# Kent Academic Repository

**Application of microRNA Targeted 3'UTRs to Repress DHFR Selection Marker Expression for Development of Recombinant Antibody Expressing CHO Cell Pools**

Lyne Jossé[1]*, Lin Zhang[2] and C Mark Smales[1]*

[1]Industrial Biotechnology Centre and School of Biosciences, University of Kent, Canterbury, Kent, CT2 7NJ, UK

[2]Pfizer Inc, 1 Burtt Road, Andover, MA 01810, US

*Joint corresponding authors

Tel no:  00 441227823746

Email: c.m.smales@kent.ac.uk or lj25@kent.ac.uk

**Abstract**

The DHFR system is used for the selection of recombinant CHO cell lines using the inhibitor MTX. During clonal selection, endogenous DHFR expression and resistance to MTX allows the selection of cells expressing sufficient DHFR to survive. Here we describe a novel vector platform for the DHFR system, whereby addition of a synthetic 3'UTR destabilizes DHFR expression. We harnessed miRs ability to negatively regulate gene expression by their near-complementary binding to the 3'UTR region of transcripts. From the literature we identified let-7f as a highly abundant, invariant miR in CHO cells. Three 3'UTR targets of the let-7f miR were then cloned in the DHFR host 3'UTR to determine the impact on gene expression (HMGA2 3'UTR sequence 1, 2 and 3). Using luciferase as a reporter we show down-regulation of luciferase activity is mediated by the nature of the 3'UTR and its ability to bind let-7f. We then applied the same 3'UTRs downstream of the DHFR gene to show this also results in reduced transcript amounts. Finally, we applied this methodology to generate stable DG44-derived cell pools expressing a model monoclonal antibody, demonstrating this approach can be used for the selection of antibody producing cells with low MTX concentrations.

**Introduction**

Cultured Chinese hamster ovary (CHO) cells are the most commonly utilized industrial mammalian cell expression system used for the production of biotherapeutic proteins [1]. When using mammalian cell expression systems such as CHO to generate recombinant proteins at an industrial scale, the usual approach is to generate and isolate stably expressing recombinant cell lines/clones whereby the gene (or genes) of interest are stably incorporated into the host genome [2]. The two most commonly utilized selection systems for the generation of recombinant CHO cell lines are the glutamine synthetase (GS) and dihydrofolate reductase (DHFR) selection systems [3]. GS catalyzes the synthesis of glutamine from glutamic acid, ammonia and ATP. In the GS selection system either a host cell where GS is knocked out is used and/or endogenous GS is inhibited by L-methionine sulfoximine (MSX) and hence only transfectants where exogenous GS is expressed are able to survive selection in glutamine free medium containing MSX [3]. The DHFR enzyme catalyzes the synthesis of tetrahydrofolic acid from dihydrofolic acid and the inhibitor of DHFR, methotrexate (MTX), is used to select for and amplify cells or clones with elevated DHFR expression [4]. The DFHR system is usually used in DHFR deficient CHO cell lines (e.g. DG44 or DUXB-11) and via multiple rounds of selection with increasing DHFR concentrations (up to 1000 nM) high-producing clones can be isolated that contain multiple copies of the DHFR and target recombinant genes [5].

Although the DHFR system is widely used industrially for the generation of high-producing recombinant cell lines, the process can be time consuming and labour intensive [6], with several rounds of MTX amplification of individual clones often used [7]. Consequently, genomic rearrangements can occur upon amplification that leads to instability in the isolated clones [5]. In order to improve the process and reduce the timelines for cell line development, a number of approaches have been reported. These include the combination of DHFR selection with codon optimization of the recombinant gene(s) of interest which when combined are reported to result in saturation of gene expression at MTX concentrations as low as 5 nM without compromising the recombinant protein yields from the isolated clones and in some cases improving product quality [5]. The combination of MTX selection with the use of a mammalian replication initiation region (IR) and a matrix attachment region (MAR), (termed the IR/MAR-DHFR fusion method) reportedly gave rise to clones with improved antibody yields compared to those achieved from either of these methods alone and the clones were more stable than those generated from the traditional DHFR/MTX approach [8].

Others have approached improving the DHFR selection stringency and system by attenuating or reducing DHFR expression. For example, Wu and colleagues used a short hairpin RNA targeted to the DHFR gene to generate recombinant IgG expressing CHO clones and found that this approach improved IgG expression by more than 100% and genomic stability in MTX free culture by 30% [9]. The use of a destabilized-DHFR selection marker

3

linked to an attenuated IRES element has also been shown to be effective in generating high yielding recombinant CHO cell lines [10]. The use of a codon de-optimized DHFR selectable marker to improve selection stringency in the presence of MTX has also been shown to result in enhanced expression of recombinant proteins, showing that reducing translational efficiency of DHFR, and hence its expression, can be used to isolate cell lines with improved stringency and higher productivities without multiple and time-consuming gene amplification steps [6].

An alternative method of tuning or attenuating translation, and hence gene expression, is *via* the manipulation of microRNAs (miRNAs or miRs). miRs were first described in *C. elegans* as an antisense RNA regulating the level of LIN-14 protein [11] and are small (20-23 nucleotides) non-coding RNA molecules that mainly act as negative regulators of gene expression by binding to target mRNAs in their 3'UTR inhibiting their translation [12]. Target mRNAs can be paired to more than one miR via the 3'UTR region, likewise a single miR can in theory target hundreds of unrelated messages and thus it is thought that the influence of miRs is wide-ranging. Indeed, the use of miR cell engineering to change the phenotype of CHO cells for improved recombinant protein and cell growth characteristics has now been reported by a number of groups (e.g. see [13-19]), this approach having the advantage of being able to engineer or manipulate multiple pathways simultaneously without placing an additional translational burden on the cell that traditional cell engineering approaches impose [20]. Here we describe the application of harnessing endogenously expressed CHO miRs to mediate DHFR gene expression and hence manipulate selection stringency to develop monoclonal antibody (mAb) expressing recombinant CHO cell pools. The approach relies on fusing the miRNA target containing 3'UTR sequences to the DHFR gene which is co-located with the heavy chain gene in a mAb plasmid expression vector. We show that using this approach, endogenous miR expression can be harnessed to enhance the stringency of recombinant gene expression selection and demonstrate this leads to generation of recombinant CHO cell pools with elevated mAb expression compared to those without the miR repression with low concentrations of MTX.

**Materials and Methods**

*CHO Cell Lines and Cell Culture Protocols*

For the general cloning of genes/isolation of RNA and genomic DNA, and for undertaking luciferase expression luminescence experiments, the commercially available CHO-S cell line was utilized (ThermoFisher). The DG44 DHFR$^{-/-}$ CHO cell line was obtained from the European Collection of Cell Cultures (ECACC, No. 05011002) (http://www.hpacultures.org.uk/collections/ecacc.jsp). CHO-S cells were routinely cultured in commercially available CD-CHO (ThermoFisher, UK) supplemented with 8 mM L-glutamine, in Erlenmeyer flasks (Corning Inc.). For routine sub-culturing, 250 mL flasks containing 50 mL

cultures were used. Cells were sub-cultured every 3 or 4 days with new cultures being seeded at 0.2 x 10$^6$ viable cells/mL in fresh media. The cells were incubated at 37°C, with shaking at 100 rpm in a 5% $CO_2$ atmosphere. The concentration and viability of cells was determined on a Vi-Cell XR (Beckman Coulter Inc, UK) instrument. DG44 cells were grown adherently under 5% $CO_2$ in air conditions at 37°C in 90% Iscove's Modified Dulbecco's Medium (IMDM), 10% (v/v) γ-irradiated, dialyzed foetal calf serum (FCS) (Lonza), supplemented with 2 mM glutamine, 0.1 mM hypoxanthine, 0.01 mM thymidine (all ThermoFisher, UK), in flat T-flasks.

*Identification of miR 3'UTR Targets using Computational Prediction Software*
The online freeware software packages Targetscan Version 5.2 (http://www.targetscan.org/), miRanda (http://www.microrna.org/) and TargetRank (http://genes.mit.edu/targetrank/) were used to search against mouse and human databases to identify/predict genes targeted by microRNAs of interest. Both software packages utilize genome coordinates for 3'UTRs using RefSeq annotations available from UCSC (http://genome.uc.sc.edu/).

*Plasmids, Primers and Cloning Strategies*
The commercially available pcDNA4/myc-HIS A plasmid was from ThermoFisher (Catalogue no. V863-20). pRLSV40L78 (containing the synthetic 8 let7A target sequence) and pLSVM3' (containing the c-Myc target sequence) were a kind gift from Prof. Bushell, MRC Toxicology Unit, University of Leicester [21]. The pPHA79407 and pPHA79408 plasmids are Pfizer Inc proprietary constructs containing the genes for the expression of a model monoclonal antibody IgG (pHA79407, heavy chain (HC) and DHFR genes; pHA79408 light chain (LC) gene). For exogenous expression of miRs, the miR precursors (primiRs) were cloned into the commercial pcDNA4/myc-HIS A vector. The human pri-miR coordinates and sequences were obtained from the miR database (http://mirbase.org/). All PCR amplified pri-miR hairpins covered at least 250 nucleotide flanking sequences on either side of the pri-miR and were amplified using the forward and reverse primers detailed in Supplementary Table 1 (restriction sites for cloning in each primer underlined). For the cloning of the human HMGA2 3'UTR fragments into the pRL vector, the sequence was retrieved from Ensembl (http://www.ensembl.org/index.html/). All primer sequences utilized for amplifying target 3'UTRs and cloning into the pRL vector are detailed in Supplementary Table 2. Primers for introducing the EcoRV and KpnI restriction sites into the Pfizer pPHA79407 plasmid were designed using Stratagene's web-based QuikChange Primer Design Program (http://www.stratagene.com/qcprimerdesign/). The 3'UTRs were cloned into the resulting plasmid termed pPHA79407 E/K as EcoRV/KpnI fragments using the primer sequences described in Supplementary Table 3. Primers used for qRT-PCR mRNA analyses are described in Supplementary Table 4. Site-directed mutagenesis was undertaken using the QuikChange Lightning Site-Directed Mutagenesis Kit

(Stratagene). Genomic DNA was isolated using the commercially available FlexiGene DNA kit (Qiagen).

*RNA Isolation and qRT-PCR Assays for the Determination of Relative mRNA Amounts*

Total RNA was isolated from cell pellets using an RNeasy Mini Kit (Qiagen). qRT-PCR was carried out in sealed low 96-well white plates (BioRad multiplate PCR plates) using the QuantiFast SYBR Green RT-PCR kit (Qiagen) and the gene specific primers designed (see Supplementary Table 4) as previously described [22]. PCR reactions were run on a BioRad DNA engine with CHROMO4 continuous fluorescence detector and analyzed using Opticon Monitor 3.1 software and the following program ($50^{\circ}$C 10 min, $95^{\circ}$C 5 min, 40 cycles of $95^{\circ}$C 10 sec, $58^{\circ}$C 30 sec, plate read and followed by a melting curve $58^{\circ}$C to $95^{\circ}$C, read every $0.5^{\circ}$C, hold 1 sec).

*Transfection of Plasmid DNA into CHO Cells*

Transient transfections were carried out using Lipofectamine LTX (ThermoFisher). Transient transfection experiments were undertaken in 6-well plates and at the time of transfection $0.8 \times 10^6$ cells/well were mixed with 2 $\mu$g of the relevant plasmid. CHO-S suspension cells were seeded immediately prior to transfection whilst DG44 adherent cells were seeded 24 h prior to transfection. Stable transfection of DG44 cells was undertaken using electroporation. Briefly, 80 μg each of both light and heavy chain plasmids, linearized using the restriction enzyme NotI, were used to transfect $1.25 \times 10^7$ viable cells in 0.4 cm pulse cuvettes (BioRad) by selecting the exponential decay programme on a Biorad Gene Pulser Xcell system (300 V, 950 μF capacitance and infinite resistance). Following electroporation, the cells were gently transferred to flat T75 flasks (final volume 20 mL) and placed in a static 5% $CO_2$ incubator at $37.0^{\circ}$C.

*Selection and Amplification of DHFR Positive Cell Pools*

Two days post-transfection, cells expressing DHFR were selected for by removing the HT supplement and growing in 90% IMDM, 10 % θ-irradiated dialyzed FBS, and grown in T-75 flasks containing 5 or 10 nM MTX. Cells were then returned to the incubator for an extra 7 to 14 days. The cells were then left for 2 weeks for colonies to grow and then amplification continued up to 6 weeks by sub-cultuirng every 3 or 4 days in the presence of the appropriate amount of MTX (5 or 10 nM) at which point cell pools were cryopreserved.

*ELISA for the Determination of IgG Amounts in Cell Culture Supernatants*

The concentration of assembled IgG in supernatant of cell pools was assessed by ELISA as previously described [23].

*Reporter Gene Analysis*

For determining *Renilla* luciferase activity we used the commercially available Renilla-Glo Luciferase Assay system (Promega). For the measurement of Gaussia luciferase activity in the cell culture supernatant we used the commercially available Gaussia Luciferase Assay kit (New England Biolabs, NEB). For both methods, the luminescence was measured after a 10 min equilibration period.

*Statistical Testing*

ANOVA analysis was used for statistical testing followed by *post hoc* Tukey's test analysis.

**Results**

*Selection of endogenous miRs and target 3'UTRs to mediate repression of DHFR expression*

Our approach to mediate repression of DHFR expression was to utilize endogenous CHO miRs that are highly expressed with little variation throughout culture, that would target specific 3'UTRs and could be introduced onto the DHFR mRNA. This would circumvent the need to engineer the miR in the cell to achieve repression of DHFR expression that might perturb cell physiology. Further, it was necessary to consider that the introduction of the exogenous DHFR with a 3'UTR targeted by a specific miR may titrate down endogenous levels of that specific miR. Therefore, for our application it was desirable to utilize miRs that are naturally abundant within the host cell.

We initially scanned the literature to identify miRs with low variability across culture conditions (*i.e.* the miR concentration was independent of the CHO cell subtype, the medium, and was maintained at high concentrations relative to other miRs throughout culture). The initial miRs were selected based upon the data published by Johnson et al who reported on the prevalence of miRs in CHO cells using high-throughput sequencing in six cDNA libraries generated from four different cell lines [18]. The authors of this study identified 350 miR/miR* sequences by homology to other species and established that there was a range of expressions of these miRs. The read counts plotted in Figure 1A, on a $\log_{10}$ scale, show the 25 miRs (from the 350) exhibiting the lowest variability in abundance compared to the other miRs identified in the study across the six libraries investigated. The nomenclature used in Figure 1A corresponds to the species that showed the best match in the miRbase reference sequence to the CHO small RNA sequence obtained by deep sequencing. Variability was assessed by calculating the ratio of the mean abundance of the miRs to the standard deviation (SD). Interestingly, members of the conserved let-7 family were highly-represented in terms

of being very abundant with low variability across the libraries (Figure 1A). Let-7f was the most abundant miR ranging from 650,000 to 1.4 million copies per cell (Figure 1A) [18]. As this miR was naturally abundant, we assumed that its expression would not need to be artificially modulated to repress let-7f specific 3'UTR target mRNAs. We therefore selected let-7f as the miR for which we would look for suitable 3'UTR targets that might be utilized when placed in the 3'UTR of appropriate transcripts (in this case DHFR) to repress their expression.

In order to identify 3'UTRs that are a target of the let-7f miR we used *in silico* prediction software. The limitation of using *in silico* prediction methods to identify miR 3'UTR targets for a specific miR is that miR:mRNA pairing programs often show very little overlap or agreement in their predictions [24-25]. However, such approaches are useful in identifying potential targets for further evaluation and hence we used the freeware software programmes TargetScan (www.targetscan.org), miRanda (www.microrna.sanger.ac.k/targets) and TargetRank (http://genes.mit.edu/targetrank/) for initial target prediction. We ran a search for human let-7f mRNA targets using TargetScan version 5.2 and compared the top 10 hits with those retrieved from miRanda and TargetRank (Table 1). From the potential targets identified, we chose the 3'UTR of HMGA2 as a potential match for let-7f binding to place in the 3'UTR of target transcripts such that their expression would be repressed by let-7f (Table 1).

The human HMGA2 3'UTR is long at 2996 bp and harbors five predicted conserved, and one predicted non-conserved, binding sites for let-7f (Figure 1B). The pPHA79407 plasmid which carries the DHFR marker is 7915 bp and the DHFR native 3'UTR is approximately 1000 bp. Replacing the latter with the full HMGA2 3'UTR would have resulted in a vector over 10 kb and could have potentially reduced the overall transfection efficiency of the plasmid. We therefore opted to clone partial HMGA2 3'UTR sequences onto the 3'UTR of DHFR. We selected three regions (as highlighted in Figure 1B), each one containing a different seed region, the nucleotides 2-7 or 2-8 of the miR 5' end that are responsible for initiating mRNA binding. As the nature of the seed impacts on the degree of regulation, an 8mer seed being the strongest and 7merA1 the weakest [25], we hypothesized that region 1 (1 x 8mer), 2 (2 x 7mer-m8, 1x7mer-A1) and 3 (1x7mer-A1) might give differential translational repression. A complete repression of DHFR expression would be unsuitable for our purposes and this was thus a further basis for utilizing partial HMAGA2 sequences rather than the full length 3'UTR. The full length 3'UTR encompasses all the identified potential regulation sites and hence might result in very high repression of target 3'UTR constructs compared to the partial sequences.

In parallel to using the HMGA2 target sequences identified via the bioinformatics approach described above, we obtained two constructs containing 3'UTR miR targets. The first construct contained the c-myc 3'UTR and the second a synthetic cassette 8let-7A which are targeted by miR34c and let-7a, respectively. Previous work has already demonstrated that

the presence of these target sequences results in the repression of *renilla* luciferase when placed downstream of the open reading frame and that the repression was relieved in the presence of the antisense oligonucleotide (anti-miR) for the appropriate miR (miR34c or let-7a [21]). These targets were therefore validated and were a positive control for the luciferase assay.

*Functional validation of repressed gene expression in CHO cell lines by the presence of 3'UTRs targeted by specific miRs*

Having identified 3'UTRs that should be targeted by endogenous high abundant miRs in CHO cells, it was necessary to validate that the presence of the 3'UTRs on specific recombinant mRNAs resulted in repression of expression at the protein level. To achieve this, we fused the target 3'UTRs onto the open reading frames of the reporter Renilla luciferase in transient expression studies. Renilla luciferase is retained in the cytoplasm of the cell as the DHFR protein would be. When the wild-type 3'UTR of the Renilla luciferase was replaced with any of the HGMA2 regions, the c-myc 3'UTR or synthetic let-7a the luciferase expression was dramatically repressed (Figure 2A). When constructs for the expression of exogenous pri-miRs to elevate the amount of miRs present were co-transfected with the luciferase constructs, no additional repression was observed (Figure 2A). This suggests that endogenous miR amounts are not limiting repression and hence repression could not be further enhanced by over-expressing the specific miRs. The repression of Renilla luciferase expression indicated that the HGMA2 region 1 and 2 3'UTRs gave rise to the largest repression in reporter gene expression compared to the control (approximately 90%) whilst the HGMA2 region 3 3'UTR harboring construct reduced the expression by approximately 75% compared to the control (Figure 2A).

To determine if the same 3'UTRs could also repress DHFR expression we sub-cloned all the 3'UTRs downstream of the DHFR marker in the pPH79407 E/K heavy chain expressing vector. To this end, the 3'UTRs (HMGA2 sub-regions 1, 2 and 3 as well as the c-myc and let-7a miR target cassettes) were amplified as EcoRV/KpnI fragments using pGL as a template and the primers described in Supplementary Table 3. Figure 2B (left-hand bar for each set) shows that for all five 3'UTR variants (HMGA2 1, 2, 3, let7a cassette and c-myc cassette), following transient transfection in CHO-S cells the amount of DHFR mRNA was reduced when compared to the control construct. The endogenous DHFR mRNA levels were negligible in comparison to that observed in the pPHA79407 E/K control 3'UTR (Figure 2B, Null). The drop in the relative amounts of the DHFR mRNA implies that 3'UTR-mediated regulation of DHFR leads to mRNA degradation, even though miR mediated repression can result in either inhibition of translation (and no change to mRNA levels) or degradation. We could not however confirm a reduction at the protein level (not shown) as we could not detect DHFR protein by

western blot analysis even in the control pPHA79407 E/K transfection experiment (although the western blot assay was shown to be working using a positive control for DHFR, data not shown). Collectively these data suggest that the 3'UTRs identified were able to lead to repression of gene expression and were thus suitable for use in repressing DHFR expression for the purpose of modulating selection stringy during stable integration of recombinant genes.

*Application of the miR target-containing 3'UTRs in the DHFR system for recombinant IgG production*

Having established that the selected 3'UTRs could result in repression of transgene expression, we then applied this to DHFR expression coupled to IgG production. pPH79407 E/K and its derivatives bearing HMGA2 1, 2, 3, let7a cassette and c-myc cassette were therefore co-transfected with the light chain vector (pPHA79408) into the DG44 (DHFR-/-) host and initially transient expression of IgG was analyzed by ELISA (Figure 3A). The ELISA data showed that 24 hour post-transfection, the amount of IgG secreted from the cells transfected with the miR target containing 3'UTR vectors was lower than that of the controls (Figure 3B). However, this trend was reversed at 48 h post-transfection (Table 2), suggesting that post-transcriptional repression of DHFR does not in the long term affect the overall expression of the heavy chain and ultimately IgG production in a transient system.

Although there was a small benefit in transient expression of IgG using the miR targeted 3'UTRs, these constructs were designed to facilitate stable cell line construction via the repression of DHFR expression. In this way, during stable cell line construction the system would select for those stable integrants that were expressing the highest amounts of DHFR, presumably due to the cassette being integrated into sites with high transcriptional activity. We therefore used the 3'UTR repressed DHFR system to construct stable pools expressing the target IgG. To generate the stable pools we transfected the different constructs into the DG44-deficient host cell line and then selected for DHFR positive cells using 5 or 10 nM MTX. All measurements were made in triplicate. Figure 3B and Table 2 show the amount of IgG1 from each of the pools that resulted using the different 3'UTR constructs as determined using ELISA. In our hands, cells originally transfected with pPHA79407 bearing HMGA2 3 3'UTR died when exposed to MTX concentration greater than 5 nM MTX. Polyclonal populations originally transfected with pPHA79407 E/K bearing HMGA2 1 3'UTR or the 8-let-7a cassette, exhibited higher IgG levels following gene amplification at both 5 nM and 10 nM MTX compared to the control. However, the change in absorbance in the ELISA reflective of IgG production between the no MTX control and the MTX containing pools was greatest in HMGA2 1 pools at 10 nM MTX. This result was consistent with our earlier observation that the HMGA2 1 3'UTR led to reduced gene expression (see luciferase in Figure 2A and DHFR in Figure 2B). The data therefore collectively suggest that the described approach which is a novel way to

elicit repression of gene expression of DHFR can be associated with MTX selection, at low concentration, to screen mAb producing polyclonal pools prior to cloning out cells.

**Discussion**

The DHFR system and use of the MTX inhibitor is widely used for the generation of recombinant cell lines when using DHFR deficient CHO host cell lines [8, 26-28]. The traditional approach to generating recombinant pools of cells and clonal populations using this approach is labor intensive and as described in the introduction, there have been a number of reports to modify the approach to reduce the time and labor required to generate cell lines. Further, whilst this approach can be used to amplify recombinant gene copy number, the MTX inhibitor of the DHFR enzyme is cytotoxic and mutagenic and its use can lead to unwanted mis-translation and mis-incorporation of amino acids into recombinant proteins [29]. As such, being able to reduce the concentration of MTX required to obtain appropriate cell lines would be advantageous.

Here we have identified from the literature highly abundant miRs in CHO cells and used this information to select a panel of 3'UTRs targeted by these endogenous and highly abundant miRs for repression of the DHFR marker. The majority of studies into miRs in CHO to date have focused upon profiling or manipulation of miR amounts (e.g. [30-31]), but here we focused upon utilizing endogenous amounts of highly expressed and conserved miRs. Selecting and exploiting the endogenous miR profile of the cell means that miR mediated repression would in theory not require the manipulation of the miR itself which could result in unwanted side-effects in terms of changing the gene expression of the CHO host. With this in mind, let-7f was selected as the endogenous miR for which target 3'-UTR s were identified to place on the *DHFR* gene as previous reports show this miR to be highly expressed across hosts/clones [18]. We would therefore not expect let-7f expression to be lost due to clonal stability as manipulation of let-7f expression was not required. This should help circumvent, as far as possible, issues around stability and lost expression. The high expression of let-7f means that even with the DHFR transcript acting as a sponge for this miR, we assumed it would not significantly deplete this miR and thus impact on other CHO cell processes reliant on let-7f expression.

Using the approach of placing let-7f target UTRs on the DHFR , the selected 3'UTRS were able to repress reporter gene expression as shown using luciferase as a model system. Although we do not have direct evidence that the repression is due to the miRs targeting the 3'UTRs, the reduction in luciferase expression and mRNA in the presence of the specific 3'UTRs suggests that this is likely to be the case. We could not use sponge vectors that 'mop-up' miRs [32] to show that this reduced the repression as the chosen miRs are among the

11

most highly abundant present and hence reducing these significantly would be exceedingly difficult.

We then applied our 3'UTRs in the DG44 DHFR null host CHO cell line and used a heavy chain vector containing the DHFR selection gene with various 3'UTRs to rescue the deficiency in folate metabolism in the presence of MTX. In this system, resistance to the folic acid analogue, methotrexate, require stable integration of the DHFR marker in highly transcribed sites and co-selection for the enhanced expression of the heavy chain of the model IgG1 used in this study. We also wanted to minimize the concentration of MTX utilized to generate antibody expressing cell pools and thus the selection procedure used 5-10 nM MTX to limit detrimental effects on cellular fitness and potential mutations. The selected repressed 3'UTRs and subsequent ELISA determination of mAb amounts observed early in the cell line development phase were compared to a standard 3'UTR used on the DHFR selection marker at Pfizer. In each case, higher ELISA readings were obtained for the selected 3'UTRs compared to the control.

The data shown in this study collectively demonstrates the feasibility of using 3'UTR manipulation of the DHFR gene to improve the time and product yield from CHO DHFR deficient pools using low concentrations of MTX for selection. Further, our data across the luciferase and antibody studies suggests that using 3'UTRs targeted by highly endogenous miRs in the CHO cell can be used to rapidly generate pools expressing a gene of interest. In particular, our data suggests that the 3'UTR/miR pair let-7f and HMGA2-1 can be used as a tool to repress the DHFR selection system in DG44 (DHFR$^{-/-}$) cells and enhance the productivity observed in surviving cells. Interestingly, the 3'UTR of the human DHFR mRNA has a natural polymorphism for the miR-24 binding site that results in MTX resistance [33]. It may therefore be possible to further manipulate the 3'UTR of the DHFR mRNA used for selection purposes in CHO cell line construction to either enhance or reduce the repression of DHFR expression in tandem with miR-24 or other miR binding sites. Use of such a miR-mediated DHFR repression approach also has the advantage over other selection methods that use drug inhibitor approaches alone in that the data here suggests lower drug concentrations can be used to obtain the same or better results and that the repression is independent of the selection system. This enables the application of this technology to other selection markers or approaches and for the user to maintain stable cell pools using lower drug concentrations, giving the methodology versatility in that it could be applied to a wide range of recombinant protein expression systems.

**Acknowledgements**

**Conflict of Interest Statement**

The authors declare no commercial or financial conflict of interest.

**REFERENCES**

1. J.Y. Kim, Y.G. Kim, G.M. Lee, *Appl. Microbiol. Biotechnol.* 2011*,* **93,** 917.

2. A.J. Porter, A.J. Racher, R. Preziosi, A.J. Dickson, *Biotechnol. Prog.* 2010, **26,** 1455.

3. M Feary, A.J. Racher, R.J. Young, C.M. Smales, *Biotechnol. Prog.* 2017, **33**, 17.

4. N. Vishwanathan, H. Le, N.M. Jacob, Y.S. Tsao, S.W. Ng, B. Loo, Z. Liu, A. Kantardjieff, W.S. Hu, *Biotechnol. Bioeng.* 2014, **111,** 518.

5. E. Kotsopoulou, H. Bosteels, Y.T. Chim, P. Pegman, G. Stephen, S.I. Thornhill, J.D. Faulkner, M. Uden, *J. Biotechnol.* 2010, **146**, 186.

6. A.D. Westwood, D.A. Rowe, H.R.G. Clarke, *Biotechnol. Prog.* 2010, **26**, 1558.

7. S.C. Jun, M.S. Kim, J.Y. Baik, S.O. Hwang, G.M. Lee, *Appl. Microbiol. Biotechnol.* 2005, **69**, 162.

8. C. Noguchi, Y. Araki, D. Miki, N. Shimizu, *PLoS One* 2012, **7**, e52990.

9. S.C. Wu, W.W.L. Hong, J.H. Liu, *Vaccine* 2008, **26**, 4969.

10. S.K. Ng, T.R. Tan, Y. Wang, D. Ng, L.T. Goh, M. Bardor, V.V. Wong, K.P. Lam *PLoS One* 2012, **7**, e52785.

11. R.C. Lee, R.L. Feinbaum, V. Ambros, *Cell* 1993, **75**, 843.

12. V.N. Kim, J. Han, M.C. Siomi, *Nat. Rev. Mol. Cell Biol.* 2009, **10**, 126.

13. N. Barron, N. Sanchez, P. Kelly, M. Clynes, *Biotechnol. Lett.* 2011, **33**, 11.

14. N. Barron, N. Kumar, N. Sanchez, P. Doolan, C. Clarke, P. Meleady, F. O'Sullivan, M. Clynes, *J. Biotechnol.* 2011, **15**1, 204.

15. P. Gammell, N. Barron, N. Kumar, M. Clynes, *J. Biotechnol.* 2007, **130**, 213.

16. M. Hackl, T. Jakobi, J. Blom, D. Doppmeier, K. Brinkrolf, R. Szczepanowski, S.H. Bernhart, C. Honer Zu Siederdissen, J.A. Bort, M. Wieser, R. Kunert, S. Jeffs, I.L. Hofacker, A. Goesmann, A. Puhler, N. Borth, J. Grillari, *J. Biotechnol.* 2011, **153**, 62.

17. J.A. Hernández Bort, M. Hackl, H. Hoflmayer, V. Jadhav, E. Harreither, N. Kumar, W. Ernst, J. Grillari, N. Borth, *Biotechnol. J.* 2012, **7**, 500.

18. K.C. Johnson, N.M. Jacob, P.M. Nissom, M. Hackl, L.H. Lee, M. Yap, W.S. Hu *Biotechnol. Bioeng.* 2011, **108**, 475.

19. A. Kantardjieff, P.M. Nissom, S.H. Chuah, F. Yusufi, N.M.Jacob, B.C. Mulukutla, M. Yap, W.S. Hu *Biotechnol. Adv.* 2009, **27**, 1028.

20. M. Hackl, N. Borth, J. Grillari, *Trends Biotechnol.* 2012, **30**, 405.

21. Y.W. Kong, I.G. Cannell, C.H. de Moor, K. Hill, P.G. Garside, T.L. Hamilton, H.A. Meijer, H.C. Dobbyn, M. Stoneley, K.A. Spriggs, A.E. Willis, M. Bushell, *Proc. Natl. Acad. Sci. U S A* 2008, **105**, 8866.

22. E.J. Mead, L.M. Chiverton, C.M. Smales, T. von der Haar, T. *Biotechnol. Bioeng.* 2009, **102**, 1593.

23. L. Jossé, C.M. Smales, M.F. Tuite, *Biotechnol. Bioeng.* 2010, **105**, 556.

24. D.P. Bartel, *Cell* 2009, **136**, 215.

25. J. Brennecke, A. Stark, R.B. Russell, S.M. Cohen, *PLoS Biol.* 2005, **3**, e85.

26. W.W. Hong, S.C. Wu, *Vaccine* 2007, **25**, 4103.

27. T. Lai, Y. Yang, S.K. Ng, *Pharmaceuticals* 2013, **6**, 579.

28. D.A. Eichholtz, S.G. Rogers, R.B. Horsch, H.J. Klee, M. Hayford, N.L. Hoffman, S.B. Bradford, C. Fink, J. Flick, K.M. O'Connell, *Somat. Cell Mol. Genet.* 1987, **13**, 67.

29. S. Yamasaki, P. Ivanov, G.F. Hu, P. Anderson, *J. Cell Biol.* 2009, **185**, 35.

30. A. Maccani, M. Hackl, C. Leitner, W. Steinfellner, A.B. Graf, N.E. Tatto, M. Karbiener, M. Scheideler, J. Grillari, D. Mattanovich, R. Kunert, N. Borth, R. Grabherr, W. Ernst, *Appl. Microbiol. Biotechnol.* 2014, **98**, 7535.

31. F. Stiefel, S. Fischer, A. Sczyrba, K. Otte, F. Hesse, *J. Biotechnol.* 2016, **225**, 31.

32. M.S. Ebert, P.A. Sharp, *Curr. Biol.* 2010, **20**, R858.

33. P.J. Mishra, R. Humeniuk, G.S. Longo-Sorbello, D. Banerjee, J.R. Bertino, *Proc. Natl. Acad. Sci. U S A* 2007, **104**, 13513.

**Figure Legends**

**Figure1. A.** The top 25 miR species (from the 350) that exhibited low variability in abundance compared to other miRs across a panel of six CHO miR libraries previously reported by Johnson et al [18]. The results are plotted on a $\log_{10}$ scale. Variability was assessed by calculating the ratio of the mean abundance of the miRs to the standard deviation (SD), n=6 **B.** let-7f binding sites onto the HMGA2 3'UTR, according to TargetScan (Version 5.2): sites marked in light shade are conserved across species while the site marked in black shade is poorly conserved.

**Figure 2. A.** Renilla luciferase activity in CHO-S cells, 24 h post transfection with plasmid constructs, using lipofection. AOVA analysis shows an overall P value of 1.85 $E^{-10}$. All data are significantly different to pRSLSV40 (Control) as determined by Tukey's Honestly Significant Difference criterion. None of the pairwise comparison (endogenous versus exogenous) show significant statistical difference. **B.** The DHFR mRNA levels were quantified by qRT-PCR for all 3'UTR variants in transiently transfected CHO-S cells. Bars represent standard deviation, samples were analyzed in triplicates.

**Figure 3. A.** The amount of actively bound lambda chain was measured by ELISA in transiently transfected DG44 (DHFR-/-) cell ines at 24 h and 48 h post transfection. **B.** For all variants, the level of actively bound lambda chain was measured by ELISA in DG44 (DHFR-/-) stable lineage (polyclonals). The polyclonal pools were passaged three-four times under selection pressure. Bars represent standard deviation, samples were analyzed in triplicates. ANOVA analysis for dataset with 5 nM MTX shows an overall P value of 7.85 $E^{-15}$ while ANOVA analysis for dataset with 10 nM MTX shows an overall P value of 0.000361. This can be explained by a greater variability within the 10 nM MTX dataset. For HMGA2 REG1 there was a significant difference between control and 10 nM MTX treatment (p value of 0.0012) while no significant difference was observed at 5 nM MTX (p value of 1).

**Table 1.** hsa-let-7f mRNA target predictions as inferred by TargetScan (Version 5.2), TargetRank and miRanda.

| Target gene prediction: **TargetScan Version 5.2** | Gene Name | Target gene prediction: **TargetRank** | Target gene prediction: **miRanda** |
|---|---|---|---|
| *HMGA2* | High mobility group AT-hook 2 | *C14orf28* | *HMGA2* |
| *C14orf28* | chromosome 14 open reading frame 28 | *HMGA2* | *IGF2BP1* |
| *LIN28B* | lin-28 homolog B (*C.elegans*) | *TRIM71* | *LIN28B* |
| *FIGNL2* | fidgetin-like 2 | *DNA2* | *CLCN5* |
| *TRIM71* | Tripartite motif-containing 71 | *SMARCAD1* | *C14orf28* |
| *IGF2BP1* | Insulin-like growth factor2 mRNA binding protein 1 | *ESR2* | *ARID3B* |
| *PUNC* | Putative neuronal cell adhesion molecule | *EDN1* | *IGFIR* |
| *ARID3B* | AT rich interactive domain 3B (BRIGHT-like) | *FIGN* | *LOC401720* |
| *FIGN* | fidgetin | *FIGNL2* | *FIGN* |
| *THRSP* | Thyroid hormone responsive (SPOT14 homolog, rat) | *CTPS2* | *PUNC* |

**Table 2**. Fold-increase in antibody production after treatment with either 5 nM or 10 nM MTX, compared to non-treated samples (n = 3).

| | Fold increase (ELISA) | |
|---|---|---|
| | 5 nM MTX | 10 nM MTX |
| CON | 1.4 | 1.5 |
| HMGA2- REG1 | 1.0 | 14.6 |
| HMGA2- REG2 | 0.8 | 3.4 |
| HMGA2- REG3 | 1.3 | |
| c-myc | 0.7 | 0.6 |
| 8-let-7A | 1.4 | 2.0 |

**A**



**B**



**Figure 1.**

**A**



**B**



**Figure 2.**

**A**



**B**



**Figure 3**