% of MT-IIA control)
MT-IIA	Control	100
MT-IIA	5 uM $CdCl_2$	1064
MT-IIA	1 uM Dexamethasone	103
MT-IIA	5 uM $CdCl_2$ + 1 uM Dexamethasone	1074
SG1	Control	32
SG1	5 uM $CdCl_2$	328
SG1	1 uM Dexamethasone	957
SG1	5 uM $CdCl_2$ + 1 uM Dexamethasone	1364
SG2	Control	36
SG2	5 uM $CdCl_2$	364
SG2	1 uM Dexamethsone	1164
SG2	5 uM $CdCl_2$ + 1 uM Dexamethasone	2324
SG3	Control	50
SG3	5 uM $CdCl_2$	596
SG3	1 uM Dexamethasone	1821
SG3	5 uM $CdCl_2$ + 1 uM Dexamethasone	3156
SG4	Control	29
SG4	5 uM $CdCl_2$	210
SG4	1 uM Dexamethasone	386
SG4	5 uM $CdCl_2$ + 1 uM Dexamethasone	1317
SG5	Control	21
SG5	5 uM $CdCl_2$	200
SG5	1 uM Dexamethasone	136
SG5	5 uM $CdCl_2$ + 1 uM Dexamethasone	1117
XG1	Control	46
XG1	5 uM $CdCl_2$	1755
XG1	1 uM Dexamethasone	275
XG1	5 uM $CdCl_2$ + 1 uM Dexamethasone	1574
XG2	Control	12
XG2	5 uM $CdCl_2$	519
XG2	1 uM Dexamethasone	394
XG2	5 uM $CdCl_2$ + 1 uM Dexamethasone	1957
XG3	Control	11
XG3	5 uM $CdCl_2$	107
XG3	1 uM Dexamethasone	36
XG3	5 uM $CdCl_2$ + 1 uM Dexamethasone	229
X-SG2	Control	84
X-SG2	5 uM $CdCl_2$	1482
X-SG2	1 uM Dexamethasone	495
X-SG2	5 uM $CdCl_2$ + 1 uM Dexamethasone	2562
X-SG3	Control	146
X-SG3	5 uM $CdCl_2$	1145
X-SG3	1 uM Dexamethasone	833
X-SG3	5 uM $CdCl_2$ + 1 uM Dexamethasone	3383
SM4	Control	393

TABLE III-continued

Promoter	Inducer	Relative CAT activity (% of MT-IIA control)
SM4	5 uM $CdCl_2$	1485
SM4	1 uM Dexamethasone	382
SM4	5 uM $CdCl_2$ + 1 uM Dexamethasone	1524

TABLE IV

Promoter	Inducer	Standardised CAT activity (CPM)
MMTV-LTR	control	1326
MMTV-LTR	Dex	135405
MMTV-LTR	Zn + Cd	225
MMTV-LTR	Zn + Cd + Dex	145416 (102X)*
BM2-MMTV	control	1078
BM2-MMTV	Dex	92899
BM2-MMTV	Zn + Cd	10827
BM2-MMTV	Zn + Cd + Dex	196614 (182X)*

*Induction ratio.

expression by said promoter and to increase the ratio of induced to basal gene expression by said promoter.

2. The promoter of claim **1** wherein said at least one constitutive element is disabled by deletion from the native promoter or insertion of an inducible element into the native promoter.

3. A synthetic inducible eukaryotic promoter for the regulation of transcription of a gene, comprising at least two different classes of inducible elements, wherein said different classes of inducible elements are selected to provide a synergistic level of expression of a gene product in a eukaryotic expression system, said promoter being derived from a native promoter and one of said different classes of inducible elements is a native inducible element and another of said different classes of inducible elements is a different inducible element provided in said native promoter, wherein said native inducible element is a metal-responsive element (MRE) and said different inducible element is at least one glucocorticoid-responsive element (GRE) and is provided in said native promoter by insertion.

4. The promoter of claim **3** wherein said inserted GRE is a synthetic molecule containing the GRE consensus sequence and having a positive strand having the nucleotide sequence:

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i i i) NUMBER OF SEQUENCES: 2

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GATCTGGTAC AGGATGTTCT AGCTACG 27

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GATCTTGCGC CCGGCCCG 18

What we claim is:

1. A synthetic inducible eukaryotic promoter for the regulation of transcription of a gene, comprising at least two different classes of inducible elements, wherein said different classes of inducible elements are selected to provide a synergistic level of expression of a gene product in a eukaryotic expression system, said promoter being derived from a native promoter containing at least one constitutive element, wherein said at least one constitutive element is functionally disabled to decrease the level of basel gene

5'-GATCTGGTACAGGATGTTCTAGCTACG-3' (SEQ ID NO: 1).

5. The promoter of claim **4** wherein a plurality of GREs is inserted in said native promoter in the form of a multimeric head-to-tail element self-ligated in the presence of BamHI and BglII.

6. The promoter of claim **3** which is derived from a native promoter containing at least one constitutive element and wherein said at least one constitutive element is functionally disabled.

7. The promoter of claim **6** wherein said at least one constitutive element is disabled by deletion of the constitutive element from the native promoter or insertion of an inducible element in the constitutive element.

8. The promoter of claim **7** wherein said at least one constitutive element is functionally disabled sufficient to provide a decreased level of basal gene expression and an increased ratio of induced gene expression.

9. The promoter of claim **8** wherein said native promoter is the hMT-IIA promoter.

10. The promoter of claim **3** wherein said native promoter is the hMT-IIA promoter.

11. The promoter of claim **8** wherein said at least one GRE is inserted at either the SacII site at base −175 or the XmaIII site at base −129 or at both the SacII and the XmaIII sites of a native hMT-IIA promoter.

12. A synthetic inducible eukaryotic promoter for the regulation of transcription of a gene, comprising at least two different classes of inducible elements, wherein said different classes of inducible elements are selected to provide a synergistic level of expression of a gene product in a eukaryotic expression system, said promoter being derived from a native promoter and one of said different classes of inducible elements is a native inducible element and another of said different classes of inducible elements is a different inducible element provided in said native promoter, wherein said native inducible element is a metal responsive element (MRE) and said different inducible element is a glucocorticoid-responsive element (GRE) and is provided in said native promoter by insertion, wherein said native promoter is the hMT-IIA promoter, wherein multiple linked GRE's are inserted into the native promoter.

13. The promoter of claim **10** wherein at least one native constitutive element is disabled.

14. The promoter of claim **13** wherein said constitutive element is disabled by deletion and/or insertion of at least one GRE therein.

15. The promoter of claim **14** wherein two native constitutive elements, AP1 and AP2, located between bases −79 to −129 of the native hMT-IIA promoter are deleted.

16. The promoter of claim **15** wherein at least one GRE sequence is inserted at the SacII site (base −175) of the native hMT-IIA promoter thereby disabling a second AP2 constitutive element at that location.

17. The promoter of claim **12** wherein two linked GRE sequences are inserted at the XmaIII site at base −129 of the hMT-IIA promoter.

18. The promoter of claim **12** wherein three linked GRE sequences are inserted at the SacII site at base −175 of the hMT-IIA promoter.

19. A synthetic inducible eukaryotic promoter for the regulation of transcription of a gene, comprising at least two different classes of inducible elements, wherein said different classes of inducible elements are selected to provide a synergistic level of expression of a gene product in a eukaryotic expression system, said promoter being derived from a native promoter and one of said different classes of inducible elements is a native inducible element and another of said different classes of inducible elements is a different inducible element provided in said native promoter, wherein said native responsive element is a glucocorticoid-responsive element (GRE) and said different inducible element is a metal responsive element (MRE) which is provided in said native promoter by insertion.

20. The promoter of claim **19** wherein said inserted MRE is a synthetic molecule containing the MRE consensus sequence and having a-positive strand having the nucleotide sequence:

5'-GATCTTGCGCCCGGCCCG-3' (SEQ ID NO: 2).

21. The promoter of claim **20** wherein a plurality of MREs is inserted into the native promoter in the form of a multimeric head-to-tail element self-ligated in the presence of BamHI and BglII.

22. The promoter of claim **19** wherein said native promoter is the MMTV-LTR promoter.

23. A synthetic inducible eukaryotic promoter for the regulation of transcription of a gene, comprising at least two different classes of inducible elements, wherein said different classes of inducible elements are selected to provide a synergistic level of expression of a gene product in a eukaryotic expression system, said promoter being derived from a native promoter and one of said different classes of inducible elements is a element is a native inducible element and another of said different classes of inducible elements is a different inducible element provided in said native promoter, wherein said native responsive element is a glucocorticoid-responsive element (GRE) and said different inducible element is a metal responsive element (MRE) which is provided in said native promoter by insertion, wherein said native promoter is the MMTV-LTR promoter, wherein at least two linked MRE's are inserted into the native promoter.

24. A vector for eukaryotic expression of a gene product, comprising a synthetic inducible eukaryotic promoter operatively connected to a gene encoding said gene product said promoter being a modified native hMT-IIA promoter as claimed in claim **10**.

25. A vector for eukaryotic expression of a gene product, comprising a synthetic inducible eukaryotic promoter operatively connected to a gene encoding said gene product, said promoter being a modified native MMTV-LTR promoter as claimed in claim **22**.

26. A eukaryotic expression system, comprising eukaryotic cells containing a vector as claimed in claim **24** or **25** for effecting induced gene expression.

27. The expression system of claim **26** wherein said eukaryotic cells are mammalian cells.

28. The expression system of claim **27** wherein said mammalian cells are selected from Vero, CHO, HeLa, RatII and epithelial cells.

* * * * *