Article:
Johari, Y.B., Scarrott, J.M. orcid.org/0000-0002-6046-7687, Pohle, T.H. et al. (4 more authors) (2022) Engineering of the CMV promoter for controlled expression of recombinant genes in HEK293 cells. Biotechnology Journal, 17 (8). e2200062. ISSN 1860-6768

https://doi.org/10.1002/biot.202200062

Received: 31 January 2022 | Revised: 7 April 2022 | Accepted: 23 April 2022

DOI: 10.1002/biot.202200062

Biotechnology Journal

RESEARCH ARTICLE

Engineering of the CMV promoter for controlled expression of recombinant genes in HEK293 cells

Yusuf B. Johari[1] | Joseph M. Scarrott[1] | Thilo H. Pohle[1] | Ping Liu[2] | Ayda Mayer[2] | Adam J. Brown[1,3] | David C. James[1,3]

[1]Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, UK

[2]Cell Line Development, REGENXBIO Inc., Rockville, Maryland, USA

[3]Syngensys Ltd., Sheffield, UK

Correspondence
David C. James, Department of Chemical and Biological Engineering, University of Sheffield, Mappin St., Sheffield S1 3JD, UK.
Email: d.c.james@sheffield.ac.uk

Present address: Yusuf B. Johari, Lonza Biologics, Cambridge, UK

Abstract

Expression of recombinant genes in HEK293 cells is frequently utilized for production of recombinant proteins and viral vectors. These systems frequently employ the cytomegalovirus (CMV) promoter to drive recombinant gene transcription. However, the mechanistic basis of CMV-mediated transcriptional activation in HEK293 cells is unknown and consequently there are no strategies to engineer CMV for controlled expression of recombinant genes. Extensive bioinformatic analyses of transcription factor regulatory elements (TFREs) within the human CMV sequence and transcription factor mRNAs within the HEK293 transcriptome revealed 80 possible regulatory interactions. Through in vitro functional testing using reporter constructs harboring discrete TFREs or CMV deletion variants we identified key TFRE components and clusters of TFREs (cis-regulatory modules) within the CMV sequence. Our data reveal that CMV activity in HEK293 cells is a function of the promoters various constituent TFREs including AhR:ARNT, CREB, E4F, Sp1, ZBED1, JunB, c-Rel, and NF-κB. We also identified critical Sp1-dependent upstream activator elements near the transcriptional start site that were required for efficient transcription and YY1 and RBP-Jκ binding sites that mediate transrepression. Our study shows for the first time that novel, compact CMV-derived promoters can be engineered that exhibit up to 50% higher transcriptional efficiency (activity per unit DNA sequence) or 14% increase in total activity compared to the wild-type counterpart.

KEYWORDS
CMV promoter, HEK293 cells, transcription factor, transcriptional regulation, transient gene expression

Abbreviations: AAV, adeno-associated virus; CHO, Chinese hamster ovary; CMV, cytomegalovirus; CRM, *cis*-regulatory modules; GFP, green fluorescent protein; HEK, human embryonic kidney; TPM, transcripts per million; ORF, open reading frame; RNA-Seq, next-generation RNA sequencing; TF, transcription factor; TFRE, transcription factor regulatory element; TGE, transient gene expression; TSS, transcriptional start site..

1 | INTRODUCTION

The human embryonic kidney 293 (HEK293) cell line is the most commonly utilized human cell line for manufacturing of therapeutic proteins and viral vectors. The cell line serves as a fitting expression host for proteins with particular requirement for human post-translational modifications,[1] and possesses integrated E1 genes required for adeno-associated virus (AAV) production.[2] Moreover, facile cell culture and transfection enable transient gene expression (TGE) methods to be effectively employed for rapid production of potential drug candidates and expression of toxic viral genes.[3,4] Accordingly, various cell and process development have been carried out to improve TGE in HEK293 cells. Examples include optimization of transfection methods,[5] and co-expression of the genes encoding cell cycle-regulators p18 and p21[6] or knock-out of the pro-apoptotic genes Bax and Bak.[7] Further efforts to boost TGE levels are exemplified via process engineering strategies by implementing mild hypothermia[8,9] as well as the addition of chemical inducers such as valproic acid[6,10] and sodium butyrate.[11] Despite these improvements, at the core of HEK293-based production systems, recombinant gene transcription is most often directed by the human cytomegalovirus immediate early (hCMV-IE) promoter,[6,8–10,12,13] although there is little mechanistic understanding of how CMV transcriptional activity can be controlled and enhanced. Engineering fundamental synthetic processes in mammalian cell factories such as HEK293 therefore remains a highly desirable objective.

hCMV-IE promoter (henceforth referred to as the CMV promoter) is a highly complex element comprising binding sites (transcription factor regulatory elements [TFREs]) for numerous ubiquitously expressed transcription factors (TFs).[14] This is not surprising considering that the promoter has evolved to function in a broad cell tropism.[15,16] However, promoter activity in any given host is regulated by a system-specific combination of interactions between the promoter's constituent TFREs and the cells repertoire of endogenous TFs.[17] Therefore, transcriptional activity of the CMV promoter is highly context-specific and cell type-dependent expression has been observed both in vivo[18,19] and in vitro[20,21]. With respect to the latter, we have for example, demonstrated that CMV-driven TGE in Chinese hamster ovary (CHO) cells was largely a function of transactivation mediated through just two discrete TFREs (NF-κB and CREB).[22] Further, the CMV promoter comprises binding sites of several transcriptional repressors such as YY1 — conferring on cytomegalovirus the ability to establish latent infection.[16,22–24] Accordingly, it is likely that the CMV promoter is fundamentally sub-optimal for use in unnatural, specific processes such as recombinant gene expression in HEK293 cells. Despite this, no previous studies have examined the specific CMV promoter interactions within the HEK293 transcriptional landscape that drive recombinant gene transcription, and the lack of this mechanistic information currently renders optimization of CMV-driven TGE in HEK293 cells intractable.

In this study, we identify the regulators of CMV-mediated TGE in HEK293 cells through mechanistic dissection of the CMV promoter. We performed an extensive bioinformatic analysis on the promoter's TFRE composition, coupled with a detailed in vitro comparative analysis of the relative influence of CMV component parts on gene expression to identify functional elements (TFRE sequences and *cis*-regulatory modules [CRMs]) that critically control promoter activity in HEK293 cells. We demonstrate, for the first time, that the wild-type CMV promoter can be re-engineered specifically for HEK293 cells to derive highly compact and transcriptionally efficient novel promoters with increased transcriptional activity.

2 | MATERIALS AND METHODS

2.1 | HEK and CHO cell cultures

Suspension-adapted HEK293 cells were provided by REGENXBIO and cultured in Dynamis medium (Thermo Fisher Scientific) supplemented with L-glutamine (Thermo Fisher Scientific). Expi293F cells (Thermo Fisher Scientific) were cultured in Expi293 Expression medium (Thermo Fisher Scientific). CHO-S cells (Thermo Fisher Scientific) were cultured in CD CHO medium (Thermo Fisher Scientific) supplemented with 8 mM L-glutamine. Cells were maintained in Erlenmeyer flasks (Corning) at 37°C, 140 rpm under 5% CO_2, 85% humidity and were sub-cultured every 3 to 4 days by seeding at 3×10^5 viable cells per mL. Cell viability and viable cell density were measured using a Vi-CELL XR (Beckman Coulter).

2.2 | Vector construction

pmaxGFP vector (Lonza) was utilized as a backbone. The CMV promoter and chimeric intron of pmaxGFP were deleted by digestion with BsrGI and KpnI, and replaced with a short DNA fragment containing EcoRI and HindIII cloning sites. A full-length hCMV-IE promoter (–550 to +48 relative to the TSS) was synthesized (Eurofins Genomics) and inserted directly upstream of the green fluorescent protein (GFP) open reading frame (ORF) of the promoter less vector backbone. A minimal CMV core promoter (–36 to +48 relative to the TSS) was also synthesized and inserted directly upstream of the GFP ORF. To create TFRE reporter plasmids, synthetic oligonucleotides containing 7 × repeat copies of the TFRE sequences in Table S1 were synthesized, PCR amplified (Q5 high-fidelity 2 × master mix; NEB), and purified (QIAquick PCR Purification kit; Qiagen). The PCR products were then digested, gel extracted (QIAquick Gel Extraction kit; Qiagen) and inserted into the cloning sites upstream of the CMV core promoter. Discrete regions of the CMV promoter sequence were PCR amplified and inserted upstream of the CMV core promoter. Mutated promoter constructs were synthesized and inserted upstream of the CMV core promoter. The CBh promoter was excised from pSpCas9(BB)-2A-GFP plasmid (Addgene) by digestion with KpnI and AgeI and inserted directly upstream of the GFP ORF. Clonally derived plasmids were purified using a QIAGEN Plasmid Plus kit (Qiagen). The sequence of all plasmid constructs was confirmed by restriction enzyme analysis and DNA sequencing (Eurofins Genomics).

2.3 | PEI-mediated transient transfection

One day before transfection, cells were sub-cultured in an Erlenmeyer flask, grown to 1×10^6 cells per mL and aliquots of 10 mL were added to each TubeSpin bioreactor tube (TPP). 8 μg of DNA and 24 μL of PEI MAX (1 mg mL^{-1}; Polysciences) were each pre-diluted in 150 μL of NaCl (150 mM; Polyplus-transfection), combined and incubated at room temperature for 4 min before being added into culture. Transfected cells were cultured for 48 h at 37°C, 230 rpm under 5% CO_2, 85% humidity.

2.4 | Measurement of recombinant GFP expression in vitro

GFP expression was quantified using a SpectraMax iD5 microplate reader (Molecular Devices) 48 h post-transfection. Prior to fluorescence read (excitation: 485 nm, emission: 535 nm), culture medium was removed by centrifugation at $200 \times g$ for 5 min. 1.5×10^6 viable cells were resuspended in 750 μL Dulbecco's phosphate-buffered saline (DPBS; Sigma) and then transferred to a 96-well microplate at 3×10^5 cells (150 μL) per well. To measure transfection efficiency, cells were analyzed using Attune Acoustic Focusing Cytometer (Thermo Fisher Scientific). Background fluorescence/absorbance was determined in cells transfected with a promoterless vector.

2.5 | In silico analysis of transcription factor regulatory elements

Genomatix Gene Regulation software (MatInspector Release 8.4 and MatBase Version 11.2; Precigen Bioinformatics Germany) was used to analyze the CMV promoter to find putative human TFREs. To capture all possible binding sites of different TF subtypes within the promoter, analysis was performed using the Individual Matrix function (rather than Matrix Families), with the Core Similarity set to 0.80 and the Matrix Similarity set to 'Optimized'. Cognate TF of each TFRE matrix (totaling 116; Table S2) was obtained from previously published studies as listed in MatBase. Selective mutation of a specific TFRE was performed by mutating at least of one the four highly conserved nucleotides (core sequence) defined in MatBase. Mutated sequence was subjected to the same analysis using the software to ensure that the specific TF binding site was removed, and that neither any overlapping TFRE was perturbed nor new TFRE was introduced.

2.6 | Analysis of HEK293 transcription factor expression

HEK293 cells were seeded at 1×10^6 viable cells per mL and cultured as described above. From day 3, cells were fed daily (1% v/v) with feed medium containing 130 g L^{-1} glucose, 29.23 g L^{-1} L-glutamine, 25 g L^{-1} arginine and 20 g L^{-1} serine. Total RNA was extracted from duplicate cultures during exponential (~5×10^6 cells per mL) and stationary phases (~1.6×10^7 cells per mL) of growth. For each sample, 3×10^6 viable cells were collected by centrifugation at $200 \times g$ for 5 min. Cell pellets were immediately resuspended in 300 μL of RNAprotect Cell Reagent and stored at −80°C. RNA-seq libraries were prepared and sequenced by GENEWIZ using an Illumina NovaSeq (Illumina). Galaxy (usegalaxy.org) and R software were used to analyze the RNA-seq data using Salmon alignment tool and human GRCh38 GTF and FASTA files from www.ensembl.org. A curated database of ~1600 human TFs was obtained from Lambert et al.[26] (see Table S3).

2.7 | Statistical analysis

Microsoft Excel 2016 (Microsoft) was used to analyze the difference between the means (GFP expression) of two promoter constructs. For samples comprising only HEK293 cells, analysis was performed using Student's *t*-test with p-value < 0.05 was considered significant. For samples comprising HEK293 and Expi293F cells, analysis was performed using Two-Factor ANOVA with Replication with p-value < 0.05 was considered significant.

3 | RESULTS

3.1 | In silico and in vitro identification of regulators of CMV promoter transcriptional activity

In order to identify potential regulatory elements in CMV capable of recombinant gene transactivation in HEK293 cells, we performed bioinformatic survey of (i) putative TFREs (binding sites) in the promoter, and (ii) the TF repertoire of HEK293 cells based on RNA-seq datasets. With regard to the latter, although gene expression analysis does not permit precise quantification of active TF levels, it provides useful information on general TF expression profile where genes with more than two transcripts per million (TPM) can be considered active.[27] Using the Genomatix search tool, 108 discrete TFREs from 74 TF families were identified in the CMV promoter at copy numbers ranging from one to six. However, the gene expression analysis indicated that 22% (24/108) of the TFREs' cognate TFs were not expressed in HEK293 cells (exponential phase $\log_2$ TPM < 1; Figure 1A). Further, as we wanted to identify key regulatory elements, we focused our search on TFs that exhibit gene expression activities in both exponential and stationary phases of culture (i.e., "context-specific" expression can extend beyond cell-type), thus eliminating an additional four TFs that were not expressed in the latter phase of culture — yielding 80 potential TFREs. We note that two TFREs may have identical or overlapping sequences within the CMV promoter (e.g., NF-κB and NFAT5). Table S2 lists the identified TFREs and their cognate TFs.

To minimize the TFRE pool for functional testing, we filtered out TFREs with substantially overlapping binding sites and selected two TFREs from each TF family — yielding a subset of 25 TFREs. Figure 1B shows the map of select TFREs in the CMV promoter. To measure the

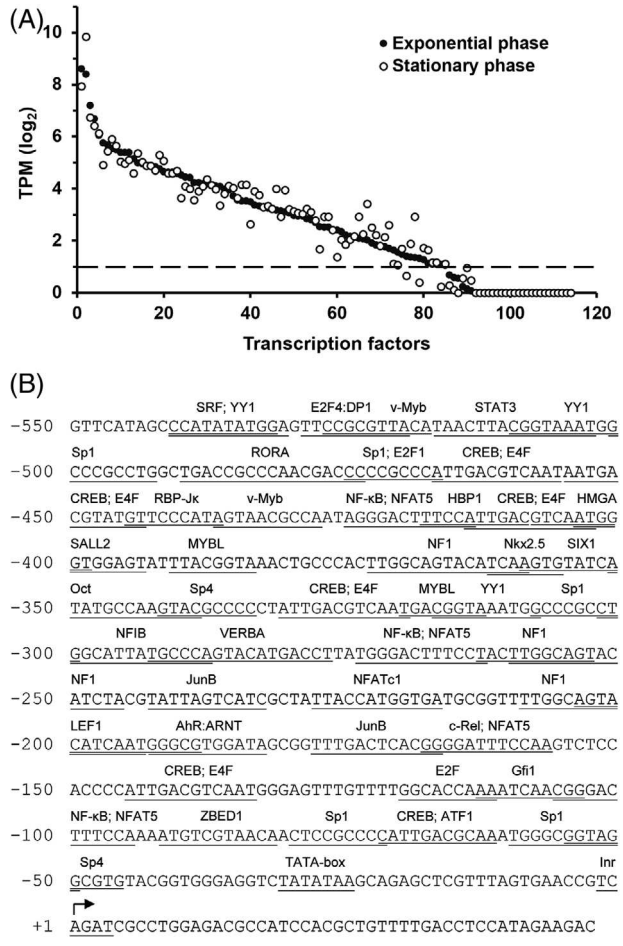

FIGURE 1 In silico identification of potential transcriptional regulators of CMV promoter activity. CMV promoter (−550 to +48 relative to the transcription start site; TSS) was surveyed for the presence of putative transcription factor regulatory elements (TFREs) using Genomatix software. One hundred and eight discrete TFREs identified in CMV promoter were subsequently analyzed for the presence of their cognate transcription factors (TFs) in HEK293 cells. (A) RNA-seq analysis of HEK293 cell transcriptome determined the relative gene expression level of TFs. Points represent the expression level (transcripts per million; TPM) of each TF sampled at exponential and stationary phases of culture. Genes with more than two transcripts per million ($\log_2$ TPM > 1) was considered as actively transcribed genes. (B) CMV promoter sequence with 25 selected TFREs for in vitro analysis. The TSS is indicated with an arrow

relative ability of TFREs to activate transcription of recombinant genes in HEK293 cells, we created a set of GFP reporter constructs that contained seven repeat copies of a specific TFRE in series, upstream of a minimal CMV core promoter (−36 to +48 relative to the TSS, containing a TATA box and an Inr motif) as previously described.[25,28] We note that the transcriptional output from a single TF binding site is often insufficient to drive detectable levels of recombinant gene expression. Optimized PEI-mediated transient transfection of plasmid DNA into suspension HEK293 cells yielded a transfection efficiency of ~ 94% with a cell viability of ~ 90% at 48 h post-transfection (measured using a vector harboring a CMV promoter). Additionally, preliminary experiments confirmed that GFP fluorescence intensity in HEK293 cell host is directly proportional to GFP mRNA levels post-transfection (Figure S1).[29] Measurement of GFP expression after transient transfection of HEK293 cells with each TFRE reporter plasmid is shown in Figure 2A. This analysis revealed eight TFREs with significantly increased expression (> 10-fold, $p < 0.01$) over basal expression from the minimal core promoter, that is, AhR:ARNT, CREB/ATF1, CREB/E4F, Sp1, ZBED1, JunB, c-Rel, and NF-κB. We note that in some instances, TFRE sequences with competing (overlapping) binding sites may be resolved by utilizing their consensus sequence (e.g., CREB and E4F; Figure 2A). Other TFRE reporter constructs displayed no obvious increase in GFP above core control level, suggesting alternative mechanisms of TF-mediated transcriptional activation or suboptimal TF binding sequences. To elucidate the latter, we tested the consensus sequence of MYBL1, Oct and E2F (selected based on a posteriori knowledge). This analysis revealed that the consensus sequences exhibited between 10 and 53-fold increase in expression over the core promoter (Figure 2A), indicating that the TFREs were essentially able to mediate activation of recombinant gene transcription in HEK293 cells using available TF activity.

In order to both confirm and further demonstrate the distinctive transcriptional landscape of HEK293 cells that influence CMV-mediated TGE, we tested the NF-κB and CREB sequences in CHO cells, as well as NF-κB p65 subunit sequence that is not present in the CMV promoter (Table S1). This analysis (Figure 2B) indicated that NF-κB and CREB were highly active in CHO cells (133% and 62% of CMV activity respectively), in line with our previous study that identified these two elements as key positive regulators of CHO cell-specific CMV promoter activity.[22] While NF-κB p65 subunit was five times more active than NF-κB in HEK293 cells, the activity was only one-fifth of that observed in CHO cells (we note that relative TFRE activities were not proportional to cognate TF expression levels). We therefore deduced that HEK293 cell-specific regulation of CMV promoter activity, in contrast to CHO, was a function of cooperative interactions amongst a broader range of TFREs.

3.2 | CMV promoter-mediated gene expression in HEK cells is regulated by proximal elements

In its natural context the CMV promoter can be divided into two modular components, the proximal and distal enhancers (Figure 3A).[14] Furthermore, TFREs often occur together in clusters as *cis*-regulatory modules (CRMs) where some elements may require interactions with adjacent or nearby TFRE partners in order to drive transcription.[30] To identify DNA sequence regions that are required for regulating gene expression in HEK293 cells, we inserted seven ~150-bp CRMs upstream of the CMV core in GFP reporter vectors (CRMs 1–7; Figure 3A). Figure 3B shows transient GFP reporter production from each CRM. CRMs from within the proximal enhancer sequence were generally more active than those from the distal, with CRM 1 alone yielding 67% of CMV's transcriptional activity. Analysis of the TFRE composition indicated that all positive regulators identified in the

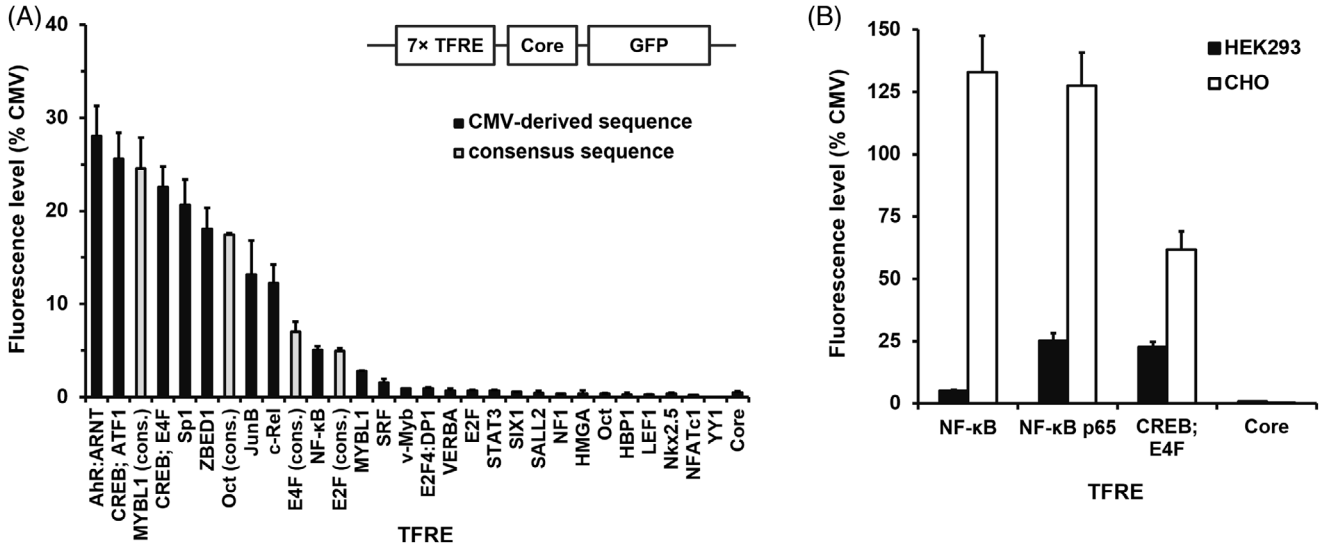

FIGURE 2 Identification of active transcription factor regulatory elements (TFREs). (A) TFRE sequence derived from the CMV promoter (black bars) or its consensus sequence (gray bars) was cloned in series (7 × copies) upstream of a minimal CMV core promoter in GFP-reporter vectors. HEK293 cells were transfected with each homotypic TFRE-reporter using polyethylenimine (PEI) and cultured in tube-spin bioreactors at 37°C. GFP expression was quantified 48 h post-transfection. (B) NF-κB p65 consensus sequence was cloned in series (7 × copies) upstream of a minimal CMV core promoter in GFP-reporter vectors and transfected into HEK293 and CHO-S cells alongside NF-κB and CREB/E4F constructs from A. Cells were cultured in tube-spin bioreactors at 37°C and GFP expression was quantified 48 h post-transfection. Data are expressed as a percentage with respect to the GFP expression of a vector containing the CMV promoter. Data shown are the mean value ± SD of two independent experiments each performed in duplicate

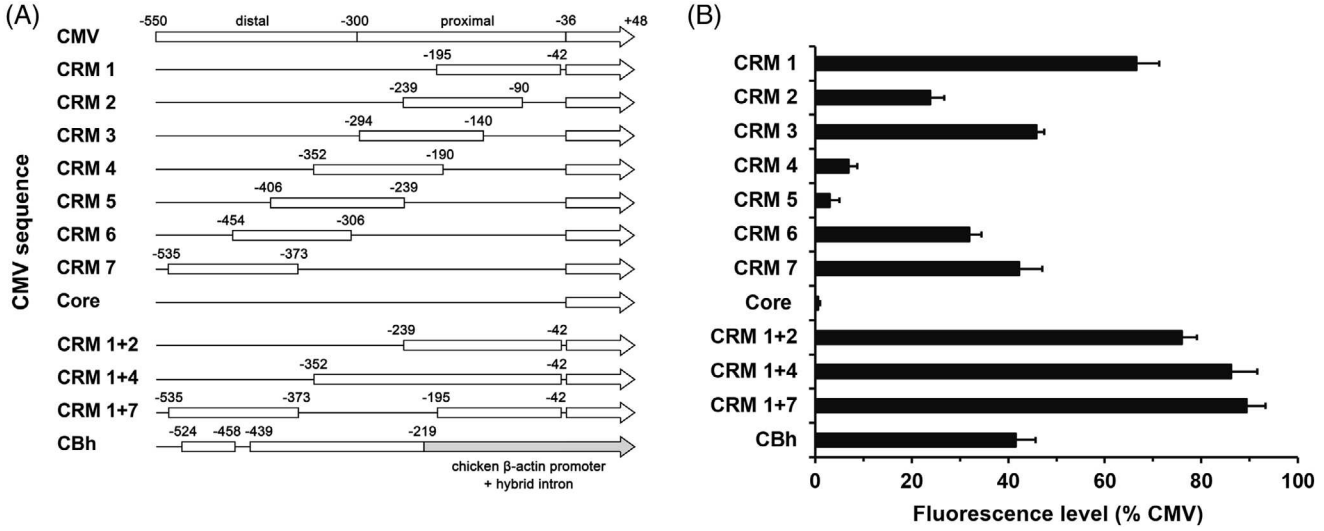

FIGURE 3 Relative transcriptional activity exhibited by CMV promoter structural elements. (A) The CMV promoter contains the proximal and distal enhancers and clusters of TFREs (*cis*-regulatory modules; CRMs). Each element was cloned upstream of a minimal CMV core promoter in GFP reporter plasmids while the CBh promoter (793 bp) was inserted directly upstream of the GFP open reading frame. (B) Reporter plasmids were transfected into HEK293 cells using PEI and cultured in tube-spin bioreactors at 37°C. GFP expression was quantified 48 h post-transfection. Data are expressed as a percentage with respect to the GFP expression of a vector containing the CMV promoter. Data shown are the mean value ± SD of two independent experiments each performed in duplicate

functional screen (Figure 2A) occurred in CRM 1, with one copy each of AhR:ARNT, CREB/ATF1, CREB/EF4, ZBED1, JunB, c-Rel, and NF-κB and two copies of Sp1. Moreover, multiple copies of CREB/E4F and Sp1 were present in CRMs 6 and 7, yielding 32% to 42% of CMV's activity. Conversely, CRMs from the middle of the CMV promoter (i.e., CRMs 4 and 5) did not display observable activity (≤ 7% of CMV). This was not unexpected considering that the constituent TFREs of these CRMs were mostly inactive in the functional screen.

Assembly of CRMs and comparison of their relative activity provided further analysis of individual CRM functions (Figure 3B). Combining CRM 1 and CRM 7 (67% and 42% of CMV activity respectively) yielded a promoter with only 90% CMV activity (CRM 1+7), suggesting a partially redundant function of the distal enhancer and/or spatial effects. On the other hand, adding inactive CRM 4 (7% CMV) onto CRM 1 significantly enhanced the transcriptional activity to 86% CMV (CRM 1+4; $p = 0.005$). This data implies a synergistic interaction of specific TFREs within the proximal enhancer. To expound this observation, we constructed an extended CRM 1 reporter vector (CRM 1+2) incorporating the NFATc1, NF1 and LEF1 binding sites. Even though these TFRE sequences were not active on their own (Figure 2A) the extended promoter displayed a 15% increase in activity ($p = 0.007$; Figure 3B), possibly via NFATc1–c-Rel interaction (−117 and −57 relative to the TSS respectively).[31] Critically, the data in Figure 3B reveal that CRM 2 exhibited 64% lower activity than CRM 1 despite a significant sequence overlap (Figure 3A), suggesting that either additional TFREs within the 5′ region of CRM 2 functioned to negatively regulate transcription, or essential regulators of CMV-mediated TGE in HEK293 were located in the 3′ region of CRM 1. With regard to the former, the apparent increase in GFP activity of CRM 1+2 compared to CRM 1 (see above) discounted the possibility of a specific transrepression effect of CRM 2. To substantiate the latter, we constructed a reporter vector utilizing a CMV enhancer/chicken β-actin hybrid (CBh) promoter (Figure 3A).[32] The promoter, comprising a practically complete CMV enhancer apart from CRM 1, exhibited only 41% of CMV's activity. Combining all observations made above, we inferred that (i) TFREs within the proximal enhancer functioned synergistically to drive transcription, and (ii) critical regulators of CMV promoter activity in HEK293 were located in the 3′ region of the proximal enhancer sequence (i.e., approximately −90 to −42 relative to the TSS).

3.3 | Sp1 binding sites near the TATA-box are essential for efficient CMV promoter-mediated gene expression in HEK293 cells

In order to specifically determine the key regulators of CMV-mediated gene expression in HEK293 cells we created CMV promoter variants with specific TFREs within −107 to −45 relative to the TSS "knocked-out". Proximal CMV (−300 to +48 relative to the TSS, ~ 84% CMV activity) rather than full-length CMV promoter was utilized for maximal impact of a single TFRE knock-out (i.e., minimal potential "noise" by other elements). Selective mutation was performed on the core sequence of a specific TFRE in order to disrupt the binding site without perturbing overlapping or introducing new TFREs (Figure 4A). Further, given the complexity of CMV promoter, we hypothesized that different HEK293 hosts (in different medium formulations) may potentially vary in their TF repertoires that could significantly influence CMV promoter regulation. In order to evaluate this, we determined the activity of the synthetic proximal CMV constructs in our standard HEK293 cell line as well as the commercially available Expi293F cell line, cultured in Dynamis medium and Expi293 Expression medium respectively. Measurement of GFP production after transient transfection of HEK293 and Expi293F cells with the knocked-out proximal CMV promoters is shown in Figure 4B. We observed that relative promoter activities were very similar in both cell lines, invalidating our hypothesis.

Our data (Figure 4B) also show that removal of NF-κB and ZBED1 binding sites, either individually or simultaneously, did not reduce GFP expression. This result is in line with the above finding (Figure 2A) that NF-κB had a very minimal activity in HEK293 cells but was not fully anticipated for the relatively active ZBED1. Utilizing the TFRE identification tool at a lower stringency, the in silico analysis identified a weak ZBED1 binding site at the mutated sequence (matrix similarity 0.734, optimal matrix threshold 0.76) suggesting that the ZBED1 mutation did not fully knock-out the TFRE. Removal of Sp1, CREB/ATF1, and Sp4 binding sites individually reduced promoter activity to ~ 62%, ~ 74%, and ~ 44% ($p < 0.008$) of that deriving from wild-type proximal CMV, respectively. Additionally, removal of the Sp1 site with CREB/ATF1 or Sp4 (mutations 3+4 and 3+6) led to further decrease in promoter activities. Critically, when the two Sp1 sites were simultaneously removed (mutations 3+5) GFP expression was reduced to the lowest level, that is, ~ 25% compared to the wild-type proximal CMV. No further reduction in promoter activity was observed when the Sp4 was mutated in conjunction with the two Sp1 sites — indicative of the Sp1's vital regulatory function. However, considering the relatively weak activity of the Sp1 homotypic promoter in Figure 2A, our data do not support the conclusion that Sp1 blocks (or any other TFRE) could support high transcriptional activity alone. We therefore deduced that the two Sp1 sites act as an upstream activator element[33] for CMV promoter-mediated transcription in HEK293 cells.

3.4 | Knock-out of repressor elements results in increased gene expression in CMV promoter variants

The above in silico analysis of regulation of CMV promoter activity by sequence elements (Figure 1) also identified two TFRE components that have previously been shown to negatively regulate transcription from the murine CMV-IE promoter in cytomegalovirus-infected mouse kidneys, YY1, and RBP-Jκ,[23] as well as Gfi1 where its overexpression has been shown to repress hCMV-IE promoter activity in mouse fibroblast cells.[34] We hypothesized that CMV promoter could be optimized for TGE by disrupting transrepression mediated by these TFREs. To evaluate the functional activity of these TFREs as regulators of CMV-mediated TGE in HEK293 cells, we synthesized CMV(-derived) promoters with repressor elements knocked-out and inserted them into GFP reporter vectors (Figure 5A,B). We note that the YY1 (three binding sites) and RBP-Jκ (one binding site) are located in the distal enhancer while the Gfi1 (one binding site) is located in the proximal enhancer. Additionally, previous studies suggested a fourth YY1 binding site at −343 to −353 relative to the TSS[14] which was identified as a weak binding sequence in this study (matrix similarity 0.889, optimal matrix threshold 0.94). GFP expression levels in HEK293 and Expi293F cells were measured 48 h post-transfection.

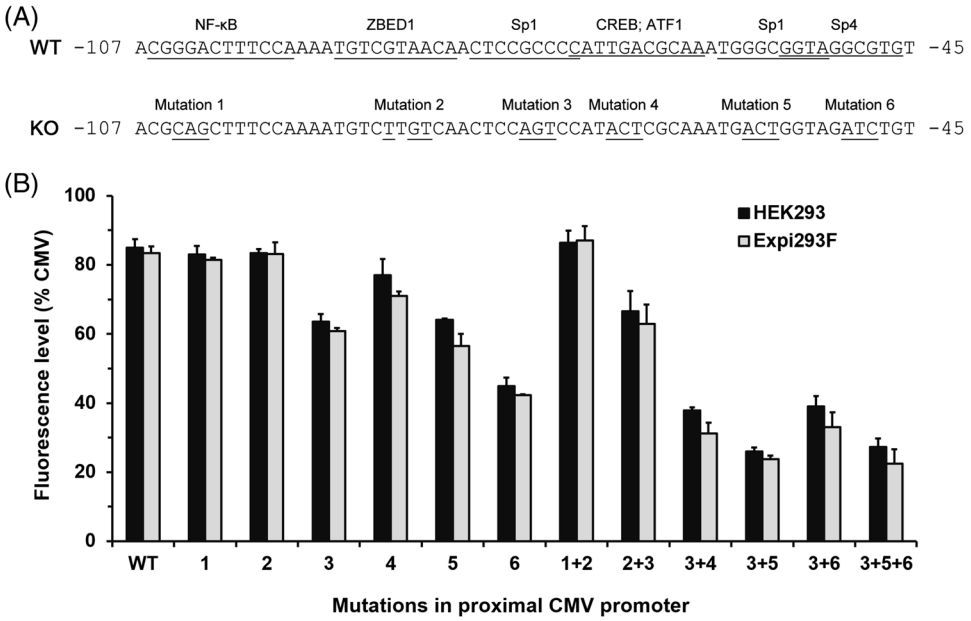

FIGURE 4 A proximal CMV promoter devoid of two Sp1 sites near the TATA box is unable to drive transcription in HEK293 cells. (A) Wild-type (WT) and mutated proximal CMV promoters (−300 to +48 relative to the TSS) with specific TFREs knocked-out (KO) were synthesized and cloned into GFP reporter vectors. Selective mutation was performed on a specific TFRE to disrupt the binding site without perturbing overlapping or introducing new TFREs. (B) The relative activity of each proximal CMV promoter construct was determined in HEK293 and Expi293F cells. Reporter plasmids were transfected into HEK293 cells using PEI and cultured in tube-spin bioreactors at 37°C. GFP expression was quantified 48 h post-transfection. Data are expressed as a percentage with respect to the GFP expression of a vector containing the full-length CMV promoter. Data shown are the mean value ± SD of two independent experiments each performed in duplicate

Removal of the repressor elements in full length CMV (promoters 1.01 and 1.02; Figure 5C) did not result in increased GFP expression compared to the wild-type control ($p > 0.603$), suggesting that the TFREs were not critical regulators affecting transcriptional activity under the conditions employed. This is possibly due to positive TF–TFRE interactions within the proximal enhancer decreasing the influence of distal enhancer-mediated processes (see above, Figure 3A,B). To further investigate the impact of proximal enhancer, we truncated the 5' region of the enhancer (promoter 2.01) which resulted in a ~ 13% decrease in transcriptional activity compared to the wild-type CMV (see Figure S2). In this promoter construct, removal of YY1 and RBP-Jκ binding sites increased the promoter activity by 12% (promoter 2.02; $p = 0.023$), indicating that the TFREs can act as negative regulators of CMV promoter in HEK293 cells. No additional increase was observed with further removal of Gfi1 (promoter 2.03) — this was not entirely unexpected considering that Gfi1 gene was lowly expressed ($\log_2$ TPM = 2.23) whereas YY1 and RBP-Jκ genes exhibited expression levels above the 90th percentile ($\log_2$ TPM = 5.51 and 5.53 respectively; Figure 1A).

To confirm the ability of YY1 and RBP-Jκ to mediate transrepression of recombinant gene transcription in HEK293 cells, we constructed CMV promoter variants with minimal proximal and distal enhancers containing only one site each of YY1, RBP-Jκ, and Gfi1 (promoters 3.01–3.03; see CRM 1+7 above). As anticipated, removal of the YY1 and RBP-Jκ binding sites increased the promoter activity by 11% ($p = 0.031$) compared to its non-mutated counterpart while no significant change was observed with further removal of Gfi1 (we note that removal of the YY1 binding alone increased the promoter activity by ~ 7% [$p = 0.053$]; data not shown). Moreover, these shorter promoter sequences displayed similar activities to promoters 2.01–2.03. In this regard, we assumed that the deletion of distal enhancer's 3' region effectively removed transrepression mediated by the fourth YY1 repressor motif (excluded in our in silico survey of CMV-constituent TFREs). Based on the above observations, we constructed promoter 4.01 that was devoid of repressor elements while retaining the active regions. This engineered CMV promoter displayed a ~ 14% increase in expression ($p = 0.005$; Figure S2) while being 25% smaller in size compared to the wild-type CMV. To illustrate the enhanced capability of the promoters in driving transcription, we calculated a "transcriptional efficiency" for each promoter as a function of transcriptional output per promoter length. This analysis indicates that promoters 3.03 and 4.01 were ~ 50% more transcriptionally efficient compared to the wild-type CMV promoter (Table S4). We conclude that the CMV promoter can be engineered for improved TGE in HEK293 via disruption of transrepression mediated by YY1 and RBP-Jκ and removal of redundant sequences.

4 | DISCUSSION

The vast majority of current HEK293 cell TGE systems utilize the CMV promoter for high-yield production of therapeutic proteins[6,8] and improved lentiviral and AAV expression vectors.[13,35,36] Our comparative transient expression analyses revealed that the CMV promoter

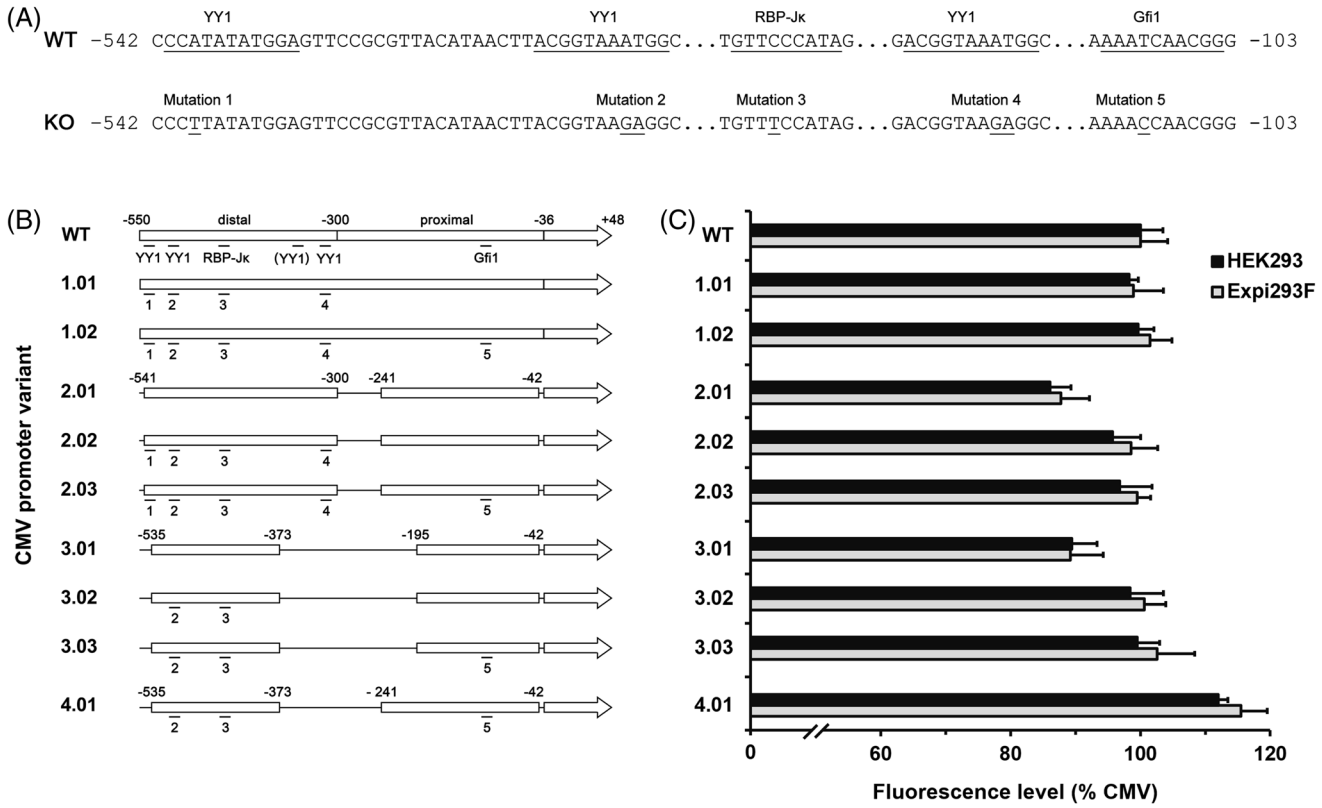

FIGURE 5 Removal of transrepression mediated by YY1 and RBP-Jκ and redundant sequences enhances CMV activity. (A) Selective mutation was performed on a specific TFRE to disrupt the binding site without perturbing overlapping or introducing new TFREs. (B) Wild-type (WT) CMV promoters (−550 to +48 relative to the TSS) and mutated CMV variants with specific TFREs knocked-out (KO) were synthesized and cloned into GFP reporter vectors. The locations of the repressor elements in CMV promoter are underlined. Numbers denote the corresponding TFRE knock-out in A. A fourth putative YYI binding site excluded by the TFRE analysis in this study is shown in bracket. (C) The relative activity of each promoter construct was determined in HEK293 and Expi293F cells. Reporter plasmids were transfected into HEK293 cells using PEI and cultured in tube-spin bioreactors at 37°C. GFP expression was quantified 48 h post-transfection. Data are expressed as a percentage with respect to the GFP expression of a vector containing the wild-type CMV promoter. Data shown are the mean value ± SD of two independent experiments each performed in duplicate

activity in HEK293 cells was a function of the promoter's various constituent TFREs including AhR:ARNT, CREB, E4F, Sp1, ZBED1, JunB, c-Rel, and NF-κB. This is a very significant and useful finding, as they form the basis of promoter engineering containing enhanced binding sites,[37] or can be directly utilized as modular building blocks to construct synthetic promoters de novo.[25,28] We further identified several sub-optimal TF binding sequences (MYBL1, Oct, E2F) which suggests an immense opportunity for maximizing CMV promoter's transcriptional output. Hundreds of TFRE motif sequence variants can be characterized simultaneously via in vitro use of high-throughput parallel screening methods, allowing determination of their optimal binding affinity. The major challenge with such functional tests is the difficulty in identifying TFREs underpinning the more complex regulation governing synergistic transactivation[30,31] which would require intricate screens of TFRE motif pairs with position-sensitive function. However, this limitation can be circumvented by using the TF decoy technology developed in this laboratory[38] to inhibit specific TFRE(s) within the CMV promoter architecture, obviating the need to characterize spatial effects between two TFRE motif pairs. Furthermore, this work, in effect, generated a novel library of promoter sequences (Figure 2) to control gene expression over a wide range. These highly compact promoters could be utilized in multigene vectors to give predictable stoichiometries (e.g., optimization of monoclonal antibody heavy chain to light chain ratio),[39] especially with non-overlapping sequences to avoid homologous recombination-mediated silencing (see below).

Bioinformatic analysis on the CMV promoter sequence indicated that Sp1 family is predominant in the promoter, in line with the notion that the element is essential for prevention of de novo methylation of CpG islands.[40] Importantly, our results show that each of the two Spl binding sites near the TATA box contributes to full activation of the CMV promoter in HEK293 cells — resembling the previous report in which mutation of these Sp1 binding sites caused inefficient CMV promoter transcription and cytomegalovirus replication in human fibroblast cells.[41] Similar transcription activation mechanism had been reported for the simian virus 40 (SV40) promoter in which Sp1 binding to its cognate sequences upstream of the TATA box enhanced the activity of RNA polymerase II.[42,43] Indeed, analyses of synthetic core promoters indicated that Sp1 binding sites, when placed

upstream of an Inr and/or TATA box, acted as an upstream activator element for efficient transcription initiation in vitro[33] and in HEK293 cells.[44] Nevertheless, our previous studies[22,25,38] as well as data in Figure 2B showed that CHO cells were able to drive efficient recombinant gene transcription in the absence of such upstream activator elements, illustrating that engineering strategies to improve CMV promoter activity have to be cell-type specific for maximum efficacy. We further conjecture that a Spl-dependent upstream activator element is a design prerequisite for construction of strong synthetic/hybrid promoters for HEK293 cells. This is in contrast to modulation at the translational level (e.g., engineering of 5'UTR elements) that is generally not cell-type dependent.[45]

Another important outcome of this study is the identification of negatively acting cellular TFs, and that a substantial proportion of the CMV sequence may be functionally redundant for recombinant gene expression in HEK293 cells. Specifically, our results showed that YY1 and RBP-Jκ-mediated transrepression of the CMV promoter could be removed by designing engineered CMV constructs with inactive cognate binding sites. It is worth noting previous studies have also shown that ERF (Ets-2 Repressor Factor) was able to repress the CMV promoter by binding to the 21 bp repeat motifs overlapping YY1 and Sp1 within the distal enhancer (see Figure S3),[24] and that the ERF gene was highly expressed in HEK293 cells ($\log_2$ TPM = 4.91; Table S2). Therefore, we postulate that the deletion of 3' region of the distal enhancer (promoters 3.03 and 4.01) effectively removed the YY1 as well as an ERF binding site, permitting a more defined, improved regulation of recombinant transcriptional activity and with relatively small promoter size. The engineered promoters may further confer additional advantages in dynamic bioprocess conditions in respond to changes in cellular transcriptional landscape. For example, differential gene expression analysis on the HEK293 transcriptomic data showed that RBP-Jκ was (slightly) upregulated from the mid-exponential to early-stationary phase ($\log_2$ fold-change = 0.362, p-adj = 0.0012), suggesting that the positive impact of RBP-Jκ knock-out would be more pronounced in long-term, fed-batch production processes for example. It also interesting to note that our bioinformatic analysis indicated that the CMV promoter did not contain the binding sites of TFs associated with the unfolded protein response (ATF4, ATF6, eIF2α, and XBP1), suggesting that CMV promoter activity is not affected by cellular stress that may result from recombinant gene overexpression.

Lastly, the data presented in this study may offer benefits to systems beyond TGE. For instance, long-term stable expression can be compromised by the occurrences of sequence features such as repeat elements (homologous recombination-mediated silencing)[46] and CpG islands (methylation-mediated silencing).[47] With regard to the former, the CMV promoter contains two copies of 21 bp repeat motif in the distal enhancer as mentioned above. Promoters 3.03 and 4.01 indirectly removed one of these repeat elements, therefore avoiding potential genetic homologous recombination events associated with gene deletion. With regard to the latter, it may be possible to reduce the number of CpG dinucleotides within the CMV promoter by mutating TFREs with no/low activities, thus minimizing the formation of methylation-mediated epigenetic silencing linked to production instability. Minimal CpG dinucleotides is also a desirable feature for gene therapy vectors in which CpG motifs have immunostimulatory effects (e.g., promoter 3.03 contains 20% less CpG dinucleotides compared to the wild-type CMV promoter; Table S4).[48] Accordingly, it should now be possible to rationally design synthetic CMV promoter variants in order to equip HEK293 cells with new transcriptional machinery optimally suited for a specific intended purpose. We anticipate that similar approaches can be used to deconstruct and reconstruct other promoters for optimal functionalities in particular cell types.

ACKNOWLEDGMENTS

This study was supported by REGENXBIO Inc., U.S.A. The authors thank Adrian Bourke (University of Sheffield) for technical assistance in RNA-seq analysis, Selase Enuameh (REGENXBIO) and Jie Li (REGENXBIO) for useful discussions.

CONFLICT OF INTERESTS

The authors have a patent application filed based on the work in this paper.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

Yusuf B. Johari https://orcid.org/0000-0001-9933-5764
Joseph M. Scarrott https://orcid.org/0000-0002-6046-7687
Thilo H. Pohle https://orcid.org/0000-0003-4437-3231
Adam J. Brown https://orcid.org/0000-0002-3290-4560
David C. James https://orcid.org/0000-0002-1697-151X

REFERENCES

1. Dumont, J., Euwart, D., Mei, B., Estes, S., & Kshirsagar, R. (2016). Human cell lines for biopharmaceutical manufacturing: History, status, and future perspectives. *Critical Reviews in Biotechnology*, *36*, 1110–1122.
2. Robert, M. A., Chahal, P. S., Audy, A., Kamen, A., Gilbert, R., & Gaillet, B. (2017). Manufacturing of recombinant adeno-associated viruses using mammalian expression platforms. *Biotechnology Journal*, *12*, 1600193.
3. Gutiérrez-Granados, S., Cervera, L., Kamen, A. A., & Gòdia, F. (2018). Advancements in mammalian cell transient gene expression (TGE) technology for accelerated production of biologics. *Critical Reviews in Biotechnology*, *38*, 918–940.
4. Abaandou, L., Quan, D., & Shiloach, J. (2021). Affecting HEK293 cell growth and production performance by modifying the expression of specific genes. *Cells*, *10*, 1667.
5. Backliwal, G., Hildinger, M., Hasija, V., & Wurm, F. M. (2008). High-density transfection with HEK-293 cells allows doubling of transient titers and removes need for a priori DNA complex formation with PEI. *Biotechnology and Bioengineering*, *99*, 721–727.
6. Backliwal, G., Hildinger, M., Chenuet, S., Wulhfard, S., De Jesus, M., & Wurm, F. M. (2008). Rational vector design and multi-pathway modulation of HEK 293E cells yield recombinant antibody titers exceeding 1 g/l by transient transfection under serum-free conditions. *Nucleic Acids Research*, *36*, e96.
7. Arena, T. A., Chou, B., Harms, P. D., & Wong, A. W. (2019). An anti-apoptotic HEK293 cell line provides a robust and high titer platform for transient protein expression in bioreactors. *MAbs*, *11*, 977–986.

8. Swiech, K., Kamen, A., Ansorge, S., Durocher, Y., Picanço-Castro, V., Russo-Carbolante, E. M., Neto, M. S., & Covas, D. T. (2011). Transient transfection of serum-free suspension HEK 293 cell culture for efficient production of human rFVIII. *Bmc Biotechnology [Electronic Resource]*, *11*, 114.
9. Lin, C. Y., Huang, Z., Wen, W., Wu, A., Wang, C., & Niu, L. (2015). Enhancing protein expression in HEK-293 cells by lowering culture temperature. *Plos One*, *10*, e0123562.
10. Kiszel, P., Fiesel, S., Voit, S., Waechtler, B., Meier, T., Oelschlaegel, T., Schraeml, M., & Engel, A. M. (2019). Transient gene expression using valproic acid in combination with co-transfection of SV40 large T antigen and human p21 CIP /p27 KIP. *Biotechnology progress*, *35*, e2786.
11. Zhao, H., Lee, K. J., Daris, M., Lin, Y., Wolfe, T., Sheng, J., Plewa, C., Wang, S., & Meisen, W. H. (2020). Creation of a high-yield aav vector production platform in suspension cells using a design-of-experiment approach. *Molecular Therapy – Methods & Clinical Development*, *18*, 312–320.
12. Román, R., Miret, J., Scalia, F., Casablancas, A., Lecina, M., & Cairó, J. J. (2016). Enhancing heterologous protein expression and secretion in HEK293 cells by means of combination of CMV promoter and IFNÎ±2 signal peptide. *Journal of Biotechnology*, *239*, 57–60.
13. Allen, J. M., Halbert, C. L., & Miller, A. D. (2000). Improved adeno-associated virus vector production with transfection of a single helper adenovirus gene, E4orf6. *Molecular Therapy*, *1*, 88–95.
14. Stinski, M. F., & Isomura, H. (2008). Role of the cytomegalovirus major immediate early enhancer in acute infection and reactivation from latency. *Medical Microbiology and Immunology*, *197*, 223–231.
15. Sinzger, C., Digel, M., & Jahn, G. (2008). Cytomegalovirus cell tropism. *Current Topics in Microbiology and Immunology*, *325*, 63–83.
16. Forte, E., Zhang, Z., Thorp, E. B., & Hummel, M. (2020). Cytomegalovirus latency and reactivation: An intricate interplay with the host immune response. *Frontiers in Cellular and Infection Microbiology*, *10*, 130.
17. Coulon, A., Chow, C. C., Singer, R. H., & Larson, D. R. (2013). Eukaryotic transcriptional dynamics: From single molecules to cell populations. *Nature Reviews Genetics*, *14*, 572–584.
18. Mella-Alvarado, V., Gautier, A., Le Gac, F., & Lareyre, J. - J. (2013). Tissue and cell-specific transcriptional activity of the human cytomegalovirus immediate early gene promoter (UL123) in zebrafish. *Gene Expression Patterns*, *13*, 91–103.
19. Vasey, D., Lillico, S., Sang, H., King, T., & Whitelaw, C. (2009). CMV enhancer-promoter is preferentially active in exocrine cells in vivo. *Transgenic Research*, *18*, 309–314.
20. Qin, J. Y., Zhang, L., Clift, K. L., Hulur, I., Xiang, A. P., Ren, B. Z., & Lahn, B. T. (2010). Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *Plos One*, *5*, e10611.
21. Xia, W., Bringmann, P., McClary, J., Jones, P. P., Manzana, W., Zhu, Y., Wang, S., Liu, Y., Harvey, S., Madlansacay, M. R., McLean, K., Rosser, M. P., MacRobbie, J., Olsen, C. L., & Cobb, R. R. (2006). High levels of protein expression using different mammalian CMV promoters in several cell lines. *Protein Expression and Purification*, *45*, 115–124.
22. Brown, A. J., Sweeney, B., Mainwaring, D. O., & James, D. C. (2015). NF-Î°B, CRE and YY1 elements are key functional regulators of CMV promoter-driven transient gene expression in CHO cells. *Biotechnology Journal*, *10*, 1019–1028.
23. Liu, X. F., Yan, S., Abecassis, M., & Hummel, M. (2008). Establishment of murine cytomegalovirus latency in vivo is associated with changes in histone modifications and recruitment of transcriptional repressors to the major immediate-early promoter. *Journal of Virology*, *82*, 10922–10931.
24. Bain, M., Mendelson, M., & Sinclair, J. (2003). Ets-2 Repressor Factor (ERF) mediates repression of the human cytomegalovirus major immediate-early promoter in undifferentiated non-permissive cells. *Journal of General Virology*, *84*, 41–49.
25. Johari, Y. B., Brown, A. J., Alves, C. S., Zhou, Y., Wright, C. M., Estes, S. D., Kshirsagar, R., & James, D. C. (2019). CHO genome mining for synthetic promoter design. *Journal of Biotechnology*, *294*, 1–13.
26. Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T. R., & Weirauch, M. T. (2018). The human transcription factors. *Cell*, *172*, 650–665.
27. Wagner, G. P., Kin, K., & Lynch, V. J. (2013). A model based criterion for gene expression calls using RNA-seq data. *Theory in Biosciences*, *132*, 159–164.
28. Johari, Y. B., Mercer, A. C., Liu, Y., Brown, A. J., & James, D. C. (2021). Design of synthetic promoters for controlled expression of therapeutic genes in retinal pigment epithelial cells. *Biotechnology and Bioengineering*, *118*, 2001–2015.
29. Ferreira, J. P., Peacock, R. W., Lawhorn, I. E., & Wang, C. L. (2011). Modulating ectopic gene expression levels by using retroviral vectors equipped with synthetic promoters. *Systems and Synthetic Biology*, *5*, 131.
30. Hardison, R., & Taylor, J. (2012). Genomic approaches towards finding cis-regulatory modules in animals. *Nature Reviews Genetics*, *13*, 469–483.
31. Pham, L. V., Tamayo, A. T., Yoshimura, L. C., Lin-Lee, Y. C., & Ford, R. J. (2005). Constitutive NF-Î°B and NFAT activation in aggressive B-cell lymphomas synergistically activates the CD154 gene and maintains lymphoma cell survival. *Blood*, *106*, 3940–3947.
32. Gray, S. J., Foti, S. B., Schwartz, J. W., Bachaboina, L., Taylor-Blake, B., Coleman, J., Ehlers, M. D., Zylka, M. J., McCown, T. J., & Samulski, R. J. (2011). Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors. *Human Gene Therapy*, *22*, 1143–1153.
33. Smale, S. T., Schmidt, M. C., Berk, A. J., & Baltimore, D. (1990). Transcriptional activation by Sp1 as directed through TATA or initiator: Specific requirement for mammalian transcription factor IID. *Proceedings of the National Academy of Sciences of the United States of America*, *87*, 4509–4513.
34. Zweidler-McKay, P. A., Grimes, H. L., Flubacher, M. M., & Tsichlis, P. N. (1996). Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. *Molecular and Cellular Biology*, *16*, 4024–4034.
35. Vink, C. A., Counsell, J. R., Perocheau, D. P., Karda, R., Buckley, S. M. K., Brugman, M. H., Galla, M., Schambach, A., McKay, T. R., Waddington, S. N., & Howe, S. J. (2017). Eliminating HIV-1 packaging sequences from lentiviral vector proviruses enhances safety and expedites gene transfer for gene therapy. *Molecular Therapy*, *25*, 1790–1804.
36. Wang, Z., Cheng, F., Engelhardt, J. F., Yan, Z., & Qiu, J. (2018). Development of a novel recombinant adeno-associated virus production system using human bocavirus 1 helper genes. *Molecular Therapy—Methods & Clinical Development*, *11*, 40–41.
37. Nong, L., Zhang, Y., Duan, Y., Hu, S., Lin, Y., & Liang, S. (2020). Engineering the regulatory site of the catalase promoter for improved heterologous protein production in Pichia pastoris. *Biotechnology Letters*, *42*, 2703–2709.
38. Brown, A. J., Mainwaring, D. O., Sweeney, B., & James, D. C. (2013). Block decoys: Transcription-factor decoys designed for in vitro gene regulation studies. *Analytical Biochemistry*, *443*, 205–210.
39. Patel, Y. D., Brown, A. J., Zhu, J., Rosignoli, G., Gibson, S. J., Hatton, D., & James, D. C. (2021). Control of multigene expression stoichiometry in mammalian cells using synthetic promoters. *ACS Synthetic Biology*, *10*, 1155–1165.
40. Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Names, A., Temper, V., Razin, A., & Cedar, H. (1994). Spl elements protect a CpG island from de novo methylation. *Nature*, *371*, 435–438.
41. Isomura, H., Stinski, M. F., Kudoh, A., Daikoku, T., Shirata, N., & Tsurumi, T. (2005). Two Sp1/Sp3 binding sites in the major immediate-early

proximal enhancer of human cytomegalovirus have a significant role in viral replication. *Journal of Virology*, *79*, 9597–9607.
42. Dynan, W. S., & Tjian, R. (1983). The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell*, *35*, 79–87.
43. Vigneron, M., Barrera-Saldana, H. A., Baty, D., Everett, R. E., & Chambon, P. (1984). Effect of the 21-bp repeat upstream element on in vitro transcription from the early and late SV40 promoters. *Embo Journal*, *3*, 2373–2382.
44. Emami, K. H., Navarre, W. W., & Smale, S. T. (1995). Core promoter specificities of the Sp1 and VP16 transcriptional activation domains. *Molecular and Cellular Biology*, *15*, 5906–5916.
45. Eisenhut, P., Mebrahtu, A., Barzadd, M. M., Thalén, N., Klanert, G., Weinguny, M., Sandegren, A., Su, C., Hatton, D., Borth, N., & Rockberg, J. (2020). Systematic use of synthetic 5′-UTR RNA structures to tune protein translation improves yield and quality of complex proteins in mammalian cell factories. *Nucleic Acids Research*, *48*, e119.
46. Jasin, M., & Rothstein, R. (2013). Repair of strand breaks by homologous recombination. *Cold Spring Harbor perspectives in biology*, *5*, a012740.
47. Kim, M., O'Callaghan, P. M., Droms, K. A., & James, D. C. (2011). A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies. *Biotechnology and Bioengineering*, *108*, 2434–2446.
48. Bessis, N., GarciaCozar, F. J., & Boissier, M. C. (2004). Immune responses to gene therapy vectors: Influence on vector function and effector mechanisms. *Gene Therapy*, *11*, S10–S17.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Johari, Y. B., Scarrott, J. M., Pohle, T. H., Liu, P., Mayer, A., Brown, A. J., & James, D. C. (2022). Engineering of the CMV promoter for controlled expression of recombinant genes in HEK293 cells. *Biotechnology Journal*, 1–11. https://doi.org/10.1002/biot.202200062