



UvA-DARE (Digital Academic Repository)

Selection markers and expression enhancing elements: novel strategies to create mammalian production cell lines

Hoeksema, F.

Publication date
2011

[Link to publication](#)

Citation for published version (APA):

Hoeksema, F. (2011). *Selection markers and expression enhancing elements: novel strategies to create mammalian production cell lines*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

CHAPTER 8

**THE CYSTATHIONINE γ -LYASE ENZYME THAT IS INVOLVED
IN THE SYNTHESIS OF L-CYSTEINE CAN BE USED AS
'METABOLIC' SELECTION MARKER IN CHO CELLS**

Femke Hoeksema, Michel Siep, John Verhees, Rik van Blokland,
Tjalling Siersma, Karien Hamer, Jan den Blaauwen and Arie. P. Otte

Submitted for publication

ABSTRACT

Stable mammalian cell lines that produce therapeutic proteins are usually selected with toxic agents such as antibiotics that are also needed for maintenance of high protein expression levels over prolonged time periods. L-cysteine is an essential amino acid for CHO cells since they miss the Cystathionine γ -lyase (CLase) enzyme that converts L-cystathionine to L-cysteine. Here we show that the CLase enzyme can be used as selection marker in CHO cells. High protein expressing clones can be selected in L-cysteine minus culture medium that lacks selection agents such as Zeocin. The CLase gene needs to be coupled to the gene of interest through an IRES sequence. However, efficient use of the CLase gene that is incorporated in the expression cassette requires that the translation stringency of the CLase mRNA is carefully chosen. When the translation stringency is too high, high protein expression levels can be obtained, but cell growth is severely hampered. We therefore designed an IRES mutant that provides high enough translation stringency for the CLase mRNA to warrant high protein expression levels, but without hampering cell growth. Our results indicate the usefulness of CLase as selection marker in CHO cells, without the need for toxic antibiotics to obtain CHO cell lines that have high protein levels that remain stable over prolonged time periods.

INTRODUCTION

STAR elements are genomic stretches of DNA that block chromatin-associated repressor activity such as Polycomb-group proteins and heterochromatin-related proteins¹. When these STAR elements are utilized to flank gene expression cassettes, gene expression levels are strongly elevated. Also, they induce the formation of many more cell colonies than plasmids without STAR elements. To reduce the number of colonies we employ a very stringent selection system to create relatively few, but highly productive mammalian cell lines². In essence, we place a selection marker with a startcodon that confers attenuated translation initiation frequency, upstream of the gene of interest with a startcodon that confers optimal translation initiation. This modified startcodon could, for instance, be a TTG instead of an ATG, resulting in selection markers such as TTG Zeocin (TTG Zeo) or TTG Neomycin (TTG Neo). From the transcribed bicistronic mRNA, the selection marker is translated at a low frequency, and the protein of interest at a high frequency. Flanking such constructs with STAR elements results in the formation of clones that express proteins at high levels. We called this selection system STAR-Select3. Unfortunately, keeping high protein expression levels in a TTG Zeo or TTG Neo selected colony at the same high level over a long time period, requires maintenance of selection pressure. This can be accomplished by keeping Zeocin or Neomycin in the culture medium, but this procedure is not favored since such agents are both toxic and expensive.

Another approach is to couple the gene of interest to a metabolic selection marker that is essential for cell survival, such as an enzyme that metabolizes one or more essential steps in a metabolic pathway. With essential is meant that the cell is not able to synthesize a specific metabolic building block itself, implying that this building block has to be present in the culture medium in order to allow the cell to survive. A well-known example is the tetrahydrofolate synthesizing *Dhfr* gene⁴⁻⁶. The DHFR protein is an enzyme in the folate pathway that converts dihydrofolate into tetrahydrofolate, which is required for the synthesis of purines (hypoxanthine), thymidylic acid (thymidine), and the amino acid glycine. CHO-DG44 cells⁷ lack the *Dhfr* gene and they therefore need glycine, hypoxanthine and thymidine in the culture medium to survive. If, however, the *Dhfr* gene is present on the expression cassette,

the cell can convert folate into tetrahydrofolate, provided that the end-products glycine, hypoxanthine and thymidine are absent and the precursor is present in the culture medium⁷. This principle has been used for many years as selection methodology to create stably transfected mammalian cell lines.

Other examples are the essential amino acids that are required for cell survival and which cannot be synthesized by mammalian cells since they lack the enzymes to synthesize these amino acids. An example is the enzyme Cystathionine γ -lyase (or CLase) that catalyzes the conversion of L-cystathionine into L-cysteine (Figure 1A). L-cysteine is an important sulfur-containing compound, it is a unique metabolic entrance for reduced sulfur into cell metabolism in most organisms, and it is required for the biosynthesis of sulfur-containing compounds such as L-methionine, thiamine, biotin, coenzyme A, and others. In addition, L-cysteine plays crucial roles in protein folding, assembly, and stability through disulfide-bond formation. L-cysteine is also needed for synthesis of glutathione, a major intracellular antioxidant⁸. Generally, mammalian cells are capable of synthesizing L-cysteine from L-methionine via the trans-sulphuration pathway⁹. However, CHO cells lack CLase, and therefore L-cysteine can be considered as an essential amino acid for CHO cells^{10,11}. As a consequence, L-cysteine has to be present in CHO culture medium to allow the cells to survive. Hence, if the gene encoding CLase is present on the expression cassette, the cell can convert L-cystathionine into L-cysteine, provided L-cysteine is absent and the non-toxic L-cystathionine is present in the culture medium. In theory, the CLase encoding gene can be used as selection marker in appropriate cell lines. Here we tested this possibility in CHO cell lines.

RESULTS

The Cystathionine γ -lyase (CLase) gene can be used to convey maintenance of stable protein expression

We tested the potential usefulness as maintenance/ selection marker of the CLase enzyme that converts L-cystathionine into L-cysteine (Figure 1A). To do that, the amino acid L-cysteine had to be removed from the culture medium. However, L-cysteine is interconvertible with L-cystine, which can occur in the culture medium (Figure 1A). Therefore, also L-cystine needs to be absent in the culture medium. As a first step we established the effects of omitting L-cysteine/ L-cystine from the culture medium on wild type CHO-K1 cells. In this experiment we added 0.1 mM L-cystathionine precursor to the culture medium. Trypan blue staining was used to monitor the viability of cells. We noted that already within three days cells started to float in the medium and that after one week 75% of the cells were dead (Figure 1B). After two weeks approximately 95% of the cells had died. In comparison, addition of 150 μ g/ml Zeocin already lead to 50% cell death after three days and >90% cell death after one week (Figure 1B). Addition of Zeocin therefore appears to be a more effective means to kill wild type CHO-K1 than removal of the essential metabolic component L-cysteine/ L-cystine. Despite that, the result indicated that depletion of L-cysteine/ L-cystine might work as selection principle for CHO-K1 cells.

We developed a plasmid harboring the human CLase gene. The modified TTG Zeocin (TTG Zeo) selection marker preceded the d2EGFP reporter gene (Figure 2). We placed the CLase gene behind the d2EGFP reporter gene, separated by an IRES sequence. A version of the IRES sequence was used that lacks the last three of the twelve ATGs. This results in less efficient translation initiation of the second open reading frame in the bicistronic mRNA¹² (see Figure 4). This IRES sequence is called a weakened IRES and we dubbed it IRES^{ATG} (Figure 3). Also modified translation initiation codons were introduced in the CLase gene to test different translation initiation stringencies. We created an ATG, GTG and TTG CLase gene.

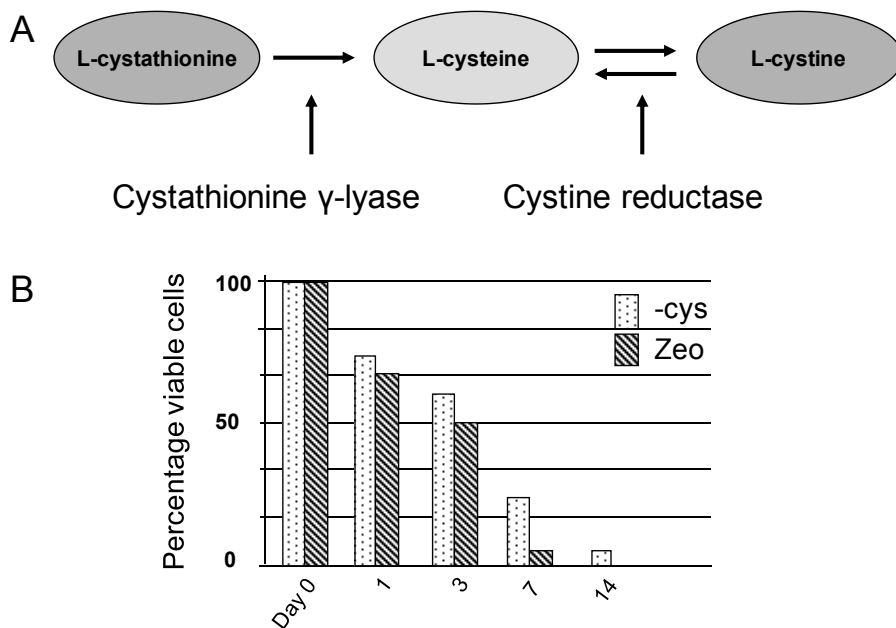


Figure 1. The metabolic pathway in which Cystathionine γ -lyase is involved and the effects of culturing in the absence of L-cysteine on CHO-K1 wild type cells.

A. The metabolic pathway of L-cysteine synthesis. Two enzymes are involved in the synthesis of L-cysteine. The Cystathionine γ -lyase (CLase) enzyme catalyzes the conversion of L-cystathionine to L-cysteine. Cystine reductase reversibly converts L-cysteine into L-cystine and vice versa. **B.** Killing curve of wild type CHO-K1 cells by Zeocin and in culture medium devoid of L-cysteine/ L-cystine. Tryphan blue staining monitored cell viability for 14 days. The percentage of viable cells is indicated.

The GTG translation initiation codon is less efficient than the ATG, and the TTG translation initiation codon is less efficient than the GTG codon^{13,14} (see Figure 4). In principle the GTG CLase derived mRNA would thus result in less translation to the CLase enzyme than the ATG CLase derived mRNA. Likewise, the TTG CLase derived mRNA would result in less translation of the CLase enzyme than the GTG CLase derived mRNA. In both cases higher bicistronic mRNA levels would be required than with an ATG CLase gene to obtain sufficient levels of the marker protein in order for the cell to survive². The expression cassette was flanked by a STAR 7/67/7 element combination cassette, thus creating the following expression cassette: STAR7/67-CMV promoter-TTG Zeo-d2EGFP-IRES ^{Δ ATG} -XGT CLase-STAR7 (Figure 2).

These plasmids were transfected to CHO-K1 cells and we selected stably transfected colonies with Zeocin (150 μ g/ml Zeocin). We isolated between six and twelve colonies. Approximately two weeks after isolation and propagation of the clones we quantified d2EGFP expression by flow cytometry (FACS). As shown in Figure 2, the average d2EGFP value in 6-12 colonies was between 300 and 400 (day 1) for the three plasmids encompassing ATG, GTG or TTG CLase, when measured in the presence of 150 μ g/ml Zeocin. To assess the influence of the presence or absence of selection pressure on the stability of protein expression, we split the cells and they were further cultured under three culturing conditions:

- Zeocin selection (150 μ g/ml Zeocin);
- No selection (no Zeocin);
- CLase selection (no Zeocin and without L-cysteine/ L-cystine, but with 0.1 mM of the precursor L-cystathionine in the medium).

After 65 days we again measured the d2EGFP values. As shown in Figure 2, the d2EGFP values remained rather stable with all three plasmids encompassing ATG, GTG or TTG *CLase* with continuous Zeocin selection pressure (day 65). Also, with all three plasmids, without any selection pressure, the average d2EGFP value in the clones dropped to below 100. However, a distinct picture emerged for the three different plasmids after the switch to *CLase* selection. d2EGFP values did not remain stable in case of the ATG *CLase*, and dropped to comparable levels as under no selection conditions. In contrast, in the case of GTG *CLase* gene, d2EGFP values increased to an average 692 (Figure 2). Finally, no colonies survived in case of the TTG *CLase* gene containing plasmid after the switch to L-cysteine/ L-cystine minus medium (Figure 2). We interpret these results as that the ATG *CLase* gene produced too much *CLase* enzyme in the bicistronic mRNA to create a stringent selection system. At the other end of the spectrum the TTG *CLase* gene produced too small amount of *CLase* enzyme to allow cell survival: the TTG *CLase* is too stringent as selection system. This indicates that the modifications that are used to modulate the different translation initiation efficiencies are rather critical.

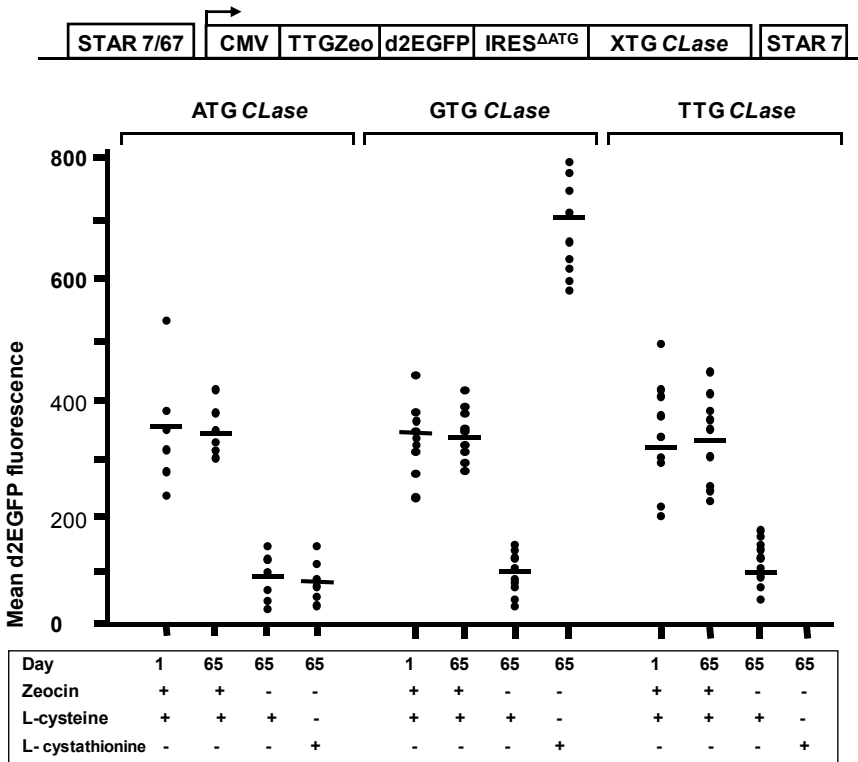


Figure 2. The use of modified *CLase* genes as maintenance marker in CHO-K1 cells.

Three *CLase* genes with different translation initiation codons, ATG, GTG and TTG are placed downstream of the d2EGFP reporter gene, separated by a weakened IRES sequence, IRES^{ATG}, as indicated. CHO-K1 cells were transfected with this plasmid and selection was performed with Zeocin. Stably transfected CHO-K1 clones were isolated and propagated before d2EGFP values were measured. First measurements were performed after selection with Zeocin, and in the presence of L-cysteine/ L-cystine in the culture medium (day 1). After this first measurement cells were split and cultured for 65 days under different selection conditions, as indicated. After 65 days d2EGFP values were determined. The mean d2EGFP values per colony are shown. The horizontal bars signify the average d2EGFP values of the colonies that are cultured under the indicated experimental conditions.

Although the GTG *CLase* gene configuration resulted in average higher d2EGFP values, we noted that the clones grew less well. When we measured the different growth rates of colonies that were subject to the different selection regimes, we found that instead of approximately one cell division per day, the GTG *CLase* derived clones divided on average once per ~ 38 hours (Figure 3). This indicates that the higher d2EGFP values that are obtained by introduction of the GTG *CLase* gene have an unwanted side effect that the growth rate in such cells is severely hampered.

Identification of higher stringency IRES sequences

The above results indicate that an ATG *CLase* gene is not stringent enough to provide maintenance of stable gene expression and that the GTG (and TTG) *CLase* gene is too stringent. In case of the GTG and TTG this results in lower cell growth (GTG) or even cell death (TTG). This could be the result of reduced translation initiation efficiency of the modified translation initiation codons. If this were the case, a translation initiation codon, less stringent than the GTG would be advisable for this selection system. Unfortunately, we do not know such codons and we, therefore, opted for a different approach. We used the wild type, ATG *CLase* configuration and attempted to further hamper the function of the IRES^{ΔATG} sequence, in order to reduce the translation initiation efficiency of the downstream gene. Previously it has been described that the introduction of specific mutations in the IRES sequence indeed reduces the translation efficiency of the downstream-located mRNA¹⁵ (Figure 4A). One class of mutations involves modification of the IRES sequence at bases 704 (A) and 705 (C)¹⁵ (Figure 4A). We introduced and tested two of the described mutations at these sites. Bases 704 (A) and 705 (C) were exchanged by a G and T, which created the IRES^{GT} (Figure 4A). Another mutant involved changing only 704 (A) into T, thus creating mutant IRES^T. To test the translation initiation efficiencies of these IRES mutations, we placed the d2EGFP reporter gene downstream of these IRES sequences. Upstream of the IRES sequences we placed the (ATG) Zeocin gene (Figure 4B). As controls we took the wild type IRES

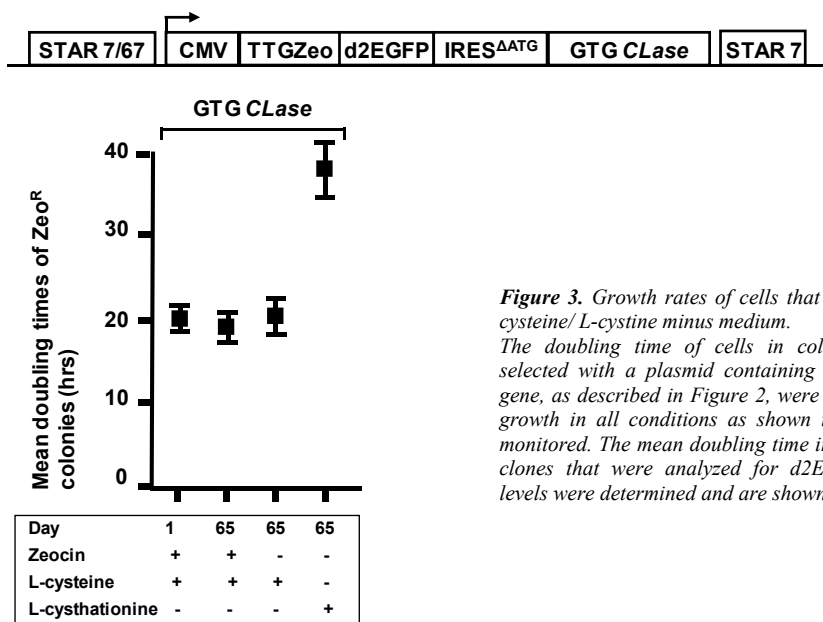


Figure 3. Growth rates of cells that are grown in L-cysteine/ L-cystine minus medium. The doubling time of cells in colonies that were selected with a plasmid containing the GTG *CLase* gene, as described in Figure 2, were determined. Cell growth in all conditions as shown in Figure 2 was monitored. The mean doubling time in hours of the 12 clones that were analyzed for d2EGFP expression levels were determined and are shown in the graph.

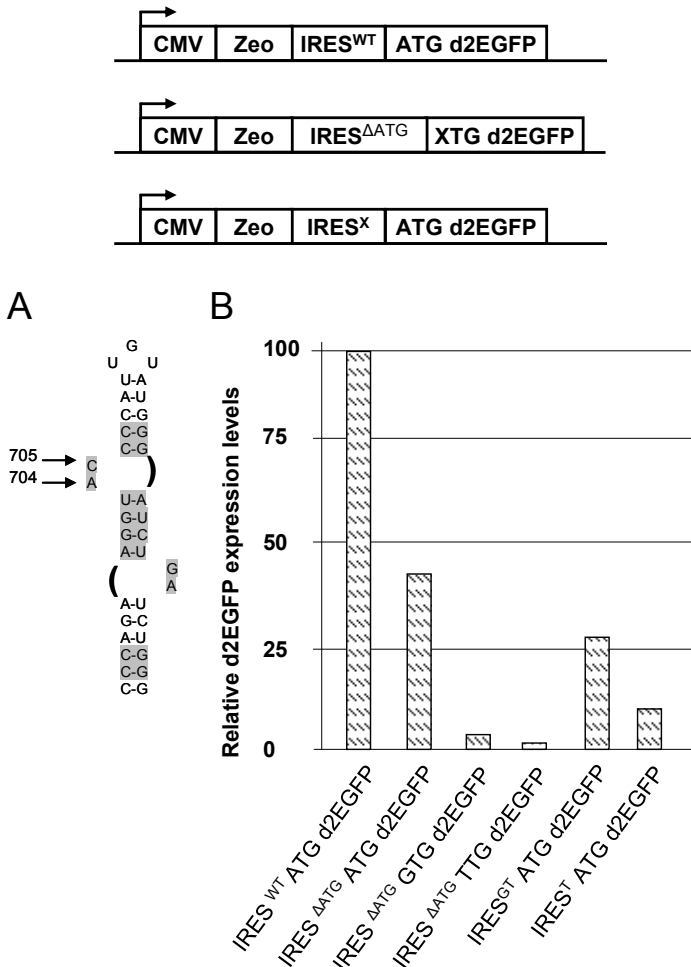


Figure 4. Transient expression of the d2EGFP gene that is placed downstream of different IRES sequences.

A. The secondary structure of the relevant portion (the distal stem-loop of the J domain) of the encephalomyocarditis virus (EMCV) IRES sequence that has been mutated is shown. Grey background indicates nucleotides that are highly conserved across all the cardiovirus and aphthovirus IRES sequences. **B.** The d2EGFP gene was placed downstream of indicated IRES sequences, as described in the text. Beside the ATG d2EGFP gene, also the GTG and TTG d2EGFP genes were used. The different plasmids were transfected to CHO-K1 cells and the transient d2EGFP expression of the transfected pool of cells was determined one day after transfection. The d2EGFP expression of the plasmid in which the ATG d2EGFP gene was placed downstream of the wild type IRES sequence was the highest of all plasmids and was arbitrarily put at 100%.

sequence^{16,17} and the ‘weakened’ IRES^{ΔATG} sequence that we used in the above-described experiments. Downstream of this IRES^{ΔATG} we placed the ATG, GTG and TTG d2EGFP to obtain the same configuration as in the experiments we described in Figure 2. We transfected the plasmids to CHO-K1 cells and measured the transient d2EGFP expression levels. As shown in Figure 4B, the different plasmids indeed displayed different transient d2EGFP expression levels. As expected, the plasmid with the wild type IRES sequence gave the highest d2EGFP expression levels, followed by the IRES^{ΔATG} ATG d2EGFP (Figure 4B). The IRES^{ΔATG} GTG d2EGFP and IRES^{ΔATG} TTG d2EGFP plasmids gave hardly any d2EGFP signal, demonstrating in a direct way the high stringency of the translation initiation that accompanies these modified translation initiation codons. Finally, both IRES^{GT} and IRES^T sequences gave less d2EGFP signal than the wild type IRES and IRES^{ΔATG} ATG d2EGFP plasmid, indicating a lower translation efficiency of the corresponding mRNA. However, these mutant IRES sequences gave a higher transient d2EGFP signal than the IRES^{ΔATG} GTG d2EGFP and IRES^{ΔATG} TTG d2EGFP plasmids (Figure 4B), suggesting a higher translation initiation efficiency than the GTG and TTG modifications.

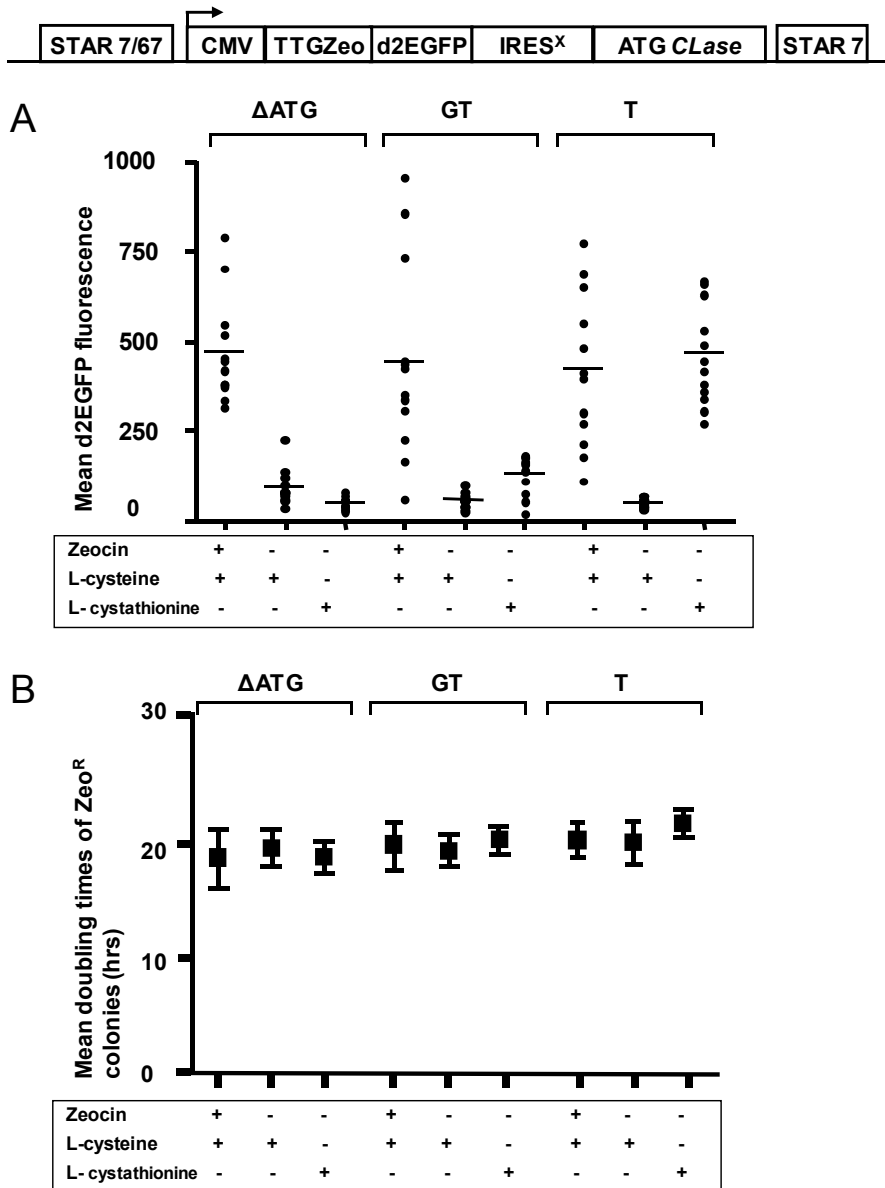


Figure 5. The use of CLase as maintenance marker in the context of different IRES mutants. *A.* Indicated are the plasmids in which the ATG d2EGFP is placed downstream of the IRES ^{Δ ATG}, IRES^{GT} and IRES^T sequences, as described in the text. CHO-K1 cells were transfected with these plasmids and selection was performed with Zeocin. Stably transfected CHO-K1 clones were isolated and propagated before d2EGFP values were determined. With all three plasmids, measurements were performed under maximum Zeocin selection pressure for 65 days, no selection (no Zeocin, but with L-cysteine/ L-cystine) and under CLase selection (no Zeocin and no L-cysteine/ L-cystine, but L-cystathionine in the medium), as indicated with pluses and minuses. The mean d2EGFP values per colony are shown. The horizontal bars signify the average d2EGFP values of the colonies that are given in the lowest line of the text box. *B.* The doubling time of cells in colonies that were selected with a plasmid containing the CLase gene downstream of a specific IRES sequence, as described in *A.*, were determined. The mean doubling time in hours of the clones that were analyzed for d2EGFP expression levels were determined and are shown in the graph.

Based on this result, we cloned the ATG *CLase* gene downstream of the mutated IRES sequences and transfected them to CHO-K1 cells (Figure 5A). As control plasmid we took the plasmid with the IRES^{ΔATG} ATG *CLase* sequence. Similar as described above, we isolated 12 stably transfected clones (150 μg/ml Zeocin). Approximately two weeks after isolation and propagation of the clones we quantified d2EGFP expression (data not shown). As before, we split the cells and they were further cultured under three culturing conditions:

- Zeocin selection (150 μg/ml Zeocin);
- No selection (no Zeocin);
- CLase selection (no Zeocin and without L-cysteine/ L-cystine, but with 0.1 mM of the precursor L-cystathionine in the medium).

After 65 days we again measured the d2EGFP values. In case of the IRES^{ΔATG} ATG *CLase* plasmid, the d2EGFP values were not stable in the absence of L-cysteine/ L-cystine and in the presence of L-cystathionine (Figure 5A, as in Figure 2). Also the plasmid with the IRES^{GT} mutant did not provide stable d2EGFP expression (Figure 5A). However, the IRES^T mutant did provide stable d2EGFP expression levels after 65 days culturing without L-cysteine/ L-cystine (Figure 5A). We also measured the growth rate of the clones transfected with the IRES^{ΔATG}, IRES^{GT} and IRES^T sequences, 65 days after splitting the cells. We found no significant differences in growth rate between any of these clones, whether they remained on Zeocin selection for 65 days, with no selection or on CLase selection (Figure 5B). These results indicate the usefulness of the *CLase* gene as maintenance marker, when used in combination with a specific IRES sequence. Apparently, the level of translation efficiency of the corresponding *CLase* mRNA is critical for providing high enough selection stringency, without hampering cell growth.

Direct selection of CHO-K1 clones by using the CLase gene

Above, we showed that the *CLase* gene can be used as maintenance marker for stable gene expression when used in the context of a specific IRES mutant sequence. We next tested whether this combination could also be used for direct selection of CHO-K1 clones in culture medium devoid of L-cysteine/ L-cystine and in the presence of the precursor L-cystathionine. We removed the TTG Zeo marker from the constructs, thus creating plasmids with the IRES^{ΔATG}, IRES^{GT} and IRES^T sequences, upstream of the ATG *CLase* gene (Figure 6). These

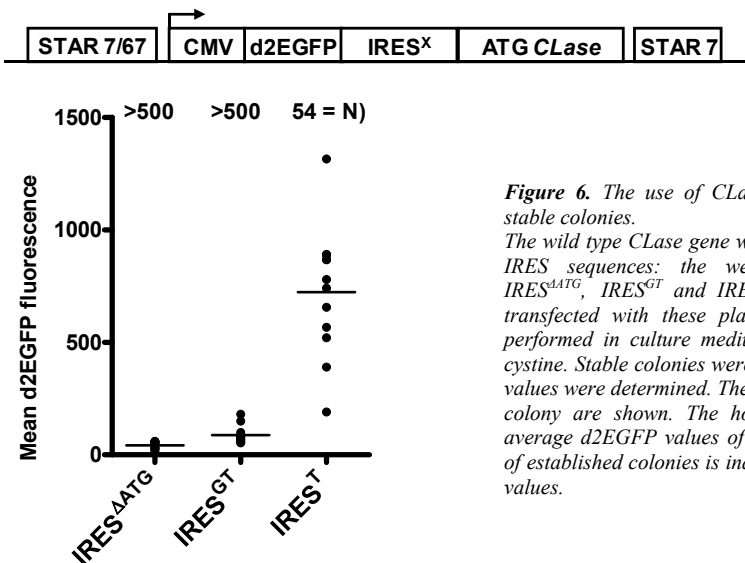


Figure 6. The use of *CLase* for direct selection of stable colonies.

The wild type *CLase* gene was placed behind different IRES sequences: the weakened IRES sequence, IRES^{ΔATG}, IRES^{GT} and IRES^T. CHO-K1 cells were transfected with these plasmids and selection was performed in culture medium without L-cysteine/ L-cystine. Stable colonies were established and d2EGFP values were determined. The mean d2EGFP values per colony are shown. The horizontal bars signify the average d2EGFP values of the colonies. The number of established colonies is indicated above the d2EGFP values.

plasmids were transfected to CHO-K1 cells and cultured on medium devoid of L-cysteine/ L-cystine and in the presence of the precursor L-cystathionine. As shown in Figure 6, more than 500 colonies formed with the plasmids containing the IRES^{ΔATG} and IRES^{GT} sequences. The d2EGFP values in the corresponding clones were also very low (Figure 6). In contrast, the plasmid containing the IRES^T sequence induced only 54 colonies. This is very similar to the average colony number that is established with a plasmid harboring the TTG Zeo d2EGFP configuration. The average d2EGFP values in these clones was 722 (Figure 6) and the clones grew well (data not shown). This differs from the TTG Zeo induced clones that were subsequently switched to L-cysteine/ L-cystine minus medium, in the context of the GTG *CLase* gene (Figure 2). After that switch, these clones displayed strongly increased d2EGFP values, but also increasing doubling times (Figure 2 and 3). Apparently, switching from a low to a higher selection stringency (concomitant with higher d2EGFP values) has an impact on cell doubling times. Instead, when clones are selected from the start in L-cysteine/ L-cystine minus medium, cells that grow too slow probably won't survive and only normally outgrowing colonies are selected. This would argue that also direct selection with the GTG *CLase* would be an option. Several attempts to achieve this were, however, unsuccessful. Apparently the IRES GTG *CLase* configuration is still more stringent than the IRES^T ATG *CLase* configuration.

We also tested the potential of the *CLase* gene as marker for direct selection with human erythropoietin (hEPO) as reporter gene. As shown in Figure 7A, plasmids with the IRES^{ΔATG} and the IRES^{GT} IRES sequences gave >500 colonies and very low specific hEPO expression levels. In contrast, the plasmid with the IRES^T sequence gave 34 colonies. The average hEPO

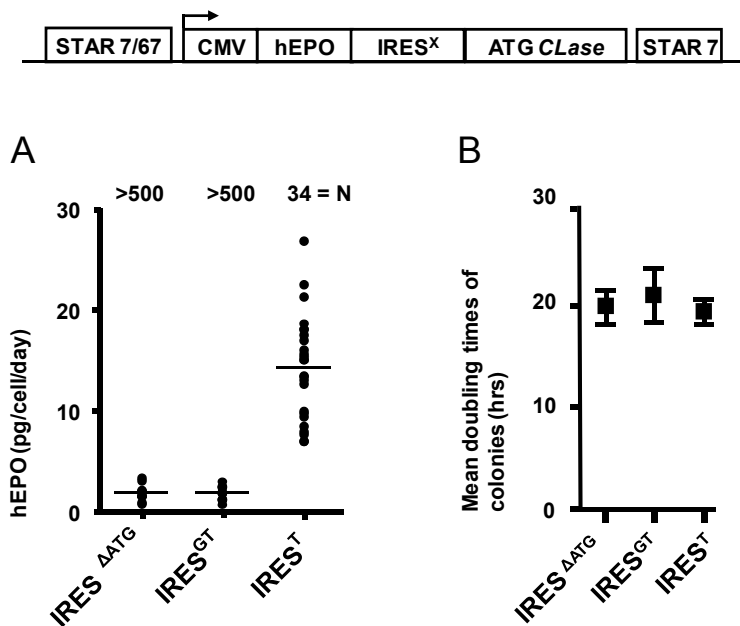


Figure 7. The use of *CLase* for direct selection of stable colonies, with hEPO as reporter gene. **A.** The wild type *CLase* gene was placed behind different IRES sequences: IRES^{ΔATG}, IRES^{GT} and IRES^T. The gene encoding for the therapeutic protein hEPO was placed upstream of these IRES sequences. CHO-K1 cells were transfected with these plasmids and selection was performed in culture medium without L-cysteine/ L-cystine. Stable colonies were established and hEPO values were determined. The mean specific hEPO activities, (pg/ cell/ day) per colony are shown. The horizontal bars signify the average hEPO values of the colonies. The number of established colonies is indicated above the hEPO values. **B.** The mean doubling time of the colonies was determined.

expression in 24 colonies was 14.3 pg/ cell/ day (Figure 7A). Also the growth rates in the respective clones was measured and no significant differences were observed, the average cell doubling time was approximately 20 hours with all three plasmids (Figure 7B).

These results indicate that the IRES^T mutant sequence provides such translation stringency that the 'wild type' ATG *CLase* can be used as marker for direct selection in CHO-K1 cells, without hampering cell growth.

DISCUSSION

In this paper we provide evidence that the CLase enzyme that converts a non-toxic precursor, L-cystathionine, into the amino acid L-cysteine can be used as an efficient selection marker in CHO-K1 cells. Usually, protein-expressing mammalian cell lines are generated by a selection system based on antibiotics. One drawback of antibiotics as selection agents is that they are expensive, particularly when they are used during large-scale production runs. A second drawback is their toxicity. Keeping antibiotics up to the end in a production run requires costly additional purification steps. Therefore, the use of a metabolic marker that does not involve toxic substances is preferable. Two metabolic markers are commonly being used for this purpose: the *Dhfr* gene and glutamine synthetase (*GS*)^{5,6,18}.

The DHFR protein is an enzyme in the folate pathway and the GS enzyme is involved in the synthesis of the essential amino acid glutamine. DHFR and GS require respectively dihydrofolate and glutamate as precursor in the culture medium that is devoid of the end products made by these enzymes. Therefore both enzymes can serve as selection marker. There are, however, restrictions to the use of these enzymes as selection marker. For instance, there are only a limited number of mammalian cells lines that lack the *Dhfr* or *GS* gene. In the case of the *Dhfr* gene these are *Dhfr* minus CHO cells, such as CHO-DG447, and it is not surprising that this a commonly used cell line in the biotechnology industry for the production of therapeutic proteins. The mouse hybridoma cell line NS0 is *GS* minus and can be properly used with the *GS* gene as a selection marker, but this cell line is less often used for the production of therapeutic proteins. These restrictions can partly be overcome by the addition of toxic inhibitors of either the DHFR or GS enzymes: respectively methotrexate and methionine sulphoxime¹⁸⁻²⁰. Addition of these inhibitors to cells that are *Dhfr*+ or *GS*+ results in the inhibition of the activity of the endogenous DHFR and GS proteins. This inhibition subsequently requires the addition of exogenously DHFR/GS enzymes, which are provided by the transfected protein expression plasmids. Unfortunately we are now in part back to where we started: the use of toxic substances that we wanted to avoid in the first place.

How does the CLase enzyme as selection marker compare to the DHFR and GS system? In the first place, unlike the GS enzyme, the CLase enzyme appears to be a useful marker for CHO cell lines, without the need for toxic inhibitors of the endogenous enzyme. This is because CHO cells are deficient in the CLase enzyme. In this regard, CLase has more in common with the DHFR enzyme. We tested other CHO cell lines, such as the commonly used CHO-DG44 cell line and found that the *CLase* gene can be used as marker for direct selection in these cells as well (data not shown). Like the DHFR marker, the CLase is less attractive for cell lines other than CHO cell lines, such as cell lines derived from humans (e.g. PER.C6)²¹. The simple reason is that these cells harbor a functional *CLase* gene. The *Dhfr* gene as selection marker is often used in combination with Neomycin as selection agent. The reason is that cells do not very rapidly die in medium lacking hypoxanthine and thymidine. Such medium is needed for DHFR as selection marker, since they are the end products of the metabolic pathway in which DHFR participates. Therefore, often Neomycin and the Neomycin resistance marker are combined with the *Dhfr* gene as selection marker to facilitate cell death of

non-transfected cells and increase the efficiency of the selection procedure. Furthermore, the DHFR selection system requires the addition of the inhibitor methotrexate that induces gene amplification^{19,20,22}. This is done to enhance the expression levels of the protein of interest in the DHFR-selected expression cassette. Finally, multiple rounds of time consuming subcloning are the standard procedure to induce and obtain high protein expressing clones with the DHFR selection system. None of these procedures appear to be needed with CLase as selection marker.

Why is the *CLase* gene successful as selection marker? The most straightforward answer is that cells rapidly die in culture medium devoid of L-cysteine/ L-cystine. This allows for a timely identification of emerging, transfected colonies, without a fast overgrowing monolayer of non-transfected cells. However, even if this were not the case, *CLase* would still be useful as maintenance marker. This requires that after a selection agent such as Zeocin has killed non-transfected cells, a subsequent switch to L-cysteine/ L-cystine minus medium is needed. Although maybe not preferable, the limited use of Zeocin for two weeks initial selection is much less expensive and toxic than keeping it present during production phases. An alternative selection strategy might be to start selection immediately in L-cysteine/ L-cystine minus medium and simultaneously add a low concentration of Zeocin.

One complicating aspect of CLase as selection marker is determining the optimal level of CLase enzyme that is needed for efficient selection. Placing the *CLase* gene downstream of the commercially available, weakened IRES^{ΔATG} does not provide sufficient selection stringency to the system to establish high protein expressing colonies. At the other end of the spectrum, employing very high stringency translation initiation codons, such as the GTG codon does induce high protein expressing colonies, but these are severely hampered in their growth. This makes the STAR-Select modifications that we previously used rather useless in employing the *CLase* gene as selection marker. Instead, we had to resort to identify mutations in the IRES sequence that reduce the ability of the IRES sequence to allow translation initiation of the open reading frame downstream of the IRES. When one such mutated IRES sequence was placed upstream of the *CLase* selection marker, the induced reporter protein expression levels were comparable with those derived from the STAR-Select system. Importantly, these colonies were not hampered in their growth rate. It remains to be seen, however, whether this particular IRES mutation always provides the 'right' selection stringency to the *CLase* system, for instance when used in other CHO cell lines. The width of the stringency spectrum is obviously not very broad and it is feasible that additional IRES mutations have to be tested to identify ones that allow the desired balance between protein expression levels and cell growth.

Another possibility to create optimal selection stringency might be to modify the CLase protein itself. One could introduce mutations in the coding sequence of the *CLase* gene that leads to reduced enzyme activity. This would also result in more stringent selection stringency and potentially avoid tinkering with translation initiation efficiencies. Future developments like that may be important for evaluating the usefulness of the *CLase* gene as selection marker in an industrial setting.

MATERIALS AND METHODS

Vector constructions

The human Cystathionine γ -lyase (*CLase*) coding region was synthesized by GENEART with three kinds of translation initiation codons (ATG, GTG, TTG). Together with the EMCV IRES (Clontech)¹² this gene was cloned into a STAR-Select vector². The IRES mutants were generated by overlap PCR. Two complementary oligonucleotides containing the mutated

base(s) were employed as primers. Together with a forward and reverse primer (spanning the 5' and 3' ends of the IRES, respectively), two separate PCRs were performed. The products were purified and a mixture of these products was used as the template in another PCR employing just the forward and reverse primers spanning the 5' and 3' ends of the IRES. MluI and BamHI restriction sites were introduced by PCR and used for cloning.

F MluI 5' AGGCACGCGTCGAGCATGCATCTAGGGCGGC 3'
 R^{GT} 5' CCATACAATGGG**AC**ACCTTCTGG 3'
 F^{GT} 5' CCAGAAGGT**GT**CCCATTTGTATGG 3'
 R^T 5' CCATACAATGGG**GA**ACCTTCTGG 3'
 F^T 5' CCAGAAGGT**TCC**CCATTTGTATGG 3'
 R BamHI 5' AGCGGATCCCGGGTTGTGGCAAGCTTATCATC 3'

The IRES mutants were sequenced. For direct selection with the CLase gene, the TTG Zeo was removed from the expression cassette.

Cell culture, transfection, and analysis of clones

CHO-K1 cells (CCL-61; American Type Culture Collection (ATCC)) were grown in HamF12 medium (Gibco), supplemented with 9.1% fetal bovine serum (FBS) (Invitrogen), 2 mM glutamine (Invitrogen), 100 U/ml penicillin (Invitrogen), and 100 µg/ml streptomycin (Invitrogen) at 37°C / 5% CO₂. For culturing in the absence of L-cysteine/ L-cystine, special HamF12 medium lacking L-cysteine/ L-cystine (designed and produced by Invitrogen) and dialyzed FBS (Invitrogen) was used. L-cystathionine (Sigma) was diluted in 0.5 M HCl and used at a final concentration in the culture media of 0.1 mM.

For transfections, 0.3·10⁶ CHO-K1 cells were seeded in 6-well culture plates 24 hours prior to transfection. Cells were transfected with 3 µg of plasmid DNA using LipofectamineTM 2000 (Invitrogen) as described by the manufacturer. Selection involved Zeocin (Invitrogen) at a concentration of 150 µg/ml. Approximately 12 days after transfection, individual colonies became visible and these were isolated and propagated in 24-well plates in medium containing Zeocin. When grown to ~70% confluence, cells were transferred to 6-well plates. Cells were continued to grow in 6-well plates for another one to two weeks before FACS analysis was performed. The d2EGFP expression levels were determined on an Epics XL Beckman Coulter flowcytometer. At this point of time, after determining d2EGFP expression levels, cells were split to the various media. Cells were continued to grow for 65 days in the various HamF12 media: medium without antibiotics; medium (containing L-cysteine/ L-cystine) with Zeocin; medium lacking L-cysteine/ L-cystine and containing 0.1 mM L-cystathionine. After 65 days d2EGFP expression levels were determined again. Values were visualized using Graphpad Prism 5 for Windows.

For determining survival rates of CHO-K1 wild-type cells, 0.2·10⁶ cells were seeded in 6-well plates in medium containing 150 µg/ml Zeocin or in medium lacking L-cysteine/ L-cystine. Daily, viability of cells was monitored by trypan blue staining for a period of two weeks. For determining growth rates of CHO-K1 subclones, 5·10⁴ cells were seeded in duplo in 6-well plates. After 24 hours, cells were counted. After another 48 hours, cells were counted again. With these numbers, doubling times were calculated.

ELISA

For hEPO measurements, equal numbers of cells (0.1·10⁶) were seeded in six-well plates three days prior to cell counting and collection of the medium. The amount of secreted hEPO was determined using an ELISA-kit (Quantikine IVD kit; R&D systems; Ref DEP00). The antibody concentration was determined by comparing optical density at 415 nm with that of the known antibody standard, as supplied in the ELISA-kit.

REFERENCES

1. Kwaks, T.H.J., Barnett, P., Hemrika, W., Siersma, T., Sewalt, R.G.A.B., Satijn, D.P.E., Brons, J.F., Van Blokland, R., Kwakman, P., Kruckeberg, A.L., Kelder, A. and Otte, A.P. Identification of anti-repressor elements that confer high and stable protein production in mammalian cells. *Nat Biotechnol* **21**, 553-8 (2003).
2. Van Blokland, H.J.M., Kwaks, T.H.J., Sewalt, R.G.A.B., Verhees, J.A., Klaren, V.N.A., Siersma, T.K., Korse, J.W.M., Teunissen, N.C., Botschuijver, S., Van Mer, C., Man, S.Y. and Otte, A.P. A novel, high stringency selection system allows screening of few clones for high protein expression. *J Biotechnol* **128**, 237-245 (2007).
3. Otte, A.P., Kwaks, T.H., Van Blokland, R.J., Sewalt, R.G., Verhees, J., Klaren, V.N., Siersma, T.K., Korse, H.W., Teunissen, N.C., Botschuijver, S., Van Mer, C. and Man, S.Y. Various expression-augmenting DNA elements benefit from STAR-Select, a novel high stringency selection system for protein expression. *Biotechnol Prog* **23**, 801-7 (2007).
4. Urlaub, G. and Chasin, L.A. Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proc Natl Acad Sci U S A* **77**, 4216-4220 (1980).
5. Kaufman, R.J. and Sharp, P.A. Amplification and expression of sequences cotransfected with a modular dihydrofolate reductase complementary DNA gene. *J Mol Biol* **159**, 601-21 (1982).
6. McCormick, F., Trahey, M., Innis, M., Dieckmann, B. and Ringold, G. Inducible expression of amplified human beta interferon genes in CHO cells. *Mol Cell Biol* **4**, 166-72 (1984).
7. Urlaub, G., Käs, E., Carothers, A.M. and Chasin, L.A. Deletion of the diploid dihydrofolate reductase locus from cultured mammalian cells. *Cell* **33**, 405-12 (1983).
8. Meister, A. and Anderson, M.E. Glutathione. *Annu Rev Biochem* **52**, 711-60 (1983).
9. Rose, W.C. and Wixom, R.L. The amino acid requirements of man. XIII. The sparing effect of cystine on the methionine requirement. *J Biol Chem* **216**, 753-73 (1955).
10. Naylor, S.L., Busby, L.L. and Klebe, R.J. Biochemical selection systems for mammalian cells: the essential amino acids. *Somatic Cell Genet* **2**, 93-111 (1976).
11. Naylor, S.L., Townsend, J.K. and Klebe, R.J. Characterization of naturally occurring auxotrophic mammalian cells. *Somatic Cell Genet* **5**, 271-7 (1979).
12. Rees, S., Coote, J., Stables, J., Goodson, S., Harris, S. and Lee, M.G. Bicistronic vector for the creation of stable mammalian cell lines that predisposes all antibiotic-resistant cells to express recombinant protein. *Biotechniques* **20**, 102-110 (1996).
13. Kozak, M. Context effects and inefficient initiation at non-AUG codons in eucaryotic cell-free translation systems. *Mol Cell Biol* **9**, 5073-80 (1989).
14. Kozak, M. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc Natl Acad Sci USA* **87**, 8301-5 (1990).
15. Clark, A.T., Robertson, M.E.M., Conn, G.L. and Belsham, G.J. Conserved nucleotides within the J domain of the encephalomyocarditis virus internal ribosome entry site are required for activity and for interaction with eIF4G. *J Virol* **77**, 12441-12449 (2003).
16. Jang, S.K., Kräusslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C. and Wimmer, E. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol* **62**, 2636-43 (1988).
17. Jackson, R.J., Howell, M.T. and Kaminski, A. The novel mechanism of initiation of picornavirus RNA translation. *Trends Biochem Sci* **15**, 477-83 (1990).
18. Bebbington, C.R., Renner, G., Thomson, S., King, D., Abrams, D. and Yarranton, G.T. High-level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker. *Bio/Technology* **10**, 169-75 (1992).
19. Kaufman, R.J. and Schimke, R.T. Amplification and loss of dihydrofolate reductase genes in a Chinese hamster ovary cell line. *Mol Cell Biol* **1**, 1069-76 (1981).
20. Nunberg, J.H., Kaufman, R.J., Schimke, R.T., Urlaub, G. and Chasin, L.A. Amplified dihydrofolate reductase genes are localized to a homogeneously staining region of a single chromosome in a methotrexate-resistant Chinese hamster ovary cell line. *Proc Natl Acad Sci USA* **75**, 5553-6 (1978).
21. Jones, D., Kroos, N., Anema, R., Van Montfort, B., Vooy, A., Van der Kraats, S., Van der Helm, E., Smits, S., Schouten, J., Brouwer, K., Lagerwerf, F., Van Berker, P., Opstelten, D.J., Logtenberg, T. and Bout, A. High-level expression of recombinant IgG in the human cell line per.c6. *Biotechnol Prog* **19**, 163-8 (2003).
22. Kaufman, R.J., Sharp, P.A. and Latt, S.A. Evolution of chromosomal regions containing transfected and amplified dihydrofolate reductase sequences. *Mol Cell Biol* **3**, 699-711 (1983).