Virology 415 (2011) 1–5

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Rapid Communication

Occupancy of chromatin organizers in the Epstein–Barr virus genome

Meghan M. Holdorf [a], Samantha B. Cooper [a,b,1], Keith R. Yamamoto [a], JJ L. Miranda [a,*]

[a] Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA
[b] Graduate Program in Biological and Medical Informatics, University of California, San Francisco, San Francisco, CA 94158, USA

ARTICLE INFO

Article history:
Received 22 January 2011
Returned to author for revision
11 February 2011
Accepted 15 April 2011
Available online 8 May 2011

Keywords:
CTCF
Chromatin
Transcription
Epstein–Barr virus
Cohesin
RNA Polymerase II

ABSTRACT

The human CCCTC-binding factor, CTCF, regulates transcription of the double-stranded DNA genomes of herpesviruses. The architectural complex cohesin and RNA Polymerase II also contribute to this organization. We profiled the occupancy of CTCF, cohesin, and RNA Polymerase II on the episomal genome of the Epstein–Barr virus in a cell culture model of latent infection. CTCF colocalizes with cohesin but not RNA Polymerase II. CTCF and cohesin bind specific sequences throughout the genome that are found not just proximal to the regulatory elements of latent genes, but also near lytic genes. In addition to tracking with known transcripts, RNA Polymerase II appears at two unannotated positions, one of which lies within the latent origin of replication. The widespread occupancy profile of each protein reveals binding near or at a myriad of regulatory elements and suggests context-dependent functions.

Introduction

The human CCCTC-binding factor, CTCF, regulates gene expression by organizing DNA in the three-dimensional space of the nucleus. CTCF binds tens of thousands of sites throughout the human genome (Barski et al., 2007; Kim et al., 2007) and is necessary for the assembly of chromatin loops that bring together distal DNA to regulate transcription (Hou et al., 2010; Majumder and Boss, 2010; Majumder et al., 2008; Mishiro et al., 2009). Globally, the human genome is organized by long-range interactions (Lieberman-Aiden et al., 2009) and binding sites are significantly enriched at these nodes (Botta et al., 2010). CTCF also organizes the double-stranded DNA genomes of herpesviruses. Binding sites have been identified on the Epstein–Barr virus (EBV) (Chau et al., 2006; Day et al., 2007; Tempera et al., 2010), herpes simplex virus 1 (Chen et al., 2007), and Kaposi's sarcoma-associated herpesvirus (KSHV) (Stedman et al., 2008). Functional studies with EBV have focused on binding sites near the major promoters of latency and suggested possible roles for CTCF in the repression, activation, and insulation of latent transcripts (Chau et al., 2006; Day et al., 2007; Tempera et al., 2010). CTCF occupancy upstream of the C promoter arguably correlates with repression of EBNA2 transcription (Chau et al., 2006; Salamon et al., 2009). Mutation of a CTCF-binding site upstream of the Q promoter resulted in loss of EBNA1 transcription accompanied by the spread of repressive histone marks and CpG methylation (Tempera et al., 2010). Based on a myriad of functions, CTCF has emerged as a central component of transcriptional regulation in human and viral genomes.

CTCF may form complexes with diverse binding partners that include transcription factors, histone modifiers, and a variety of chromatin regulators (Wallace and Felsenfeld, 2007; Zlatanova and Caiafa, 2009). The CTCF protein consists of eleven zinc fingers (Filippova et al., 1996) and two flanking unstructured termini (Martinez and Miranda, 2010); both types of protein segments are capable of recruiting cofactors directly by molecular recognition (Dunker et al., 2002; Gamsjaeger et al., 2007). Specifically, the architectural complex cohesin and RNA Polymerase II (RNAP II) may play functional roles in CTCF-dependent transcriptional control. CTCF is necessary for cohesin positioning and colocalization throughout the human, mouse, and KSHV genomes (Parelho et al., 2008; Stedman et al., 2008; Wendt et al., 2008). Depletion of cohesin perturbs gene expression by the disruption of DNA looping and long-range interactions between transcriptional control elements (Hadjur et al., 2009; Mishiro et al., 2009; Nativio et al., 2009). RNAP II is thought to interact with CTCF directly through protein–protein interactions, resulting in colocalization at a subset of CTCF-binding sites (Chernukhin et al., 2007). Stalling of RNAP II tends to occur at sites of CTCF and cohesin colocalization (Wada et al., 2009), and a large fraction of CTCF sites are found within actively transcribed genes (Barski et al., 2007). The diversity and genome-wide distribution of the myriad of possible CTCF assemblies, given significant functional implications, remains actively investigated.

We initially hoped to use profiling of protein occupancy to determine the composition of CTCF complexes by identifying which proteins bind the same DNA at high resolution. Colocalization would

* Corresponding author.
E-mail address: miranda@cmp.ucsf.edu (JJ L. Miranda).
[1] Present address: Illumina, Inc., Hayward, CA 94545.

0042-6822/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2011.04.004

strongly suggest molecular assembly and cooperation, perhaps characterizing the structural and functional heterogeneity or uniformity of complexes genome-wide. Indeed, we were able to do so and determined that CTCF colocalizes with cohesin but not RNAP II. In the midst of identifying binding sites, however, we also uncovered protein occupancy in unexpected regions of the EBV genome. We will discuss the functional implications of CTCF and cohesin colocalization as well as RNAP II binding outside regions of known transcription.

Results

CTCF and cohesin colocalization

CTCF binds specific sites throughout the latent EBV genome. We identified fifteen sites of CTCF occupancy at high resolution across the entire unique sequence of the EBV genome using chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) (Fig. 1, Table 1). An overrepresented sequence motif, similar to that found in the human genome (Kim et al., 2007), is detected in every region of occupancy (data not shown). During the life cycle of EBV, only a handful of regulatory elements control expression of a few latent genes, but many more genes are transcribed upon lytic activation (Kieff and Rickinson, 2007). We find CTCF-binding sites proximal to many latent regulatory elements. Occupancy is observed near the C, W, and Q promoters that drive messages encoding for the EBNA proteins. Enrichment is also detected close to start sites of the EBER-1 non-coding RNA and RPSM1 BART transcript, as well as the LMP-2A and LMP-1 messages. We also find, however, CTCF-binding sites proximal to many lytic genes. For example, occupancy is detected near the promoter of BZLF1, an immediate early transactivator critical for the switch to the productive cycle (Biggin et al., 1987). Both

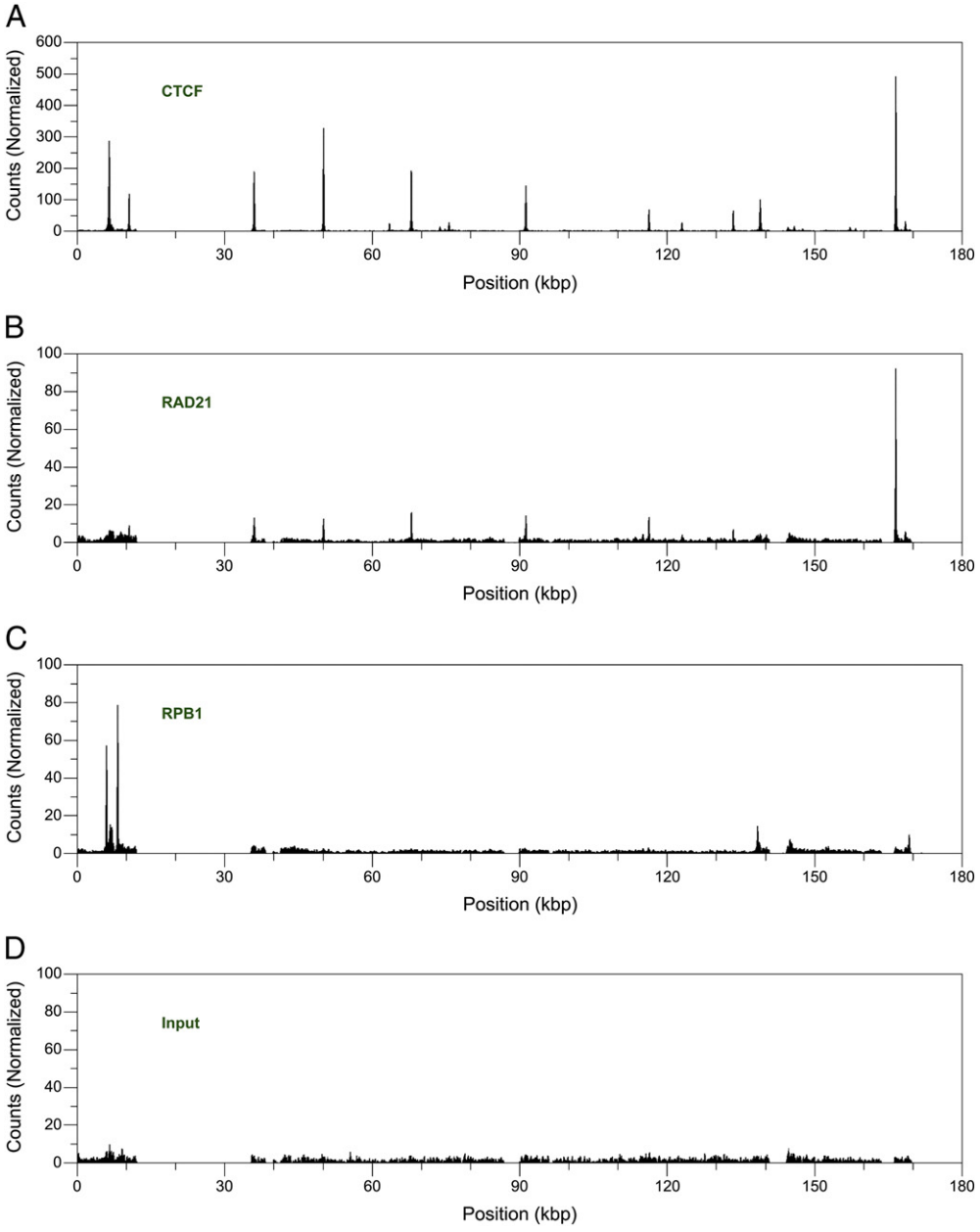

Fig. 1. Occupancy of CTCF, cohesin, and RNA Polymerase II in the Epstein–Barr virus genome. Chromatin immunoprecipitation profiles of (A) CTCF, (B) the RAD21 subunit of cohesin, and (C) the RPB1 subunit of RNA Polymerase II in the latent Epstein–Barr virus genome of Raji cells. The profile of (D) input DNA serves as a control reference. Occupancy is calculated as enrichment over the background baseline.