Title: Novel sequence for improving expression of nucleic acid

The invention relates to the field of molecular biology and biotechnology. More specifically the present invention relates to means and methods for improving the production of one or more nucleic acids that may encode proteins in a host cell.

Proteins can be produced in various host cells for a wide range of applications in biology and biotechnology, for instance as biopharmaceuticals. Methods for such production are well established, and generally entail the expression in a host cell of a nucleic acid (also referred to as 'transgene') encoding the protein of interest.

One problem associated with the expression of transgenes is that it is unpredictable, stemming from the high likelihood that the transgene will become inactive due to gene silencing (McBurney et al., 2002), and therefore many host cell clones have to be tested for high expression of the transgene. Furthermore, once such a clone has been established, the expression of the transgene is often not stable and silencing of transgene expression during prolonged host cell cultivation is a commonly observed phenomenon. In vertebrate cells it can be caused by formation of heterochromatin at the transgene locus, which prevents transcription of the transgene (Whitelaw et al, 2001). Transgene silencing is stochastic; it can occur shortly after integration of the transgene into the genome, or only after a number of cell divisions. This results in heterogeneous cell populations after prolonged cultivation, in which some cells continue to express high levels of recombinant protein while

others express low or undetectable levels of the protein (Martin & Whitelaw, 1996, McBurney et al., 2002). A cell line that is used for heterologous protein production is derived from a single cell, yet is often scaled up to, and maintained for long periods at, cell densities in excess of ten million cells per ml in cultivators of 1,000 liters or more. These large cell populations (10^{14} - 10^{16} cells) are prone to serious declines in productivity due to transgene silencing (Migliaccio et al., 2000, Strutzenberger et al., 1999).

One possibility to overcome the problems described above is to employ so-called STAR sequences in the expression system. A variety of such STAR sequences have been described in WO 03/004704. These sequences can be used to improve predictability, yield and/or stability of protein production using expression constructs in several types of host cells (WO 03/004704; Kwaks et al, 2003).

International publication WO 03/106674 describes the use of STAR sequences for expressing nucleic acid encoding recombinant proteins in several cell lines.

International publication WO 03/106684 describes the use of STAR sequences for expressing nucleic acid encoding multimeric proteins, such as antibodies.

The present invention aims at providing alternative means and methods for improving protein production.

Summary of the invention

In one aspect, the invention provides a recombinant nucleic acid molecule comprising a nucleic acid sequence having anti-repressor activity selected from the group consisting of: a) SEQ. ID. NO. 66, b) fragments of SEQ. ID. NO. 66 wherein said fragments have anti-repressor activity, c) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences

have anti-repressor activity; and d) the complement to any one of a) to c); said recombinant nucleic acid molecule further comprising an expression cassette, said expression cassette comprising a heterologous promoter linked to a nucleic acid of interest, and wherein said nucleic acid of interest preferably encodes all or part of a protein of interest. In a preferred embodiment, said nucleic acid sequence having anti-repressor activity is situated upstream of said promoter in said expression cassette, and in particularly preferred embodiments thereof, said sequence having anti-repressor activity and said promoter are separated by less than 2 kb. In certain embodiments, said nucleic acid encoding a protein of interest is present in a multicistronic gene further encoding a selectable marker gene. In certain embodiments, said molecule further comprises at least one other sequence having anti-repressor activity, said at least one other sequence being chosen from a) any of SEQ. ID. NOs. 1-65, b) fragments of any of SEQ. ID. NOs. 1-65, wherein said fragments have anti-repressor activity, c) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences have anti-repressor activity, and d) the complement to any one of a) to c). The invention also provides a recombinant nucleic acid molecule comprising an expression cassette comprising: 5′- anti-repressor sequence A - promoter - nucleic acid encoding all or part of a protein of interest - anti-repressor sequence B - 3′ wherein anti-repressor sequences A and B may be the same or different and are chosen from the group consisting of (i) any of SEQ. ID. NOs. 1-65, (ii) fragments of any of SEQ. ID. NOs. 1-65, wherein said fragments have anti-repressor activity, (iii) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences

have anti-repressor activity, and (iv) the complement to any one of (i) to (iii), characterized in that said expression cassette further comprises between said anti-repressor sequences A and B a sequence having anti-repressor activity chosen from the group consisting of a) SEQ. ID. NO. 66, b) fragments of SEQ. ID. NO. 66 wherein said fragment has anti-repressor activity, c) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences have anti-repressor activity; and d) the complement to any one of a) to c).

The invention further provides a cell comprising a molecule according to the invention. Preferably said cell is a mammalian cell, and in certain embodiments said cell is a CHO cell.

The invention further provides a method for producing a protein of interest, comprising culturing a cell comprising a recombinant nucleic acid molecule encoding the protein of interest to express said nucleic acid encoding the protein of interest in said cell, characterized in that said recombinant nucleic acid molecule comprises a nucleic acid sequence having anti-repressor activity selected from the group consisting of: a) SEQ. ID. NO. 66, b) fragments of SEQ. ID. NO. 66 wherein said fragments have anti-repressor activity, c) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences have anti-repressor activity; and d) the complement to any one of a) to c). In a preferred embodiment, the method further comprises isolating said protein of interest. In preferred embodiments, said nucleic acid sequence having anti-repressor activity is situated upstream of a promoter that controls expression of the protein of interest. In a preferred embodiment thereof,

said sequence having anti-repressor activity and said promoter are separated by less than 2 kb. In certain embodiments, said nucleic acid encoding a protein of interest is present in a multicistronic gene further encoding a selectable marker gene. In certain embodiments, said cell is a mammalian cell, e.g. a CHO cell.

The invention also provides for the use of a nucleic acid sequence having anti-repressor activity selected from the group consisting of: a) SEQ. ID. NO. 66, b) fragments of SEQ. ID. NO. 66 wherein said fragment has anti-repressor activity, c) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences have anti-repressor activity; and d) the complement to any one of a) to c), for increasing expression of a nucleic acid of interest.

In yet another aspect, the invention provides a method for generating a host cell expressing two polypeptides of interest, the method comprising: a) introducing into a host cell one or more nucleic acid molecules, the nucleic acid molecule or molecules together comprising: (i) a promoter functionally linked to a sequence encoding a first polypeptide of interest and a first selectable marker gene, and (ii) a promoter functionally linked to a sequence encoding a second polypeptide of interest and a second selectable marker gene, and (iii) at least one sequence having anti-repressor activity, chosen from the group consisting of (a) any one of SEQ. ID. NOs. 1-66, (b) fragments of any one of SEQ. ID. NOs. 1-66, said fragments having anti-repressor activity, (c) sequences that are at least 70% identical to (a) or (b) and having anti-repressor activity, and (d) the complement of any one of (a)-(c);

b) selecting a host cell by selecting essentially simultaneously for expression of said first and second selectable marker genes. The invention also provides a method for expressing two polypeptides of interest, the method comprising: culturing a host cell obtained by the method according to this aspect of the invention to express said first and second polypeptides, and optionally isolating said polypeptides. In one embodiment, said sequence having anti-repressor activity is selected from the group consisting of a) SEQ. ID. NO. 66, b) fragments of SEQ. ID. NO. 66 wherein said fragment has anti-repressor activity, c) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences have anti-repressor activity; and d) the complement to any one of a) to c). In certain embodiments, said two polypeptides are part of a multimeric protein.

Detailed description of the invention

Anti-repressor sequences

Sequences having anti-repressor activity as used herein are sequences that are capable of at least in part counteracting the repressive effect of HP1 or HPC2 proteins when these proteins are tethered to DNA. Sequences having anti-repressor activity (sometimes also referred to as anti-repressor sequences or anti-repressor elements herein) suitable for the present invention, have been disclosed in WO 03/004704, incorporated herein by reference, and were coined "STAR" sequences therein (wherever a sequence is referred to as a STAR sequence herein, this sequence has anti-repressor activity according to the invention). As a non-limiting example, the sequences of 65 anti-repressor elements, named STAR1-

65 (see WO 03/004704), are presented herein as SEQ. ID. NOs. 1-65, respectively.

According to the invention, a functional fragment or derivative of a given anti-repressor element is considered equivalent to said anti-repressor element, when it still has anti-repressor activity. The presence of such anti-repressor activity can easily be checked by the person skilled in the art, for instance by the assay described below. Functional fragments or derivatives can easily be obtained by a person skilled in the art of molecular biology, by starting with a given anti-repressor sequence, and making deletions, additions, substitutions, inversions and the like (see e.g. WO 03/004704). A functional fragment or derivative also comprises orthologs from other species, which can be found using the known anti-repressor sequences by methods known by the person skilled in the art (see e.g. WO 03/004704). Hence, the present invention encompasses fragments of the anti-repressor sequences, wherein said fragments still have anti-repressor activity. For fragments of a given sequence, percent identity refers to that portion of the reference native sequence that is found in the fragment. The invention also encompasses sequences that are at least 70% identical in nucleotide sequence to said sequences having anti-repressor activity or to functional fragments thereof having anti-repressor activity, as long as these sequences that are at least 70% identical still have the anti-repressor activity according to the invention. Preferably, said sequences are at least 80% identical, more preferably at least 90% identical and still more preferably at least 95% identical to the reference native sequence or functional fragment thereof.

Sequences having anti-repressor activity according to the invention can be obtained by various methods, including but not limited to the cloning from the human genome or from the genome of another organism, or by for instance amplifying known anti-repressor sequences directly from such a genome by using the knowledge of the sequences, e.g. by PCR, or can in part or wholly be chemically synthesized.

Sequences having anti-repressor activity, and functional fragments or derivatives thereof, are structurally defined herein by their sequence and in addition are functionally defined as sequences having anti-repressor activity, which can be determined with the assay described below.

Any sequence having anti-repressor activity according to the present invention should at least be capable of surviving the following functional assay (see WO 03/004704, example 1, incorporated herein by reference). Human U-2 OS cells (ATCC HTB-96) are stably transfected with the pTet-Off plasmid (Clontech K1620-A) and with nucleic acid encoding a LexA-repressor fusion protein containing the LexA DNA binding domain and the coding region of either HP1 or HPC2 (Drosophila Polycomb group proteins that repress gene expression when tethered to DNA; the assay works with either fusion protein) under control of the Tet-Off transcriptional regulatory system (Gossen and Bujard, 1992). These cells are referred to below as the reporter cells for the anti-repressor activity assay. A reporter plasmid, which provides hygromycin resistance, contains a polylinker sequence positioned between four LexA operator sites and the SV40 promoter that controls the zeocin resistance gene. The sequence to be tested for anti-repressor activity can be cloned in said polylinker. Construction of a suitable

reporter plasmid, such as pSelect, is described in example 1 and Fig. 1 of WO 00/004704. The reporter plasmid is transfected into the reporter cells, and the cells are cultured under hygromycin selection (25 µg/ml; selection for presence of the reporter plasmid) and tetracycline repression (doxycycline, 10 ng/ml; prevents expression of the LexA-repressor fusion protein). After 1 week of growth under these conditions, the doxycycline concentration is reduced to 0.1 ng/ml (or lower) to induce the LexA-repressor gene, and after 2 days zeocin is added to 250 µg/ml. The cells are cultured for 5 weeks, until the control cultures (transfected with empty reporter plasmid, i.e. lacking a cloned anti-repressor sequence in the polylinker) are killed by the zeocin (in this control plasmid, the SV40 promoter is repressed by the LexA-repressor fusion protein that is tethered to the LexA operating sites, resulting in insufficient zeocin expression in such cells to survive zeocin selection). A sequence has anti-repressor activity according to the present invention if, when said sequence is cloned in the polylinker of the reporter plasmid, the reporter cells survive the 5 weeks selection under zeocin. Cells from such colonies can still be propagated onto new medium containing zeocin after the 5 weeks zeocin selection, whereas cells transfected with reporter plasmids lacking anti-repressor sequences cannot be propagated onto new medium containing zeocin. Any sequence not capable of conferring such growth after 5 weeks on zeocin in this assay, does not qualify as a sequence having anti-repressor activity, or functional fragment or functional derivative thereof according to the present invention. As an example, known boundary sequences such as those tested by Van der Vlag et al (2000), including Drosophila scs (Kellum and Schedl, 1991), 5'-HS4 of the chicken β-globin

locus (Chung et al, 1993, 1997) or Matrix Attachment Regions (MARs) (Phi-Van et al., 1990), do not survive this assay.

In addition, it is preferred that the anti-repressor sequence or functional fragment or derivative thereof confers a higher proportion of reporter over-expressing clones when flanking a reporter gene (e.g. luciferase, GFP) which is integrated into the genome of U-2 OS or CHO cells, compared to when said reporter gene is not flanked by anti-repressor sequences, or flanked by weaker repression blocking sequences such as Drosophila scs. This can be verified using for instance the pSDH vector, or similar vectors, as described in example 1 and Fig. 2 of WO 03/004704.

Anti-repressor elements of the invention can have at least one of three consequences for production of protein: (1) they increase the predictability of identifying host cell lines that express a protein at industrially acceptable levels; (2) they result in host cell lines with increased protein yields; and/or (3) they result in host cell lines that exhibit more stable protein production during prolonged cultivation. Each of these attributes is discussed in more detail below:
(1) Increased predictability: Integration of transgene expression cassettes can occur at random positions throughout the host cell genome. However, much of the genome is transcriptionally silent heterochromatin. When the expression cassettes include anti-repressor elements flanking the transgene, the position of integration has a reduced effect on expression. The anti-repressor elements impair the ability of adjacent heterochromatin to silence the transgene. Consequently, the proportion of transgene-containing host cells with acceptable expression levels is increased. Indeed, incorporation of anti-repressor

sequences results in the establishment of up to 10 times more colonies, compared to the same transgene lacking anti-repressor sequences, when the same amount of DNA is transfected.

(2) Yield: The levels of protein expression in primary populations of recombinant host cells, directly after transgene integration, have been surveyed. The expression level of individuals in the populations varies. However, when the transgenes are protected by anti-repressor elements, the variability is reduced. This reduced variability is most conspicuous in that fewer clones are recovered that have low levels of expression. Furthermore, the populations with anti-repressor elements commonly have individuals with strikingly high expression. These high-yielding individuals are favourable for production of proteins, either for harvesting purposes or for purposes of changing the phenotype of the cell or an organism comprising such cells.

(3) Increased stability: anti-repressor elements increase the stability of transgenes in recombinant host cell lines by ensuring that the transgenes are not transcriptionally silenced during prolonged cultivation. Comparative trials show that, under conditions in which transgenes that are not protected by anti-repressor elements are progressively silenced (5 - 25 passages in cultivation), anti-repressor element-protected transgenes continue to be expressed at high levels. This is an advantage during industrial production of proteinaceous molecules, during which cell cultivation continues for prolonged periods, from a few weeks to many months. Similarly, stability of expression over prolonged periods may be advantageous in plants or animals with altered

phenotypes as a consequence of recombinant expression of anti-repressor protected transgenes.

STAR67

The present invention provides a novel sequence having anti-repressor acitivity, which was coined STAR67 (SEQ. ID. NO. 66). This anti-repressor sequence already strongly increases expression when it is placed only upstream of a promoter driving expression of a gene of interest, as is evident from the examples provided herein, whereas hitherto known potent anti-repressor sequences such as STAR6 or STAR7 appear to provide much less advantage in this configuration (they function especially well when flanking a transgene, i.e. when present both upstream and downstream of the transgene). Hence, STAR67 provides an alternative to already known anti-repressor sequences, and when placed upstream of a promoter appears to have increased benefit when compared to such known anti-repressor sequences. STAR67 is not an enhancer-blocker, in contrast to other anti-repressor sequences tested for this property (Kwaks et al, 2003), providing another difference between STAR67 and other anti-repressor sequences disclosed before. Also, STAR67 may operate in a bi-directional manner, as shown in example 7. According to the invention the presence of STAR67 sequence in an expression cassette provides improved predictability and/or yield and/or stability of expression. It is demonstrated herein that STAR67 is functional in combination with various promoters, and in different cell lines.

Expression cassette

A nucleic acid molecule according to the invention, comprising STAR67, may be in any format, e.g. as DNA

fragment, optionally present on a cloning vector such as a plasmid, preferably an expression vector, and can be used for instance for cloning purposes using standard recombinant DNA technology. In preferred embodiments of the invention, nucleic acid comprises an expression cassette, which is useful to express sequences of interest, for instance in host cells.

An 'expression cassette' as used herein is a nucleic acid sequence comprising at least a promoter functionally linked to a sequence of which expression is desired, which sequence of which expression is desired preferably is an open reading frame encoding all or part of a protein of interest. The promoter is preferably a heterologous promoter with respect to said nucleic acid of interest, whereby a heterologous promoter is defined as a promoter which is not the natural promoter of said sequence of interest. In other words, some form of human intervention, e.g. molecular cloning, has been used at any point in time to make the functional combination of a heterologous promoter with a nucleic acid of interest, and it is readily understood in this context that a heterologous promoter can be derived from the same or from a different organism as the sequence of interest. Preferably, an expression cassette further contains transcription termination and polyadenylation sequences. Other regulatory sequences such as enhancers may also be included. The expression units according to the invention further comprise at least one anti-repressor sequence, such as STAR67. In certain preferred embodiments said anti-repressor sequence, preferably STAR67, is placed upstream of said promoter, preferably such that less than 2kb are present between the 3' end of the anti-repressor sequence and the start of the promoter sequence. In preferred embodiments, less than 1 kb, more preferably

less than 500 nucleotides (nt), still more preferably less than about 200, 100, 50, or 30 nt are present between the 3' end of the anti-repressor sequence and the start of the promoter sequence. In certain preferred embodiments, the anti-repressor sequence is cloned directly upstream of the promoter, resulting in only about 0-20 nt between the 3' end of the anti-repressor sequence and the start of the promoter sequence.

To obtain expression of nucleic acid sequences encoding recombinant protein, it is well known to those skilled in the art that sequences capable of driving such expression, can be functionally linked to the nucleic acid sequences encoding the protein, resulting in recombinant nucleic acid molecules encoding a recombinant protein in expressible format. In general, the promoter sequence is placed upstream of the sequences encoding the protein of interest. Useful expression vectors are available in the art, e.g. the pcDNA and pEF vector series of Invitrogen, pMSCV and pTK-Hyg from BD Sciences, pCMV-Script from Stratagene, etc.

Where the sequence encoding the polypeptide of interest is properly inserted with reference to sequences governing the transcription and translation of the encoded polypeptide, the resulting expression cassette is useful to produce the protein of interest, referred to as expression. Sequences driving expression may include promoters, enhancers and the like, and combinations thereof. These should be capable of functioning in the host cell, thereby driving expression of the nucleic acid sequences that are functionally linked to them. The person skilled in the art is aware that various promoters can be used to obtain expression of a gene of interest in host cells. Promoters can be constitutive or regulated, and can be obtained from various sources, including

viruses, prokaryotic, or eukaryotic sources, or artificially designed. Expression of nucleic acids of interest may be from the natural promoter or derivative thereof or from an entirely heterologous promoter (Kaufman, 2000). Some well-known and much used promoters for expression in eukaryotic cells comprise promoters derived from viruses, such as adenovirus, e.g. the E1A promoter, promoters derived from cytomegalovirus (CMV), such as the CMV immediate early (IE) promoter (referred to herein as the CMV promoter) (obtainable for instance from pcDNA, Invitrogen), promoters derived from Simian Virus 40 (SV40) (Das et al, 1985), and the like. Suitable promoters can also be derived from eukaryotic cells, such as methallothionein (MT) promoters, elongation factor 1α (EF-1α) promoter (Gill et al., 2001), ubiquitin C or UB6 promoter (Gill et al., 2001; Schorpp et al, 1996), actin promoter, an immunoglobulin promoter, heat shock promoters, and the like. Some preferred promoters for obtaining expression in eukaryotic cells, which are suitable promoters in the present invention, are the CMV-promoter, a mammalian EF1-alpha promoter, a mammalian ubiquitin promoter such as a ubiquitin C promoter, or a SV40 promoter (e.g. obtainable from pIRES, cat.no. 631605, BD Sciences). Testing for promoter function and strength of a promoter is a matter of routine for a person skilled in the art, and in general may for instance encompass cloning a test gene such as lacZ, luciferase, GFP, etc. behind the promoter sequence, and test for expression of the test gene. Of course, promoters may be altered by deletion, addition, mutation of sequences therein, and tested for functionality, to find new, attenuated, or improved promoter sequences.

An expression cassette according to the invention may be monocistronic, bicistronic or multicistronic. The

term "bicistronic gene," is defined as a gene capable of providing a RNA molecule that encodes two proteins/polypeptides. The term "monocistronic gene" is defined as a gene capable of providing a RNA molecule that encodes one protein/polypeptide. The term "multicistronic" is defined as a gene capable of providing a RNA molecule that encodes two or more proteins/polypeptides, and a bicistronic gene is therefore encompassed within the definition of a multicistronic gene. A "gene" as used in the present invention can comprise chromosomal DNA, cDNA, artificial DNA, combinations thereof, and the like, and could also be in the form of other nucleic acid, e.g. RNA. In certain embodiments, a protein expression unit comprises a multicistronic gene. Units comprising several cistrons can be transcribed as a single mRNA. Translation of the second and further coding regions present on that RNA can be achieved in various ways, including the use of translation reinitiation sites or internal ribosome entry sites, the latter of which is preferred. One advantage of bi- or multi-cistronic units can be an easy selection of clones expressing a protein of interest, by placing the nucleic acid encoding a selectable marker protein downstream of nucleic acid encoding a protein or polypeptide of interest.

For the production of multimeric proteins, two or more expression cassettes can be used. This embodiment has proven to give good results, e.g. for the expression of the heavy and light chain of antibodies. According to the invention, at least one of the expression cassettes, but preferably each of them, should comprise a STAR sequence. In another embodiment, the different subunits or parts of a multimeric protein are present on a single expression cassette.

Instead of or in addition to the presence of an anti-repressor sequence placed upstream of a promoter in an expression cassette, it has proven highly beneficial to provide an anti-repressor sequence on both sides of an expression cassette, such that expression cassette comprising the transgene is flanked by two anti-repressor sequences, which in certain embodiments are essentially identical to each other. Of course, this can also be done with STAR67, to obtain an expression cassette flanked by two STAR67 sequences. Alternative positions of a single STAR67 sequence in an expression cassette, e.g. behind the transgene, preferably behind transcriptional termination and polyadenylation signal, with the 3′ end of the STAR67 sequence facing the transgene, are also possible.

An expression cassette according to the invention may optionally comprise a selection marker gene. The term "selection marker or selectable marker" is typically used to refer to a gene and/or protein whose presence can be detected directly or indirectly in a cell, for example a gene and/or a protein that inactivates a selection agent and protects the host cell from the agent's lethal or growth-inhibitory effects (e.g. an antibiotic resistance gene and/or protein). Another possibility is that said selection marker induces fluorescence or a color deposit (e.g. green fluorescent protein and derivatives, luciferase, lacZ, alkaline phosphatase, etc.). In certain embodiments, a selection marker used for the invention is zeocin, and for selecting a second expression cassette puromycin is used. The person skilled in the art will know that other selection markers are available and can be used, e.g. neomycin, blasticidin, puromycin, bleomycin, hygromycin, dhfr, etc.

The term "selection" is typically defined as the process of using a selection marker/selectable marker and a selection agent to identify host cells with specific genetic properties (e.g. that the host cell contains a transgene integrated into its genome). It is clear to a person skilled in the art that numerous combinations of selection markers are possible. An example of a possible antibiotic is provided above. The one antibiotic that is particularly advantageous is zeocin, because the zeocin-resistance protein (zeocin-R) acts by binding the drug and rendering it harmless. Therefore it is easy to titrate the amount of drug that kills cells with low levels of zeocin-R expression, while allowing the high-expressors to survive. All other antibiotic-resistance proteins in common use are enzymes, and thus act catalytically (not 1:1 with the drug). When a two-step selection is performed it is therefore advantageous to use an antibiotic resistance protein with this 1:1 binding mode of action. Hence, the antibiotic zeocin is a preferred selection marker. For convenience the zeocin antibiotic can in a two-step selection method combined with for instance puromycin, blasticidin or hygromycin, which may for instance be present in a monocistronic gene.

It is also possible to combine an antibiotic selection marker with a selection marker which provides induction of fluorescence or which provide a colour deposit. Different promoters can be used as long as they are functional in the used cell.

In certain embodiments an expression cassette is provided with a (weak) Internal Ribosome Binding Site (IRES) as an example of a protein translation initiation site with a reduced translation efficiency, e.g. between the open reading frame of the protein of interest and the

selection marker open reading frame. Translation of proteins from IRES elements is less efficient than cap-dependent translation: the amount of protein from IRES-dependent open reading frames (ORFs) ranges from less than 20% to 50% of the amount from the first ORF (Mizuguchi et al., 2000). Furthermore, mutation of IRES elements can attenuate their activity, and lower the expression from the IRES-dependent ORFs to below 10% of the first ORF (Lopez de Quinto & Martinez-Salas, 1998, Rees et al., 1996). When the IRES-dependent ORF encodes a selectable marker protein, its low relative level of translation means that high absolute levels of transcription must occur in order for the recombinant host cell to be selected. Therefore, selected recombinant host cell isolates will by necessity express high amounts of the transgene mRNA. Since the recombinant protein is translated from the cap-dependent ORF, it can be produced in abundance resulting in high product yields.

Conventional expression systems are DNA molecules in the form of a recombinant plasmid or a recombinant viral genome. The plasmid or the viral genome is introduced into (eukaryotic host) cells and preferably integrated into their genomes by methods known in the art. In preferred embodiments, the present invention also uses these types of DNA molecules to deliver its improved transgene expression system. A preferred embodiment of the invention is the use of plasmid DNA for delivery of the expression system. A plasmid contains a number of components: conventional components, known in the art, are an origin of replication and a selectable marker for propagation of the plasmid in bacterial cells; a selectable marker that functions in eukaryotic cells to identify and isolate host cells that carry an integrated transgene expression system; nucleic acid encoding the

protein of interest, whose high-level transcription is brought about by a promoter that is functional in eukaryotic cells (e.g. the human cytomegalovirus major immediate early promoter/enhancer, pCMV (Boshart et al., 1985); and transcriptional terminators (e.g. the SV40 polyadenylation site (Kaufman & Sharp, 1982) for the transgene of interest and the selectable marker. The vector used can be any vector that is suitable for cloning DNA and that can be used for transcription of a nucleic acid of interest. When host cells are used it is preferred that the vector is an integrating vector. Alternatively, the vector may be an episomally replicating vector.

Some non-limiting, schematic representations of possible configurations of expression cassettes are provided in Fig. 1. This is the configuration of the DNA elements of the expression cassettes in the plasmid as well as after integration into the genome. Construct A contains an expression unit that encompasses an open reading frame encoding a protein (Gene). This is upstream of the attenuated EMCV IRES (Martinez-Salas et al 1999; Mizuguchi et al 2000; Rees et al 1996), and of the open reading frame encoding the zeocin resistance selectable marker protein (zeo). The gene cassette has the SV40 transcriptional terminator at their 3' ends (**t**). This bicistronic transgene is transcribed at high levels from the CMV promoter. Construct B is derived from construct A, but STAR67 is now cloned upstream of the CMV promoter. Construct C is derived from construct B, but now two anti-repressor elements (in this case STAR7) are cloned to flank the entire cassette.

Combination of anti-repressor sequences in expression cassette

It is shown herein that the combination of a first anti-repressor element upstream of a promoter and flanking the expression cassette by two other anti-repressor sequences provides superior results. In particular, when an SV40 promoter is used in CHO cells, expression is already very high in the absence of anti-repressor sequences, but is considerably enhanced when STAR67 is placed upstream of said promoter, and the whole expression cassette is flanked by two anti-repressor elements, such as STAR6, STAR7, or STAR 40.

It is therefore an object of the present invention to provide a recombinant nucleic acid molecule comprising an expression cassette comprising (from 5' to 3'): anti-repressor sequence A - promoter - nucleic acid encoding a protein of interest - anti-repressor sequence B, characterized in that said expression cassette further comprises a anti-repressor sequence C between said anti-repressor sequences A and B. In a preferred embodiment, said anti-repressor sequence C is present upstream of said promoter, in a configuration as described above under the heading 'expression cassette'. The expression cassette between the anti-repressor sequences A and B may further comprise the elements as described above for expression cassettes, e.g. transcription terminator sequence, polyadenylation signal, selection marker gene, enhancer, etc., and may be monocistronic, bicistronic or multicistronic as described above.

Two or all three of anti-repressor sequences A, B and C may be the same, or all three may be different. Anti-repressor sequences A and B may the same or different as anti-repressor sequence C. In certain embodiments, anti-repressor sequence A and B are (essentially) identical to

each other. In one embodiment, anti-repressor sequence C is STAR67, or a functional fragment or derivative thereof. Anti-repressor sequences A and B can be any anti-repressor sequence, and in certain embodiments comprise one of SEQ. ID. NOs. 1-65, or functional fragments or derivatives thereof. In certain embodiments, the expression cassette contains a multicistronic gene, and in preferred embodiments thereof said multicistronic gene comprises a sequence encoding a protein of interest and a selection marker gene. Alternatively, a selection marker gene is present under control of a separate promoter.

In certain embodiments, a fourth anti-repressor sequence D may be present between STAR sequences A and B. In such an embodiment, said anti-repressor sequence D is preferably positioned downstream of the nucleic acid encoding the protein of interest. Again, this anti-repressor sequence D may be the same or different from the other anti-repressor sequences in the recombinant nucleic acid molecule, it can be any anti-repressor sequence, and in certain embodiments is chosen from any one of SEQ. ID. NOs. 1-66, or functional fragments or derivatives thereof.

As at least some anti-repressor sequences can be directional (WO 00/004704), the anti-repressor sequences flanking the expression cassette (anti-repressor sequences A and B) may beneficially placed in opposite direction with respect to each other, such that the 3′ end of each of these anti-repressor sequences is facing inwards to the expression cassette (and to each other). Hence, in preferred embodiments, the 5′ side of an anti-repressor element faces the DNA/chromatin of which the influence on the transgene is to be diminished by said anti-repressor element. For an anti-repressor sequence

upstream of a promoter in an expression cassette, the 3' end faces the promoter. The sequences of the anti-repressor elements (SEQ. ID. NOs. 1-66) in the sequence listing are given in 5' to 3' direction, unless otherwise indicated.

Cells according to the invention

Nucleic acid molecules comprising STAR sequences and/or expression cassettes according to the present invention can be used for improving expression of nucleic acid, preferably in host cells. The terms "cell"/"host cell" and "cell line"/"host cell line" are respectively typically defined as a cell and homogeneous populations thereof that can be maintained in cell culture by methods known in the art, and that have the ability to express heterologous or homologous proteins. A host cell according to the present invention preferably is a eukaryotic cell, more preferably a mammalian cell, such as a rodent cell or a human cell or fusion between different cells. In certain non-limiting embodiments, said host cell is a U-2 OS osteosarcoma, CHO (Chinese hamster ovary), HEK 293, HuNS-1 myeloma, WERI-Rb-1 retinoblastoma, BHK, Vero, non-secreting mouse myeloma Sp2/0-Ag 14, non-secreting mouse myeloma NS0, NCI-H295R adrenal gland carcinomal or a PER.C6® cell.
In certain embodiments of the invention, a host cell is a cell expressing at least E1A, and preferably also E1B, of an adenovirus. As non-limiting examples, such a cell can be derived from for instance human cells, for instance from a kidney (example: HEK 293 cells, see Graham et al, 1977), lung (e.g. A549, see e.g. WO 98/39411) or retina (example: HER cells marketed under the trade mark PER.C6®, see US patent 5,994,128), or from amniocytes (e.g. N52.E6, described in US patent 6,558,948), and similarly

from other cells. Methods for obtaining such cells are described for instance in US patents 5,994,128 and US 6,558,948. PER.C6® cells for the purpose of the present application means cells from an upstream or downstream passage or a descendent of an upstream or downstream passage of cells as deposited under ECACC no. 96022940. It has been previously shown that such cells are capable of expression of proteins at high levels (e.g. WO 00/63403, and Jones et al, 2003).

Such host cells expressing the desired protein according to the invention can be obtained by introduction of a nucleic acid molecule, preferably in the form of an expression cassette according to the invention, into the cells. In an alternative embodiment, the STAR67 sequence is targeted for integration into a chromosomal region to improve the expression of a gene of interest that is already integrated into the genome, e.g. a naturally occurring gene, optionally under control of a heterologous promoter that was targeted upstream of said gene to regulate expression of said gene.
Preferably the host cells are from a stable clone that can be selected and propagated according to standard procedures known to the person skilled in the art. A culture of such a clone is capable of producing recombinant protein of interest. Cells according to the invention preferably are able to grow in suspension culture in serum-free medium.

A protein of interest according to the invention can be any protein, and may be a monomeric protein or a multimeric protein. A multimeric protein comprises at least two polypeptide chains. Non-limiting examples of a protein of interest according to the invention are enzymes, hormones, immunoglobulin chains, therapeutic

proteins like anti-cancer proteins, blood coagulation proteins such as Factor VIII, multi-functional proteins, such as erythropoietin, diagnostic proteins, or proteins or fragments thereof useful for vaccination purposes, all known to the person skilled in the art.

In certain embodiments, an expression cassette of the invention encodes an immunoglobulin heavy or light chain or an antigen binding part, derivative and/or analogue thereof. In a preferred embodiment a protein expression unit according to the invention is provided, wherein said protein of interest is an immunoglobulin heavy chain. In yet another preferred embodiment a protein expression unit according to the invention is provided, wherein said protein of interest is an immunoglobulin light chain. When these two protein expression units are present within the same (host) cell a multimeric protein and more specifically an antibody is assembled. Hence, in certain embodiments, the protein of interest is an immunoglobulin, such as an antibody, which is a multimeric protein. Preferably, such an antibody is a human or humanized antibody. In certain embodiments thereof, it is an IgG, IgA, or IgM antibody. An immunoglobulin may be encoded by the heavy and light chains on different expression cassettes, or on a single expression cassette, wherein the gene encoding each individual chain may each be under control of a separate promoter in monocistronic transcription units, or alternatively both chains may be encoded on a multicistronic gene.

The protein of interest may be from any source, and in certain embodiments is a mammalian protein, an artificial protein (e.g. a fusion protein or mutated protein), and preferably is a human protein.

Obviously, the anti-repressor sequences and configurations of the expression cassettes of the present invention may also be used when the ultimate goal is not the production of a protein, but the RNA itself, for instance for producing increased quantities of RNA from an expression cassette, which may be used for purposes of regulating other genes (e.g. RNAi, antisense RNA), gene therapy, in vitro protein production, etc.

Method for producing protein of interest

In one aspect, the invention provides a method for expressing a sequence of interest, preferably encoding a protein of interest, by providing a host cell with a nucleic acid molecule or expression cassette according to the invention, culturing the cell and expressing the sequence of interest.

The term "expression" is typically used to refer to the production of a specific RNA product or products, or a specific protein or proteins, in a cell. In the case of RNA products, it refers to the process of transcription. In the case of protein products, it refers to the processes of transcription, translation and optionally post-translational modifications (e.g. glycosylation, disulfide bond formation, etc.). In the case of secreted proteins, it refers to the processes of transcription, translation, and optionally post-translational modification, followed by secretion.

Introduction of the nucleic acid that is to be expressed in a cell, can be done by one of several methods, which as such are known to the person skilled in the art, also dependent on the format of the nucleic acid to be introduced. Said methods include but are not limited to transfection, infection, injection, transformation, and the like.

Culturing a cell is done to enable it to metabolize, and/or grow and/or divide and/or produce recombinant proteins of interest. This can be accomplished by methods well known to persons skilled in the art, and includes but is not limited to providing nutrients for the cell. The methods comprise growth adhering to surfaces, growth in suspension, or combinations thereof. Culturing can be done for instance in dishes, roller bottles or in bioreactors, using batch, fed-batch, continuous systems such as perfusion systems, and the like. In order to achieve large scale (continuous) production of recombinant proteins through cell culture it is preferred in the art to have cells capable of growing in suspension, and it is preferred to have cells capable of being cultured in the absence of animal- or human-derived serum or animal- or human-derived serum components.

The conditions for growing or multiplying cells (see e.g. Tissue Culture, Academic Press, Kruse and Paterson, editors (1973)) and the conditions for expression of the recombinant product are known to the person skilled in the art. In general, principles, protocols, and practical techniques for maximizing the productivity of mammalian cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach (M. Butler, ed., IRL Press, 1991).

In a preferred embodiment, the expressed protein is collected (isolated), either from the cells or from the culture medium or from both. It may then be further purified using known methods, e.g. filtration, column chromatography, etc, by methods generally known to the person skilled in the art.

Novel selection method

The present invention discloses a novel method for generating host cells expressing two polypeptides of interest, which method gives a surprisingly good result, in that almost no or no colonies appear without the use of an anti-repressor sequence, while the presence of an anti-repressor sequence does lead to clones surviving the selection, and these clones in general express the two polypeptides of interest at high levels (example 7, Fig. 11). The invention therefore provides a method for generating a host cell expressing two polypeptides of interest, the method comprising: a) introducing into a host cell one or more nucleic acid molecules, the nucleic acid molecule or molecules together comprising: (i) a promoter functionally linked to a sequence encoding a first polypeptide of interest and a first selectable marker gene, and (ii) a promoter functionally linked to a sequence encoding a second polypeptide of interest and a second selectable marker, and (iii) at least one anti-repressor sequence, chosen from the group consisting of (a) any one of SEQ. ID. NO. 1-66, (b) fragments of any one of SEQ. ID. NOs. 1-66, said fragments having anti-repressor activity, (c) sequences that are at least 70% identical to (a) or (b) and having anti-repressor activity, and (d) the complement of any one of (a)-(c); b) selecting a host cell by essentially simultaneously selecting for expression of said first and second selectable marker genes. The selected host cells can be conveniently used for the expression of said two polypeptides, by culturing said selected host cells. In one embodiment, two nucleic acid molecules, one containing (i) and the other containing (ii), are used for introduction into the host cell. In this embodiment, each nucleic acid molecule preferably contains at least one said anti-repressor sequence. In another embodiment,

a single nucleic acid molecule containing both (i) and (ii) is used, which must then contain at least one said anti-repressor sequence. One advantage of having the coding sequences for both polypeptides on a single nucleic acid molecule is that only one nucleic acid preparation is required before the nucleic acid is introduced into the cells. Furthermore, in such an embodiment a single integration event is sufficient for the coding sequences of both polypeptides, thereby eliminating the possibility that the nucleic acid encoding the first polypeptide is integrated on a different location or with a different copy number than that encoding the second polypeptide. In preferred embodiments, said two polypeptides may form part of a multimeric protein, such as an immunoglobulin. The method is therefore particularly suited for the expression of antibodies. The two promoters may be the same or different, and could be any promoters as described supra. When the embodiment where a single nucleic acid molecule contains both (i) and (ii) is used, the direction of the two transcription units on the single nucleic acid molecule may for instance be both the same, or in opposite directions facing each other, or in opposite directions each pointing outward. In the latter case, the anti-repressor sequence may be placed between the two promoters (e.g. see Fig. 11B). One suitable anti-repressor sequence for this aspect of the invention is SEQ. ID. NO. 66 (see example 7), but the other anti-repressor sequences may function as well. This can be easily tested by the person skilled in the art, using the disclosure herein, and hence also the use of other anti-repressor sequences according to this aspect of the invention is encompassed by the scope of the invention. Furthermore, anti-repressor sequences may be added to

flank one or both transcription units. The marker genes must each be different, and may for instance be selected from the marker genes as described supra. In certain embodiments, one selectable marker is zeocin, and the other selectable marker is for instance puromycin. However, it will be clear that other combinations may be used, and are within the scope of this aspect of the invention. Expression of the markers should preferably be dependent from the expression of the polypeptides of interest. This can be established by linking the expression of the marker genes to that of the polypeptides of interest for instance by using multicistronic genes, e.g. with IRES sequences, as discussed supra. "Essentially simultaneous" selection means that at least part of the time, preferably at least one day, more preferably at least two days, still more preferably at least 5 days, both selection agents are present, which can be brought about by having both selection agents added to the culture medium at the same time or adding culture medium comprising both selection agents (e.g. example 7). It will be clear that adding the second selection agent only after a few hours or days to medium wherein the first selection agent is already present, will still result in essentially simultaneous selection within the meaning of the invention. This situation is distinguished from the situation where cells are first selected with the first selection agent, and once colonies are formed the cells are selected on new medium whereto the second selection agent and not the first selection agent has been added (consecutive selection, e.g. example 5 and WO 03/106684). By selecting for both selectable markers essentially simultaneously, it is disclosed herein surprisingly that a fast and strong selection is provided, whereby many low-producing

clones are eliminated, resulting in a strongly decreased workload for finding a clone with desirable expression characteristics. This is particularly advantageous for the biotechnological industry, where many, usually hundreds or even thousands, of colonies have to be screened before a desired clone having high expression levels is identified.

In addition, the system provides the possibility to test parts of anti-repressor elements, at least of STAR67 but also of other anti-repressor sequences when they function in this setting, for functionality. This easy screen, which provides an almost or even complete black and white difference in many cases, therefore will contribute to identifying functional parts or derivatives from anti-repressor sequences, and/or aid in characterizing novel anti-repressor sequences. The screen could even be used to find novel anti-repressor and/or expression-enhancing sequences (since this type of screen is not literally selecting for anti-repressor activity, but rather for expression-enhancing activity, it is also possible that sequences having the latter but not the former activity will be found). It is therefore another aspect of the present invention to provide a method for identifying an anti-repressor and/or expression-enhancing sequences, or a functional fragment or derivative thereof, comprising: a) introducing into a host cell one or more nucleic acid molecules, the nucleic acid molecule or molecules together comprising: (i) a promoter functionally linked to a first selectable marker gene, and (ii) a promoter functionally linked to a second selectable marker, and (iii) at least one sequence to be tested for anti-repressor and/or expression-enhancing activity; b) selecting a host cell by selecting essentially simultaneously for expression of said first

and second selectable marker genes. Host cells that survive this screen are likely to contain a functional anti-repressor and/or expression-enhancing element. This can then easily be analysed further, e.g. by determining its sequence and/or by subjecting it to other functional tests for anti-repressor activity as described supra. In one embodiment, both (i) and (ii) are on one nucleic acid molecule, e.g. with diverging promoters (the first and second selectable marker genes facing outward), between which the to be tested sequence can be cloned for analysis. Preferably, said selectable marker genes are not highly expressed, e.g. by using relatively weak promoters directly driving the selectable marker genes, which relatively weak promoters are known to and/or can be identified with routine assays by the person skilled in the art, or for instance by using a strong promoter for transcription and an IRES, preferably a relatively weak IRES as described supra, for translation of the marker gene. The marker genes may also conveniently be placed downstream of nucleic acid sequences encoding a polypeptide and linked to the expression thereof in a multicistronic gene, e.g. by an IRES (e.g. example 7, Fig. 11 for possible configuration). Such polypeptides could again be polypeptides of interest, so that directly useful clones for expression of said polypeptides of interest are obtained, but could for screening purposes also be markers of which expression can easily be quantified, e.g. GFP or a derivative thereof, SEAP, lacZ, luciferase, etc. In such a combination, the expression of the polypeptides could directly be quantified to pick out anti-repressor and/or expression-enhancing elements or fragments thereof that provide sufficient colonies during the simultaneous selection, and also provide high transcription levels. When 'random' nucleic acid (i.e.

nucleic acid not yet known to have anti-repressor or expression-enhancing activity), e.g. in the form of genome fragments, is used as the sequence to be tested for anti-repressor and/or expression-enhancing activity, novel anti-repressor and/or expression-enhancing sequences can be found. When known anti-repressor sequences are tested, this assay can be used to characterize them further. When fragments of known anti-repressor sequences are tested, the assay will provide functional fragments of such known anti-repressor sequences.

The practice of this invention will employ, unless otherwise indicated, conventional techniques of immunology, molecular biology, microbiology, cell biology, and recombinant DNA, which are within the skill of the art. See e.g. Sambrook, Fritsch and Maniatis, Molecular Cloning: A Laboratory Manual, 2nd edition, 1989; Current Protocols in Molecular Biology, Ausubel FM, et al, eds, 1987; the series Methods in Enzymology (Academic Press, Inc.); PCR2: A Practical Approach, MacPherson MJ, Hams BD, Taylor GR, eds, 1995; Antibodies: A Laboratory Manual, Harlow and Lane, eds, 1988.

The invention is further explained in the following examples. The examples do not limit the invention in any way. They merely serve to clarify the invention.

Examples

Example 1: Construction of STAR67 vectors

A novel anti-repressor sequence was isolated using a genetic screen as described in WO 03/004704, and this novel sequence was coined STAR67 (SEQ. ID. NO. 66). The effects of STAR67 on expression of transgenes in

mammalian cell lines were tested. Here we describe the construction of the various constructs.

Materials and Methods

Three plasmids were created (Fig. 1):

A) CMV-d2EGFP-ires-Zeo (CMV Control),

B) STAR67-CMV-d2EGFP-ires-Zeo (CMV-STAR67),

C) STAR7-STAR67-CMV-d2EGFP-ires-Zeo-STAR7 (CMV-STAR67 7/7)

The construction of construct A is described below. Plasmid pd2EGFP (Clontech 6010-1) was modified by insertion of a linker at the *Bsi*WI site to yield pd2EGFP-link. The linker, made by annealing oligonucleotides GTACGGATATCAGATCTTTAATTAAG (SEQ. ID. NO. 67) and GTACCTTAATTAAAGATCTGATAT (SEQ. ID. NO. 68), introduced sites for the *Pac*I, *Bgl*II, and *Eco*RV restriction endonucleases. This created the multiple cloning site MCSII for insertion of STAR elements. Then primers GATCAGATCTGGCGCGCCATTTAAATCGTCTCGCGCGTTTCGGTGATGACGG (SEQ. ID. NO. 69) and (AGGCGGATCCGAATGTATTTAGAAAAATAAACAAATAGGGG (SEQ. ID. NO. 70) were used to amplify a region of 0.37 kb from pd2EGFP, which was inserted into the *Bgl*II site of pIRES (Clontech 6028-1) to yield pIRES-stuf. This introduced sites for the *Asc*I and *Swa*I restriction endonucleases at MCSI, and acts as a "stuffer fragment" to avoid potential interference between STAR elements and adjacent promoters. pIRES-stuf was digested with *Bgl*II and *Fsp*I to liberate a DNA fragment composed of the stuffer fragment, the CMV promoter, the IRES element (flanked by multiple cloning sites MCS A and MCS B), and the SV40

polyadenylation signal. This fragment was ligated with the vector backbone of pd2EGFP-link produced by digestion with *Bam*HI and *Stu*I, to yield pIRES-link.

The open reading frame of the zeocin-resistance gene was inserted into *Bam*HI/*Not*I sites downstream of the pIRES as follows: the zeocin-resistance ORF was amplified by PCR with primers GATCGGATCCTTCGAAATGGCCAAGTTGACCAGTGC (SEQ. ID. NO. 71) and AGGCGCGGCCGCAATTCTCAGTCCTGCTCCTC (SEQ. ID. NO. 72) from plasmid pCMV/zeo (Invitrogen, cat.no. V50120), digested with *Bam*HI and *Not*I, and ligated with *Bam*HI/*Not*I-digested pIRES-link to yield pIRES-link-zeo. The d2EGFP reporter ORF was introduced into pIRES-link-zeo by amplification of pd2EGFP (Clontech 6010-1) with primers GATCGAATTCTCGCGAATGGTGAGCAAGCAGATCCTGAAG (SEQ. ID. NO. 73) and AGGCGAATTCACCGGTGTTTAAACTTACACCCACTCGTGCAGGCTGCCCAGG (SEQ. ID. NO. 74), and insertion of the *Eco*RI-digested d2EGFP cassette into the *Eco*RI site in the pIRES-link-zeo plasmid. This created construct A, CMV-d2EGFP-IRES-Zeo (CMV Control).

STAR67 was cloned upstream of the CMV promoter, in the AscI site (about 15 nt remaining between STAR67 and the promoter). This created construct B, STAR67-CMV-d2EGFP-ires-Zeo (CMV-STAR67).

STAR67 was also tested in the context of a cassette that contains also the STAR7 (SEQ. ID. NO. 7), cloned directionally in the 5′ SalI and XbaI sites and 3′ BglII and PacI sites to flank the entire cassette with STAR7. This is construct C (CMV-STAR67 7/7).

Example 2: STAR67 enhances the expression level from CMV,

EF1α and UB6 promoters in stably transfected CHO cells. We tested whether the presence of STAR67 adjacent to the CMV, EF1α and UB6 promoters influences the expression level of these promoters in CHO cells. The constructs A and B (Fig. 1) described in Example 1 are used for this purpose, modified for the respective promoters:

1 CMV-d2EGFP-ires-Zeo (CMV Control)

2 STAR67-CMV-d2EGFP-ires-Zeo (CMV-STAR67)

3 EF1α-d2EGFP-ires-Zeo (EF1α Control)

4 STAR67-EF1α-d2EGFP-ires-Zeo (EF1α-STAR67)

5 UB6-d2EGFP-ires-Zeo (UB6 Control)

6 STAR67-UB6-d2EGFP-ires-Zeo (UB6-STAR67).

Materials and Methods

The UB6 and EF1α promoters were exchanged for the CMV promoter in the plasmids described in Fig. 1. The UB6 promoter was cloned as follows. A DNA 0.37 kb stuffer from the pd2EGFP plasmid was amplified by PCR, as described in example 1, using primers identified by SEQ. ID. NOs. 69 and 70. The resulting DNA stuffer was cloned in the BglII site of pUB6/V5-His [Invitrogen V250-20], creating pUB6-stuf. From pUB6-stuf an AscI-SacI fragment was cloned into CMV-d2EGFP-IRES-Zeo, from which the CMV promoter was removed.

The EF1α promoter was amplified by PCR with pEF1α/V5-His [Invitrogen V920-20] as template, using primers GATCGGCGCGCCATTTAAATCCGAAAAGTGCCACCTGACG (SEQ. ID. NO. 79) and AGGCGGGACCCCCTCACGACACCTGAAATGGAAG (SEQ. ID. NO. 80). The PCR fragment was cloned in the AscI and PpuMI sites of

CMV-d2EGFP-IRES-Zeo, from which the CMV promoter was removed.

The Chinese Hamster Ovary cell line CHO-K1 (ATCC CCL-61) was cultured in HAMS-F12 medium + 10% Fetal Calf Serum containing 2 mM glutamine, 100 U/ml penicillin, and 100 micrograms/ml streptomycin at 37°C/5% CO_2. Cells were transfected with the plasmids using SuperFect (QIAGEN) as described by the manufacturer. Briefly, cells were seeded to culture vessels and grown overnight to 70-90% confluence. SuperFect reagent was combined with plasmid DNA at a ratio of 6 microliters per microgram (e.g. for a 10 cm Petri dish, 20 micrograms DNA and 120 microliters SuperFect) and added to the cells. After overnight incubation the transfection mixture was replaced with fresh medium, and the transfected cells were incubated further. After overnight cultivation, cells were trypsinized and seeded into fresh culture vessels with fresh medium. After another overnight incubation zeocin was added to a concentration of 50 µg/ml and the cells were cultured further. After another three days the medium was replaced by fresh medium containing zeocin (100 µg/ml) and cultured further. When individual colonies became visible (approximately ten days after transfection) medium was removed and replaced with fresh medium without zeocin. Individual clones were isolated and transferred to 24-well plates in medium without zeocin. One day after isolation of the colonies zeocin was added to the medium. Expression of the d2EGFP reporter gene was assessed approximately 3 weeks after transfection. d2EGFP expression levels in the colonies were measured after periods of two weeks. After the initial two weeks after transfection when the first d2EGFP measurements were performed, the colonies were

cultured in medium without zeocin or other antibiotics. This continued for the remainder of the experiment.

Results

Fig. 2 shows that transfection of the construct that contains STAR67 cloned upstream of the CMV promoter resulted in a number of CHO colonies that express significantly higher levels of d2EGFP protein, as compared to the "empty" control without STAR67, CMV Control. The average of the d2EGFP signal in the 10 colonies transfected with the CMV Control plasmid was 34, when measured 25 days after transfection. 60 days after transfection the average of the d2EGFP signal in these 10 colonies was reduced to 13, indicating that expression is not stable over time. In comparison, the average of the d2EGFP signal in the 10 colonies transfected with the STAR67-CMV plasmid was 42 when measured after 25 days and 32 measured 60 days after transfection. Hence 60 days after transfection, a STAR67-encompassing CMV construct conveyed a factor 2.5 higher CMV promoter driven expression level of the reporter protein in stably transfected clones. Importantly, 25 days after transfection, after the first measurement, selection pressure was removed by culturing the colonies in medium without zeocin. Hence, colonies containing the STAR67 construct are more stable over time in the absence of selection pressure than colonies that do not contain a STAR67 construct.

Fig. 3 shows that transfection of the construct that contains STAR67 cloned upstream of the EF1α- promoter resulted in a number of CHO colonies that express significantly higher levels of d2EGFP protein, as compared to the "empty" control without STAR67, EF1α

Control. The average of the d2EGFP signal in the 10 colonies transfected with the EF1α Control plasmid was 31, when measured 25 days after transfection. 60 days after transfection the average of the d2EGFP signal in these 10 colonies was 26. In comparison, the average of the d2EGFP signal in the 10 colonies transfected with the EF1α-STAR67 plasmid was 60 when measured after 25 days and 76 measured 60 days after transfection. Hence, both after 25 and 60 days after transfection, a STAR67-encompassing EF1α construct conveyed a factor 2.9 higher EF1α promoter driven expression level of the reporter protein in stably transfected clones.

Fig. 4 shows that transfection of the construct that contains STAR67 cloned upstream of the UB6 promoter resulted in a number of CHO colonies that express significantly higher levels of d2EGFP protein, as compared to the "empty" control without STAR67, UB6 Control. The average of the d2EGFP signal in the 10 colonies transfected with the UB6 Control plasmid was 51, when measured 25 days after transfection. 60 days after transfection the average of the d2EGFP signal in these 10 colonies was 29, indicating that the expression was not stable over time. In comparison, the average of the d2EGFP signal in the 10 colonies transfected with the UB6-STAR67 plasmid was 218 when measured after 25 days and 224 measured 60 days after transfection. Hence, 25 days after transfection, a STAR67-encompassing UB6 construct conveyed a factor 4.3 higher UB6 promoter driven expression level of the reporter protein in stably transfected clones. After 60 days this factor was 7.7, due to instability of expression in the control colonies and stability in the UB6-STAR67 colonies. Importantly, 25 days after transfection, after the first measurement,

selection pressure was removed by culturing the colonies in medium without zeocin. Hence, colonies containing the STAR67 construct are more stable over time in the absence of selection pressure than colonies that do not contain a STAR67 construct.

In conclusion, STAR67 increases expression from three different, unrelated promoters.

Example 3: STAR67 enhances the expression level from CMV, EF1α and UB6 promoters in stably transfected PER.C6 cells.

We tested whether the presence of STAR67 adjacent of the CMV, EF1α and UB6 promoters influences the expression level of these promoters in another cell type than CHO cells, namely human PER.C6 cells. The same constructs as in example 1 were used.

Materials and Methods

Transfection, culturing and analysis of PER.C6 cells

PER.C6® cells were cultured in DMEM medium + pyridoxine + 9% Foetal Bovine Serum (Non-Heat Inactivated), 8.9 mM $MgCl_2$ 100 U/ml penicillin, and 100 micrograms/ml streptomycin at 37°C/10% CO_2. Cells were transfected with the plasmids using Lipofectamine 2000 (Invitrogen) as described by the manufacturer. Briefly, cells were seeded to 6-wells and grown overnight to 70-90% confluence. Lipofectamine reagent was combined with plasmid DNA at a ratio of 15 microliters per 3 microgram (e.g. for a 10 cm Petri dish, 20 micrograms DNA and 120 microliters Lipofectamine) and added after 30 minutes incubation at 25°C to the cells. After 6-hour incubation the transfection mixture was replaced with fresh medium, and the transfected cells were incubated further. After overnight cultivation, cells were trypsinized and seeded

(1:15, 1:30, 1:60, 1:120 dilutions) into fresh petri dishes (90 mm) with fresh medium with zeocin added to a concentration of 100 µg/ml and the cells were cultured further. When colonies became visible, individual clones were isolated by scraping and transferred to 24-well plates in medium with zeocin. When grown to ~70% confluence, cells were transferred to 6-well plates. Stable colonies were expanded for 2 weeks in 6-well plates before the d2EGFP signal was determined on a XL-MCL Beckman Coulter flowcytometer. The mean of the d2EGFP signal was taken as measure for the level of d2EGFP expression. Colonies were measured for a second time after 2 weeks. Thereafter colonies were further cultured in the absence of zeocin.

Results

Fig. 5 shows that transfection of the construct that contains STAR67 cloned upstream of the CMV promoter resulted in a number of PER.C6 colonies that express significantly higher levels of d2EGFP protein, as compared to the "empty" control without STAR67, CMV Control. The average of the d2EGFP signal in the 10 colonies transfected with the CMV Control plasmid was 37, when measured 30 days after transfection. 60 days after transfection the average of the d2EGFP signal in these 10 colonies was reduced to 14, indicating that expression was not stable over time. In comparison, the average of the d2EGFP signal in the 10 colonies transfected with the STAR67-CMV plasmid was 101 when measured after 30 days and 45 measured 60 days after transfection. Hence 60 days after transfection, a STAR67-encompassing CMV construct conveyed a factor 3.2 higher CMV promoter driven expression level of the reporter protein in stably transfected clones.

Fig. 6 shows that transfection of the construct that contains STAR67 cloned upstream of the EF1α- promoter resulted in a number of PER.C6 colonies that express significantly higher levels of d2EGFP protein, as compared to the "empty" control without STAR67, EF1α Control. The average of the d2EGFP signal in the 10 colonies transfected with the EF1α Control plasmid was 5, when measured 30 days after transfection. 60 days after transfection the average of the d2EGFP signal in these 10 colonies was 6. In comparison, the average of the d2EGFP signal in the 10 colonies transfected with the EF1α-STAR67 plasmid was 25 when measured after 30 days and 20 measured 60 days after transfection. Hence, both after 30 and 60 days after transfection, a STAR67-encompassing EF1α construct conveyed a factor 4 higher EF1α promoter driven expression level of the reporter protein in stably transfected clones.

Fig. 7 shows that transfection of the construct that contains STAR67 cloned upstream of the UB6 promoter resulted in a number of PER.C6 colonies that express significantly higher levels of d2EGFP protein, as compared to the "empty" control without STAR67, UB6 Control. The average of the d2EGFP signal in the 10 colonies transfected with the UB6 Control plasmid was 4, when measured 30 days after transfection. 60 days after transfection the average of the d2EGFP signal in these 10 colonies was 2. In comparison, the average of the d2EGFP signal in the 10 colonies transfected with the UB6-STAR67 plasmid was 27 when measured after 30 days and 18 measured 60 days after transfection. Hence, both after 30 and 60 days after transfection, a STAR67-encompassing UB6

construct conveyed a factor 7 to 9 fold higher UB6 promoter driven expression level of the reporter protein in stably transfected clones.

Hence placing STAR67 upstream of the promoter resulted in significantly higher protein expression levels in comparison with STAR67-less constructs, also in PER.C6 cells. Hence STAR67 functions in different, unrelated cell types.

Example 4: Novel configuration of STAR67 combined with other anti-repressor elements to enhance the SV40 promoter in CHO cells.

We tested whether the presence of STAR67 adjacent of the SV40 promoter influences the expression level of this promoter, either alone or in combination with another anti-repressor element, in this example STAR7. The constructs that were used for this purpose (see Fig. 8), are:

1 SV40-d2EGFP-ires-Zeo (SV40 Control)
2 STAR67-SV40-d2EGFP-ires-Zeo (SV40-STAR67)
3 STAR7-SV40-d2EGFP-ires-Zeo-STAR7 (SV40-STAR7/7)
4 STAR7- STAR67-SV40--d2EGFP-ires-Zeo-STAR7 (SV40-STAR67 7/7)

Materials and Methods

The SV40 promoter was amplified by PCR with pIRES as template using primers TTGGTTGGGGCGCGCCGCAGCACCATGGCCTGAAATAACCTCTGAAAGAGG (SEQ. ID. NO. 81) and TTGGTTGGGAGCTCAAGCTTTTTGCAAAAGCCTAGGCCTCCAAAAAAGCCTCCTC (SEQ. ID. NO. 82). The PCR fragment was cloned in the AscI and SacI sites of CMV-d2EGFP-IRES-Zeo, from which the CMV promoter was removed.

CHO cells were transfected, colonies were isolated and propagated and analysed as in Example 2.

Results

Fig. 8 shows that transfection of the construct that either contains STAR67 cloned upstream of the SV40 promoter (SV40-STAR67) or STAR7 cloned to flank the entire construct (SV40-STAR 7/7) did not result in CHO colonies that express significantly higher levels of d2EGFP protein, as compared to the "empty" control without anti-repressor elements (SV40 Control). The average of the d2EGFP signal in the 18 colonies transfected with the SV40 Control plasmid was 86, when measured 40 days after transfection. In comparison, the average of the d2EGFP signal in the 18 colonies transfected with the SV40-STAR67 plasmid was 82 when measured after 40 days and the average of the d2EGFP signal in the 18 colonies transfected with the SV40-STAR 7/7 plasmid was 91 when measured after 40 days. Hence, no significant effect of these anti-repressor elements on the SV40 promoter in CHO cells was observed.

It appears that the expression levels from the SV40 promoter in these cells are already quite high, and even much higher than those observed with the CMV promoter, which was considered to be a very strong promoter. This high background expression in the absence of an anti-repressor element using the SV40 promoter in CHO cells, may explain why no significant effect of STAR67 alone, or STAR7 flanking the transgene, was observed.

However, the average of the d2EGFP signal in the 18 colonies transfected with the SV40-STAR67 7/7 plasmid was 209 when measured after 40 days colonies, which is a factor 2.4 higher than the average of the 18 control

colonies (86). Hence, when the STAR67 element is used in combination with another anti-repressor element, this results in a number of stably transfected CHO colonies that show significantly higher d2EGFP expression levels.

Therefore in this novel configuration (5'- STAR sequence A - STAR sequence C - promoter - nucleic acid encoding a protein of interest - STAR sequence B - 3', wherein in the present example STAR sequences A and B are STAR 7 and STAR sequence C is STAR67) STAR elements appear to function even better than in hitherto disclosed configurations.

We have done experiments in which the flanking STAR7 elements were replaced by flanking STAR6 (SEQ. ID. NO. 6) elements or by flanking STAR4 (SEQ. ID. NO. 4) elements, in combination with STAR67 upstream of the SV40 promoter (SV40-STAR67 6/6 and SV40-STAR67 4/4, respectively, using the same nomenclature as above), and observed improved expression also with these combinations. This proves that the flanking STAR7 elements can be exchanged for other STAR sequences, and still the improvement of the novel configuration of the expression cassette with the STAR sequences is observed.

Example 5: A combination of STAR67 and STAR7 enhance UB6-driven antibody expression levels in stably transfected CHO cells.

In example 4 we showed that the combination of STAR67 and STAR7 enhanced the expression levels of d2EGFP protein in CHO cells. Here we tested whether the combination of STAR67 and STAR7 could be used for the production of an antibody. We chose an antibody against the EpCAM molecule (Huls et al, 1999) as test protein and used the UB6 promoter.

Materials and Methods

Plasmids

The heavy chain cDNA (HC-EpCAM) was cloned in a construct encompassing the UB6 promoter. The HC-EpCAM was coupled to the Zeocin resistance gene by an IRES sequence. The light chain cDNA (LC-EpCAM) was also cloned in a construct encompassing the UB6 promoter. The LC-EpCAM was coupled to the puromycin resistance gene by an IRES sequence. Together these two plasmids represent the UB6 (HC+LC) Control (Fig. 9)

To test the effects of STAR67 and STAR7, STAR67 was cloned in both HC+LC constructs, upstream of the UB6 promoters. STAR7 was cloned to flank the entire cassettes, both at the 5' and 3' end (Fig. 9). These two plasmids represent STAR7-STAR67-UB6 (HC+LC) STAR7.

Transfection and culture of CHO cells

The Chinese Hamster Ovary cell line CHO-K1 (ATCC CCL-61) was transfected and cultured as in example 2, using zeocin (100 µg/ml) and puromycin (2.5 µg/ml) for selection. One day after transfection, zeocin was added to the culture medium. When first colonies became visible (approximately 7 days after addition of zeocin) the culture medium was removed and replaced with culture medium containing puromycin. After approximately 7 days, colonies were isolated and transferred to 24 wells plates, in culture medium containing zeocin only.

Results

Fig. 9 shows that that transfection of antibody constructs that contain STAR67 cloned upstream of the UB6 promoter and two STAR7 elements cloned to flank the entire cassettes resulted in a number of CHO colonies

that express significantly higher levels of EpCAM antibody (measured by ELISA using an anti-human IgG antibody) as compared to the "empty" control without STAR67 and STAR7, UB6 (HC+LC) Control. The average of the EpCAM production in the 18 colonies transfected with the UB6 (HC+LC) Control plasmid was 2.7 pg/cell/day, when measured 25 days after transfection. Selection agents zeocin and puromycin were removed after 25 days. 45 days after transfection the average of the EpCAM production in these 18 colonies was 2.7 pg/cell/day. In comparison, the average of the EpCAM production signal in the 18 colonies transfected with the STAR7-STAR67-UB6 (HC+LC)-STAR7 plasmid was 6.7 when measured after 25 days and 7.7 pg/cell/day, when measured 45 days after transfection. Hence, both after 30 and 45 days after transfection, a STAR67/STAR7-encompassing UB6 construct conveyed a factor 2.5 to 2.9 fold higher UB6 promoter driven EpCAM expression level in stably transfected CHO clones.

Hence placing STAR67 upstream of the promoter and two STAR7 elements to flank the cassettes resulted in significantly higher EpCAM antibody expression levels in CHO cells, in comparison with STAR67/STAR7-less constructs.

Example 6: STAR67 is not an enhancer blocker, whereas STAR6 and STAR7 are.

All hitherto known STAR elements that were tested for that property, including STAR6 and STAR7, are enhancer blockers (WO 03/004704, Kwaks et al, 2003). Enhancer blocker activity is tested by placing a STAR element between a strong enhancer and a promoter. Here we tested whether also STAR67 is an enhancer blocker.

Materials and Methods

The d2EGFP gene was PCR-amplified using primers TTGGTTGGTCATGAATGGTGAGCAAGGGCGAGGAGCTGTTC (SEQ.ID.NO. 75) and ATTCTCTAGACTACACATTGATCCTAGCAGAAGCAC (SEQ.ID.NO. 76) and cloned into plasmid pGL3-promoter (Promega) using the NcoI and XbaI restriction sites to replace the Luciferase gene to create plasmid pGL3-promoter-GFP. A linker (created by annealing oligo's CGATATCTTGGAGATCTACTAGTGGCGCGCCTTGGGCTAGCT (SEQ.ID.NO. 77) and GATCAGCTAGCCCAAGGCGCGCCACTAGTAGATCTCCAAGATATCGAGCT (SEQ.ID.NO. 78), was cloned in the SacI and BglII sites to create multiple cloning sites. The original BglII site was destroyed upon ligation of the linker, creating a new unique BglII site within the linker DNA. The SV40-enhancer was cut from plasmid pGL3-basic (Promega) using BsaBI and BamHI and cloned into the pGL3-linker-promoter-GFP using the EcoRV and BglII sites creating plasmid pGL3-enhancer-promoter-GFP. The STAR40 element (SEQ. ID. NO. 40) was placed upstream of the SV40 enhancer using KpnI and SacI sites to prevent action of the enhancer on upstream sequences. Finally, anti-repressor elements STAR6, STAR7 and STAR67 were placed in between the SV40 enhancer and the SV40 minimal promoter using the SpeI and AscI restriction sites.

Transfection and culture of CHO cells

The Chinese Hamster Ovary cell line CHO-K1 (ATCC CCL-61) is transfected as in example 2. One day after the transfection, d2EGFP levels are measured on an Epics XL flowcytometer (Beckman Coulter)

Results

Fig. 10 shows that STAR67 is not an enhancer blocker, whereas STAR6 and STAR7 act enhancer blockers in the same assay. STAR6, STAR7 and STAR67 were cloned between the SV40 enhancer and a minimal SV40-promoter upstream of the d2EGFP gene. When no STAR element was cloned between the enhancer and the promoter strong transcriptional activation occurred (arbitrarily set at 100%). When STAR6 or STAR7 was placed between the enhancer and the promoter, transcription dropped to background levels, indicating that STAR6 and STAR7 are potent enhancer blockers. In contrast, when STAR67 was cloned between the enhancer and the promoter relative transcription levels were still 80% of the control, indicating that STAR67 is not a good enhancer blocker, this in contrast with STAR6 and STAR 7, as well as other anti-repressor elements, as previously described (WO 03/004704, Kwaks et al, 2003).

Example 7: STAR67 enhances UB6 and CMV-driven antibody expression levels in stably transfected CHO cells.

In example 5 we showed that the combination of STAR67 and STAR7 enhanced the expression levels of EpCAM antibody in CHO cells, in the context of two distinct plasmids, which contained the heavy and light chains. In this example we tested whether STAR67 could be used for the production of EpCAM antibody when both heavy and light chains are placed on one plasmid. We used simultaneous selection for each selectable marker.

Materials and Methods

Plasmids

The heavy chain cDNA (HC-EpCAM) is under the control of the UB6 promoter and coupled to the Zeocin resistance gene by an IRES sequence. The light chain cDNA (LC-EpCAM)

is under control of the CMV promoter and coupled to the puromycin resistance gene by an IRES sequence. Basically these are the constructs used in example 5. These two expression cassettes were placed on one plasmid, in such a manner that transcription of the two expression units had opposite directions. In the control plasmid the UB6 and CMV promoters were separated by a stuffer of 500 bp (EpCAM Control) (Fig. 10). In another plasmid STAR67 was placed between the UB6 and CMV promoter (EpCAM STAR67) (Fig. 10).

Transfection and culture of CHO cells

The Chinese Hamster Ovary cell line CHO-K1 (ATCC CCL-61) was transfected and cultured as in example 2, using zeocin (100 µg/ml) and puromycin (2.5 µg/ml) for selection. In example 5, consecutive selection for both selection markers was used. In contrast, here both selection agents were present in the culture medium simultaneously. One day after transfection, zeocin and puromycin were added to the culture medium. The selection medium was present until colonies were isolated (approximately 14 days after transfection). After colonies were isolated and transferred to 24-wells plates, the cells were cultured in the presence of zeocin and puromycin.

Results

Fig. 11 shows that that transfection of the antibody construct that contains STAR67 cloned between the UB6 and CMV promoters resulted in a number of CHO colonies that express EpCAM antibody (measured by ELISA using an anti-human IgG antibody). The average of the EpCAM production in the 19 colonies transfected with the EpCAM STAR67

plasmid was 9.8 pg/cell/day, when measured 25 days after transfection.

In contrast, surprisingly no colonies survived of the transfection with the EpCAM Control plasmid. When selection was performed with either zeocin or puromycin alone, EpCAM Control colonies survived. However, when the selection pressure was increased by placing selection pressure on both the heavy and light chain, these conditions allowed only colonies to survive that have a STAR67 present in the transfected plasmid.

The results also show that incorporation of STAR67 has a beneficial effect on two promoters, the UB6 and CMV promoters, that are placed upstream and downstream of one STAR67 element. This indicates that STAR67 may operate in a bi-directional fashion.

The difference between the EpCAM control and EpCAM STAR67 plasmid is black-white in the sense that only transfection of EpCAM STAR67 results in the establishment of colonies, when the selection pressure is high. This opens an opportunity to use this plasmid configuration for identifying the region in STAR67 that is responsible for mediating this effect. Smaller, overlapping portions of STAR67 are placed between the UB6 and CMV promoters, driving the EpCAM molecule. When a portion of STAR67 is functional, colonies will survive when both zeocin and puromycin are simultaneously used as selection agent. When a portion of STAR67 is not functional, no colonies will survive under identical selection conditions.

Hence placing STAR67 upstream of the promoter resulted in significantly higher EpCAM antibody expression levels in CHO cells, in comparison with constructs lacking such anti-repressor elements.

A similar experiment is performed with the SV40 promoter to drive expression of the heavy and light chains of the anti-EpCAM antibody. In other experiments, STAR67 is exchanged for other STAR sequences. In further experiments, the orientation of the two expression units encoding the heavy and light chains is changed such that both are in the same direction. In another experiment, the expression units for the heavy and light chains are placed on different nucleic acid molecules respectively (as in example 5), and the resulting clones are simultaneously selected for both selectable markers. In another experiment, the host cell type is varied. Combinations of these variations are made.

DESCRIPTION OF FIGURES

Fig. 1. Schematic diagram of the invention.
A) bicistronic gene containing (from 5′ to 3′) a transgene (encoding for example the d2EGFP gene), an IRES, and a selectable marker (zeo, conferring zeocin resistance) under control of the CMV promoter. The expression unit has the SV40 transcriptional terminator at its 3′ end (**t**). The name of the construct is CMV-d2EGFP-ires-Zeo (CMV Control).
B) construct as in A, but now STAR 67 is cloned upstream of the CMV promoter. The name of the construct is STAR67-CMV-d2EGFP-ires-Zeo (CMV-STAR67).
C) construct as in B, but upstream and downstream STAR7 elements are cloned to flank the entire construct. The name of the construct is STAR7-STAR67-CMV-d2EGFP-ires-Zeo-STAR7 (CMV-STAR67 7/7)

Fig. 2. STAR67 improves CMV driven d2EGFP expression in

CHO cells.
The mean of the d2EGFP signal for 10 independent stable colonies is plotted for the indicated constructs in CHO cells. A) CMV Control; B) STAR67-CMV. X(10): average d2EGFP expression levels of the 10 colonies. See example 2 for details.

Fig. 3. STAR67 improves EF1α driven d2EGFP expression in CHO cells. Same as Fig. 2, but now with EF1α promoter. A) EF1α Control; B) STAR67- EF1α.

Fig. 4. STAR67 improves UB6 driven d2EGFP expression in CHO cells. Same as Figs. 2 and 3, but now with UB6 promoter. A) UB6 Control; B) STAR67-UB6.

Fig. 5. STAR67 improves CMV driven d2EGFP expression in PER.C6 cells. Same as Fig. 2, but now in PER.C6 cells. A) CMV Control; B) STAR67-CMV.

Fig. 6. STAR67 improves EF1α driven d2EGFP expression in PER.C6 cells. Same as Fig. 3, but now in PER.C6 cells. A) EF1α Control; B) STAR67- EF1α.

Fig. 7. STAR67 improves UB6 driven d2EGFP expression in PER.C6 cells. Same as Fig. 4, but now in PER.C6 cells. A) UB6 Control; B) STAR67-UB6.

Fig. 8. STAR67 improves SV40 driven d2EGFP expression in CHO-K1 cells in combination with another STAR element. Similar as Fig. 2, but now using the SV40 promoter in CHO cells, and the indicated constructs. The mean of the d2EGFP signal is plotted for 18-20 independent stable colonies 60 days after transfection. A) SV40 Control, SV40-STAR67, SV40-STAR7/7; B) SV40 Control and SV40-

STAR67 7/7. See example 4 for details.

Fig. 9. STAR67 improves UB6 driven expression levels of the EpCAM antibody in CHO-K1 cells. See example 5 for details. A) constructs without STAR elements; B) constructs with STAR67 upstream of promoter and flanking STAR7 elements. The anti-EpCAM antibody concentration is presented as pg/cell/day. X(18): average production level of the 18 colonies.

Fig. 10. STAR67 is not an enhancer blocker, whereas STAR 6 and 7 are. See example 6 for details.

Fig. 11. STAR67 enhances UB6 and CMV-driven antibody expression levels in stably transfected CHO cells. See example 7 for details. A) single DNA molecule containing anti-EpCAM heavy chain (HC) and light chain (LC), each behind a promoter, and each linked to a different selectable marker gene (simultaneous selection was used for both markers): construct without STAR elements. No colonies were found; B) same construct with STAR67 between the two promoters. The anti-EpCAM antibody concentration is presented as pg/cell/day. X(19): average production level of the 19 colonies.

References

Boshart, M, Weber, F, Jahn, G, Dorsch-Hasler, K, Fleckenstein, B, and Schaffner, W. (1985) A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus *Cell 41*, 521-530.

Chung JH, Whiteley M, and Felsenfeld G. (1993) A 5′ element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. *Cell* 74: 505-514.

Chung JH, Bell AC, Felsenfeld G. (1997). Characterization of the chicken beta-globin insulator. *Proc Natl Acad Sci USA* 94: 575-580.

Das, GC, Niyogi, SK, and Salzman, NP. (1985) SV40 promoters and their regulation *Prog Nucleic Acid Res Mol Biol 32*, 217-236.

Gill DR, Smyth SE, Goddard CA, Pringle IA, Higgins CF, Colledge WH, and Hyde SC. (2001) Increased persistence of lung gene expression using plasmids containing the ubiquitin C or elongation factor 1α promoter. Gene Therapy 8: 1539-1546.

Gossen M, and Bujard H. (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci USA* 89: 5547-5551.

Graham FO, Smiley J, Russell W and Nairn R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J. Gen. Virol. 36, 59-72.

Huls GA, Heijnen IAFM, Cuomo ME, Koningsberger JC, Wiegman L, Boel E, van der Vuurst-de Vries A-R, Loyson SAJ, Helfrich W, van Berge Henegouwen GP, van Meijer M, de Kruif J, Logtenberg T. (1999). A recombinant, fully human monoclonal antibody with antitumor activity constructed from phage-displayed antibody fragments. Nat Biotechnol. 17, 276-281.

Jones D, Kroos N, Anema R, Van Montfort B, Vooys A, Van Der Kraats S, Van Der Helm E, Smits S, Schouten J, Brouwer K, Lagerwerf F, Van Berkel P, Opstelten D-J, Logtenberg T, Bout A (2003) High-level expression of recombinant IgG in the human cell line PER.C6. Biotechnol. Prog. 19: 163-168.

Kaufman, RJ. (2000) Overview of vector design for mammalian gene expression *Mol Biotechnol 16*, 151-160.

Kaufman, RJ, and Sharp, PA. (1982) Construction of a modular dihydrofolate reductase cDNA gene: analysis of signals utilized for efficient expression *Mol Cell Biol 2*, 1304-1319.

Kellum R, and Schedl P. (1991) A position-effect assay for boundaries of higher order chromosomal domains. *Cell* 64: 941-950.

Kwaks TH, Barnett P, Hemrika W, Siersma T, Sewalt RG, Satijn DP, Brons JF, van Blokland R, Kwakman P, Kruckeberg AL,

Kelder A, Otte AP. (2003) Identification of anti-repressor elements that confer high and stable protein production in mammalian cells. *Nat Biotechnol 21*, 553-558. Erratum in: Nat Biotechnol 21, 822 (2003).

Phi-Van L, Von Kreis JP, Ostertag W, and Strätling WH. (1990) The chicken lysozyme 5′ matrix attachment region increases transcription from a heterologous promoter in heterologous cells and dampens position effects on the expression of transfected genes. Mol. Cell. Biol. 10: 2302-2307.

Lopez de Quinto, S, and Martinez-Salas, E. (1998) Parameters influencing translational efficiency in aphthovirus IRES- based bicistronic expression vectors *Gene 217*, 51-56.

Martin, DI, and Whitelaw, E. (1996) The vagaries of variegating transgenes *Bioessays 18*, 919-923.

Martinez-Salas, E. (1999) Internal ribosome entry site biology and its use in expression vectors *Curr Opin Biotechnol 10*, 458-464.

McBurney, MW, Mai, T, Yang, X, and Jardine, K. (2002) Evidence for repeat-induced gene silencing in cultured Mammalian cells: inactivation of tandem repeats of transfected genes *Exp Cell Res 274*, 1-8.

Migliaccio, AR, Bengra, C, Ling, J, Pi, W, Li, C, Zeng, S, Keskintepe, M, Whitney, B, Sanchez, M, Migliaccio, G, and Tuan, D. (2000) Stable and unstable transgene integration sites in the human genome: extinction of the Green Fluorescent Protein transgene in K562 cells *Gene 256*, 197-214.

Mizuguchi, H, Xu, Z, Ishii-Watabe, A, Uchida, E, and Hayakawa, T. (2000) IRES-dependent second gene expression is significantly lower than cap- dependent first gene expression in a bicistronic vector *Mol Ther 1*, 376-382.

Rees, S, Coote, J, Stables, J, Goodson, S, Harris, S, and Lee, MG. (1996) Bicistronic vector for the creation of stable mammalian cell lines that predisposes all antibiotic-resistant cells to express recombinant protein *Biotechniques 20*, 102-104, 106, 108-110.

Schorpp, M, Jager, R, Schellander, K, Schenkel, J, Wagner, EF, Weiher, H, and Angel, P. (1996) The human ubiquitin C promoter directs high ubiquitous expression of transgenes in mice *Nucleic Acids Res 24*, 1787-8.

Strutzenberger, K, Borth, N, Kunert, R, Steinfellner, W, and Katinger, H. (1999) Changes during subclone development and ageing of human antibody- producing recombinant CHO cells *J Biotechnol 69*, 215-26.

Van der Vlag, J, den Blaauwen, JL, Sewalt, RG, van Driel, R, and Otte, AP. (2000) Transcriptional repression mediated by polycomb group proteins and other chromatin-associated repressors is selectively blocked by insulators *J Biol Chem 275*, 697-704.

Whitelaw, E, Sutherland, H, Kearns, M, Morgan, H, Weaving, L, and Garrick, D. (2001) Epigenetic effects on transgene expression *Methods Mol Biol 158*, 351-68.

Claims

1. A recombinant nucleic acid molecule comprising a nucleic acid sequence having anti-repressor activity selected from the group consisting of:
a) SEQ. ID. NO. 66,
b) fragments of SEQ. ID. NO. 66 wherein said fragment has anti-repressor activity,
c) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences have anti-repressor activity; and
d) the complement to any one of a) to c);
said recombinant nucleic acid molecule further comprising an expression cassette, said expression cassette comprising a heterologous promoter linked to a nucleic acid of interest, and wherein said nucleic acid of interest preferably encodes all or part of a protein of interest.

2. A molecule according to claim 1, wherein said nucleic acid sequence having anti-repressor activity is situated upstream of said promoter in said expression cassette.

3. A molecule according to claim 2, wherein said sequence having anti-repressor activity and said promoter are separated by less than 2 kb.

4. A molecule according to any one of the preceding claims, wherein said nucleic acid encoding a protein of interest is present in a multicistronic gene further encoding a selectable marker gene.

5. A molecule according to any one of the preceding claims, further comprising at least one other sequence having anti-repressor activity, said at least one other sequence being chosen from
a) any of SEQ. ID. NOs. 1-65,
b) fragments of any of SEQ. ID. NOs. 1-65, wherein said fragments have anti-repressor activity, and
c) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences have anti-repressor activity, and
d) the complement to any one of a) to c).

6. A recombinant nucleic acid molecule comprising an expression cassette comprising:
5'- anti-repressor sequence A - promoter - nucleic acid encoding all or part of a protein of interest - anti-repressor sequence B - 3' wherein anti-repressor sequences A and B may be the same or different and are chosen from the group consisting of
(i) any of SEQ. ID. NOs. 1-65,
(ii) fragments of any of SEQ. ID. NOs. 1-65, wherein said fragments have anti-repressor activity,
(iii) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences have anti-repressor activity, and
(iv) the complement to any one of (i) to (iii),
characterized in that said expression cassette further comprises between said anti-repressor sequences A and B a sequence having anti-repressor activity chosen from the group consisting of
a) SEQ. ID. NO. 66,
b) fragments of SEQ. ID. NO. 66 wherein said fragment has anti-repressor activity,

c) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences have anti-repressor activity; and
d) the complement to any one of a) to c).

7. A cell comprising a molecule according to any one of claims 1-6.

8. A cell according to claim 7, which is a mammalian cell.

9. A cell according to claim 8, which is a CHO cell.

10. A method for producing a protein of interest, comprising culturing a cell comprising a recombinant nucleic acid molecule encoding the protein of interest to express said nucleic acid encoding the protein of interest in said cell,
characterized in that
said recombinant nucleic acid molecule comprises a nucleic acid sequence having anti-repressor activity selected from the group consisting of:
a) SEQ. ID. NO. 66,
b) fragments of SEQ. ID. NO. 66 wherein said fragment has anti-repressor activity,
c) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences have anti-repressor activity; and
d) the complement to any one of a) to c).

11. A method according to claim 10, further comprising isolating said protein of interest.

12. A method according to claim 10 or 11, wherein said nucleic acid sequence having anti-repressor activity is situated upstream of a promoter that controls expression of the protein of interest.

13. A method according to claim 12, wherein said sequence having anti-repressor activity and said promoter are separated by less than 2 kb.

14. A method according to any one of claims 10 to 13, wherein said nucleic acid encoding a protein of interest is present in a multicistronic gene further encoding a selectable marker gene.

15. A method according to any one of claims 10 to 14, wherein said cell is a mammalian cell.

16. A method according to claim 15, wherein said cell is a CHO cell.

17. Use of a nucleic acid sequence having anti-repressor activity selected from the group consisting of:
a) SEQ. ID. NO. 66,
b) fragments of SEQ. ID. NO. 66 wherein said fragment has anti-repressor activity,
c) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences have anti-repressor activity; and
d) the complement to any one of a) to c),
for increasing expression of a nucleic acid of interest.

18. A method for generating a host cell expressing two polypeptides of interest, the method comprising:
a) introducing into a host cell one or more nucleic acid molecules, the nucleic acid molecule or molecules together comprising:
(i) a promoter functionally linked to a sequence encoding a first polypeptide of interest and a first selectable marker gene, and
(ii) a promoter functionally linked to a sequence encoding a second polypeptide of interest and a second selectable marker gene, and
(iii) at least one sequence having anti-repressor activity, chosen from the group consisting of (a) any one of SEQ. ID. NOs. 1-66, (b) fragments of any one of SEQ. ID. NOs. 1-66, said fragments having anti-repressor activity, (c) sequences that are at least 70% identical to (a) or (b) and having anti-repressor activity, and (d) the complement of any one of (a)-(c);
b) selecting a host cell by selecting essentially simultaneously for expression of said first and second selectable marker genes.

19. A method for expressing two polypeptides of interest, the method comprising:
culturing a host cell obtained by the method according to claim 18 to express said first and second polypeptides, and optionally isolating said polypeptides.

20. A method according to claim 18 or claim 19, wherein said sequence having anti-repressor activity is selected from the group consisting of
a) SEQ. ID. NO. 66,

b) fragments of SEQ. ID. NO. 66 wherein said fragment has anti-repressor activity,
c) sequences that are at least 70% identical in nucleotide sequence to a) or b) wherein said sequences have anti-repressor activity; and
d) the complement to any one of a) to c).

21. A method according to any one of claims 18-20, wherein said two polypeptides are part of a multimeric protein.