

## The CCCTC-binding factor CTCF represses hepatitis B virus Enhancer I and regulates viral transcription

D'Arienzo, Valentina; Ferguson, Jack; Giraud, Guillaume; Chapus, Fleur; Harris, James; Wing, Peter; Claydon, Adam; Begum, Sophia; Zhuang, Xiaodong; Balfe, Peter; Testoni, Barbara; McKeating, Jane A; Parish, Jo

DOI:

[10.1111/cmi.13274](https://doi.org/10.1111/cmi.13274)

License:

Creative Commons: Attribution (CC BY)

*Document Version*

Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*

D'Arienzo, V, Ferguson, J, Giraud, G, Chapus, F, Harris, J, Wing, P, Claydon, A, Begum, S, Zhuang, X, Balfe, P, Testoni, B, McKeating, JA & Parish, J 2020, 'The CCCTC-binding factor CTCF represses hepatitis B virus Enhancer I and regulates viral transcription', *Cellular Microbiology*, vol. 2020, e13274.  
<https://doi.org/10.1111/cmi.13274>

[Link to publication on Research at Birmingham portal](#)

### General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

### Take down policy





While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

RESEARCH ARTICLE

WILEY

# The CCCTC-binding factor CTCF represses hepatitis B virus enhancer I and regulates viral transcription

Valentina D'Arienzo<sup>1</sup> | Jack Ferguson<sup>2</sup> | Guillaume Giraud<sup>3</sup> | Fleur Chapus<sup>3</sup> | James M. Harris<sup>1</sup> | Peter A. C. Wing<sup>1</sup> | Adam Claydon<sup>2</sup>  | Sophia Begum<sup>2</sup> | Xiaodong Zhuang<sup>1</sup>  | Peter Balfe<sup>1</sup> | Barbara Testoni<sup>3</sup> | Jane A. McKeating<sup>1</sup>  | Joanna L. Parish<sup>2</sup> 

<sup>1</sup>Nuffield Department of Medicine, University of Oxford, Oxford, UK

<sup>2</sup>Institute of Cancer and Genomic sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

<sup>3</sup>CRCL INSERM and Cancer Research Center of Lyon (CRCL), Lyon, France

## Correspondence

Joanna L. Parish, Institute of Cancer and Genomic Sciences, University of Birmingham, Edgbaston B15 2TT, UK.  
Email: j.l.parish@bham.ac.uk

## Funding information

Agence Nationale de Recherches sur le Sida et les Hépatites Virales, Grant/Award Number: ECTZ75178; Medical Research Council, Grant/Award Numbers: MR/N023498/1, MR/R022011/1, MR/T015985/1; Wellcome Trust, Grant/Award Number: 200838/Z/16/Z

## Abstract

Hepatitis B virus (HBV) infection is of global importance with over 2 billion people exposed to the virus during their lifetime and at risk of progressive liver disease, cirrhosis and hepatocellular carcinoma. HBV is a member of the *Hepadnaviridae* family that replicates via episomal copies of a covalently closed circular DNA (cccDNA) genome. The chromatinization of this small viral genome, with overlapping open reading frames and regulatory elements, suggests an important role for epigenetic pathways to regulate viral transcription. The chromatin-organising transcriptional insulator protein, CCCTC-binding factor (CTCF), has been reported to regulate transcription in a diverse range of viruses. We identified two conserved CTCF binding sites in the HBV genome within enhancer I and chromatin immunoprecipitation (ChIP) analysis demonstrated an enrichment of CTCF binding to integrated or episomal copies of the viral genome. siRNA knock-down of CTCF results in a significant increase in pre-genomic RNA levels in de novo infected HepG2 cells and those supporting episomal HBV DNA replication. Furthermore, mutation of these sites in HBV DNA minicircles abrogated CTCF binding and increased pre-genomic RNA levels, providing evidence of a direct role for CTCF in repressing HBV transcription.

## KEYWORDS

CTCF, epigenetics, HBV, transcription insulation, transcription regulation

## 1 | INTRODUCTION

Hepatitis B virus (HBV) infection is one of the world's unconquered infections with an estimated 2 billion people exposed to the virus in their lifetime. HBV replicates in hepatocytes, and chronic infection

can result in progressive liver disease, cirrhosis and hepatocellular carcinoma. HBV is a member of the *Hepadnaviridae* family and is classified into eight genotypes, A-H (McNaughton et al., 2019), that can associate with different clinical outcomes and treatment responses (Kao, Chen, Lai, & Chen, 2000). The HBV genome is a small, partially double-stranded relaxed circular DNA (rcDNA) of approximately 3.2 Kb. Following HBV entry into hepatocytes via the liver-specific bile-acid transporter, sodium taurocholate co-transporting polypeptide

Valentina D'Arienzo and Jack Ferguson have equally contributed to the work.  
Jane A. McKeating and Joanna L. Parish are joint senior authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Cellular Microbiology* published by John Wiley & Sons Ltd.

(NTCP) (Ni et al., 2014), rcDNA is released into the nucleus and is repaired by host DNA repair pathways to generate covalently closed circular DNA (cccDNA). This cccDNA persists in the nucleus as a long-lived nucleosome-associated minichromosome (Lythgoe, Lumley, McKeating, & Mathews, 2020) and is the transcriptional template for all viral RNAs (Hong, Kim, & Guo, 2017).

The HBV genome is transcribed by the host RNA polymerase II (RNA pol II) complex from four promoters (basal core promoter, BCP, Sp1, Sp2 and Xp) (Hong et al., 2017) that results in six major viral RNAs of increasing length with heterogeneous 5' ends and a common polyadenylation signal (Stadelmayer et al., 2020). These RNAs include: pre-core (preC) that encodes e antigen (HBeAg); pre-genomic (pgRNA) that is translated to yield core protein (HBcAg) and polymerase; preS1, preS2 and S RNAs encoding the surface envelope glycoproteins and X transcript for the multi-functional x protein (HBx). Two viral enhancers regulate viral transcription: Enhancer I (EnhI) is located upstream of and partially overlaps the X promoter (Xp) and directs BCP activity (Hu & Siddiqui, 1991), which stimulates the production of both preC and pgRNAs. Enhancer II (EnhII) overlaps a large portion of BCP and stimulates activity of the distal Sp1 and Sp2 promoters, as well as the BCP and Xp (Yuh & Ting, 1990). The BCP encodes a negative regulatory element (NRE) that overlaps with Enh II (Sun et al., 2001) and has been reported to repress promoter activity. Encapsidated pgRNA is reverse-transcribed by the viral polymerase to generate new DNA genomes that can be re-imported to the nucleus to maintain the cccDNA pool or are enveloped and secreted as infectious particles (Urban, Schulze, Dandri, & Petersen, 2010), highlighting an essential role for the BCP in regulating viral replication.

HBV cccDNA is assembled into nucleosomes by cellular histones to form episomal chromatin (Levrero et al., 2009). The viral DNA is enriched with active epigenetic histone modifications including trimethylation of lysine 4 (H3K4Me3) and acetylation of lysine 27 on histone 3 (H3K27Ac) but devoid of repressive marks, such as trimethylation of lysine 27 on histone 3 (H3K27Me3) (Flecken et al., 2019; Tropberger et al., 2015). The overlap of active histone marks with RNA pol II occupancy suggests that viral transcription is regulated by epigenetic modification. In support of this, treating de novo infected primary human hepatocytes with inhibitors of the histone acetyltransferase p300/CBP reduces HBV RNA levels (Tropberger et al., 2015). Although the mechanisms underlying the epigenetic regulation of HBV cccDNA are not fully understood, several epigenetic modifiers are recruited to HBV cccDNA by HBx. As such, HBx behaves as a transcriptional regulator of both viral and cellular promoters (Guerrieri et al., 2017), and although HBx cannot bind to DNA directly, it can associate with components of the basal transcription machinery, transcription factors and transcriptional co-activators (Tang, Oishi, Kaneko, & Murakami, 2006). HBx coordinates the recruitment of the CBP/p300 and PCAF histone acetyl transferases (HAT) to cccDNA while facilitating the exclusion of histone deacetylases (HDACs) HDAC1 and Sirtuin 1 (Sirt1), resulting in hyper-acetylation of cccDNA (Belloni et al., 2009; Chong et al., 2020). HBV transcription is dependent on an array of ubiquitous and liver-specific cellular transcription factors, including the liver-specific hepatocyte

nuclear factors, 1 and 4 (HNF-1/4), and ubiquitously expressed octamer-binding protein 1 and specificity protein 1 (reviewed in Oropeza et al., 2020; Turton, Meier-Stephenson, Badmalia, Coffin, & Patel, 2020).

The genomes of metazoans are organised into megabase-sized regions termed topologically associated domains (TADs) that provide regulatory segmentation required for appropriate gene expression and replication. TADs are separated by regions enriched in binding sites of the ubiquitously expressed CCCTC-binding factor (CTCF), which stabilises chromatin loops by anchoring cohesin rings at the base of the loops (Rowley & Corces, 2018). Such spatial organisation can create epigenetic boundaries that separate transcriptionally active and inactive chromatin domains and control cis-regulatory elements, such as transcriptional enhancers. CTCF binds to tens of thousands of either ubiquitous or cell type specific consensus binding sites within the human genome, regulating both tissue-specific and developmental changes in gene expression (Chen, Tian, Shu, Bo, & Wang, 2012).

The occupancy of specific CTCF binding sites is dictated by chromatin accessibility and local epigenetic status (Kumar & Bucher, 2016). In addition to the organisation of chromatin domains, CTCF can function as a transcriptional repressor, or activator, by direct association with promoter proximal elements. CTCF was shown to act as a transcriptional repressor of the *c-myc* oncogene by creating a roadblock to RNA pol II (Filippova et al., 1996). Conversely, CTCF can physically associate with transcriptional regulators, such as the general transcription factor, TFIID, to promote recruitment of the cyclin-dependent kinase 8, resulting in stimulation of RNA pol II activity (Pena-Hernandez et al., 2015). CTCF regulates the transcription (up or down) of evolutionarily distinct DNA viruses (Pentland & Parish, 2015) including: Kaposi sarcoma-associated herpesvirus; Epstein-Barr virus and herpes simplex virus (Chen et al., 2014; Kang, Wiedmer, Yuan, Robertson, & Lieberman, 2011; Lang et al., 2017; Washington, Musarrat, Ertel, Backes, & Neumann, 2018). We have demonstrated that CTCF recruitment to the human papillomavirus (HPV) genome negatively regulates early promoter usage via host cell differentiation-specific stabilisation of an epigenetically repressed chromatin loop (Paris et al., 2015; Pentland et al., 2018). However, a role in HBV transcription regulation has not yet been reported, herein we show that CTCF binds HBV DNA and acts as a repressor of viral transcription.

## 2 | RESULTS

### 2.1 | CTCF binds HBV DNA at conserved sites within enhancer elements

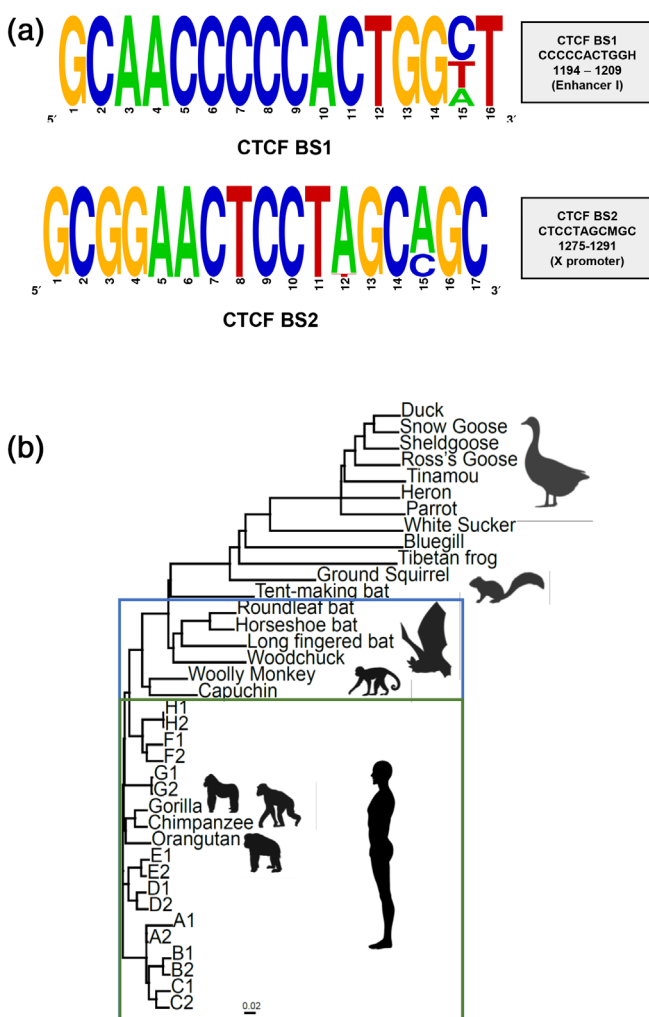
To investigate a role for CTCF in the HBV life cycle, we first assessed whether the viral genome encodes any CTCF binding sites, using an open access database (<http://insulatordb.uthsc.edu/>), to identify putative binding sites. Screening >7,000 HBV sequences available in the HBV database (Hayer et al., 2013) identified two CTCF binding sites (BS) between nucleotides 1194–1209 in EnhI (CTCF BS1) and

1275–1291 in the Xp (CTCF BS2). Importantly, these binding sites are conserved among all HBV genotypes (Figure 1a). The *Hepadnaviridae* family includes a number of related viruses that infect other species, including birds, mammals, fish, reptiles and amphibians. Inspection of reference sequences from distinct *Hepadnaviridae* showed that both consensus CTCF binding sites are conserved in viruses infecting primates and the majority of mammals and bats but are absent from viruses infecting birds, fish or amphibians, demonstrating evolutionary conservation of both CTCF binding sites (Figure 1b).

To assess whether these putative motifs can bind CTCF, we selected two independent lines, HepG2.2.15 (Sells et al., 1987) and HepAD38 (Ladner et al., 1997), that carry integrated copies of HBV genomes, maintain cccDNA and generate infectious virus. We isolated

chromatin from nuclear fractions to limit contamination of cytoplasmic rcDNA and performed anti-CTCF chromatin immunoprecipitation (ChIP) followed by quantitative PCR (ChIP-qPCR). Primers were selected to amplify 100–200 base pair regions of the viral genome to provisionally identify CTCF binding sites. We show low-level CTCF binding above the control IgG across the viral DNA with a significant enrichment in the Xp in both cell lines (Figure 2a), consistent with our motif scanning results. These data are in line with earlier reports for CTCF binding cellular target genes (Zhang et al., 2017). Analysing histone modifications of HBV chromatin purified from HepG2.2.15 cells showed minimal evidence for the repressive H3K27Me3, which is in agreement with previous reports (Flecken et al., 2019; Tropberger et al., 2015) (Figure 2b). In contrast, ChIP for histone marks associating with active transcription, including H4Ac and H3K4Me3, identified these epigenetic marks throughout the viral genome, with an enrichment in the BCP and Xp regions (Figure 2b).

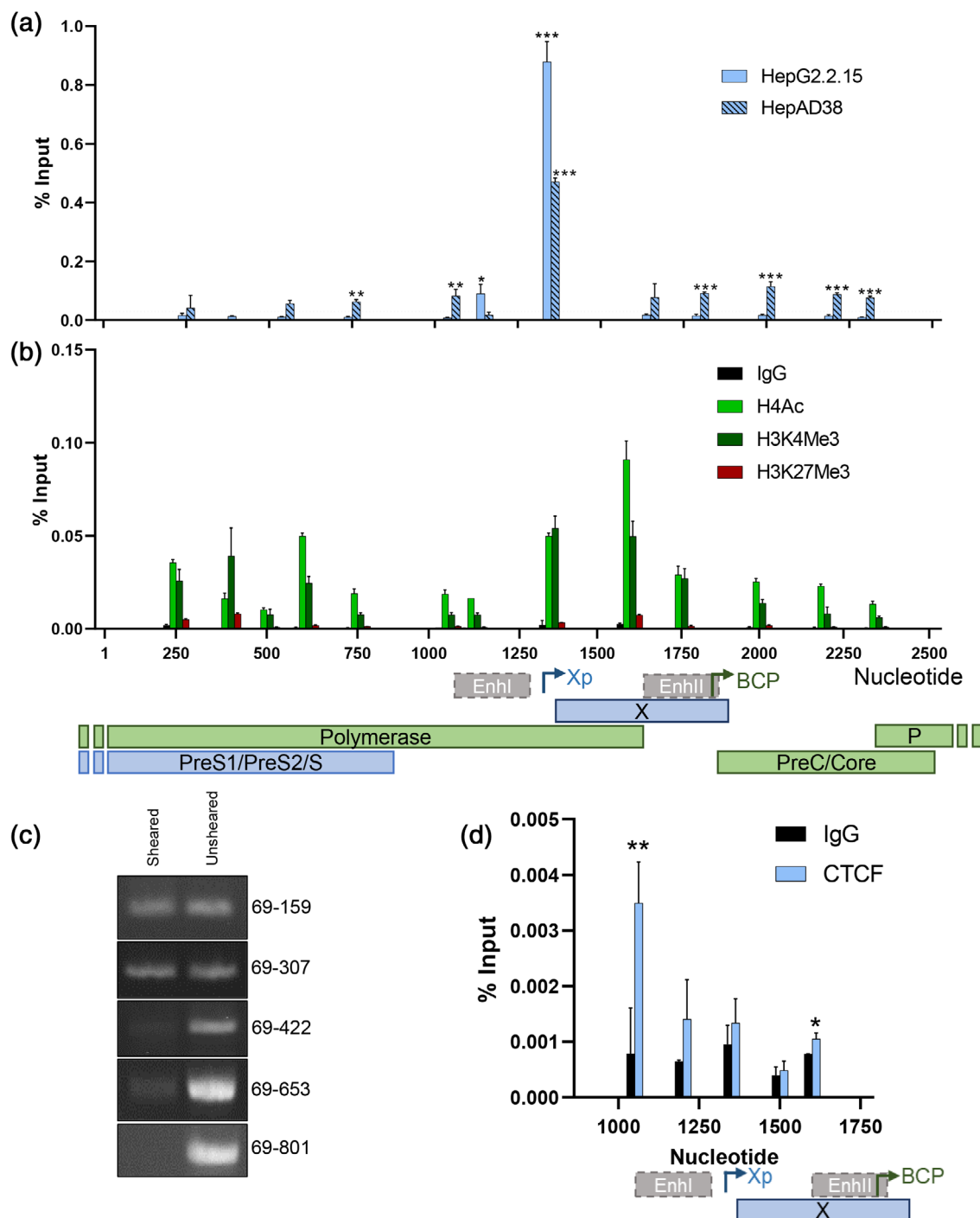
Since both HepG2.2.15 and HepAD38 cell lines carry integrated viral genomes and cccDNA, we are unable to discriminate CTCF binding between these forms of viral DNA. We studied HepG2 cells expressing an episomal copy of HBV DNA (HepG2-HBV-Epi) (Lucifora et al., 2014) to evaluate whether CTCF can bind episomal non-integrated copies of HBV DNA. Our initial experiments assessed whether our sonication protocol sheared HBV DNA by PCR amplification using primers for viral targets of increasing length pre- and post-sonication. While the unsonicated chromatin yielded a series of PCR products of increasing length, only amplicons below 238 base pairs were detected in the sonicated material (Figure 2c). Amplicons over 353 base pairs were barely visible in the sonicated samples, demonstrating effective shearing of episomal HBV genomes. ChIP of sheared chromatin, isolated from HepG2-HBV-Epi nuclear extracts, showed CTCF bound to the EnhI region (Figure 2d). We noted relatively lower ChIP of viral DNA from the HepG2-HBV-Epi cells compared to HepG2.2.15 or HepAD38 cells, which may reflect differences in the epigenetic status of the viral DNA in these model systems. Our observation that CTCF binds EnhI, the major transcriptional regulatory element of the BCP and Xp, suggests that CTCF regulates its activity.



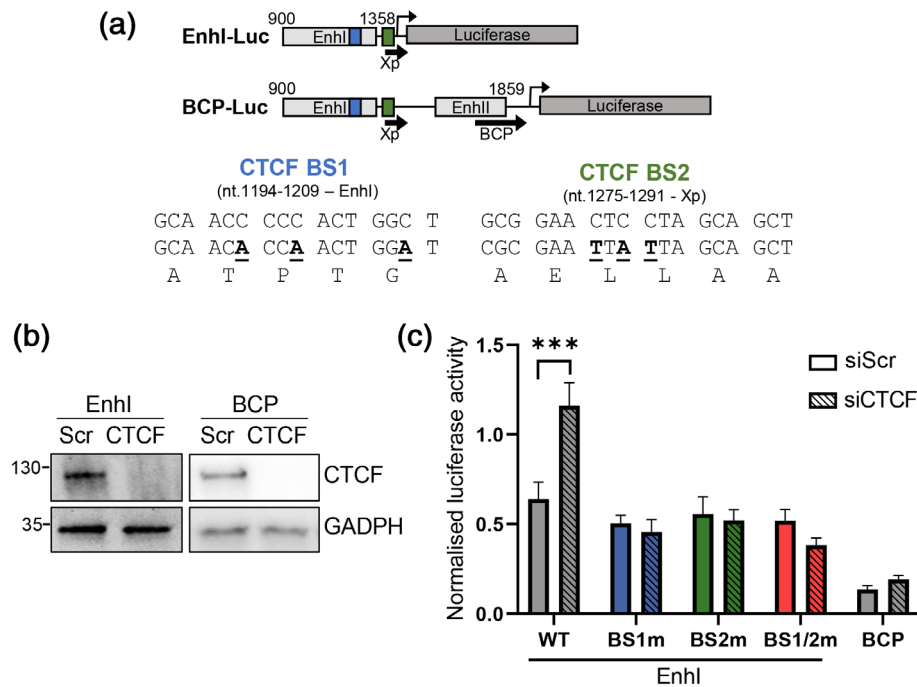
**FIGURE 1** Identification of conserved CTCF binding sites in HBV genomes and diverse *Hepadnaviridae*. (a) Conservation of CTCF BS among 7,313 HBV sequences (HBVdb.fr). All sites, except where indicated, are >98% conserved. (b) Neighbour-joining phylogenetic tree of members of the *Hepadnaviridae* (adapted from McNaughton et al., 2019). The green box shows viral genomes that encode both CTCF BS1 and 2 (all human and old world primate viruses), whereas the blue box shows viral genomes encoding only CTCF BS1 (new world monkeys, woodchucks and all bats except the tent making bat)

## 2.2 | CTCF represses HBV enhancer I

To analyse the role of CTCF in regulating HBV enhancer activity, we used promoter constructs encoding Firefly luciferase under the control of EnhI and Xp (nt 900–1358) or the BCP (nt 900–1859) (Figure 3a) (Ko, Lee, Windisch, & Ryu, 2014). We silenced CTCF in HepG2-NTCP using an siRNA Smartpool (Figure 3b) and transfected the viral promoter plasmids along with a *Renilla* luciferase control plasmid and measured the activity after 72 hr. Transient knock-down of CTCF protein significantly increased HBV EnhI activity (Figure 3c). To assess whether the putative CTCF directly regulated EnhI, we introduced silent mutations into the pEnhI-Luc to abrogate CTCF binding (Schmidt et al., 2012), without altering the polymerase protein sequence as this would adversely affect subsequent experiments with intact HBV genomes (Figure 3a). Mutation of either CTCF BS1 (BS1m)



**FIGURE 2** CTCF associates with HBV DNA and is enriched at viral Enhancer I and X promoter. (a) Association of CTCF with HBV DNA in HepG2.2.15 and HepAD38 cells was analysed by ChIP-qPCR and presented as % Input recovery. Statistical significance shows comparison of CTCF-specific ChIP with maximal recovery using IgG control (dotted line). (b) The distribution of histone modifications (H3K4Me3, H3K27Me3 and H4Ac) in HepG2.2.15 cells by ChIP-qPCR. (c) Chromatin shearing in HepG2-HBV-Epi cells was assessed by PCR of sonicated versus non-sonicated chromatin. Amplicons were generated with a constant sense primer (anneals at nt 69) and anti-sense primers binding at increasing distance from the sense primer (nt 159, 307, 422, 653 and 801). Amplification of HBV DNA was assessed by SyBr green staining of bands separated by electrophoresis. (d) Association of CTCF was assessed by ChIP-qPCR. (A, B and D) Data shown are the mean  $\pm$  SEM of three technical repeats and are representative of three biological repetitions. *p* values were determined using a paired *t* test. \*denotes *p* < .05, \*\*denotes *p* < .01, \*\*\*denotes *p* < .001. Annotation of HBV genome features, including open reading frames, enhancers and selected promoters, is shown below the histograms



**FIGURE 3** CTCF represses HBV Enhancer I activity. (a) Depiction of HBV genome regions cloned upstream of Firefly luciferase in transcriptional reporter plasmids and mutagenesis strategy of CTCF BS1 and BS2 showing viral enhancers, Xp and BCP, and CTCF BS 1 (blue) and CTCF BS 2 (green). (b) Western blot showing depletion of CTCF after siRNA transfection in pEnhI-Luc and pBCP-Luc transfected HepG2-NTCP cells. (c) Firefly luciferase activity normalised to *Renilla* luciferase expression in HepG2-NTCP cells co-transfected with pGL3-basic, pEnhI-Luc or pBCP-Luc and either scrambled (Scr) or CTCF-specific siRNA duplexes. Normalised luciferase activity in HepG2-NTCP cells transfected pEnhI-Luc containing mutations in CTCF binding site 1 (BS1m) or 2 (BS2m) or a combination of both (BS1/2m). Data shown are the mean  $\pm$  SEM of three independent repetitions.  $p$  values were determined by the Sidak's ANOVA multiple comparisons test. \*\*\*denotes  $p < .001$

or BS2 (BS2m) in isolation or in combination (BS1/2m) abrogated the increase in EnhI activity after CTCF depletion (Figure 3c). Silencing CTCF protein showed a minimal effect on the BCP activity, suggesting that CTCF represses EnhI but this effect may be blunted in the presence of an NRE and overall reduced activity in the full transcriptional reporter construct (Figure 3c). Together, these data suggest that CTCF binds to both motifs within EnhI to directly repress its activity.

### 2.3 | Silencing CTCF increases HBV preC/pgRNA levels

To determine the effect of CTCF depletion on viral transcripts, we selected to use the HepG2-HBV-Epi cells as we previously demonstrated CTCF binding to the viral genome in these cells. We confirmed effective knock-down of CTCF at the protein and RNA level 72 hr post-siRNA transfection (Figure 4a). We observed a significant increase in total HBV transcripts and preC/pgRNA levels after CTCF depletion (Figure 4b,c). To determine whether the observed increase in preC/pgRNA levels was due to an alteration of the HBV epigenome after CTCF depletion, we measured H4Ac modification of viral DNA as this was previously reported to associate with HBV transcription (Pollicino et al., 2006). Silencing of CTCF in HepG2-HBV-Epi cells increased H4Ac abundance within the viral enhancers, BCP Xp and

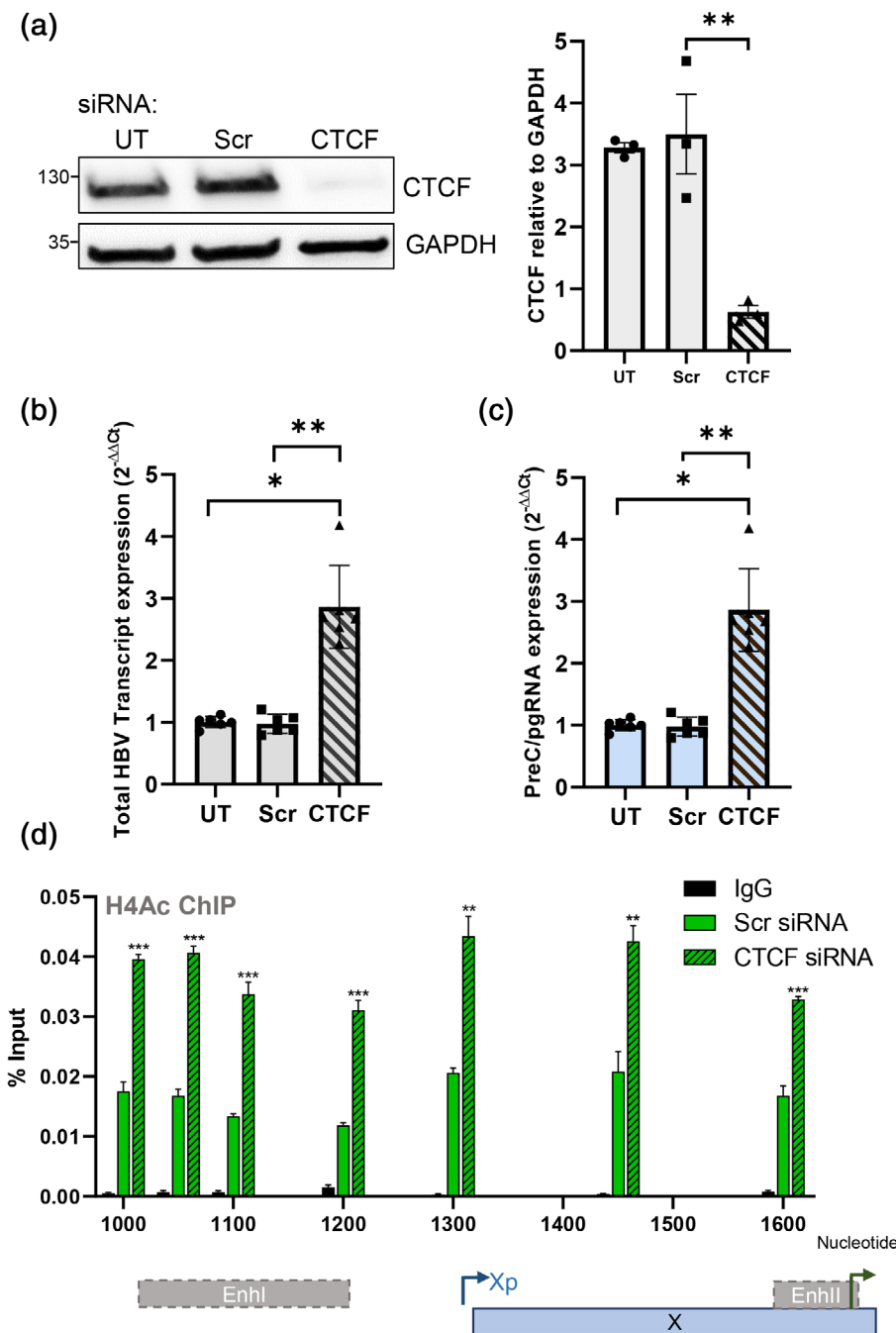
BCP, suggesting that CTCF regulates the epigenetic status of HBV cccDNA (Figure 4d).

To extend our studies and to validate a role for CTCF in repressing viral transcription during a de novo infection, we silenced CTCF in HBV infected HepG2-NTCP cells (Figure 5a). Efficient depletion of CTCF was demonstrated by western blotting (Figure 5a), and viral RNAs were analysed by RT-qPCR. In agreement with our earlier data with HepG2-HBV-Epi cells, CTCF depletion in this de novo infection model increased total transcripts and preC/pgRNA levels (Figure 5b,c). We previously reported a qPCR technique to quantify the relative abundance of HBV RNAs (D'Arienzo et al., 2019), and we used this method to assess the effect of CTCF silencing on the pattern of viral RNAs, showing no significant differences in the pattern of preS1, preS2 and HBx RNAs (Figure 5d). These data support a model where CTCF represses HBV cccDNA transcription, the major transcriptional template in de novo infected HepG2-NTCP cells. Taken together, our findings provide evidence that CTCF represses the BCP activity and hence preC/pgRNA levels.

### 2.4 | Mutation of CTCF binding sites within HBV enhancer I increases transcription

To demonstrate a direct role for CTCF binding to, and regulating, cccDNA transcription, we utilised the HBV minicircle (mcHBV)





**FIGURE 4** CTCF represses preC/pgRNA transcription from HBV cccDNA. (a) HepG2-HBV-Epi cells were untransfected (UT) or transfected with scrambled (Scr) or CTCF-specific siRNA duplexes and incubated for 72 hr. CTCTF depletion was assessed by western blotting and quantification in three independent experiments shown. (b) Total viral RNA and (c) preC/pgRNA levels were quantified by qPCR as previously described (D'Arienzo et al., 2019). Data are the mean  $\pm$  SD of two independent experiments performed in triplicate. *p* values were determined by the Kruskal-Wallis ANOVA multiple group comparison. (d) Enrichment of H4Ac marks was assessed by ChIP-qPCR and shown as % Input recovery. *p* values were determined using a paired *t* test. \*denotes  $p < .05$ , \*\*denotes  $p < .01$ , \*\*\*denotes  $p < .001$

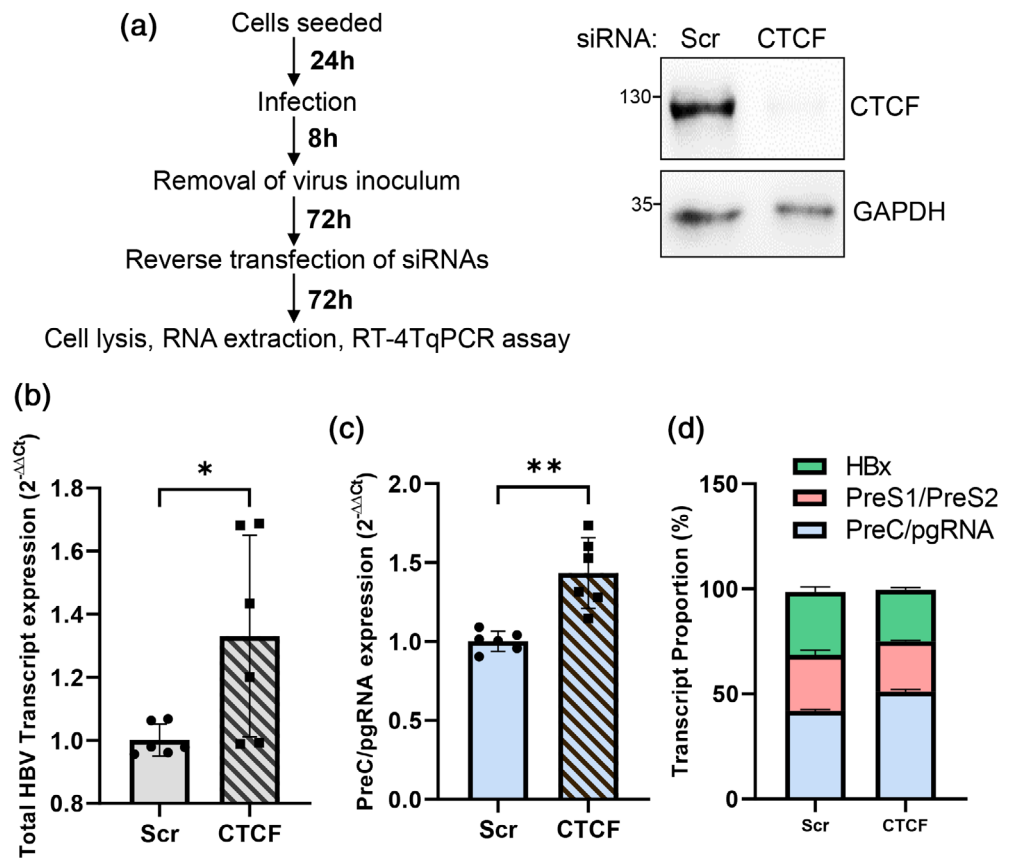
technology, as a model to study cccDNA transcription and replication (Yan et al., 2017). We mutated CTCF BS1 and BS2 alone or in combination in the mCHBV as described in Figure 3a. HepG2-NTCP cells were transfected with wild-type mCHBV (WT) or mutant mCHBV; BS1m, BS2m or BS1/2m, and harvested 3 days after transfection (Figure 6a). Analysis of CTCF binding by ChIP revealed that mutation of BS1 or BS2 alone significantly reduced CTCF binding by over >75% with the combined mutation resulting in an almost complete loss of CTCF-mCHBV complexes (Figure 6b). qPCR analysis showed a significant increase in preC/pgRNA levels when either or both of the CTCF BS were mutated (Figure 6c). However, no differences were observed in HBV mcDNA levels, confirming comparable transfection

efficiencies (Figure 6d). These data provide evidence of direct recruitment of CTCF to HBV DNA and show a repressive role for CTCF in regulating HBV transcription.

### 3 | DISCUSSION

In this study, we identified two CTCF-binding motifs within transcription regulatory elements, EnhI and Xp, of the HBV genome. We demonstrate CTCF binding to HBV DNA by ChIP-qPCR in the region of these binding sites using various model systems that bear both integrated genomes and a cccDNA pool, or cells exclusively expressing

**FIGURE 5** CTCF represses HBV preC/pgRNA transcription in de novo infected HepG2-NTCP cells. (a) HBV infected HepG2-NTCP were transfected with scrambled (Scr) or CTCF-specific siRNA duplexes and cultured for 72 hr and CTCF depletion assessed by western blotting. (b) Total HBV transcript abundance, (c) preC/pgRNA levels and (d) the relative proportion of individual HBV transcripts were analysed by PCR as previously described (D'Arienzo et al., 2019). Data are the mean  $\pm$  SD of two independent experiments performed in triplicate. *p* values were determined using the Mann-Whitney *U* test (two group comparisons). \*denotes *p* < .05, \*\*denotes *p* < .01



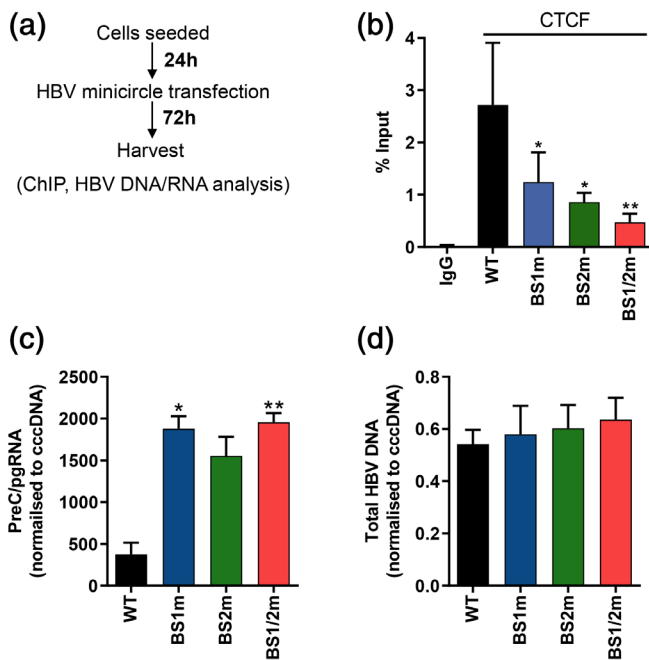
episomal copies of viral DNA. Our sonication method sheared cccDNA-like episomes and demonstrated CTCF binding to EnhI, albeit slightly upstream of the peak enrichment of CTCF binding in the integrant lines. This altered location of CTCF binding could reflect differential usage of CTCF binding sites in the different model systems or could reflect less efficient shearing of cccDNA-like molecules compared to integrated HBV DNA. Nonetheless, our ChIP-qPCR experiments allowed provisional mapping of CTCF binding sites that were confirmed by mutagenesis studies using promoter reporter constructs and mCHBV DNA. Importantly, these CTCF binding sites are conserved among all HBV genotypes and across the wider *Hepadnaviridae* family, consistent with an evolutionary conserved role in the replication of these viruses. Finally, we show a role for CTCF to repress HBV transcription.

Using several complementary HBV replication models, we show that siRNA depletion of CTCF and mutation of CTCF binding sites significantly increased preC/pgRNA levels, consistent with a role for CTCF in repressing viral transcription. To understand the mechanism of CTCF action, we used transcriptional reporter assays and found that silencing CTCF significantly increased EnhI activity. Furthermore, mutating the CTCF BS within EnhI attenuated this phenotype, confirming a direct role for CTCF in regulating EnhI. However, analysis of the full BCP, containing both EnhI and EnhII, revealed that the phenotype of CTCF silencing was lost. It is likely that the attenuation of BCP activity after CTCF silencing is explained by the dominant repressive effects of the NRE within

EnhII, highlighting the context-dependent activity of CTCF in regulating HBV. However, increased activity of the BCP is observed after CTCF silencing in cells containing the full viral episome, which may reflect differential chromatinization and epigenetic modification of the transcriptional reporters as compared to the full viral episome. Alternatively, the transcriptional elements in isolation are no longer subject to regulation by distal elements contained within the intact episome. While the transcriptional reporters used in this study provide a useful tool in the initial characterisation of CTCF function in HBV EnhI modulation, the results obtained in the context of a chromatinised viral episome may better reflect the role of CTCF in the HBV infection cycle.

To confirm a direct role of CTCF in repressing HBV transcription, we transfected HepG2-NTCP cells with mCHBV mutated in the CTCF BSs. Although the extent to which we could mutate CTCF BS was limited, to maintain the amino acid sequence of the polymerase, we observed a significant reduction of CTCF binding to mCHBV, lacking either BS1 or BS2, or both sites mutated in combination. These studies identify CTCF BSs within the viral genome and confirm CTCF association with HBV DNA. Consistent with the increased preC/pgRNA levels observed in two HBV replication model systems after CTCF depletion, we observed a significant increase in preC/pgRNA when CTCF BS1 was mutated. A similar increase in preC/pgRNA was observed when CTCF BS2 was mutated, although this did not reach statistical significance. While the mutation of both BS showed a significant increase in preC/pgRNA abundance, suggesting





**FIGURE 6** Mutation of CTCF binding sites in HBV mcDNA results in increased preC/pgRNA levels. (a) HepG2-NTCP cells were transfected with wild type HBV mcDNA (WT) or mcDNA with CTCF binding 1 (BS1m) or 2 (BS2m) or both sites mutated in combination (BS1/2m). (b) Cells were harvested 72 hr post transfection and CTCF binding analysed by ChIP-qPCR and presented as % of enrichment relative to input chromatin (% input). (c) preC/pgRNA and levels were quantified by qRT-PCR and normalised to cccDNA amount per cell and (d) total HBV DNA levels were quantified by qRT-PCR and normalised to cccDNA amount per cell to determine mcHBV transfection efficiency. Data are the mean  $\pm$  SEM of at least three independent experiments. *p* values were determined using the Kruskal–Wallis ANOVA multiple group comparison. \*denotes *p* < .05, \*\*denotes *p* < .01

these sites do not function in a synergistic manner within this model system.

Aberrant reverse transcription of pgRNA can generate double-stranded linear DNA that can integrate into the host genome (Tu, Budzinska, Shackel, & Urban, 2017). This integration step is not part of the productive HBV life cycle and occurs at a low frequency (<1 copy per diploid host genome in infected tissues) (Podlaha et al., 2019). Since integrated copies of HBV DNA generally lack a functional basal core promoter and associated CTCF binding sites, we would anticipate a minimal role for CTCF in directly regulating integrant derived transcripts. HBV integration can cause host genomic instability leading to tumour progression through tumour suppressor gene inactivation and/or oncogene activation (Zhao et al., 2016). Oncogenic integration events are thought to provide a growth advantage to cells, inducing tumourigenesis. HBV integration occurs at random sites, although a preference for integration within regions of open chromatin has been reported (Furuta et al., 2018). It will be interesting to determine whether integration of HBV DNA into the host results in an alteration of local chromatin interactions

and host cell gene regulation by the insertion of a virally encoded CTCF binding site(s), as reported for the human retrovirus, HTLV-1 (Melamed et al., 2018). Such genomic rearrangements could have an impact on host cell gene expression and contribute to HBV-driven carcinogenesis.

Analysing the epigenetic status of HBV DNA in HepG2.2.15 hepatoma cells revealed a lack of the repressive H3K27Me3 and enrichment of epigenetic marks associated with active transcription in the BCP and Xp regions, downstream of the CTCF binding sites. We noted a similar enrichment of H4Ac in episomal DNA in HepG2-HBV-Epi cells. These findings are consistent with previous reports studying the epigenetic status of HBV cccDNA in various model systems and liver biopsy samples (Flecken et al., 2019; Tropberger et al., 2015). Silencing of CTCF resulted in an increase in H4Ac abundance in HBV cccDNA, which associates with increased HBV preC/pgRNA levels. CTCF has been reported to directly repress transcription via recruitment of the Sin3/histone deacetylase compressor complex resulting in reduced histone acetylation (Lutz et al., 2000) that may explain these observations.

Taken together, these findings suggest that CTCF represses HBV transcription by insulating the BCP from the upstream enhancer element. Enh1 is an important regulator of all HBV promoters and is essential for viral transcription (Hu & Siddiqui, 1991). In support of this, HBV-transgenic mice, lacking Enh1, are defective in virion production (Guidotti, Matzke, Schaller, & Chisari, 1995). The repression of Enh1 by CTCF is likely to have a significant impact on the virus life cycle and reduce particle genesis and may therefore limit cccDNA pools. To assess whether infection perturbs CTCF expression, we analysed a publically available Affymetrix microarray database from chronic HBV infected patients (Zhou et al., 2017). We observed comparable CTCF transcript levels in normal and chronic HBV infected samples (Figure S1) with no evidence for HBV infection to perturb intra-hepatic CTCF transcript levels.

Analysis of the genomic distribution of CTCF BS in the human genome suggests a similar enhancer-blocking activity of CTCF as numerous CTCF binding loci are situated between known transcriptional enhancers and associated promoter elements (Xie et al., 2007). Such enhancer-blocking activity has been extensively characterised at imprinted loci, such as the insulin-like growth factor 2 (IGF2)/H19 locus and in development at the  $\beta$ -globin locus (Bell, West, & Felsenfeld, 1999; Chung, Whiteley, & Felsenfeld, 1993). CTCF regulates herpes simplex virus differential transcriptional programmes during the lytic and latent phases of the viral life cycle through its enhancer-blocking activity (Washington et al., 2018). Our previous work with HPV showed that CTCF repressed transcription by stabilising an epigenetically repressed chromatin loop between the viral proximal enhancer and a distal CTCF binding site. However, this repression was not associated with direct binding of CTCF to the HPV enhancer, suggesting that HBV and HPV have evolved fundamentally different mechanisms of CTCF-dependent transcriptional repression.

## 4 | EXPERIMENTAL PROCEDURES

### 4.1 | Cell lines and antibodies

HepG2.2.15 (Sells, Chen, & Acs, 1987), HepAD38 (Ladner et al., 1997), HepG2-HBV-Epi (Lucifora et al., 2014) and HepG2-NTCP cells were maintained in Dulbecco's Modified Eagles Medium (DMEM, #31966) supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, 1 mM sodium pyruvate, 50 U/ml penicillin/streptomycin and non-essential amino acids (all reagents from Invitrogen, UK). All cells were maintained in a 5% CO<sub>2</sub> atmosphere at 37°C. HepG2-HBV-Epi cells were kept at low passage to limit HBV DNA integration. The following primary antibodies were used: anti-CTCF (#61311), anti-H3K4Me3 (#39915), anti-H3K27Me3 (#39155) and anti-H4Ac (#39925) were all purchased from Active Motif (UK) and anti-GAPDH (SC-32233) was purchased from Santa Cruz.

### 4.2 | ChIP and quantitative PCR

HepG2.215, HepAD38 cells or HepG2-HBV-Epi cells were fixed with 1% formaldehyde (Sigma Aldrich) for 10 min at room temperature before quenching with 125 mM glycine. The cells were washed with ice cold PBS, containing EDTA-free protease inhibitors (Roche) and 5 mM sodium butyrate, and frozen at –80°C. Pellets were re-suspended in ChIP lysis buffer (Active Motif) supplemented with protease inhibitors and incubated on ice for 30 min. Cells were dounced 30 times using the tight pestle to release nuclei and centrifuged at 2500g for 10 min at 4°C. The supernatant was removed and discarded. Nuclei were re-suspended in shearing buffer (Active Motif) pulse sonicated using a Sonics Vibra Cell CV18 sonicator fitted with a micro-probe at 25% amplitude for 15 min on ice using 30 s on/off cycles. Chromatin samples were cleared by centrifugation and stored at –80°C.

Sonication of HBV cccDNA was evaluated by conventional PCR amplification of increasing amplicon size using a constant sense primer and anti-sense primers described in Table 1. Phenol-chloroform extracted DNA from HepG2-HBV-Epi cells before and after sonication was quantified using a NanoDrop ND-1000 spectrophotometer. PCR reactions included 100 ng DNA, MyTaq Red PCR Mix (Bioline) and 200 nM sense/anti-sense primers and amplification following 35 cycles of 95°C, 15 s; 55°C, 15 s; 72°C, 30 s assessed by agarose gel electrophoresis. Products were visualised using SyBr Green Safe dye (Invitrogen).

For ChIP, sonicated lysates were clarified by centrifugation at 16,000g for 10 min and CTCF or histone complexes immunoprecipitated with 5–8 µg antibody using a ChIP-IT® Express Chromatin Immunoprecipitation kit, including Protein A magnetic beads as per manufacturer's instructions (Active Motif, USA). The input and immunoprecipitated DNA were quantified by real-time PCR using a Stratagene MX3500P PCR System. The values were calculated as %

recovery respective to input DNA signals. All oligonucleotide sequences are listed in Table 1.

### 4.3 | siRNA transfection

Cells were trypsinized to reverse transfect with 25 nM of CTCF-specific or scrambled TARGETplus Smartpool siRNAs (Horizon, USA) using DharmaFECT4 (20% of amount recommended by the manufacturer's protocol; ThermoFisher). Cells with no siRNA (un-treated; UT) were also assayed to assess lethality of CTCF depletion.

### 4.4 | SDS-PAGE and western blots

Cells were lysed in urea lysis buffer (8 M urea, 150 mM NaCl, 20 mM Tris, pH 7.5, 0.5 M β-mercaptoethanol) supplemented with protease inhibitor cocktail (Roche) and sonicated for 10 s at 20% amplitude using a Sonics Vibra Cell sonicator fitted with a micro-probe. Following quantification of protein concentration by Bradford assay, samples were diluted in Lameli buffer before incubating at 95°C for 5 min. Proteins were separated on a 10% polyacrylamide gel and transferred to PVDF membranes (Amersham). The membranes were blocked in TBS-T, 5% skimmed milk, and proteins were detected using specific primary (diluted at 1:1000) and HRP-secondary antibodies (ThermoFisher, diluted at 1:10,000). Protein bands were detected using Pierce SuperSignal West Pico chemiluminescent substrate kit (Pierce) and images were collected using a Fusion FX Imaging system (PiqLab).

### 4.5 | HBV transcription reporter assays

1 × 10<sup>5</sup> HepG2-NTCP cells were seeded in collagen-coated 24-well plates. Immediately following cell seeding, transfection mixes were added, containing 100 ng of either pGL3b-Enhl, pGL3b-BCP or pGL3b-basic, 25 ng *Renilla* luciferase control plasmid (pCMV-*Renilla*), 25 nM scrambled or CTCF-specific siRNA and 1.5 µl Lipofectamine RNAiMAX™ (ThermoFisher) in 100 µl OptiMEM (ThermoFisher). Cells were incubated at 37°C, 5% CO<sub>2</sub> for 72 hr before being washed with PBS and 200 µl Passive Lysis buffer (Promega) added to each well. Samples were incubated at RT for 30 min with gentle rocking. Lysates were cleared by centrifugation and 20 µl of each added to a white 96-well microtiter plate. FireFly and *Renilla* luciferase activity were detected using the Dual-Luciferase® Reporter Assay (Promega) using a GloMAX®-Multi Detection system (Promega). 50 µl reagent added at a speed of 200 µl/s followed by mixing and 2 s delay. Integration time was 10 s with 1 read/well for Firefly luciferase detection. The same protocol was used for subsequent *Renilla* luciferase detection. Normalised luciferase activity was calculated by dividing Firefly luciferase activity by *Renilla* luciferase activity.

**TABLE 1** Detailing all primer sequences used

Primer pair	Forward (5'–3')	Reverse (5'–3')
T1	GGGGAACATAAGTCTCTAGCTACC	TTTAGGCCCATATTAGTGTGACA
T2	CAAGGTAGGAGCTGGAGCATTTC	GAGGCAGGAGGCGGATTTG
T3	CTCCAGTTTCAGGAACAGTAAACCC	AGGAATCCTGATGTGATGTTCTCC
T4	ACGGGGCGCACCTCTCTTTA	GTGAAGCGAAGTGCACACGG
β-actin	CCAACCGCGAGAAGATGA	CCAGAGGCGTACAGGGATAG
<b>Mutagenesis primers (Luc)</b>		
CTCF BS1	CTGGATGGGGCTTGGTCATGCGC	TTGGTGTGCGTCAGCAAACACTTGG
CTCF BS2	AGCAGCTTGTGTTTGTCTCGCAGC	AATAATCCCGCAGTATGGATCGG
<b>Mutagenesis primers (pMC-HBV)</b>		
CTCF BS1	GTGTTTGTCTGACGCAACACCAACT GGATGGGGCTTGGTC	GACCAAGCCCCATCCAGTTGGTGTGCGT CAGCAAACAC
CTCF BS2	GCCGATCCATACTGCGGAATTATT AGCAGCTTGTGTTTGTCTCGCAGCAGG	CCTGCTGCGAGCAAAACAAGCTGCTAATA ATTCGCGAGTATGGATCGGC
<b>HBV ChIP primers</b>		
178–307	TTCTAGGACCCCTTCTCGT	GGCCAAGACACACGGTAGTT
254–428	TCGTGGTGGACTTCTCTCAA	TGAGGCATAGCAGCAGGAT
346–422	TCCTGTCTCCAACTTGTC	AGCAGCAGGATGAAGAGGAA
462–562	GTGCCCCGTTTGTCTCTAATTC	GGAGGGATACATAGAGTTCTCTGA
518–653	GCCGAACCTGCATGACTACT	GCCGAACCTGCATGACTACT
718–801	CCCACTGTTTGGCTTTCAGT	CAGCGGTAAAAAGGGACTCA
995–1108	ACGAATTGTGGGTCTTTTGG	GTTGGCGAGAAAGTGAAAGC
1089–1154	GCTTTCACCTTCTCGCCAAC	AACGGGGTAAAGGTTTCAGGT
1305–1438	AGCAGGTCTGGAGCAAACAT	GACGGGACGTAAACAAAGGA
1581–1693	GTGCACTTCGCTTCACCTCT	GGTCGTTGACATTGCAGAGA
1738–1837	GGAGTTGGGGGAGGAGATTA	GGCAGAGGTGAAAAGTTGC
1901–2054	GCATGGACATCGACCCCTTAT	TGAGGTGAACAATGCTCAGG
2112–2297	CTGGGTGGGTGTTAATTTGG	TAAGCTGGAGGAGTGCGAAT
2279–2392	TTCGCACTCCTCCAGCTTAT	GAGGCGAGGGAGTTCTCTTT
2983–3133	ACAAGGTAGGAGCTGGAGCA	GTAGGCTGCCTTCTGTCTG
<b>HBV cccDNA shearing</b>		
F69–R159	CTCCAGTTTCAGGAACAGTAAACCC	AGGAATCCTGATGTGATGTTCTCC
R307		GGCCAAGACACACGGTAGTT
R422		AGCAGCAGGATGAAGAGGAA
R653		GCCGAACCTGCATGACTACT
R801		CAGCGGTAAAAAGGGACTCA

#### 4.6 | HBV de novo infection

Purified HBV was produced from HepAD38 cells as previously reported (Ko et al., 2018). HepG2-NTCP cells were seeded on collagen-coated plasticware and infected with HBV at an MOI of 250 genome equivalents per cell in the presence of 4% PEG 8,000. Viral inoculum was removed 8 hr after infection, by extensive washing with PBS, and cells were maintained in DMSO-free DMEM/10% FBS.

the samples were treated with RNase-free DNase I (14 Kunitz units/rxn, Qiagen) for 30 min at RT. RNA concentration and quality were assessed using a NanoDrop 1,000 spectrophotometer (ThermoFisher) and 2,100 Bioanalyzer (Agilent). cDNA synthesis was performed with 0.25–1 µg of RNA in a 20 µl total reaction volume using a random hexamer/oligo dT strand synthesis kit as per the manufacturer's instructions (10 min at 25°C; 15 min at 42°C; 15 min at 48°C; SensiFast, Bio-line). All oligonucleotide sequences are listed in Table 1.

#### 4.7 | RNA isolation for cDNA synthesis

Total cellular RNA was extracted using an RNeasy mini kit (Qiagen) following the manufacturer's protocol. To remove any residual HBV DNA,

#### 4.8 | Quantitative PCR of HBV transcripts

All PCR reactions were performed using a SYBR green real-time PCR protocol (qPCR BIO SyGreen, PCR Biosystems) in a Lightcycler 96™

instrument (Roche). The amplification conditions were: 95°C for 2 min (enzyme activation), followed by 45 cycles of amplification (95°C for 5 s; 60°C for 30 s). A melting curve analysis was performed on the completed reactions to assess specificity and purity of the amplicons (95°C for 10 s; 60°C for 60 s; followed by gradual heating from 60 to 97°C at 1°C/s). DNase-treated RNA samples that had not been reverse-transcribed were amplified to verify the absence of residual DNA contamination. All oligonucleotide sequences are listed in Table 1.

#### 4.9 | HBV mcDNA purification and transfection into cells

The plasmid, pMC-HBV, contains the 1.0 HBV genome (awy) and has been previously described (Yan et al., 2017). CTCF BS1 and CTCF BS2 were mutated by site-directed PCR mutagenesis using the primers detailed in Table 1 and Prime Star Max (Takara) mutagenesis kit following the manufacturer's protocols and confirmed by sequencing. ZCY10P3S2T competent bacteria (System Bioscience) were then transformed with the pMC-HBV (WT, BS1m, BS2m or BS1/2m) and a single colony amplified in Terrific Broth overnight at 37°C. Two volumes of LB medium supplemented with 0.04 N NaOH and 0.02% L-Arabinose were added to the culture and further incubated for 8 hr at 37°C. Plasmid DNA was extracted using the Nucleobond Xtra Maxi kit according to the manufacturer's protocol (Macherey-Nagel) and digested with *NdeI* (New England Biolabs) for 2 hr at 37°C and plasmid-safe DNase (System Bioscience) overnight at 37°C. After purification, plasmid DNA was assessed by agarose gel electrophoresis to check for elimination of the parental plasmid. HepG2-NTCP cells at 80–90% confluency were transfected with the pMC-HBV plasmids using TransIT-2020 (Mirus) according to the manufacturer's protocol in DMEM supplemented with 5% FBS, 1% Glutamax and 1% sodium pyruvate. The following day, cells were washed once with PBS and cultured for 72 hr in DMEM supplemented with 5% FBS, 1% Glutamax, 1% sodium pyruvate and 1% penicillin/streptomycin.

#### 4.10 | HBV nucleic acid quantification from mCHBV-transfected cells

Total DNA was extracted using MasterPure™ Complete DNA Purification Kit (Epicentre). Total RNA was extracted using ExtractAll TRI-Reagent (Sigma Aldrich), precipitated in isopropanol, washed in ethanol and re-suspended in RNase-free water. Extracted RNA was digested with RNase-free DNase I (Qiagen) and cDNA synthesised using SuperScript III reverse transcriptase (Invitrogen). cccDNA was quantified after *Exol* + *ExoIII* endonuclease (Epicentre) digestion of total extracted DNA for 2 hr at 37°C, followed by 20 min inactivation at 80°C. Real-time qPCR for total HBV DNA and cccDNA was performed using an Applied QuantStudio 7 machine (BioSystem) and TaqMan Advanced Fast Master Mix. Total HBV DNA was quantified using the TaqMan assay Pa03453406\_s1; cccDNA specific primers and probes were: forward 5'-CCGTGTGCACTTCGCTTCA-3'; reverse

5'-GCACAGCTTGGAGGCTTGA-3' TaqMan probe [6FAM]CATGGAGACCACCGTGAACGCCC[BBQ] (Testoni et al., 2019). Serial dilutions of a plasmid containing an HBV monomer (pHBV-EcoRI) served as quantification standard for total HBV DNA and cccDNA. The number of cellular genomes was determined by using the  $\beta$ -globin TaqMan assay Hs00758889\_s1 (ThermoFisher). preC/pgRNA was quantified using the following primers and probe: forward 5'-GGAGTGTGGATTGCACTCCT-3'; reverse 5'-AGATTGAGATCTCTGCGAC-3' and TaqMan probe [6FAM]AGGCAGGTCCCCTAGAA GAAGAACTCC[BBQ] (Testoni et al., 2019). Relative amount was normalised over the expression of housekeeping gene GUSB (Hs99999908\_m1).

#### 4.11 | Chromatin immunoprecipitation from infected HepG2-NTCP cells and mCHBV-transfected cells

Seventy two hours after de novo infection or mCHBV transfection, cells were washed twice with PBS and cross-linked with 1% formaldehyde for 10 minutes at 37°C. After 5 min of quenching with 125 mM glycine at 37°C, cells were washed twice with PBS, centrifuged for 5 min at 300g and incubated with Nuclear Lysis Buffer (5 mM PIPES, 85 mM KCl, 0.5% NP-40) for 30 min on ice to isolate nuclei. The lysate was then dounced 10 times and centrifuged for 5 min at 800g at 4°C. Nuclear membranes were then broken by two cycles of sonication 30 s ON, 30 s OFF on a Bioruptor (Diagenode). Debris were pelleted 10 min at 11,000g at 4°C. The supernatant was diluted 10 times with RIPA buffer (10 mM Tris-HCl pH 7.5, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-deoxycholate) supplemented with Complete Mini EDTA-free protease inhibitor (Roche Diagnostics) and 1 mM PMSF and pre-cleared for 2 hr at 4°C by adding magnetic Protein G Dynabeads (Life Technologies). Beads were discarded and 1  $\mu$ g of anti-CTCF antibody (Diagenode #C15410210) or isotype-matched negative control were added to the chromatin. After an overnight incubation at 4°C, magnetic Protein G Dynabeads and samples were incubated for 2 hr at 4°C with agitation. Beads were washed five times with RIPA buffer, once with TE buffer and re-suspended in Elution buffer (20 mM Tris-HCl pH 7.5, 5 mM EDTA, 50 mM NaCl, 1% SDS, 50  $\mu$ g/ml proteinase K). Chromatin was reverse cross-linked by incubation at 68°C for 2 hr and purified by phenol:chloroform:isoamyl alcohol 25:24:1 (Life Technologies) extraction and ethanol precipitation. cccDNA was quantified using the primers and probes listed above (Testoni et al., 2019).

#### ACKNOWLEDGEMENTS

We acknowledge Stephan Urban (University of Heidelberg, Germany) for providing HepG2-NTCP cells; Ulrike Protzer (TUM, Germany) for providing HepG2-HBV-Epi cells and Wang-Shick Ryu (Yonsei University, South Korea) for HBV promoter-transcription reporter plasmids. We thank Claudia Orbegozo Rubio for expert technical assistance and Chunkyu Ko for advice on HBV biology. BT laboratory is funded by the French Agence nationale de recherche sur le sida et les hépatites

virale (ANRS) ECTZ75178. JAM laboratory is funded by Wellcome Trust IA 200838/Z/16/Z and MRC project grant MR/R022011/1. JLP laboratory is funded by MRC project grants MR/R022011/1, MR/T015985/1 and MR/N023498/1. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

## CONFLICT OF INTEREST

None of the authors have any conflict of interest.

## AUTHOR CONTRIBUTIONS

**Valentina D'Arienzo, Jack Ferguson, Guillaume Giraud, Fleur Chapus, Peter A.C. Wing, James M. Harris, Adam Claydon, Sophia Begum and Xiaodong Zhuang:** Conducted experiments. **Peter Balfe:** Performed genetic analysis. **Barbara Testoni:** Designed experiments and edited manuscript. **Jane A. McKeating and Joanna L. Parish:** Designed experiments and co-wrote manuscript.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

Adam Claydon  <https://orcid.org/0000-0002-8070-0722>

Xiaodong Zhuang  <https://orcid.org/0000-0002-6870-9003>

Jane A. McKeating  <https://orcid.org/0000-0002-7229-5886>

Joanna L. Parish  <https://orcid.org/0000-0002-7111-4211>

## REFERENCES

- Bell, A. C., West, A. G., & Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*, 98(3), 387–396. [https://doi.org/10.1016/s0092-8674\(00\)81967-4](https://doi.org/10.1016/s0092-8674(00)81967-4)
- Belloni, L., Pollicino, T., De Nicola, F., Guerrieri, F., Raffa, G., Fanciulli, M., ... Levrero, M. (2009). Nuclear HBx binds the HBV minichromosome and modifies the epigenetic regulation of cccDNA function. *Proceedings of the National Academy of Sciences of the United States of America*, 106(47), 19975–19979. <https://doi.org/10.1073/pnas.0908365106>
- Chen, H., Tian, Y., Shu, W., Bo, X., & Wang, S. (2012). Comprehensive identification and annotation of cell type-specific and ubiquitous CTCF-binding sites in the human genome. *PLoS One*, 7(7), e41374. <https://doi.org/10.1371/journal.pone.0041374>
- Chen, H. S., Martin, K. A., Lu, F., Lupey, L. N., Mueller, J. M., Lieberman, P. M., & Tempera, I. (2014). Epigenetic deregulation of the LMP1/LMP2 locus of Epstein-Barr virus by mutation of a single CTCF-cohesin binding site. *Journal of Virology*, 88(3), 1703–1713. <https://doi.org/10.1128/JVI.02209-13>
- Chong, C. K., Cheng, C. Y. S., Tsoi, S. Y. J., Huang, F. Y., Liu, F., Fung, J., ... Wong, D. K. (2020). HBV X protein mutations affect HBV transcription and association of histone-modifying enzymes with covalently closed circular DNA. *Scientific Reports*, 10(1), 802. <https://doi.org/10.1038/s41598-020-57637-z>
- Chung, J. H., Whiteley, M., & Felsenfeld, G. (1993). A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell*, 74(3), 505–514. [https://doi.org/10.1016/0092-8674\(93\)80052-g](https://doi.org/10.1016/0092-8674(93)80052-g)
- D'Arienzo, V., Magri, A., Harris, J. M., Wing, P. A. C., Ko, C., Rubio, C. O., ... McKeating, J. A. (2019). A PCR assay to quantify patterns of HBV transcription. *The Journal of General Virology*. <https://doi.org/10.1099/jgv.0.001373>
- Filippova, G. N., Fagerlie, S., Klenova, E. M., Myers, C., Dehner, Y., Goodwin, G., ... Lobanenko, V. V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Molecular and Cellular Biology*, 16(6), 2802–2813. <https://doi.org/10.1128/mcb.16.6.2802>
- Flecken, T., Meier, M. A., Skewes-Cox, P., Barkan, D. T., Heim, M. H., Wieland, S. F., & Holdorf, M. M. (2019). Mapping the heterogeneity of histone modifications on hepatitis B virus DNA using liver needle biopsies obtained from chronically infected patients. *Journal of Virology*, 93(9), 1–21. <https://doi.org/10.1128/JVI.02036-18>
- Furuta, M., Tanaka, H., Shiraishi, Y., Unida, T., Imamura, M., Fujimoto, A., ... Nakagawa, H. (2018). Characterization of HBV integration patterns and timing in liver cancer and HBV-infected livers. *Oncotarget*, 9(38), 25075–25088. <https://doi.org/10.18632/oncotarget.25308>
- Guerrieri, F., Belloni, L., D'Andrea, D., Pediconi, N., Le Pera, L., Testoni, B., ... Levrero, M. (2017). Genome-wide identification of direct HBx genomic targets. *BMC Genomics*, 18(1), 184. <https://doi.org/10.1186/s12864-017-3561-5>
- Guidotti, L. G., Matzke, B., Schaller, H., & Chisari, F. V. (1995). High-level hepatitis B virus replication in transgenic mice. *Journal of Virology*, 69(10), 6158–6169.
- Hayer, J., Jadeau, F., Deleage, G., Kay, A., Zoulim, F., & Combet, C. (2013). HBVdb: A knowledge database for hepatitis B virus. *Nucleic Acids Research*, 41(Database issue), D566–D570. <https://doi.org/10.1093/nar/gks1022>
- Hong, X., Kim, E. S., & Guo, H. (2017). Epigenetic regulation of hepatitis B virus covalently closed circular DNA: Implications for epigenetic therapy against chronic hepatitis B. *Hepatology*, 66(6), 2066–2077. <https://doi.org/10.1002/hep.29479>
- Hu, K. Q., & Siddiqui, A. (1991). Regulation of the hepatitis B virus gene expression by the enhancer element I. *Virology*, 181(2), 721–726. [https://doi.org/10.1016/0042-6822\(91\)90906-r](https://doi.org/10.1016/0042-6822(91)90906-r)
- Kang, H., Wiedmer, A., Yuan, Y., Robertson, E., & Lieberman, P. M. (2011). Coordination of KSHV latent and lytic gene control by CTCF-cohesin mediated chromosome conformation. *PLoS Pathogens*, 7(8), e1002140. <https://doi.org/10.1371/journal.ppat.1002140>
- Kao, J. H., Chen, P. J., Lai, M. Y., & Chen, D. S. (2000). Hepatitis B genotypes correlate with clinical outcomes in patients with chronic hepatitis B. *Gastroenterology*, 118(3), 554–559. [https://doi.org/10.1016/s0016-5085\(00\)70261-7](https://doi.org/10.1016/s0016-5085(00)70261-7)
- Ko, C., Chakraborty, A., Chou, W. M., Hasreiter, J., Wettengel, J. M., Stadler, D., ... Protzer, U. (2018). Hepatitis B virus genome recycling and de novo secondary infection events maintain stable cccDNA levels. *Journal of Hepatology*, 69(6), 1231–1241. <https://doi.org/10.1016/j.jhep.2018.08.012>
- Ko, C., Lee, S., Windisch, M. P., & Ryu, W. S. (2014). DDX3 DEAD-box RNA helicase is a host factor that restricts hepatitis B virus replication at the transcriptional level. *Journal of Virology*, 88(23), 13689–13698. <https://doi.org/10.1128/JVI.02035-14>
- Kumar, S., & Bucher, P. (2016). Predicting transcription factor site occupancy using DNA sequence intrinsic and cell-type specific chromatin features. *BMC Bioinformatics*, 17(Suppl 1), 4. <https://doi.org/10.1186/s12859-015-0846-z>
- Ladner, S. K., Otto, M. J., Barker, C. S., Zaifert, K., Wang, G. H., Guo, J. T., ... King, R. W. (1997). Inducible expression of human hepatitis B virus (HBV) in stably transfected hepatoblastoma cells: A novel system for screening potential inhibitors of HBV replication. *Antimicrobial Agents and Chemotherapy*, 41(8), 1715–1720. <https://doi.org/10.1128/Aac.41.8.1715>
- Lang, F., Li, X., Vladimirova, O., Hu, B., Chen, G., Xiao, Y., ... Zhou, J. (2017). CTCF interacts with the lytic HSV-1 genome to promote viral



- transcription. *Scientific Reports*, 7, 39861. <https://doi.org/10.1038/srep39861>
- Levero, M., Pollicino, T., Petersen, J., Belloni, L., Raimondo, G., & Dandri, M. (2009). Control of cccDNA function in hepatitis B virus infection. *Journal of Hepatology*, 51(3), 581–592. <https://doi.org/10.1016/j.jhep.2009.05.022>
- Lucifora, J., Xia, Y., Reisinger, F., Zhang, K., Stadler, D., Cheng, X., ... Protzer, U. (2014). Specific and nonhepatotoxic degradation of nuclear hepatitis B virus cccDNA. *Science*, 343(6176), 1221–1228. <https://doi.org/10.1126/science.1243462>
- Lutz, M., Burke, L. J., Barreto, G., Goeman, F., Greb, H., Arnold, R., ... Renkawitz, R. (2000). Transcriptional repression by the insulator protein CTCF involves histone deacetylases. *Nucleic Acids Research*, 28(8), 1707–1713. <https://doi.org/10.1093/nar/28.8.1707>
- Lythgoe, K. A., Lumley, S. F., McKeating, J. A., & Mathews, P. C. (2020). Estimation of hepatitis B virus cccDNA persistence in chronic infection using within-host evolutionary rates. *MedRxiv*. Retrieved from <https://doi.org/10.1101/2020.02.04.20020362>
- McNaughton, A. L., D'Arienzo, V., Ansari, M. A., Lumley, S. F., Littlejohn, M., Revill, P., ... Mathews, P. C. (2019). Insights from deep sequencing of the HBV genome-unique, tiny, and misunderstood. *Gastroenterology*, 156(2), 384–399. <https://doi.org/10.1053/j.gastro.2018.07.058>
- Melamed, A., Yaguchi, H., Miura, M., Witkover, A., Fitzgerald, T. W., Birney, E., & Bangham, C. R. (2018). The human leukemia virus HTLV-1 alters the structure and transcription of host chromatin in cis. *eLife*, 7, 1–20. <https://doi.org/10.7554/eLife.36245>
- Ni, Y., Lempp, F. A., Mehrle, S., Nkongolo, S., Kaufman, C., Falth, M., ... Urban, S. (2014). Hepatitis B and D viruses exploit sodium taurocholate co-transporting polypeptide for species-specific entry into hepatocytes. *Gastroenterology*, 146(4), 1070–1083. <https://doi.org/10.1053/j.gastro.2013.12.024>
- Oropeza, C. E., Tarnow, G., Sridhar, A., Taha, T. Y., Shalaby, R. E., & McLachlan, A. (2020). The regulation of HBV transcription and replication. *Advances in Experimental Medicine and Biology*, 1179, 39–69. [https://doi.org/10.1007/978-981-13-9151-4\\_3](https://doi.org/10.1007/978-981-13-9151-4_3)
- Paris, C., Pentland, I., Groves, I., Roberts, D. C., Powis, S. J., Coleman, N., ... Parish, J. L. (2015). CCCTC-binding factor recruitment to the early region of the human papillomavirus 18 genome regulates viral onco-gene expression. *Journal of Virology*, 89(9), 4770–4785. <https://doi.org/10.1128/JVI.00097-15>
- Pena-Hernandez, R., Marques, M., Hilmi, K., Zhao, T., Saad, A., Alaoui-Jamali, M. A., ... Witcher, M. (2015). Genome-wide targeting of the epigenetic regulatory protein CTCF to gene promoters by the transcription factor TFII-I. *Proceedings of the National Academy of Sciences of the United States of America*, 112(7), E677–E686. <https://doi.org/10.1073/pnas.1416674112>
- Pentland, I., Campos-Leon, K., Cotic, M., Davies, K. J., Wood, C. D., Groves, I. J., ... Parish, J. L. (2018). Disruption of CTCF-YY1-dependent looping of the human papillomavirus genome activates differentiation-induced viral oncogene transcription. *PLoS Biology*, 16(10), e2005752. <https://doi.org/10.1371/journal.pbio.2005752>
- Pentland, I., & Parish, J. L. (2015). Targeting CTCF to control virus gene expression: A common theme amongst diverse DNA viruses. *Viruses*, 7(7), 3574–3585. <https://doi.org/10.3390/v7072791>
- Podlaha, O., Wu, G., Downie, B., Ramamurthy, R., Gaggar, A., Subramanian, M., ... Jiang, Z. (2019). Genomic modeling of hepatitis B virus integration frequency in the human genome. *PLoS One*, 14(7), e0220376. <https://doi.org/10.1371/journal.pone.0220376>
- Pollicino, T., Belloni, L., Raffa, G., Pediconi, N., Squadrito, G., Raimondo, G., & Levero, M. (2006). Hepatitis B virus replication is regulated by the acetylation status of hepatitis B virus cccDNA-bound H3 and H4 histones. *Gastroenterology*, 130(3), 823–837. <https://doi.org/10.1053/j.gastro.2006.01.001>
- Rowley, M. J., & Corces, V. G. (2018). Organizational principles of 3D genome architecture. *Nature Reviews. Genetics*, 19(12), 789–800. <https://doi.org/10.1038/s41576-018-0060-8>
- Schmidt, D., Schwalie, P. C., Wilson, M. D., Ballester, B., Goncalves, A., Kutter, C., ... Odom, D. T. (2012). Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell*, 148(1–2), 335–348. <https://doi.org/10.1016/j.cell.2011.11.058>
- Sells, M. A., Chen, M. L., & Acs, G. (1987). Production of hepatitis B virus particles in Hep G2 cells transfected with cloned hepatitis B virus DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 84(4), 1005–1009. <https://doi.org/10.1073/pnas.84.4.1005>
- Stadelmayer, B., Diederichs, A., Chapus, F., Rivoire, M., Neveu, G., Alam, A., ... Zoulim, F. (2020). Full-length 5'RACE identifies all major HBV transcripts in HBV-infected hepatocytes and patient serum. *Journal of Hepatology*, 73, 40–51. <https://doi.org/10.1016/j.jhep.2020.01.028>
- Sun, C. T., Lo, W. Y., Wang, I. H., Lo, Y. H., Shiou, S. R., Lai, C. K., & Ting, L. P. (2001). Transcription repression of human hepatitis B virus genes by negative regulatory element-binding protein/SON. *The Journal of Biological Chemistry*, 276(26), 24059–24067. <https://doi.org/10.1074/jbc.M101330200>
- Tang, H., Oishi, N., Kaneko, S., & Murakami, S. (2006). Molecular functions and biological roles of hepatitis B virus x protein. *Cancer Science*, 97(10), 977–983. <https://doi.org/10.1111/j.1349-7006.2006.00299.x>
- Testoni, B., Lebosse, F., Scholtes, C., Berby, F., Miaglia, C., Subic, M., ... Zoulim, F. (2019). Serum hepatitis B core-related antigen (HBcrAg) correlates with covalently closed circular DNA transcriptional activity in chronic hepatitis B patients. *Journal of Hepatology*, 70(4), 615–625. <https://doi.org/10.1016/j.jhep.2018.11.030>
- Tropberger, P., Mercier, A., Robinson, M., Zhong, W., Ganem, D. E., & Holdorf, M. (2015). Mapping of histone modifications in episomal HBV cccDNA uncovers an unusual chromatin organization amenable to epigenetic manipulation. *Proceedings of the National Academy of Sciences of the United States of America*, 112(42), E5715–E5724. <https://doi.org/10.1073/pnas.1518090112>
- Tu, T., Budzinska, M. A., Shackel, N. A., & Urban, S. (2017). HBV DNA integration: Molecular mechanisms and clinical implications. *Viruses*, 9(4), 1–18. <https://doi.org/10.3390/v9040075>
- Turton, K. L., Meier-Stephenson, V., Badmalia, M. D., Coffin, C. S., & Patel, T. R. (2020). Host transcription factors in hepatitis B virus RNA synthesis. *Viruses*, 12(2), 1–29. <https://doi.org/10.3390/v12020160>
- Urban, S., Schulze, A., Dandri, M., & Petersen, J. (2010). The replication cycle of hepatitis B virus. *Journal of Hepatology*, 52(2), 282–284. <https://doi.org/10.1016/j.jhep.2009.10.031>
- Washington, S. D., Musarrat, F., Ertel, M. K., Backes, G. L., & Neumann, D. M. (2018). CTCF binding sites in the herpes simplex virus 1 genome display site-specific CTCF occupation, protein recruitment, and insulator function. *Journal of Virology*, 92(8), e00156–e00118. <https://doi.org/10.1128/JVI.00156-18>
- Xie, X., Mikkelsen, T. S., Gnirke, A., Lindblad-Toh, K., Kellis, M., & Lander, E. S. (2007). Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proceedings of the National Academy of Sciences of the United States of America*, 104(17), 7145–7150. <https://doi.org/10.1073/pnas.0701811104>
- Yan, Z. P., Zeng, J., Yu, Y. J., Xiang, K. L., Hu, H., Zhou, X., ... Gao, L. (2017). HBVcircle: A novel tool to investigate hepatitis B virus covalently closed circular DNA. *Journal of Hepatology*, 66(6), 1149–1157. <https://doi.org/10.1016/j.jhep.2017.02.004>
- Yuh, C. H., & Ting, L. P. (1990). The genome of hepatitis B virus contains a second enhancer: Cooperation of two elements within this enhancer is required for its function. *Journal of Virology*, 64(9), 4281–4287.



- Zhang, B., Zhang, Y., Zou, X., Chan, A. W., Zhang, R., Lee, T. K.-W., ... Ko, B. C. (2017). The CCCTC-binding factor (CTCF)-forkhead box protein M1 axis regulates tumour growth and metastasis in hepatocellular carcinoma. *The Journal of Pathology*, 243(4), 418–430. <https://doi.org/10.1002/path.4976>
- Zhao, L. H., Liu, X., Yan, H. X., Li, W. Y., Zeng, X., Yang, Y., ... Wang, H. Y. (2016). Genomic and oncogenic preference of HBV integration in hepatocellular carcinoma. *Nature Communications*, 7, 12992. <https://doi.org/10.1038/ncomms12992>
- Zhou, W., Ma, Y., Zhang, J., Hu, J., Zhang, M., Wang, Y., ... Liu, J. (2017). Predictive model for inflammation grades of chronic hepatitis B: Large-scale analysis of clinical parameters and gene expressions. *Liver International*, 37(11), 1632–1641. <https://doi.org/10.1111/liv.13427>

## SUPPORTING INFORMATION

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

**How to cite this article:** D'Arienzo V, Ferguson J, Giraud G, et al. The CCCTC-binding factor CTCF represses hepatitis B virus enhancer I and regulates viral transcription. *Cellular Microbiology*. 2020;e13274. <https://doi.org/10.1111/cmi.13274>