Citation: Molecular Therapy—Nucleic Acids (2013) 2, e118; doi:10.1038/mtna.2013.47

www.nature.com/mtna

Genomic *cis*-acting Sequences Improve Expression and Establishment of a Nonviral Vector

Claudia Hagedorn[1], Michael N Antoniou[2], and Hans J Lipps[1]

The vector pEPI was the first nonviral and episomally replicating vector. Its functional element is an expression unit linked to a chromosomal scaffold/matrix attached region (S/MAR). The vector replicates autonomously with low copy number in various cell lines, is mitotically stable in the absence of selection over hundreds of generations, and was successfully used for the efficient generation of genetically modified pigs. Since it is assumed that establishment of the vector is a stochastic event and strongly depends on the nuclear compartment it reaches after transfection, it is of great interest to identify genomic sequences that guide DNA sequences into certain nuclear compartments. Here we inserted genomic *cis*-acting sequences into pEPI and examined their impact on transgene expression, long-term stability, and vector establishment. We demonstrated that a ubiquitous chromatin-opening element (UCOE) mediated enhanced transgene expression, while an insulator sequence (cHS4) increased establishment efficiency, presumably via an additional interaction with the nuclear matrix. Thus, besides being a promising alternative to currently used viral vectors in gene therapeutic approaches, pEPI may also serve as a tool to study nuclear compartmentalization; identification of genomic *cis*-acting sequences that are involved in nuclear organization will contribute to our understanding of the interplay between transgene expression, plasmid establishment, and nuclear architecture.

Molecular Therapy—Nucleic Acids (2013) **2**, e118; doi:10.1038/mtna.2013.47; published online 27 August 2013
Subject Category: Gene insertion, deletion & modification Gene vectors

Introduction

The first nonviral and episomally replicating vector, pEPI, was constructed in the late 1990s and is based solely on chromosomal elements.[1] The functional element of this vector type is an expression unit linked to a chromosomal scaffold/matrix attached region (S/MAR).[2,3] Deletion of this transcription unit or insertion of a transcription termination site upstream of the S/MAR results in integration or vector loss, suggesting that an ongoing transcription into the S/MAR is essential for its episomal replication and maintenance.[3,4] The S/MAR-based pEPI vector replicates autonomously with low copy number in all cells tested, including primary cells, was successfully used for the efficient generation of genetically-modified pigs, and is mitotically stable in the absence of selection over hundreds of generations.[5–7] Variations of the pEPI vector harboring a human β-globin gene transcription unit including the locus control region was fully functional in erythroid cells demonstrating its gene therapy potential.[8]

In stably established cell clones, pEPI replicates once per cell cycle during early S-phase[9] and associates with early replicating foci.[7] However, a variation of transgene expression is always observed.[6] In flow cytometric analyses 30–60% of transfected cells show detectable expression of the transgene EGFP, although the vector is retained as an episome in all cells. Moreover, the frequency of which the vector establishes in transfected cells is slow: about 95% of all vector molecules that reach the nucleus after transfection are lost during the following cell divisions.[7] It is assumed that vector establishment is a stochastic event and depends on the nuclear compartment the vector reaches,[7] similar to all other eukaryotic autonomous replicons.[10–12] Since nonviral S/MAR containing vectors represent an exciting alternative to virus-based vectors, considerable efforts have been undertaken for their improvement (reviewed in ref. 13). Here we focus on chromosomal, *cis*-acting elements that may be capable to improve efficiency and transgene expression of S/MAR-based vectors. We introduced a novel class of genetic element into pEPI-EGFP recently designated as a ubiquitous chromatin-opening elements (UCOE). These elements are responsible for establishing a transcriptionally competent open chromatin structure at regions of the genome that consist only of ubiquitously expressed housekeeping genes. The UCOE spanning the dual, divergently transcribed promoters of the human *TBP-PSMB1* and *HNRPA2B1-CBX3* housekeeping genes was shown to mediate stable expression of transgenes even when integrated in heterochromatic regions and reduce variegation effects.[14] In combination with the immediate early promoter/enhancer of the human cytomegalovirus (hCMV), the UCOE derived from the *HNRPA2B1-CBX3* locus (A2UCOE) mediates improved levels of expression and proportion of cells expressing the transgene at detectable levels.[15] Another *cis*-acting element we examined is the 5′-HS4 chicken-β-globin insulator (cHS4), which has

[1]Institute of Cell Biology, University of Witten/Herdecke, ZBAF, Witten, Germany; [2]Department of Medical and Molecular Genetics, Gene Expression and Therapy Group, King's College London School of Medicine, Guy's Hospital, London, UK. Correspondence: Claudia Hagedorn, Institute of Cell Biology/ZBAF, University of Witten/Herdecke, Stockumer Str. 10, D-58453 Witten, Germany. E-mail: claudia.hagedorn@uni-wh.de
Keywords: episomal vector; establishment efficiency; gene therapy; mitotic stability; nuclear matrix
Received 23 May 2013; accepted 21 June 2013; advance online publication 27 August 2013. doi:10.1038/mtna.2013.47

been considered as an ideal candidate for use in gene transfer applications due to its barrier and enhancer-blocking function.[16] Two initial studies employed the cHS4 1.2 kb insulator sequence in gammaretroviral vectors flanking the transgene and demonstrated a significantly increased fraction of cells expressing the respective transgenes.[17,18] Following these initial reports, a number of studies have been conducted using retroviral vectors flanked with either the whole or parts of the 1.2 kb cHS4 insulator, including gammaretroviral vectors and lentiviral vectors. Overall, it was demonstrated that inclusion of the 1.2 kb cHS4 fragment increased the probability and/or stability of vector transgene expression and reduced position-effect variegation.[17–20] Besides border and enhancer-blocking functions, there is also evidence that suggests insulators are involved in chromatin organization and can anchor active domains to the nuclear periphery, close to the nuclear pore.[21–24]

Within this study both elements, A2UCOE and cHS4, were inserted into the prototype pEPI-EGFP, and plasmid behavior (episomal replication, establishment efficiency, long-term maintenance, and transgene expression) was examined. We demonstrate here that both elements have a remarkable impact on transgene expression and establishment efficiency of S/MAR based vectors, without interfering with the episomal behavior of S/MAR vectors; and may therefore be suitable for further studies with gene therapeutic focus.

Results

pEPI-UCOE and pEPI-HS4opp are maintained episomally and undergo cellular replication

Due to its low establishment efficiency and transgene expression, the *cis*-acting elements A2UCOE and cHS4 were introduced into the episomal vector pEPI-EGFP (**Figure 1a**) in an effort to improve both establishment efficiency (number of cells in which pEPI is stably maintained) and transgene expression. The A2UCOE element was inserted upstream of the CMV promoter resulting in pEPI-UCOE (**Figure 1b**), while the cHS4 element was inserted downstream and in opposite direction of the S/MAR element,[25] resulting in pEPI-HS4opp (**Figure 1c**). Episomal maintenance of these pEPI derivates was ascertained by Southern blot analysis using DNA from HIRT extracts digested with a restriction enzyme that cuts the plasmid only once and either *Mbo*I or *Dpn*I, respectively. Complete digestion with *Mbo*I, which is inhibited by *dam* DNA methylation, indicates cellular replication and was detected for all constructs (**Figure 1d**). In no case, we observed hybridization to chromosomal DNA.

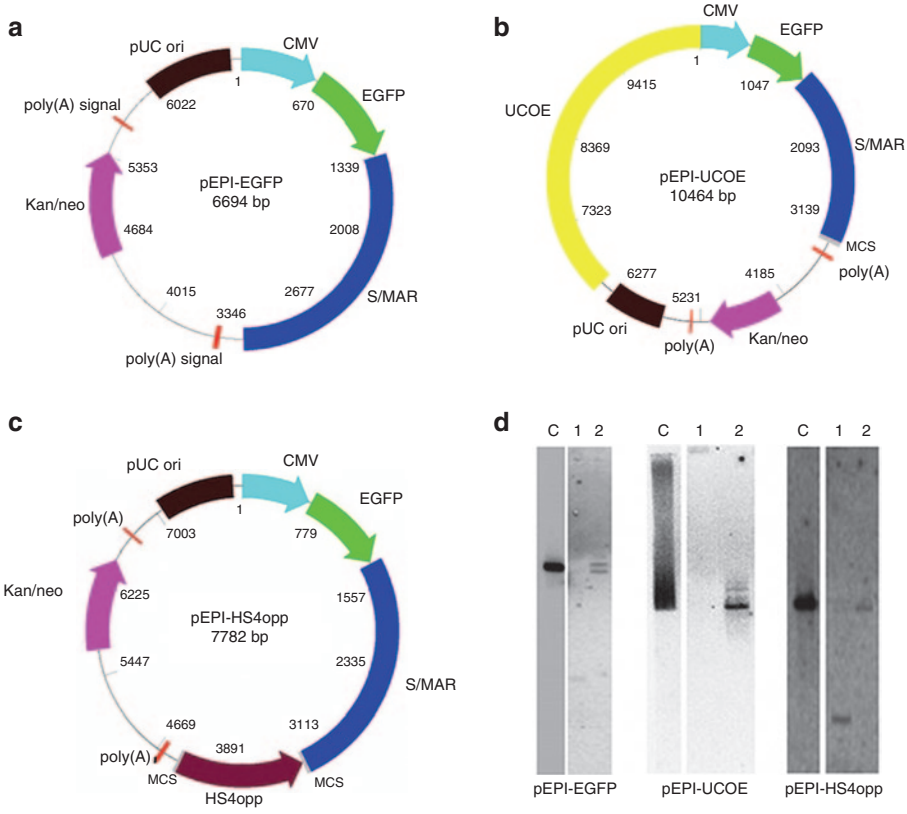

Figure 1 pEPI-EGFP and derivatives. (**a**) Original episomal vector pEPI-EGFP harboring an active transcription unit running into the S/MAR sequence. (**b**) The ubiquitous chromatin-opening element (UCOE) from the *HNRPA2B1-CBX3* locus was introduced upstream the CMV promoter resulting in pEPI-UCOE. (**c**) The chicken β-globin insulator sequence cHS4 was introduced downstream and in opposite direction to the S/MAR resulting in pEPI-HS4opp. (**d**) Episomal maintenance and cellular replication was ascertained in Southern blot analysis. Complete digestion with *Mbo*I indicates cellular replication (lane 1), as shown for all three pEPI derivatives. C: Plasmid DNA as control; 1: DNA digested with *Mbo*I and plasmid single cutter; 2: DNA digested with *Dpn*I and plasmid single cutter. In no case, we detected hybridization to chromosomal DNA.

A2UCOE and cHS4 increase transgene expression and establishment of the episomal vector pEPI

CHO cells were transfected with pEPI-EGFP, pEPI-UCOE, and pEPI-HS4opp and plasmid establishment was determined using a colony-forming assay as previously described.[26] Establishment of the original pEPI-EGFP was compared with its derivates, pEPI-UCOE and pEPI-HS4opp (**Figure 2a**). The original pEPI-EGFP vector and pEPI-UCOE established episomally with a similar efficiency (pEPI-EGFP: 9.5 ± 7.25%; pEPI-UCOE: 11.2 ± 9.04%). However, for pEPI-HS4opp establishment efficiency was significantly increased (25.3 ± 13.98%; $P = 0.028$). After an initial selection period (14 days, 400 µg/ml G418), cells were cultured in the absence of the antibiotic and transgene expression (number of EGFP expressing cells and mean fluorescent intensity (FI)) was determined in flow cytometry analysis over a time period of at least 60 days. In a mixed population, all pEPI plasmids were stably maintained in the absence of drug selective pressure (**Figure 2b**,**c**). In addition, a mixed pEPI-UCOE population exhibited not only an increased fraction of EGFP-expressing cells (**Figure 2b**; day 20: pEPI-EGFP 9.5 ± 4.8%, pEPI-UCOE 53.2 ± 21.2%, $P = 0.025$; day 60: pEPI-EGFP 13.7 ± 10.5%, pEPI-UCOE 53% ± 23.2, $P = 0.052$), but also a significantly increased expression level (FI) when compared with pEPI-EGFP and pEPI-HS4opp (**Figure 2c**; day 20: pEPI-EGFP 305.5 ± 2.7, pEPI-UCOE 684 ± 68.6, $P = 0.016$; day 60: pEPI-EGFP 283 ± 75, pEPI-UCOE 583.5 ± 83, $P = 0.064$).

We next generated and analyzed single-cell clones CHO/pEPI-EGFP#14, CHO/pEPI-UCOE#3, and CHO/pEPI-HS4opp#1 as described previously.[26] During generation of single-cell clones, we observed a high variation in percentage of EGFP-expressing cells and mean FI in different CHO clones for pEPI-EGFP (0.27–96% of expressing cells with FI 96–905) and pEPI-HS4opp (0–97% of expressing cells with FI 0–788). In contrast, clones containing pEPI-UCOE displayed a more homogenous pattern of EGFP-expressing cells (18–56%) and expression level (FI 201–848). Of each construct, we chose a clone showing the best combination of high percentage of expressing cells and expression level, respectively. The number of plasmid copies per cell for each clone was determined using quantitative PCR and specific primers within the pEPI-vectors (**Figure 3a**). The original pEPI-EGFP plasmid and its derivative pEPI-HS4opp established in CHO cells with five to ten copies/cell (CHO/pEPI-EGFP#14: 5.85 ± 3.1; CHO/pEPI-HS4opp#1: 4.54 ± 0.07), while CHO/pEPI-UCOE#3 established at a slightly higher copy number (CHO/pEPI-UCOE#3: 12.3 ± 2.6). Long-term maintenance of each single clone was analyzed at regular periodic intervals using flow cytometry. Clone CHO/pEPI-EGFP#14 showed 89.9%, CHO/pEPI-UCOE#3 30.5%, and CHO/pEPI-HS4opp#1 97.5% EGFP-expressing cells (**Figure 3b**). Again, clone CHO/pEPI-UCOE#3 showed a significantly higher level of expression compared with CHO/pEPI-EGFP#14 and CHO/pEPI-HS4opp#1 (CHO/pEPI-EGFP#14 FI 145; CHO/pEPI-UCOE#3 FI 527; CHO/pEPI-HS4opp#1 FI 76). Normalization of the FI to episome copy numbers resulted in a twofold increased FI for pEPI-UCOE over pEPI-EGFP, and 2.5-fold increased FI for pEPI-UCOE over pEPI-HS4opp. Concerning the high expression level of pEPI-UCOE, it has been shown recently that the A2UCOE is marked with histone modifications associated with actively transcribed chromatin.[27] Therefore, we examined the histone H3 acetylation status across the expression cassette of pEPI-UCOE and its corresponding sequences in pEPI-EGFP using CHO/pEPI-UCOE#3 and CHO/pEPI-EGFP#14, respectively (**Figure 4**). The level of histone H3 and acetylated H3K9/14 (H3K9/14ac) detected in pEPI-EGFP were set to 1 and used as a reference against which could be compared the values obtained for pEPI-UCOE. Surprisingly, in CHO/pEPI-UCOE#3 we found a distinctly reduced level of H3K9/14ac in the promoter region (3′CMV, 0.51-fold), spanning the transcription start site of EGFP (TSS, 0.25-fold), and in the EGFP gene (5′EGFP, 0.29-fold; 3′EGFP, 0.38-fold) compared with pEPI-EGFP. Remarkably, these reduced H3K9/14ac levels were accompanied with a reduced level of overall histone H3 (3′CMV, 0.47-fold; TSS, 0.53-fold; 5′EGFP, 0.55-fold; 3′EGFP, 0.62-fold), indicating an open chromatin structure accompanied with a lower number of nucleosomes. However, looking in the 3′ region of the A2UCOE (upstream the CMV promoter) we found an elevated overall level of H3

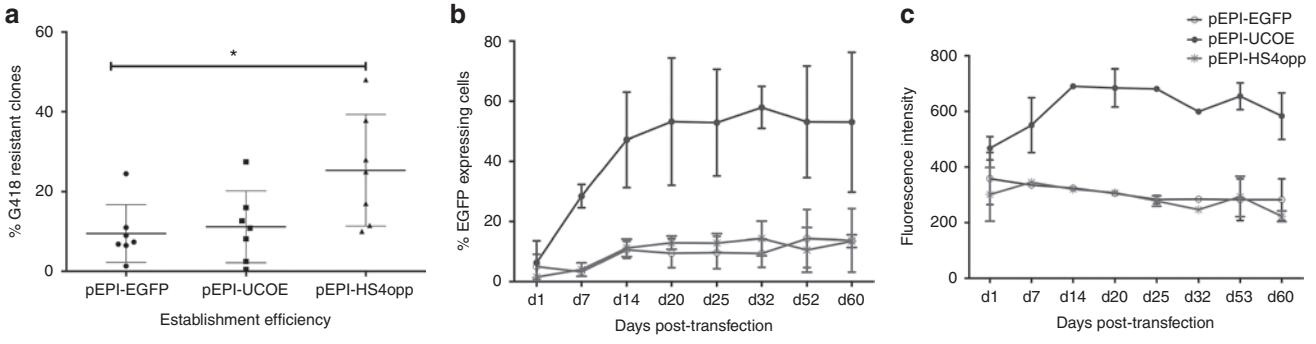

Figure 2 cHS4 enhances episomal establishment of and the A2UCOE increases transgene expression from the pEPI plasmid in CHO cells. (**a**) Establishment efficiency was determined as described.[26] pEPI-EGFP vector and its derivative pEPI-UCOE established with a similar efficiency (pEPI-EGFP: 9.5 ± 7.25%; pEPI-UCOE: 11.2 ± 9.04%) while pEPI-HS4opp showed significantly increased establishment efficiency (25.3 ± 13.8%; $P = 0.028$). Results are from seven independent transfections. (**b**) Number of EGFP-expressing cells and (**c**) fluorescent intensity (FI) of a transfected, mixed cell population maintained in the absence of G418 selective pressure: cells carrying pEPI-UCOE, displayed an increased level of transgene expression, while pEPI-HS4opp showed a similar transgene expression level to pEPI-EGFP (day 60: % EGFP-expressing cells:pEPI-EGFP 13.7 ± 10.5%, pEPI-UCOE 53 ± 23.2%; FI: pEPI-EGFP 283 ± 75, pEPI-UCOE 583.5 ± 83). Results are from three independent transfections. *$P < 0.05$. UCOE, ubiquitous chromatin-opening element.

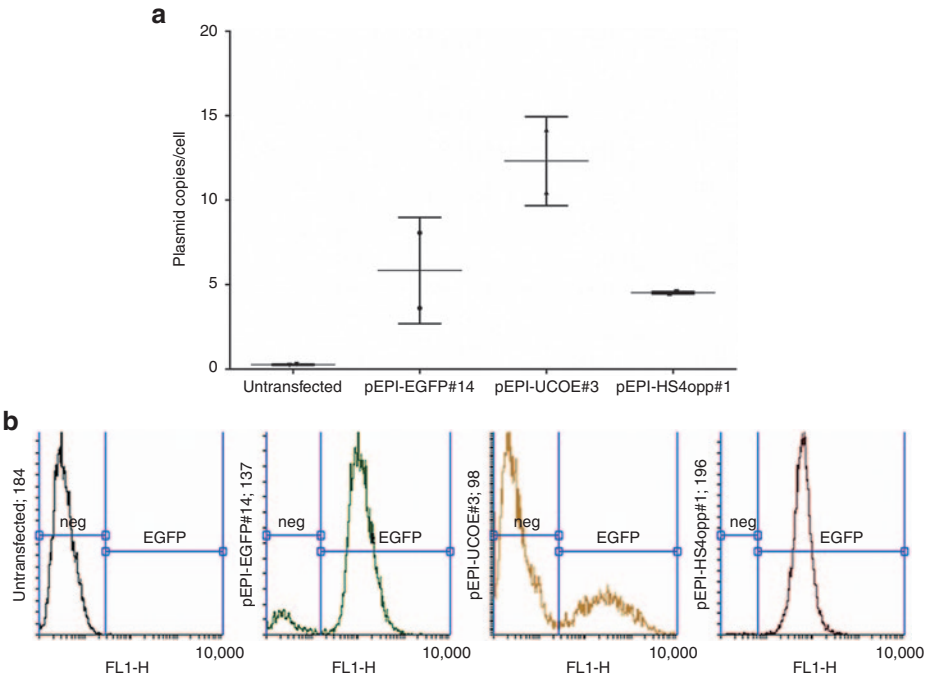

Figure 3 Analysis of single cell clone. (**a**) CHO/pEPI-EGFP#14 established with a copy number of 5.85±3.1, CHO/pEPI-UCOE#3: 12.3±2.6, and CHO/pEPI-HS4opp#1: 4.54±0.07 plasmid copies per cell. (**b**) Single clones showed stable EGFP expression in the absence of selection: pEPI-EGFP#14 89.9% (FI 145), pEPI-UCOE#3 30.5% (FI 527), and pEPI-HS4opp#1 97.5% (FI 76). FI, fluorescent intensity; UCOE, ubiquitous chromatin-opening element.

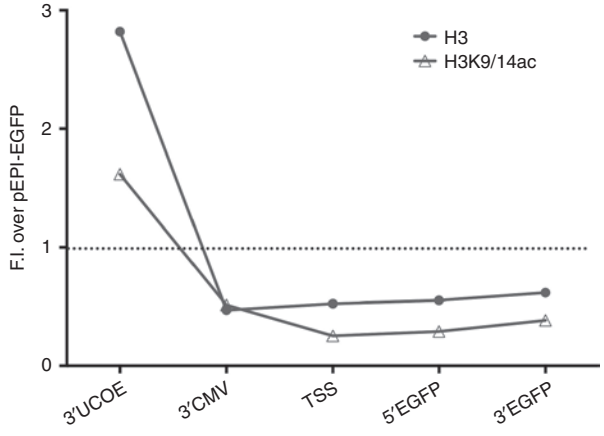

Figure 4 Increased H3 and H3K9/14ac levels on pEPI-UCOE. Higher expression of pEPI-UCOE are due to elevated levels of acetylated H3K9/14 in the 3′UCOE region (1.6 over pEPI-EGFP) propagating an open chromatin environment represented by reduced H3 level on promoter, TSS and EGFP gene (0.5 over pEPI-EGFP). UCOE, ubiquitous chromatin-opening element.

(2.8-fold) and H3K9/14ac (1.7-fold), compared with the corresponding sequence in pEPI-EGFP.

The cHS4 element interacts with the nuclear matrix enhancing S/MAR matrix interaction

Since insulator sequences are thought to be involved in chromatin organization, we suggest that the increased establishment efficiency of pEPI-HS4opp may be due to an interaction of cHS4 with the nuclear matrix. To this hypothesis we performed a nuclear fractionation procedure with the stably established clones CHO/pEPI-EGFP#14 and CHO/pEPI-HS4opp#1,[28,29] to determine if the episomal vectors they contain are associated with the nuclear matrix.. In these experiments, chromatin was digested with five restriction enzymes that do not cut in pEPI-EGFP and pEPI-HS4opp, together with either *Eco*RI (linearizes both plasmids), *Xmi*I (releasing S/MAR), or *Xma*I (releasing cHS4). After nuclear fractionation, samples were treated with Proteinase K and purified DNA was subjected to quantitative PCR to trace the complete vectors as well as their S/MAR and cHS4 depleted variants in the nuclear matrix fraction. The matrix fraction contains matrix proteins and the core filament network. In its complete version (*Eco*RI linearized, with S/MAR), pEPI-EGFP resides in the matrix fraction;[28] the detected amount of linearized vectors in the matrix fraction was set to 1 and relative vector amounts of the depleted variants are plotted in **Figure 5**. Upon depletion of the S/MAR with *Xmi*I the amount of pEPI-EGFP vector DNA in the matrix fraction is decreased by 16% (0.84±0.1-fold over linearized pEPI-EGFP **Figure 5a**). In contrast, the amount of pEPI-HS4opp remained nearly unchanged upon depletion of the S/MAR (0.96±0.13-fold over linearized pEPI-HS4opp). To test whether the cHS4 element is involved in the association with the nuclear matrix, we released the cHS4 element using *Xma*I; pEPI-HS4opp vector DNA was again traced in the matrix fraction. Upon depletion of the S/MAR element the amount of pEPI-HS4opp vector DNA in the matrix fraction remained nearly unchanged (0.83-fold over linearized pEPI-HS4opp), while there was a marked reduction of 35% of pEPI-HS4opp DNA after depletion of the cHS4 element (0.65-fold over linearized pEPI-HS4opp **Figure 5b**).

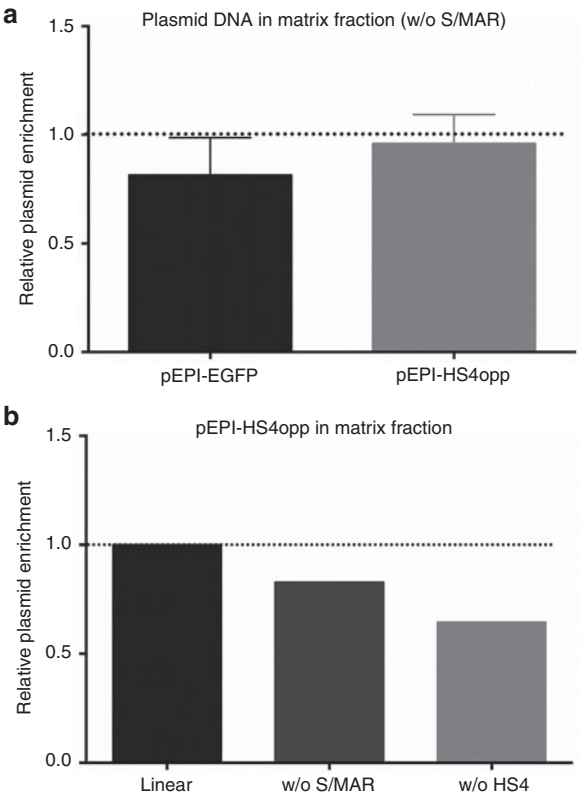

Figure 5 Enrichment of pEPI-HS4opp in the nuclear matrix fraction. DNA from the nuclear matrix fraction derived from stable clones CHO/pEPI-EGFP#14 and CHO/pEPI-HS4opp#1 was restriction enzyme digested with five noncutting enzymes and either *Eco*RI (linearization), *Xmi*I (releasing S/MAR) or *Xma*I (releasing cHS4). Quantitative PCR was performed using primers corresponding to the neomycin-resistance gene, with the amount of amplified linearized vector DNA in the matrix fraction arbitrarily set to 1. (**a**) Upon digestion with *Xmi*I, the amount of pEPI-EGFP in the matrix fraction decreased (0.84 ± 0.1), while the amount of pEPI-HS4opp remained unchanged (0.96 ± 0.13). (**b**) Releasing the cHS4 element from pEPI-HS4opp using *Xma*I led to a decreased amount of vector DNA in the matrix fraction (without S/MAR 0.83; without HS4 0.64) indicating that the cHS4 element enhanced the matrix association of the S/MAR.

Using a similar approach, endogenous DNA in nuclear matrix preparations was measured by preserving native attachment sites[30,31] using CHO/pEPI-HS4opp#1 cells. Proteins were stripped from nuclei with a mild detergent and subsequently DNA was digested with DNAse I. Remaining DNA fragments were precipitated and used for quantitative PCR analyses with total DNA extracted from CHO/pEPI-HS4opp#1 serving as reference and set arbitrarily to 1. Here, the S/MAR sequence in pEPI-HS4opp could be amplified, while the CMV promoter sequence was depleted by the digestion (data not shown; 3′CMV 0.071 ± 0.05, 3′S/MAR 0.302 ± 0.19, P = 0.055). Within the 1.2 kb cHS4 element, the 3′HS4 fragment could be amplified, while the 5′HS4 fragment was amplified at very low levels and apparently reduced by digestion (data not shown; 3′HS4 0.29 ± 0.35, 5′HS4 0.021 ± 0.01 P = 0.18). A protein that mediates many insulator functions and is involved in a broad range of chromatin regulatory events[32–35] is the CCCTC-binding factor (CTCF). Interactions of CTCF with the cHS4 have been shown previously.[31,36] Using chromatin immunoprecipitation (ChIP), we examined binding of CTCF to pEPI-HS4opp plasmid DNA. We detected a low signal for the 3′CMV and 3′S/MAR region at background levels (1.85 ± 2.5% of input DNA), and a slightly increased signal for the 3′HS4 region (2.5 ± 0.44% of input DNA). For the 5′HS4 region a distinct signal for CTCF binding was detectable (11.5 ± 9% of input DNA; **Figure 6**).

Discussion

The first nonviral and episomal plasmid vector was reported by Piechaczek *et al.*[1] The core unit being essential for its episomal behavior consists of an active transcription unit running into a scaffold/matrix attachment region (S/MAR). Its nonviral origin and episomal behavior, even in the absence of selective pressure, predestine pEPI as a promising alternative for gene therapeutic approaches. However, a major factor limiting its use in gene therapy is the low efficiency with which the vector establishes as a replicating episomal entity following delivery into the desired target cells. Nearly 95% of all plasmids are lost during subsequent cell division with only 1–5% of the originally transfected vector molecules able to establish in the nucleus and stably replicate episomally. Since the functionality of the vector depends on the nuclear compartment it reaches after transfection, it would be desirable to identify genomic sequences that are involved in genome organization and thus directing DNA (plasmid) sequences into certain, preferentially active, nuclear compartments.

Within the last decade, multiple modifications of the prototypical pEPI vector have been applied for its use *in vivo* and *in vitro* (reviewed in ref. 13). Nevertheless, all these modifications provided little or no improvement in transgene expression, copy number per cell, and/or episomal establishment efficiency. In this study, we focused on improving pEPI vector function with respect to transgene expression and episome establishment efficiency. By inserting *cis*-active sequences, namely the cHS4 insulator and a UCOE, we demonstrated increased establishment via the cHS4 element and increased transgene expression and copy number by utilizing the A2UCOE. The A2UCOE region used in this study was derived from a methylation-free CpG island spanning the divergently transcribed promoters of the housekeeping genes *HNRPA2B1* and *CBX3*. The A2UCOE is not a classical enhancer and was shown to confer resistance to silencing of the CMV promoter in a lentiviral vector system when inserted upstream the expression cassette.[37,38] Within recent years, the *HNRPA2B1-CBX3*A2UCOE region has been successfully employed in the construction of expression vectors that have markedly enhanced and improved the generation of cell lines for therapeutic protein production and retroviral/lentiviral vectors for human gene therapy applications.[37–40] Consistent with these observations we detected increased EGFP expression levels (mean FI) up to twofold over non-A2UCOE containing vectors (pEPI-EGFP and pEPI-HS4opp; **Figure 2b**,**c**), demonstrating an A2UCOE-mediated effect on gene expression in an episomal vector system. Our demonstration of A2UCOE functionality within a replicating episomal vector context suggests that this genomic element can

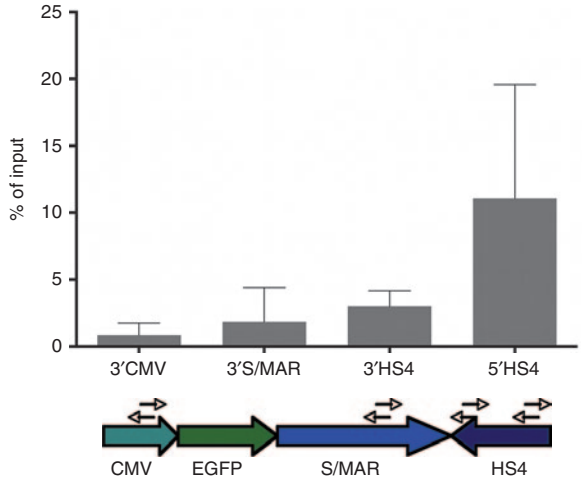

Figure 6 CTCF binding to pEPI-HS4opp plasmid DNA. ChIP was performed on CHO/pEPI-HS4#1 using a specific antibody against CTCF. Immunoprecipitated DNA was subjected to quantitative PCR and normalized to IgG; primer positions are indicated (arrows). A distinct signal of CTCF binding to the 5′-region and a somewhat weaker signal for the 3′-region of the cHS4 element were detectable. ChIP, chromatin immunoprecipitation; CTCF, CCCTC-binding factor.

not only confer resistance to silencing and elevated expression on silencing-prone viral promoters such as CMV used in this study,[41] but also on other ubiquitously active or tissue specific promoters as has been demonstrated within lentiviral vectors.[39,42] However, we assumed that this increased expression is associated with histone modifications typical for actively transcribed chromatin. We therefore performed ChIP analyses using antibodies against histone H3 and its counterpart acetylated at lysine residues 9 and 14 (H3K9/14ac). Acetylated H3 is most often found in promoter/enhancer regions[43–45] and H3K9/14ac was shown to be crucial for recruiting TFIID to IFN-γ locus.[46] We found both H3 and H3K9/14ac levels decreased in the CMV promoter and EGFP coding region of pEPI-UCOE when compared with pEPI-EGFP (**Figure 4**), which was initially surprising. However, since constantly active transcription is marked by an open chromatin structure, this might explain the overall reduced amount of histone H3 and thus a reduced level of H3K9/14ac. Consistently with this suggestion we found an increased level of H3K9/14ac in the 3′A2UCOE region. Similar results were previously reported by Lindahl Allen and Antoniou,[27] showing that the endogenous *HNRPA2B1-CBX3* A2UCOE region is characterized by an extended length of unmethylated DNA contributing to its chromatin-opening function, which then promotes efficient transcription[15] in association with elevated levels of active histone modification marks (H3 acetylation; H3K4 tri-methylation) and overall depletion of nucleosomes at the *CBX3* and *HNRPA2B1* promoters.[27]

Another *cis*-acting element we functionally evaluated in pEPI was the cHS4 1.2 kb insulator element from the chicken β-globin locus control region. Based on the observation of Moreno *et al.*[25] of increased homogeneity of transgene expression employing this element and since the cHS4 is known to interact with the nuclear matrix,[31] the cHS4 insulator was inserted downstream and in opposite orientation with respect to the S/MAR, resulting in pEPI-HS4opp (**Figure 1c**). In contrast to pEPI-UCOE, cells transfected with pEPI-HS4opp showed moderate transgene expression similar to cells transfected with pEPI-EGFP (**Figure 2b**,**c**), but a significantly elevated rate of episomal plasmid establishment (**Figure 2a**). Rate of establishment or establishment efficiency describes the frequency with which S/MAR containing vector molecules are stably and episomally maintained after transfection and inherited during subsequent cell divisions, even in the absence of selective pressure. We determined establishment efficiency of our constructs using a colony-forming assay.[26] We observed low establishment efficiencies for pEPI-EGFP (9.5 ± 7.25%) and pEPI-UCOE (11.2 ± 9.04%), meaning that ~90% of all initially transfected molecules are lost. It is widely accepted that an active transcription running into the S/MAR is mandatory for episomal maintenance of S/MAR based vectors.[3,4] But low establishment efficiencies of pEPI-UCOE suggest that high transcription rates do not affect efficiency of the initial establishment event. However, pEPI-HS4opp was found to establish with an increased efficiency of 25.3 ± 13.8% (P = 0.028). It has been shown that pEPI plasmid vectors interact with the nuclear matrix via an S/MAR–SAF-A interaction and thereby are stably maintained.[28] Furthermore, it is known that the cHS4 insulator element associates with the nuclear matrix through its interaction with CTCF.[31] Here we were able to demonstrate that: (i) pEPI-EGFP accumulated in the matrix fraction with depletion of the S/MAR led to a reduction of pEPI-EGFP in the nuclear matrix fraction[28] (**Figure 5**); (ii) pEPI-HS4opp also accumulated in the matrix fraction but independently of the S/MAR, while release of the 1.2 kb cHS4 element led to a decrease of pEPI-HS4opp in the matrix fraction (**Figure 5**), and (iii) CTCF was found to associate with the 5′ half of cHS4 and with a somewhat weaker binding detectable in the 3′ region (**Figure 6**).[24,31] Therefore, our data suggest that the elevated establishment efficiency of pEPI-HS4opp is due to the presence of cHS4 interacting with the matrix and thus increasing the probability of becoming "anchored" onto this nuclear structure and being established with higher efficiency. However, since it has been convincingly described before that an ongoing transcription running into the S/MAR is sufficient for episomal replication,[3] we assume that a construct harboring the cHS4 element alone would not be able to replicate episomally and thus would not be mitotically stable the absence of selection.

A pEPI derivative combining both the A2UCOE and cHS4 elements was also constructed and found to be maintained episomally, showing elevated (A2UCOE-mediated) EGFP expression but did not establish with an efficiency comparable with that of pEPI-HS4opp, containing cHS4 alone (data not shown), which may be due to its increased size (~11.5 kb). Nevertheless, to date S/MAR based vectors represent one, if not the only, attractive alternative to viral vectors for gene therapy applications. Here we show that the limiting factors of S/MAR based vectors, low and/or unstable transgene expression and low replicating episomal establishment efficiency were markedly improved by insertion of chromosomal *cis*-acting elements. The detailed functional mechanisms, especially those underlying the pEPI-HS4opp–nuclear matrix interaction, need to be further investigated. We believe that knowledge of the correlation between nuclear architecture, transgene expression, and plasmid establishment will have crucial impact on

our understanding of nuclear organization and may also lead to improved vectors for nonviral gene therapy.

Materials and methods

Construction of the pEPI derivates. The functional element essential for episomal replication and maintenance of pEPI is a promoter-transgene-S/MAR unit. The original pEPI-EGFP harbors a CMV-EGFP-S/MAR transcription unit. Starting with the pEPI(UCOE-hcr-intron-PDX), a kind gift from J. Fabre (King's College London, London, UK),[47] pEPI-UCOE was constructed as follows: pEPI(UCOE-hcr-intron-PDX), which contains the 1.5 kb core region of the *HNRPA2B1-CBX3* UCOE[14] was digested with *Ase*I to replace the hcr-intron-PDX unit by a CMV-EGFP fragment from pEPI-EGFP (**Figure 1a**) resulting in pEPI-UCOE (**Figure 1b**). pEPI-HS4opp was constructed by cloning the 1.2 kb cHS4 insulator fragment from pNI-CD (kindly provided by G. Felsenfeld, NIH, USA) into pEPI-EGFP downstream the S/MAR in opposite direction[25] using the In-Fusion Cloning System (Clontech, St Germain-en-Laye, France) and primers SS 5′-TAGATCCGGTGGATCCGAGCTCACGGGGACA GCC-3′ and AS 5′-CGCGGGCCCGGGATCCAATATTCTCAC TGACTCCGTCCT-3′.

Vectors, transfection and cell culture conditions. The vectors pEPI-EGFP, pEPI-HS4opp, and pEPI-UCOE were transfected into CHO cells (German Resource Centre for Biological Material, DSMZ, Braunschweig, Germany) using FuGene HD transfection reagent (Roche, Mannheim, Germany) or Amaxa Cell Line Nucleofector Kit T (Lonza, Cologne, Germany). Transfections were carried out with freshly prepared plasmid DNA and only preparations with ≥90% supercoiled pDNA were used for transfection.[48] Cells were maintained in Ham's F12 Medium (PAN Biotech, Passau, Germany) supplemented with 10% fetal bovine serum (PAA, Pasching, Germany), penicillin (10,000 units/ml)/streptomycin (10 mg/ml) (PAA), and 50 µg/ml partricin (Biochrom, Berlin, Germany). At 24 hours, post-transfection, transient transfection efficiency was determined by fluorescence microscopy and flow cytometry analysis. Generally, transfection efficiency was 50–70%. Stably transfected cells were selected in the presence of 400 µg/ml G418 for at least 14 days and plasmid maintenance of the pEPI derivates was determined by flow cytometry analysis analyses over a time period of 60 days in the absence of selection.

Flow cytometry. Transfected cells were rinsed with PBS to remove medium and centrifuged for 6 minutes at 666×*g*. Analyses were performed with Becton Dickinson FACSCalibur multicolor flow cytometer (BD Biosciences, Heidelberg, Germany) equipped with a 488 nm laser capable of detecting and distinguishing fluorescence emissions and BD CellQuest analysis software. The percentage of EGFP positive cells as well as EGFP fluorescence intensity of each sample was determined. The BD CellQuest analysis software provides values for the mean FI of EGFP based on the number of fluorochromes contained in each cell captured by the flow cytometer. Therefore, the FI can be used as an indicator of transgene expression level in a distinct cell population.

Colony-forming assay. To determine replicating episome establishment efficiency of the different pEPI derivates, colony-forming assays were performed.[26] Briefly, 24 hours post-transfection efficiency was determined as described above and 50, 100, and 200 EGFP positive cells were seeded into a 150 mm cell culture dish in the presence of 400 µg/ml G418, respectively. After 14 days of selection, G418 resistant colonies were fixed with 1% formaldehyde/PBS for 15 minutes at room temperature and subsequently stained with 0.5% crystal violet/25% methanol for 10 minutes at room temperature. Afterwards, plates were rinsed with ddH_2O to remove excessive staining solution and colonies were counted. Establishment efficiency of a respective vector was calculated using the ratio of resistant colonies to seeded cells.

Southern blot analysis. Total genomic DNA[49] or low molecular weight DNA from a Hirt extract[50] was either isolated from a selected G418 resistant cell population or individual clones. In order to confirm replicative episomal maintenance of pEPI constructs, total genomic DNA or DNA from a Hirt extract were codigested with either *Dpn*I and a plasmid single-cutter or *Mbo*I and a plasmid single-cutter (*Pci*I for pEPI-EGFP and pEPI-UCOE; *Hind*III for pEPI-HS4opp). While *Dpn*I requires bacterial methylation for restriction, *Mbo*I specifically digests eukaryotic methylated DNA. Digested DNA was subjected to Southern blot analysis as previously described[1,51] using α-^{32}P[dCTP]-labeled linearized pEPI plasmid DNA lacking the S/MAR element as a probe. Complete *Mbo*I digestion indicates cellular replication.[49,51]

Quantitative PCR. Quantitative PCR analysis was performed using a Light Cycler instrument (Roche) and the FastStart DNA Master^{PLUS} SYBER Green I reaction mix (Roche). The reaction volume was 10 µl containing 0.3 µmol/l of each primer. PCR reactions were performed at 40 cycles using standard settings as recommended by the manufacturer. Plasmid copies were quantified using a single copy housekeeping gene (DHFR; SS 5′-GTTGGAGGCAGTTCCGTTT A-3′; AS 5′-AGCTGGGATAATGTGCTGCT-3′; 101 bp product) as reference gene.

Nuclear fractionation and in vivo *matrix assay.* Nuclear fractionation was performed as before.[28,29] Briefly, cells were detached by treatment with trypsin. An aliquot of 10^6 cells per reaction were washed once with cold PBS followed by incubation in cytoskeleton buffer (10 mmol/l PIPES, 300 mmol/l saccharose, 100 mmol/l NaCl, 3 mmol/l $MgCl_2$, 1 mol/l EGTA; 4 minutes on ice) and centrifugation (1000×*g*, 4 °C, 3 minutes); supernatant contained soluble cyto- and nucleoplasmic proteins. Cells were then incubated in extraction buffer (10 mmol/l PIPES, 300 mmol/l saccharose, 250 mmol/l ammonium sulphate, 3 mmol/l $MgCl_2$, 1 mol/l EGTA; 4 minutes on ice) and collected by centrifugation as above; supernatant contained soluble nuclear components including histone H1. Digestion was performed in digestion buffer (10 mmol/l PIPES, 300 mmol/l saccharose, 50 mmol/l NaCl, 3 mmol/l $MgCl_2$, 1 mol/l EGTA; 3 hours, 37 °C) using five enzymes that do not cut within pEPI-EGFP and pEPI-HS4opp (*Spe*I, *Not*I, *Eco*RV, *Xho*I, and *Pvu*I) together with either *Eco*RI (linearizing), *Xmi*I (S/MAR depletion), or *Xma*I

(cHS4 depletion). Subsequent centrifugation left the complete matrix in the pellet fraction while supernatant contained DNA and histones. After extraction with 2 mol/l NaCl buffer (10 mmol/l PIPES, 300 mmol/l saccharose, 2 mol/l NaCl, 3 mmol/l $MgCl_2$, 1 mol/l EGTA; 4 minutes on ice) matrix-associated proteins, which are not part of the core filament were located in the supernatant; matrix proteins, and the core filament network (the "matrix fraction") were resuspended in 1× TE/0.1% SDS and subjected to Proteinase K digestion, followed by phenol-chloroform extraction and DNA precipitation. Primers used for quantitative PCR were derived from the neomycin gene (SS 5′-GGAGAGGCTATTCGGCTATGAC-3′; AS 5′-CGTCAAGAAGGCGATAGAAGGC-3′).

Low ionic strength matrices from 10^6 CHO cells (CHO/pEPI-HS4opp#1) were prepared as described elsewhere.[30,31] Briefly, cells were detached by trypsin treatment, washed twice in cold 1× PBS and resuspended in 8 ml isolation buffer (3.75 mmol/l Tris-HCl, pH 7.4; 20 mmol/l KCl; 0.5 mmol/l EDTA/KOH; 0.05 mmol/l spermidine, 0.125 mmol/l spermin; 0.5% Triton-X100; protease inhibitor cocktail (Roche)). Subsequently, cells were homogenized by 15 strokes of a loose fitting pestle Potter homogenizer. Nuclei were pelleted (900×*g*, 10 minutes, 4 °C) and washed three-times in isolation buffer; washed nuclei were resuspended in 600 µl isolation buffer without EDTA and incubated for 20 minutes in a 37 °C waterbath. Extraction buffer (5 mmol/l HEPES/NaOH, pH 7.4; 2 mmol/l KCl; 2 mmol/l EDTA; 0.25 mmol/l spermidine; 0.5% Triton-X100; 25 mmol/l lithium 3,5-diiodosalicylate) was added slowly to a final volume of 7 ml. After 5 minutes of incubation at room temperature, nuclei were recovered by centrifugation at 2400×*g*, 20 minutes at room temperature and washed three-times in 1× digestion buffer (10×: 400 mmol/l Tris-HCl, 100 mmol/l NaCl, 60 mmol/l $MgCl_2$; 10 mmol/l $CaCl_2$; pH 7.9). Digestion with RNase-free DNase (Roche) was performed with 100 U DNase for at least 3 hours at room temperature. Matrices were then centrifuged for 5 minutes at 1500×*g*, washed four-times with digestion buffer, resuspended in 1× TE/0.1% SDS, and subjected to Proteinase K digestion, followed by phenol-chloroform extraction and DNA precipitation. Quantitative PCRs were performed using a Light Cycler instrument (Roche) and the FastStart DNA Master^{PLUS} SYBER Green I reaction mix (Roche). The reaction volume was 10 µl containing 0.3 µmol/l of each primer. Following primer pairs were used: 3′ CMV (SS 5′-TACATCAATGGGCGTGGATA-3′, AS 5′-GGCGGAGTTGTTACGACATT-3′), 3′ S/MAR (SS 5′-ACGGCAAAGCCATTTTCATA-3′, AS 5′-AAAAACATGGAGGAGCGTTG-3′), 3′ HS4 (SS 5′-AGCAGTGTCTGTGGCCTTTT-3′, AS 5′-ACCTGTTCAGTGCGTCACC-3′), 5′ HS4 (SS 5′-TTGGATCTGGATGAGCACTG-3′, AS 5′-GGGAGGTGTGGGAGGTTT-3′).

ChIP assays. ChIP experiments were performed as previously described.[52] Cross-linking of confluent cell cultures was performed with 1% paraformaldehyde/PBS (pH 7.4) for 10 minutes at room temperature and stopped by adding glycine to a final concentration of 0.25 mol/l. Cells were washed twice with ice-cold PBS, scraped into a 15 ml tube and centrifuged (900×*g*, 15 minutes, 4 °C). Cells were then washed twice in Triton-X buffer (0.25% Triton-X100; 10 mmol/l EDTA; 10 mmol/l Tris/HCl, pH 8.1; 10 mmol/l NaCl; 1× protease inhibitor cocktail (PIC, Roche)), followed by resuspension in SDS lysis buffer (1% SDS; 5 mmol/l EDTA; 50 mmol/l Tris/HCl, pH 8.1; 1× PIC (Roche)), and chromatin sheared to an average size of 200–1200 bp by sonication on ice. Sonicated chromatin was centrifuged (14,000×*g*, 15 minutes, 4 °C) and the supernatant used for immunoprecipitation. Protein G Dyna Beads (Invitrogen) were prepared by washing twice with 0.1% Triton-X100/PBS followed by two washing steps with TSE I buffer (0.1% SDS; 1% Triton-X100; 2 mmol/l EDTA; 20 mmol/l Tris/HCl, pH 8.1; 150 mmol/l NaCl; 1× PIC (Roche)). Beads and respective antibodies (3 µg) were preincubated by rotation in 1 ml dilution buffer (1% Triton-X100; 2 mmol/l EDTA; 20 mmol/l Tris/HCl, pH 8.1; 150 mmol/l NaCl; 1× PIC (Roche)) for 30 minutes at room temperature, followed by 2.5 hours at 4 °C. Chromatin was then added and incubated over night at 4 °C while mixing by rotation. A specified aliquot of the non-precipitated chromatin was saved to determine the quantity of input DNA. On the following day beads were washed once with dilution buffer, twice with ice-cold TSE I buffer, twice with buffer III (0.25 mmol/l LiCl; 1% NP-40; 1% deoxycholate; 1 mmol/l EDTA; 10 mmol/l Tris/HCl, pH 8.1; 1× PIC (Roche)) and once with TE wash buffer (2 mmol/l EDTA; 10 mmol/l Tris/HCl, pH 8.1; 1× PIC (Roche)). Between each washing step, beads were incubated rotation at 4 °C for 10 minutes. Following washing, beads were suspended in Elution buffer (1% SDS; 10 mmol/l EDTA; 50 mmol/l Tris/HCl, pH 8.1) and incubated for 5 hours at 65 °C and 1400 rpm to reverse the paraformaldehyde crosslinks. Proteinase K digestion was then performed overnight at 55 °C and immunoprecipitated DNA extracted using phenol:chloroform extraction and ethanol precipitation. Subsequently, DNA was disolvedin 50 µl ddH_2O, treated with RNase A, and subjected to quantitative PCR (Light Cycler, Roche).

The following antibodies were used for ChIP: anti-CTCF (Millipore, Cat-No 17-10044), anti-H3 (Abcam, ab1791), and anti-H3K9/K14ac (Santa Cruz, sc8655). Following primers were used for quantitative PCR: 3′ CMV (SS 5′-TACATCAATGGGCGTGGATA-3′, AS 5′-GGCGGAGTTGTTACGACATT-3′), 3′ S/MAR (SS 5′-ACGGCAAAGCCATTTTCATA-3′, AS 5′-AAAAACATGGAGGAGCGTTG-3′), 3′ HS4 (SS 5′-AGCAGTGTCTGTGGCCTTTT-3′, AS 5′-ACCTGTTCAGTGCGTCACC-3′), 5′ HS4 (SS 5′-TTGGATCTGGATGAGCACTG-3′, AS5′-GGGAGGTGTGGGAGGTTT-3′), EGFP-TSS (SS5′-TACATCAATGGGCGTGGATA-3′, AS 5′-GTTTACGTCGCCGTCCAG-3′), 5′ EGFP (SS 5′-ACGTAAACGGCCACAAGTTC-3′, AS 5′-AAGTCGTGCTGCTTCATGT-3′), 3′ EGFP (SS 5′-CGACCACTACCAGCAGAACA-3′, AS 5′-GAACTCCAGCAGGACCATGT-3′).

Acknowledgments. This work was supported by the Deutsche Forschungsgemeinschaft (SPP1230) and UWH Research Support. Claudia Hagedorn is supported by the Peter and Traudl Engelhorn foundation. The authors declare that there is no conflict of interest.

1. Piechaczek, C, Fetzer, C, Baiker, A, Bode, J and Lipps, HJ (1999). A vector based on the SV40 origin of replication and chromosomal S/MARs replicates episomally in CHO cells. *Nucleic Acids Res* **27**: 426–428.
2. Giannakopoulos, A, Stavrou, EF, Zarkadis, I, Zoumbos, N, Thrasher, AJ and Athanassiadou, A (2009). The functional role of S/MARs in episomal vectors as defined

by the stress-induced destabilization profile of the vector sequences. *J Mol Biol* **387**: 1239–1249.
3. Stehle, IM, Scinteie, MF, Baiker, A, Jenke, AC and Lipps, HJ (2003). Exploiting a minimal system to study the epigenetic control of DNA replication: the interplay between transcription and replication. *Chromosome Res* **11**: 413–421.
4. Rupprecht, S, Hagedorn, C, Seruggia, D, Magnusson, T, Wagner, E, Ogris, M *et al.* (2010). Controlled removal of a nonviral episomal vector from transfected cells. *Gene* **466**: 36–42.
5. Manzini, S, Vargiolu, A, Stehle, IM, Bacci, ML, Cerrito, MG, Giovannoni, R *et al.* (2006). Genetically modified pigs produced with a nonviral episomal vector. *Proc Natl Acad Sci USA* **103**: 17672–17677.
6. Papapetrou, EP, Ziros, PG, Micheva, ID, Zoumbos, NC and Athanassiadou, A (2006). Gene transfer into human hematopoietic progenitor cells with an episomal vector carrying an S/MAR element. *Gene Ther* **13**: 40–51.
7. Stehle, IM, Postberg, J, Rupprecht, S, Cremer, T, Jackson, DA and Lipps, HJ (2007). Establishment and mitotic stability of an extra-chromosomal mammalian replicon. *BMC Cell Biol* **8**: 33.
8. Sgourou, A, Routledge, S, Spathas, D, Athanassiadou, A and Antoniou, MN (2009). Physiological levels of HBB transgene expression from S/MAR element-based replicating episomal vectors. *J Biotechnol* **143**: 85–94.
9. Schaarschmidt, D, Baltin, J, Stehle, IM, Lipps, HJ and Knippers, R (2004). An episomal mammalian replicon: sequence-independent binding of the origin recognition complex. *EMBO J* **23**: 191–201.
10. Hagedorn, C, Lipps, HJ, and Rupprecht, S (2010). The epigenetic regulation of autonomous replicons. *BioMolecular Concepts* **1**: 17–30.
11. Leight, ER and Sugden, B (2001). Establishment of an oriP replicon is dependent upon an infrequent, epigenetic event. *Mol Cell Biol* **21**: 4149–4161.
12. Leight, ER, Sugden, B and Light, ER (2001). The cis-acting family of repeats can inhibit as well as stimulate establishment of an oriP replicon. *J Virol* **75**: 10709–10720.
13. Hagedorn, C, and Lipps, HJ (2011). pEPI for gene therapy. Non viral episomes and their application in somatic gene therapy. *J Cell Sci Ther* **S6**: 001.
14. Antoniou, M, Harland, L, Mustoe, T, Williams, S, Holdstock, J, Yague, E *et al.* (2003). Transgenes encompassing dual-promoter CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing. *Genomics* **82**: 269–279.
15. Williams, S, Mustoe, T, Mulcahy, T, Griffiths, M, Simpson, D, Antoniou, M *et al.* (2005). CpG-island fragments from the HNRPA2B1/CBX3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells. *BMC Biotechnol* **5**: 17.
16. Emery, DW (2011). The use of chromatin insulators to improve the expression and safety of integrating gene transfer vectors. *Hum Gene Ther* **22**: 761–774.
17. Emery, DW, Yannaki, E, Tubb, J and Stamatoyannopoulos, G (2000). A chromatin insulator protects retrovirus vectors from chromosomal position effects. *Proc Natl Acad Sci USA* **97**: 9150–9155.
18. Rivella, S, Callegari, JA, May, C, Tan, CW and Sadelain, M (2000). The cHS4 insulator increases the probability of retroviral expression at random chromosomal integration sites. *J Virol* **74**: 4679–4687.
19. Emery, DW, Yannaki, E, Tubb, J, Nishino, T, Li, Q and Stamatoyannopoulos, G (2002). Development of virus vectors for gene therapy of beta chain hemoglobinopathies: flanking with a chromatin insulator reduces gamma-globin gene silencing in vivo. *Blood* **100**: 2012–2019.
20. Ramezani, A, Hawley, TS and Hawley, RG (2003). Performance- and safety-enhanced lentiviral vectors containing the human interferon-beta scaffold attachment region and the chicken beta-globin insulator. *Blood* **101**: 4717–4724.
21. Francastel, C, Magis, W and Groudine, M (2001). Nuclear relocation of a transactivator subunit precedes target gene activation. *Proc Natl Acad Sci USA* **98**: 12120–12125.
22. Gerasimova, TI, Byrd, K and Corces, VG (2000). A chromatin insulator determines the nuclear localization of DNA. *Mol Cell* **6**: 1025–1035.
23. Ishii, K, Arib, G, Lin, C, Van Houwe, G and Laemmli, UK (2002). Chromatin boundaries in budding yeast: the nuclear pore connection. *Cell* **109**: 551–562.
24. Yusufzai, TM, Tagami, H, Nakatani, Y and Felsenfeld, G (2004). CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell* **13**: 291–298.
25. Moreno, R, Martinez, I, Petriz, J, González, JR, Gratacós, E and Aran, JM (2009). Boundary sequences stabilize transgene expression from subtle position effects in retroviral vectors. *Blood Cells Mol Dis* **43**: 214–220.
26. Hagedorn, C, Baiker, A, Postberg, J, Ehrhardt, A, and Lipps, HJ (2012). A colony-forming assay for determining the establishment efficiency of S/MAR-containing nonviral episomal expression vectors. *Cold Spring Harb Protoc* **2012**(6): 706–708.
27. Lindahl Allen, M and Antoniou, M (2007). Correlation of DNA methylation with histone modifications across the HNRPA2B1-CBX3 ubiquitously-acting chromatin open element (UCOE). *Epigenetics* **2**: 227–236.
28. Baiker, A, Maercker, C, Piechaczek, C, Schmidt, SB, Bode, J, Benham, C *et al.* (2000). Mitotic stability of an episomal vector containing a human scaffold/matrix-attached region is provided by association with nuclear matrix. *Nat Cell Biol* **2**: 182–184.
29. Nickerson, JA, Krockmalnic, G, and Penman, S. *Cell Biology, a Laboratory Manual*. Academic press: London, UK; 10, 1998.
30. Mirkovitch, J, Mirault, ME, and Laemmli, UK (1984). Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell* **39**: 223–232.
31. Yusufzai, TM and Felsenfeld, G (2004). The 5'-HS4 chicken beta-globin insulator is a CTCF-dependent nuclear matrix-associated element. *Proc Natl Acad Sci USA* **101**: 8620–8624.
32. Baniahmad, A, Steiner, C, Köhne, AC and Renkawitz, R (1990). Modular structure of a chicken lysozyme silencer: involvement of an unusual thyroid hormone receptor binding site. *Cell* **61**: 505–514.
33. Lobanenkov, VV, Nicolas, RH, Adler, VV, Paterson, H, Klenova, EM, Polotskaja, AV *et al.* (1990). A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene* **5**: 1743–1753.
34. Ohlsson, R, Bartkuhn, M and Renkawitz, R (2010). CTCF shapes chromatin by multiple mechanisms: the impact of 20 years of CTCF research on understanding the workings of chromatin. *Chromosoma* **119**: 351–360.
35. Ohlsson, R, Lobanenkov, V and Klenova, E (2010). Does CTCF mediate between nuclear organization and gene expression? *Bioessays* **32**: 37–50.
36. Bell, AC, West, AG and Felsenfeld, G (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**: 387–396.
37. Zhang, F, Frost, AR, Blundell, MP, Bales, O, Antoniou, MN and Thrasher, AJ (2010). A ubiquitous chromatin opening element (UCOE) confers resistance to DNA methylation-mediated silencing of lentiviral vectors. *Mol Ther* **18**: 1640–1649.
38. Zhang, F, Thornhill, SI, Howe, SJ, Ulaganathan, M, Schambach, A, Sinclair, J *et al.* (2007). Lentiviral vectors containing an enhancer-less ubiquitously acting chromatin opening element (UCOE) provide highly reproducible and stable transgene expression in hematopoietic cells. *Blood* **110**: 1448–1457.
39. Brendel, C, Müller-Kuller, U, Schultze-Strasser, S, Stein, S, Chen-Wichmann, L, Krattenmacher, A *et al.* (2012). Physiological regulation of transgene expression by a lentiviral vector containing the A2UCOE linked to a myeloid promoter. *Gene Ther* **19**: 1018–1029.
40. Knight, S, Zhang, F, Mueller-Kuller, U, Bokhoven, M, Gupta, A, Broughton, T *et al.* (2012). Safer, silencing-resistant lentiviral vectors: optimization of the ubiquitous chromatin-opening element through elimination of aberrant splicing. *J Virol* **86**: 9088–9095.
41. Scharfmann, R, Axelrod, JH and Verma, IM (1991). Long-term *in vivo* expression of retrovirus-mediated gene transfer in mouse fibroblast implants. *Proc Natl Acad Sci USA* **88**: 4626–4630.
42. Pfaff, N, Lachmann, N, Ackermann, M, Kohlscheen, S, Brendel, C, Maetzig, T *et al.* (2013). A ubiquitous chromatin opening element prevents transgene silencing in pluripotent stem cells and their differentiated progeny. *Stem Cells* **31**: 488–499.
43. Liang, G, Lin, JC, Wei, V, Yoo, C, Cheng, JC, Nguyen, CT *et al.* (2004). Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proc Natl Acad Sci USA* **101**: 7357–7362.
44. Myers, FA, Chong, W, Evans, DR, Thorne, AW and Crane-Robinson, C (2003). Acetylation of histone H2B mirrors that of H4 and H3 at the chicken beta-globin locus but not at housekeeping genes. *J Biol Chem* **278**: 36315–36322.
45. Myers, FA, Evans, DR, Clayton, AL, Thorne, AW and Crane-Robinson, C (2001). Targeted and extended acetylation of histones H4 and H3 at active and inactive genes in chicken embryo erythrocytes. *J Biol Chem* **276**: 20197–20205.
46. Agalioti, T, Chen, G and Thanos, D (2002). Deciphering the transcriptional histone acetylation code for a human gene. *Cell* **111**: 381–392.
47. Cim, A, Sawyer, G, Zhang, X, Su, H, Collins, L, Jones, P, *et al.* (2010). Nonviral delivery of the rat PDX1 gene to rat liver for the in vivo transdifferentiation of liver cells to pancreatic beta-cells. *Human Gene Therapy* **21**: 1.
48. Hagedorn, C, Baiker, A, Postberg, J, Ehrhardt, A and Lipps, HJ (2012). Handling S/MAR vectors. *Cold Spring Harb Protoc* **2012**: 657–663.
49. Sambrook, J, and Russell, DW. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press: New York, NY, USA, 2001.
50. Hirt, B (1967). Selective extraction of polyoma DNA from infected mouse cell cultures. *J Mol Biol* **26**: 365–369.
51. Southern, EM (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98**: 503–517.
52. Rupprecht, S and Lipps, HJ (2009). Cell cycle dependent histone dynamics of an episomal non-viral vector. *Gene* **439**: 95–101.