

RESEARCH ARTICLE

HIV chromatin is a preferred target for drugs that bind in the DNA minor groove

Clayton K. Collings^{1,2}, Donald W. Little III³, Samuel J. Schafer⁴, John N. Anderson^{5*}

1 Department of Pediatric Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA, United States of America, **2** Broad Institute of MIT and Harvard, Cambridge, MA, United States of America, **3** University of Michigan Medical School, Ann Arbor, MI, United States of America, **4** Department of Reproductive and Developmental Sciences, University of British Columbia, Vancouver, BC, Canada, **5** Department of Biological Sciences, Purdue University, West Lafayette, IN, United States of America

* andersjn@purdue.edu

OPEN ACCESS

Citation: Collings CK, Little DW, III, Schafer SJ, Anderson JN (2019) HIV chromatin is a preferred target for drugs that bind in the DNA minor groove. PLoS ONE 14(12): e0216515. <https://doi.org/10.1371/journal.pone.0216515>

Editor: Yamini Dalal, National Cancer Institute, UNITED STATES

Received: April 18, 2019

Accepted: December 3, 2019

Published: December 30, 2019

Copyright: © 2019 Collings et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are provided within the manuscript and its supporting information files. Accessions numbers for HIV-1 and HTLV-1 are given in the methods, and location of the MGWs for the human genome is referenced in the methods. In the supporting information files, MGWs (output from DNashapeR) for the HIV-1 and HTLV-1 genomes are given, and the MGWs for all pentamers are also given. AnTm and AT tract positions in HIV-1 genome (>= 4bp) are provided. Finally, the HaeIII and RsaI cut sites of HIV Fragment 1 along with the sequence itself are given.

Abstract

The HIV genome is rich in A but not G or U and deficient in C. This nucleotide bias controls HIV phenotype by determining the highly unusual composition of all major HIV proteins. The bias is also responsible for the high frequency of narrow DNA minor groove sites in the double-stranded HIV genome as compared to cellular protein coding sequences and the bulk of the human genome. Since drugs that bind in the DNA minor groove disrupt nucleosomes on sequences that contain closely spaced oligo-A tracts which are prevalent in HIV DNA because of its bias, it was of interest to determine if these drugs exert this selective inhibitory effect on HIV chromatin. To test this possibility, nucleosomes were reconstituted onto five double-stranded DNA fragments from the HIV-1 *pol* gene in the presence and in the absence of several minor groove binding drugs (MGBDs). The results demonstrated that the MGBDs inhibited the assembly of nucleosomes onto all of the HIV-1 segments in a manner that was proportional to the A-bias, but had no detectable effect on the formation of nucleosomes on control cloned fragments or genomic DNA from chicken and human. Nucleosomes preassembled onto HIV DNA were also preferentially destabilized by the drugs as evidenced by enhanced nuclease accessibility in physiological ionic strength and by the preferential loss of the histone octamer in hyper-physiological salt solutions. The drugs also selectively disrupted HIV-containing nucleosomes in yeast as revealed by enhanced nuclease accessibility of the in vivo assembled HIV chromatin and reductions in superhelical densities of plasmid chromatin containing HIV sequences. A comparison of these results to the density of A-tracts in the HIV genome indicates that a large fraction of the nucleosomes that make up HIV chromatin should be preferred in vitro targets for the MGBDs. These results show that the MGBDs preferentially disrupt HIV-1 chromatin in vitro and in vivo and raise the possibility that non-toxic derivatives of certain MGBDs might serve as a novel class of anti-HIV agents.

Funding: This work was supported by the Showalter Research Trust at Purdue and the Purdue Department of Biological Sciences. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

The tremendous genetic variation of HIV-1 is responsible for the appearance of drug and antibody resistant forms of HIV-1 that appear during infection, and this genetic swarm is a major obstacle to the treatment of AIDS and to the development of an HIV vaccine. Viral genetic heterogeneity also produces variation in virulence, replication rates, cell tropisms and other properties [1,2]. A drug designed to combat HIV should optimally target a viral feature that is conserved in this heterogeneous viral population, is critical for viral survival, acts on the integrated proviral genome, and is not found or is significantly depleted in the host cell. Although substantial progress has been made during the past three decades in the development of anti-HIV drugs which has culminated in the use of combination antiviral drug therapy for treatment of viremia, HIV-drug-resistant mutants have been found for all nucleoside analogs and nonnucleoside and protease inhibitors used in the treatment of HIV-infected individuals. The emergence of multidrug-resistant variants of HIV-1 is also a concern for the treatment of the disease in the future [3,4].

The major challenge in current HIV research is the development of methods to eliminate the replication competent provirus that has integrated as a double stranded DNA molecule within the host cell chromatin. This viral reservoir is resistant to the host's immune system and is resistant to the current drugs used to treat viremia. Upon session of antiviral drug therapy, competent proviral sequences are activated giving rise to new rounds of replication and the resumption of active infection. Little progress has been made during the past to eliminate or inactivate the replication competent provirus and there is no universal agreement as to the best strategies that should be used to approach the problem in the future [5,6].

The HIV genome and the genomes of other lentiviruses display an unusual variation in nucleotide composition being deficient in C but not G and rich in A but not U [7,8]. We have previously shown that this nucleotide compositional bias is responsible for the unusual composition of HIV proteins which are rich in the polar amino acids encoded by A-rich codons and depleted in amino acids encoded by C-rich codons [8]. The extreme amino acid compositions are characteristic of all major proteins of the viruses including the polypeptides that make up the hypervariable viral envelope and the conserved *pol* region that codes for reverse transcriptase and integrase. The global nature of these effects makes it likely that the variation in protein composition caused by the biased nucleotide frequencies is an important factor in determining the characteristic phenotype of HIV [8]. The A-bias may also be a driving force that promotes enhance genetic variation of HIV and to increase the DNA curvature of the integrated viral genome that is caused by oligo A-tracts arranged in a 10 bp periodicity [8,9]. As a group, these observations suggest that the conserved A-bias of the HIV genome is critical for viral phenotype and thus the overall A-bias may represent a novel and fixed target for anti-HIV agents.

The minor groove binding drugs (MGBDs) are of interest in clinical medicine because of their antiviral, antimicrobial and antitumor activities [10–12]. These ligands preferentially bind to AT-rich DNA sequences that are 4 or more bp in length. The biologically important targets of these drugs are poorly understood although many studies have focused on their abilities to inhibit the interaction of regulatory proteins to short oligonucleotide-length AT-rich sequences. However, the importance of these targets in dictating the biologically relevant targets of these drugs is questionable because of the high frequency of isolated AT-rich oligonucleotide length sequences in the human genome. We previously reported a novel specificity of the MGBDs that requires much longer regions of DNA containing multiple closely spaced oligo A-tracts [13,14]. It was demonstrated that the drugs inhibited the assembly of such unusual DNA segments that are at least 100 bp long into nucleosomes. This inhibitory action

may be relevant to HIV since the entire viral genome is A-rich and consequently nucleosomes containing HIV sequences may be targets for the MGBDs that are preferred over host sequences. The studies described in this report were designed to evaluate this possibility.

Materials and methods

Bioinformatics studies

For the bioinformatics studies, sequences and gene coordinates for the HIV-1 and HTLV-1 genomes, linked to GenBank accession numbers AF033819.3 and J02029.1, respectively, were acquired from NCBI. These sequences were used as input into the R software package, DNASHapeR [15], to obtain predictions for their minor groove widths (MGWs) at bp resolution (S1 and S2 Files). MGWs for the hg19 version of the human genome were obtained from the GBshape database [16]. MGWs for all pentamers were determined by inputting all possible 4,194,304 11-mer DNA sequences into DNASHapeR. Regardless of surrounding sequence context, all central pentamers and their reverse complements possessed the same predicted MGWs (S3 File).

Profiles of AnTm occurrence and narrow MGW frequency were computed using in-house Perl scripts, and all plotted values represent the centers of 201 or 1001 bp sliding windows. For Fig 1A and 1C, tetranucleotide contents were computed within 201 bp sliding windows across sequences in 1 bp steps. For Figs 1B and 2C, tetranucleotide contents were computed within 1001 bp sliding windows across sequences in 1 bp steps. For Fig 2D, narrow MGW frequencies were computed within 1001 bp sliding windows across sequences using an MGW cutoff of 3.63 Angstroms. For Fig 2E, narrow MGW frequencies were computed in 201 bp sliding windows across sequences using an MGW cutoff of 3 Angstroms.

For Fig 2F, the numbers of narrow MGWs were computed in all 201 bp sliding windows across the human genome, human cDNA, and the HIV *pol* gene using an MGW cutoff of 3 Angstroms. Note that for the whole human genome, 100 bp steps were used while 1 bp steps were used for human cDNA and the HIV *pol* gene. For Fig 2G, the center locations of all non-overlapping AnTm sequences in the HIV-1 and human genomes were determined, where $n + m \geq 4$ (S4 and S5 Files). In other words, an AnTm sequence like AAAATTTT would only be counted as one AnTm sequence, and the centers of AnTm sequences would define their locations. For odd-length AnTm sequences, the central nucleotide would designate its center location, and for even-length AnTm sequences, the internal central position was equal to its length divided by 2. After the center locations of all non-overlapping AnTm sequences were determined, the numbers of AnTm sequences were computed in all 201 bp sliding windows in 1 bp steps across the human genome, human cDNA, and the HIV *pol* gene. At the same time, the pairwise distances between non-overlapping AnTm sequences were also computed.

For Fig 2F and 2G, the significance of the differences in the fractional distributions were evaluated by Kolmogorov-Smirnov tests in R or by traditional methods. For Fig 2F, narrow MGW frequencies in all 201 bp windows for the 3 reference sequences were input into R, and the “ks.test” function revealed high D-statistic values and P values less than 2.2×10^{-16} for all comparisons. For Fig 2G, D-statistics were determined by calculating the maximum differences between corresponding cumulative distribution functions and were compared to D-critical values. D-critical values were calculated using the KS method’s empirical formula, which incorporates sample sizes and a coefficient that is a function of selected α value. For all comparisons, the results showed that the D-statistics were considerably greater than the D-critical values (Fig 2G). We opted to use this traditional method because the AnTm dataset size for the human genome was too large for input in R as 1 bp sliding windows steps were used.

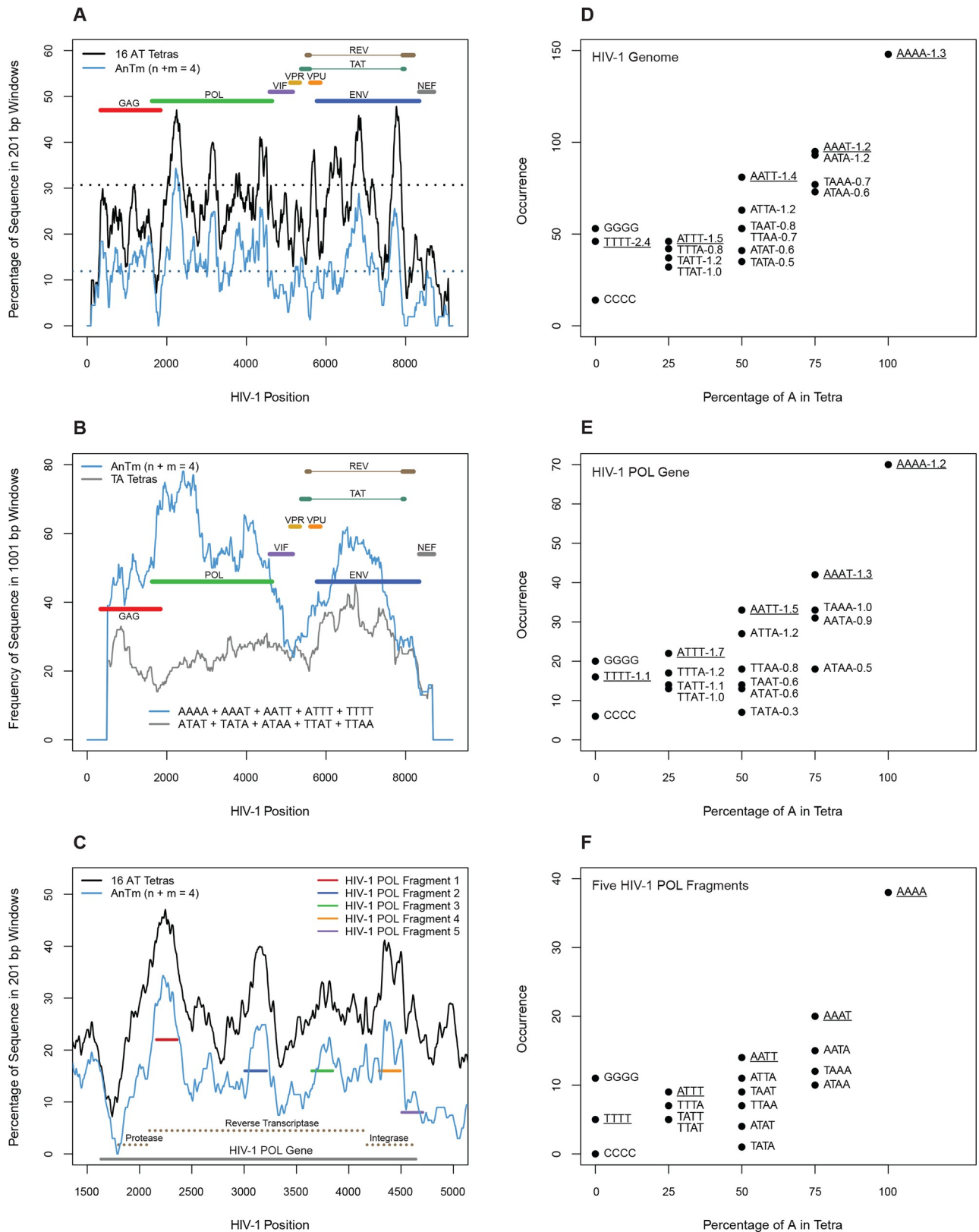


Fig 1. AT-tract density in HIV DNA. A) The percentages of the 16 AT tetranucleotides and AnTm sequences (n + m = 4, no TA steps) were computed within 201 bp segments across the HIV-1 genome. The dotted lines at 11.9% and 30.7% represent the AnTm and AT tetranucleotide

densities displayed by HIV-1 *pol* fragment 5, respectively (Table 2). B) AnTm and select TA tetranucleotide frequencies were computed within 1001 bp sliding windows along the HIV-1 genome. C) The percentages of the 16 AT tetranucleotides and AnTm sequences were computed within 201 bp segments surrounding the HIV-1 *pol* gene, and the positions of the 5 HIV-1 *pol* fragments used in biochemical studies are displayed. D-F) Occurrences of all AT tetranucleotide sequences in the coding strand of the HIV-1 DNA genome (F), the HIV *pol* region (E), and the five *pol* fragments (F) described in Table 2 are plotted against the tetranucleotide percentage of A in each region. The numbers adjacent to the tetranucleotide sequences in Fig 1D and 1E are the ratios of the occurrences in the natural to the randomized HIV sequences.

<https://doi.org/10.1371/journal.pone.0216515.g001>

In Fig 2H, the frequencies of non-overlapping AnTm sequences were normalized by the ratio of the total number of non-overlapping AnTm sequences in the reference sequence over the length of the reference sequence. For example, for the human genome AnTm profile in Fig 2H, the AnTm occurrences at distances between ± 4 to ± 100 bps were divided by $(89861037 / 3095677412)$. Note that the smallest possible distance between two non-overlapping AnTm sequences ($n + m \geq 4$) is 4 bps. However, we included a value of 1 for the distance at 0 in the plots.

DNA preparation and characterization

DNA preparation and characterization of Plasmid PHRT25 [17] containing the HIV-1 *pol* gene from isolate BH-5 was used as a template in the PCR to amplify HIV fragments 1–5 studied in Fig 3. The nucleotide (NT) positions of the fragments in the HIV genome (with primers locally aligned to AF033819.3 coordinates) are 2162–2360, 3006–3213, 3647–3846, 4290–4492 and 4507–4708, respectively. The sequences of the top strand primers for fragments 1–5 were: 5' -CAATGGCCATTGACAGAAG, 5' -CAGAAAACAGAGAGATTC, 5' -GTCAATCAAATAATAGAGCAG, 5' -TTAAGACAGCAGTACAAATGGC, 5' -AGGTGAAGGGGCAGTAGT. The sequences of the bottom strand primers were: 5' -CCCAGAAGTCTTGAGT, 5' -TGACATGCTCTGTGTTA, 5' -CTAGCCATTGCTCTCCA, 5' -TGCTGGTCCCTTCCAAAGTG, 5' -CTATAAAACCATCCCC. PCR amplification of the 2 KB EcoRI *pol* fragment used the bottom strand primer for fragment 4 and the sequence of the top strand primer was: 5' GGCCGG AATTCAATGGCCATTGACAGAAG. The amplified fragment was then cut at NTs 1973 and 4004 with EcoRI to give rise to the 2 KB fragment which was inserted into the EcoRI site of plasmid P-2 as described below.

Two parent yeast shuttle vectors were used. P-1 is a pUC18 derivative containing yeast ARS 1 in the EcoRI site and the URA 3 gene in the HindIII site [18]. P-2 is a pUC18 derivative with a 627 bp fragment from yeast ARS 2 in the Xho I site and the URA 3 gene in the HindIII site. HIV fragment 1 was blunt ended into the SmaI site of P-1 to give rise to P-1-H1 or inserted into the EcoRI site of P-2 to give P-2-H1. The 5 S rDNA sequence was inserted into the EcoRI site of P-2 to give P-2-5S rDNA. The 2KB PCR amplified EcoRI fragment from HIV *pol* (1973–4004 bp) was inserted into the EcoRI site of P-2 to give P-2-H-2 KB. To generate low copy number plasmids in yeast, a 289 bp Bam H1 fragment containing yeast CEN III was inserted into the Bam H1 site of each of the plasmids in the P-2 series. DNA size markers consisting of 20bp, 100bp and 1 KB ladder sequences were end-labeled and used to calibrate the gels in Figs 4 and 5.

In vitro studies with synthetic (Syn # 61 and 67) and natural (5S rRNA gene from sea urchin) control fragments were prepared as described previously [19]. Genomic DNAs from chicken erythrocytes and HeLa cells were digested with a mixture of HaeIII and Hinf I and fragments which comigrated with 200 bp standards were eluted from 1.4% agarose gels and used for the studies in Fig 3A and Fig 4B. DNA fragments were uniformly labeled with [α - 32 P] dATP or uniquely end-labeled by using one 32 P-end-labeled primer in the PCR. Genomic DNAs were end-labeled by standard procedures. DNA fragments were purified by

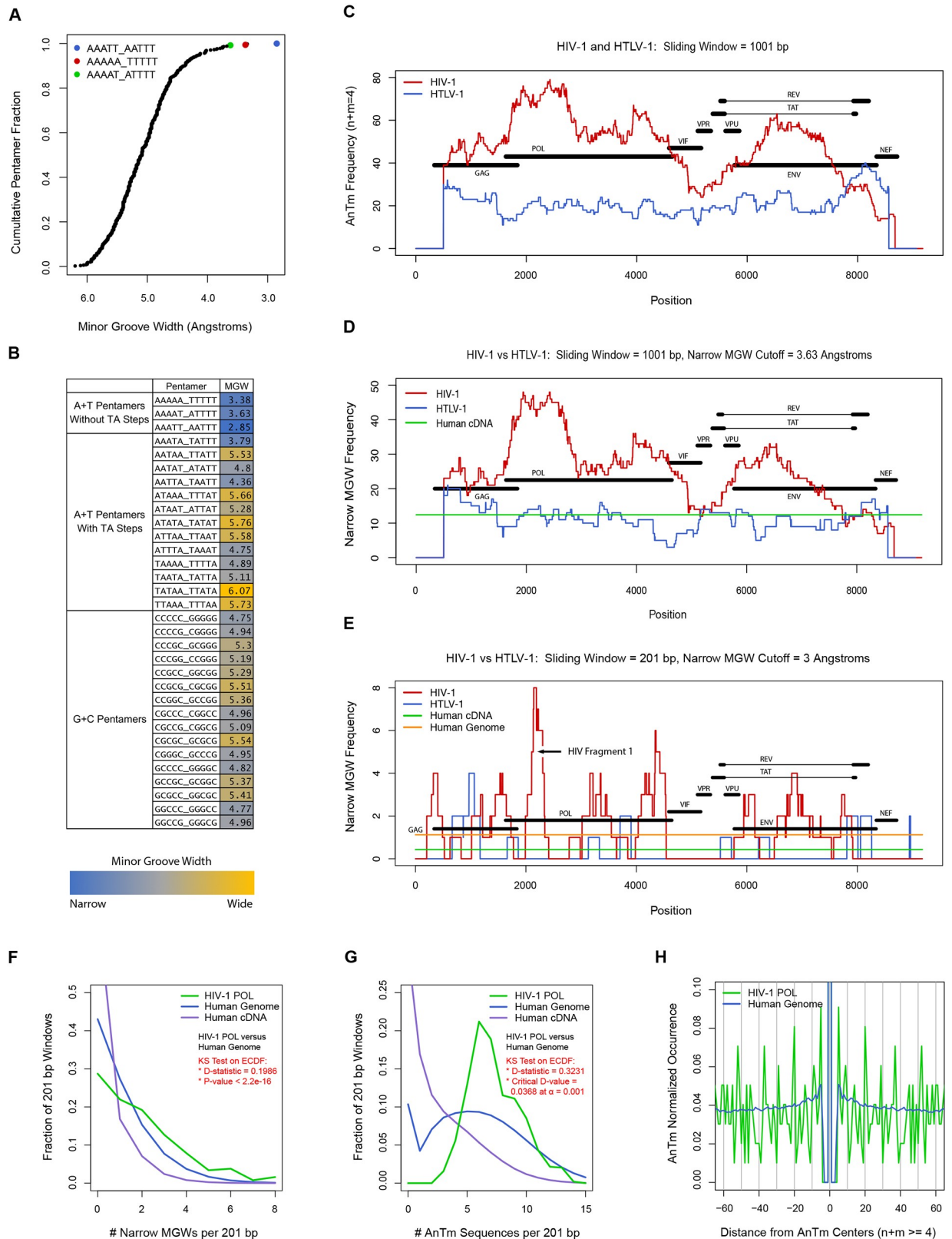


Fig 2. Analysis of AnTm occurrence and narrow minor groove width frequency in the HIV-1 genome. A) The cumulative fractional distribution of all 512 unique pentamers was plotted as a function of decreasing minor groove width. B) The minor groove widths for A+T

pentamers with and without TA steps and for G+C pentamers are displayed in a level plot. C) Occurrences of AnTm sequences in 1001 bp sliding windows are plotted along the HIV-1 and HTLV-1 genomes. D) Narrow MGW frequencies in 1001 bp sliding windows are plotted along the HIV-1 and HTLV-1 genomes, and the average narrow MGW frequency in human cDNA in 1001 bp sliding windows was also plotted. E) Narrow MGW frequencies in 201 bp sliding windows were plotted along the HIV-1 and HTLV-1 genomes, and the average narrow MGW frequencies in human cDNA and genomic DNA in 201 bp sliding windows were also plotted. F) The fractional distributions of narrow MGW frequencies in 201 bp sliding windows were plotted for the human genome, human cDNA, and the HIV-1 *pol* gene. G) The fractional distributions of non-overlapping AnTm sequences in 201 bp sliding windows were plotted for the human genome, human cDNA, and the HIV-1 *pol* gene. H) The pairwise distances between the centers of non-overlapping AnTm sequences were computed for every AnTm sequence in the HIV-1 *pol* gene and in the human genome.

<https://doi.org/10.1371/journal.pone.0216515.g002>

electrophoresis through agarose gels and appropriate bands eluted from gels by the crush-soak method.

In Vitro Minor Groove Binding Drug (MGBD) studies

MGBDs described in Table 1 were from Sigma and drug solutions in 10 mM Tris-HCl (pH 7.5) containing 10 mM NaCl were stored in the dark at -70°C . Nucleosomes were reconstituted onto the DNA fragments by the standard exchange-salt dilution method (1.0 M NaCl to 0.1 M NaCl) using H1/H5 deficient nucleosomes from chicken erythrocytes as a source of core histones [19]. For translational mapping studies with HIV fragment 1, the reconstituted nucleosome was digested with exonuclease III and analyzed as detailed previously [19]. For the restriction enzyme studies in Fig 4B, digestions were carried out for 40 min at 37°C in 10 mM MgCl₂ containing 40 mM NaCl and the fragments analyzed on 8% native PA gels (S6 and S7 Files). The nitrocellulose filter binding assay was carried out as described previously [20] using the NaCl concentrations for sample application and washing that were the same as those used in Fig 3C. The relative binding affinities of DAPI for HIV and control DNAs were determined by fluorescence-enhancement at 465 nM as detailed previously [13].

In Vivo MGBD studies in yeast

Studies using yeast- *S. cerevisiae* strain J17 (MAT a,his2,ade1,trip1,met14,ura3-52) was used as the host for plasmid transformation. The isogenic derivative containing the rad 52 deletion was obtained from K.S. Bloom [21]. Plasmid DNAs were used to transform yeast to uracil prototrophy and transformed cells were grown at 37°C in synthetic medium lacking uracil [18]. For chromatin studies, cultures (100–200 ml) were diluted to 0.2 A 660/ml, drugs were added as indicated and the cells grown overnight to late log phase corresponding to an A660/ml of 1.3–2.0. DAPI (4 μM) reduced growth rates by approximately 20% while concentrations > 10 μM caused severe growth retardation. Berenil at concentrations up to 150 μM was without effect on the generation time. Spheroplasts were isolated essentially as described by Kent et al. [22] except Lyticase (Sigma) was used in place of Glusulase and spermidine was omitted from the MNase digestion buffer. Spheroplasts were permeabilized with NP-40 and digested with MNase at 37°C . for 6 minutes. For indirect-end labeling, DNA obtained from MNase digested spheroplasts was secondarily digested with Ava II and fragments were electrophoresed on 1.0% agarose gels, transferred to Zeta probe membranes (Bio-Rad) in 0.4 M NaOH and the membranes hybridized with multiprime labeled 200–300 bp probes that abutted the Ava II restriction sites. Gels were calibrated using plasmid restriction fragments which are not shown in the figures. For the experiments in the top panels of Fig 5A–5I, blots were probed with HIV sequences or with a 190 bp MspI fragment from pUC18 which is located 734 bp from the HIV DNA. Blots from chloroquine gels were probed with pUC18. For analysis of DNA topoisomers, DNA was prepared from 10 ml cultures as described previously [18] and electrophoresed on 0.9–1.1% agarose gels in Tris-phosphate-EDTA in the presence of 1.5 $\mu\text{g}/\text{ml}$ of

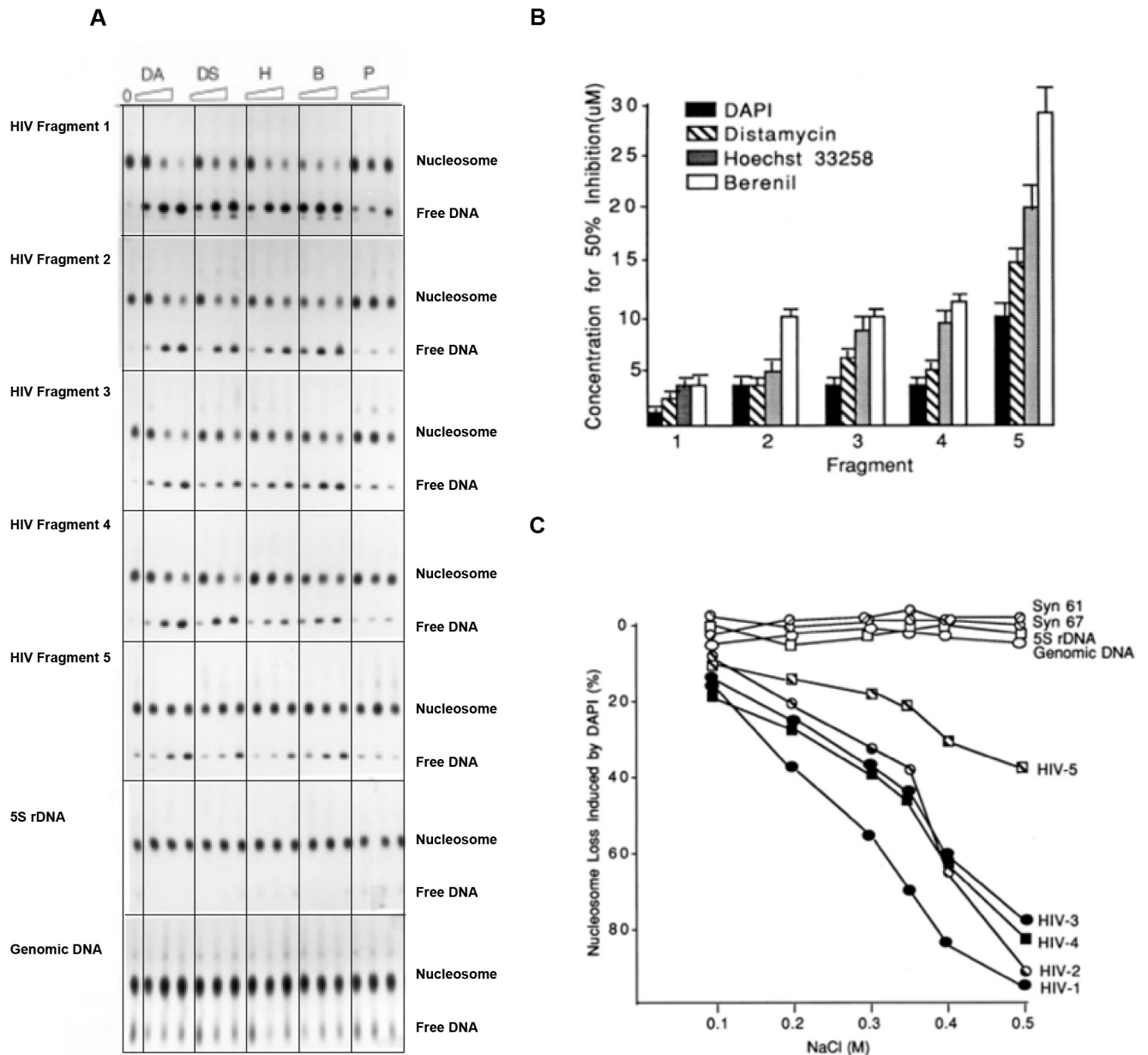


Fig 3. Selective inhibition and effectiveness of nucleosome assembly on HIV *pol* fragments by MGBDs. A) Labeled HIV fragments 1–5, the 5 S rDNA sequence and genomic DNA from HeLa cells were incubated in 1 M NaCl with increasing amounts of MGBDs (1, 5 and 10 uM for DAPI (DA), Hoechst (H) and distamycin (DS); 2, 10 and 20 uM for berenil (B) and pentamidine (P)) prior to addition of 300 ng of erythrocyte nucleosomes. The salt concentration was reduced by stepwise dilution and the products were electrophoresed on PA-glycerol gels in order to resolve nucleosomal and free DNA. B) Mean (+/- S.E.M.) micromolar concentrations of drugs that caused a 50% inhibition of nucleosomes formation was determined from densitometric scans of the results of 4–6 experiments of the type shown in Fig 3A. C) Selective destabilization by DAPI on nucleosomes preassembled on HIV fragments were quantified. Labeled HIV fragments 1–5, synthetic fragments 61 and 67, the 5 S rDNA sequence, and genomic DNA from chicken were assembled into nucleosomes as in Fig 3A and then incubated in the absence and in the presence of 10 uM DAPI at the indicated NaCl concentrations for 1 hour at 37°C. The percentage of the nucleosomes that were lost in the presence of the drug was then determined following electrophoretic analysis. With all sequences, more than 90% of the nucleosomes were retained in the absence of drug in all salt levels.

<https://doi.org/10.1371/journal.pone.0216515.g003>

chloroquine [23]. Under these conditions, the bulk of the topoisomers are negatively supercoiled. Values for drug-induced loss of superhelical turns similar to those given in Fig 5J were obtained when 90 ug/ml chloroquine was used. Under these conditions, the bulk of the

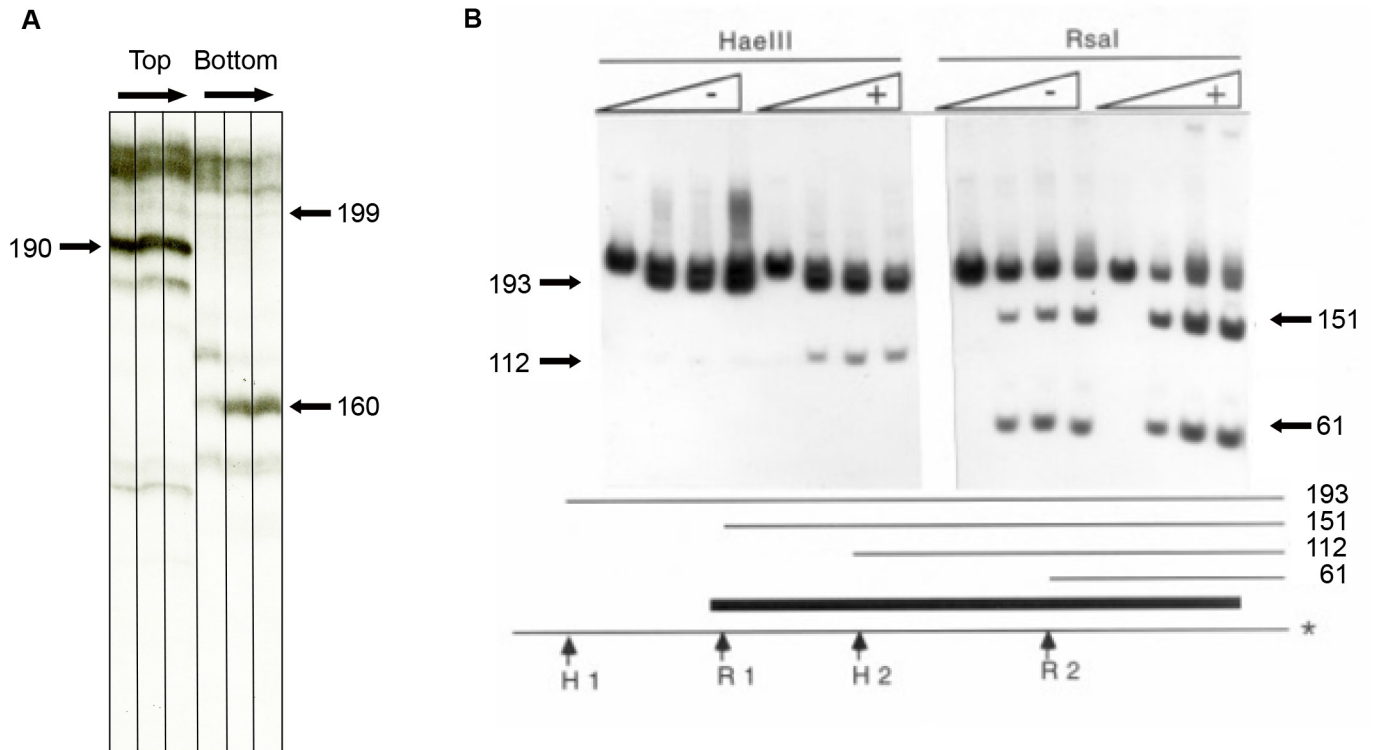


Fig 4. Effects of DAPI on the restriction enzyme accessibility of reconstituted mononucleosomes. A) Nucleosomes containing top or bottom end labeled strands of HIV fragment 1 were incubated with 100 units/ml of ExoIII at room temperature for 2 min to 15 minutes prior to separation of the DNA on the sequencing gel. The two major bands correspond to the borders of the positioned nucleosome [19]. B) A schematic diagram of this nucleosome (dark bar) and the restriction sites for HaeIII (H1 & H2) and RsaI (R1 & R2) are shown. In the gel, a preassembled nucleosome containing this sequence labeled on the bottom strand (*) was incubated for 40 min at 37 C in 40 mM NaCl and 10 mM MgCl₂ in the absence (-) and in the presence (+) of 10 uM DAPI. HaeIII (left 8 lanes; 0, 50, 150, 375 U/ml) and RsaI (right 8 lanes; 0, 50, 150, 375 U/ml) were then added and the reaction terminated after an additional 40 min.

<https://doi.org/10.1371/journal.pone.0216515.g004>

topoisomers were positively supercoiled, reflecting the differences in negative supercoiling [24]. Plasmid copy number was determined as described previously [18]. Densitometric scans from multiple exposures of films were used to obtain quantitative estimates of band intensities for result in Figs 3–5. Nearly all experiments in this paper were carried out at least three times, and the results in Figs 3B and 5J are presented as means \pm S.E.M. of at least three individual experiments with the exception of the results for berenil in Fig 5J where means were from two experiments which were within 10% of each other.

Results

Patterns of DNA sequence and shape in the HIV-1 genome

There are several characteristics of the HIV-1 genome which should make it a preferred target for the MGBDs. As shown in Fig 1A, approximately one fourth of the nucleotides in the genome consist of AT tetranucleotides, the minimal binding-site size for most of the MGBDs. The AT tetranucleotide content is nearly 2-fold greater than that of the human genome (24.2% versus 13.5%). The coding strands of these AT oligonucleotides ($> = 4$ bp) are A-rich (59.9% A) reflecting the compositional bias of the genome [7,8]. The AT tetranucleotide and the AnTm tetranucleotide frequencies ($n + m = 4$, no TA steps) as well as the A-bias are strongest in the *pol* and *env* regions and are reduced only in regions of overlapping reading frames and in the long terminal repeats (Fig 1A & 1B, Table 2).

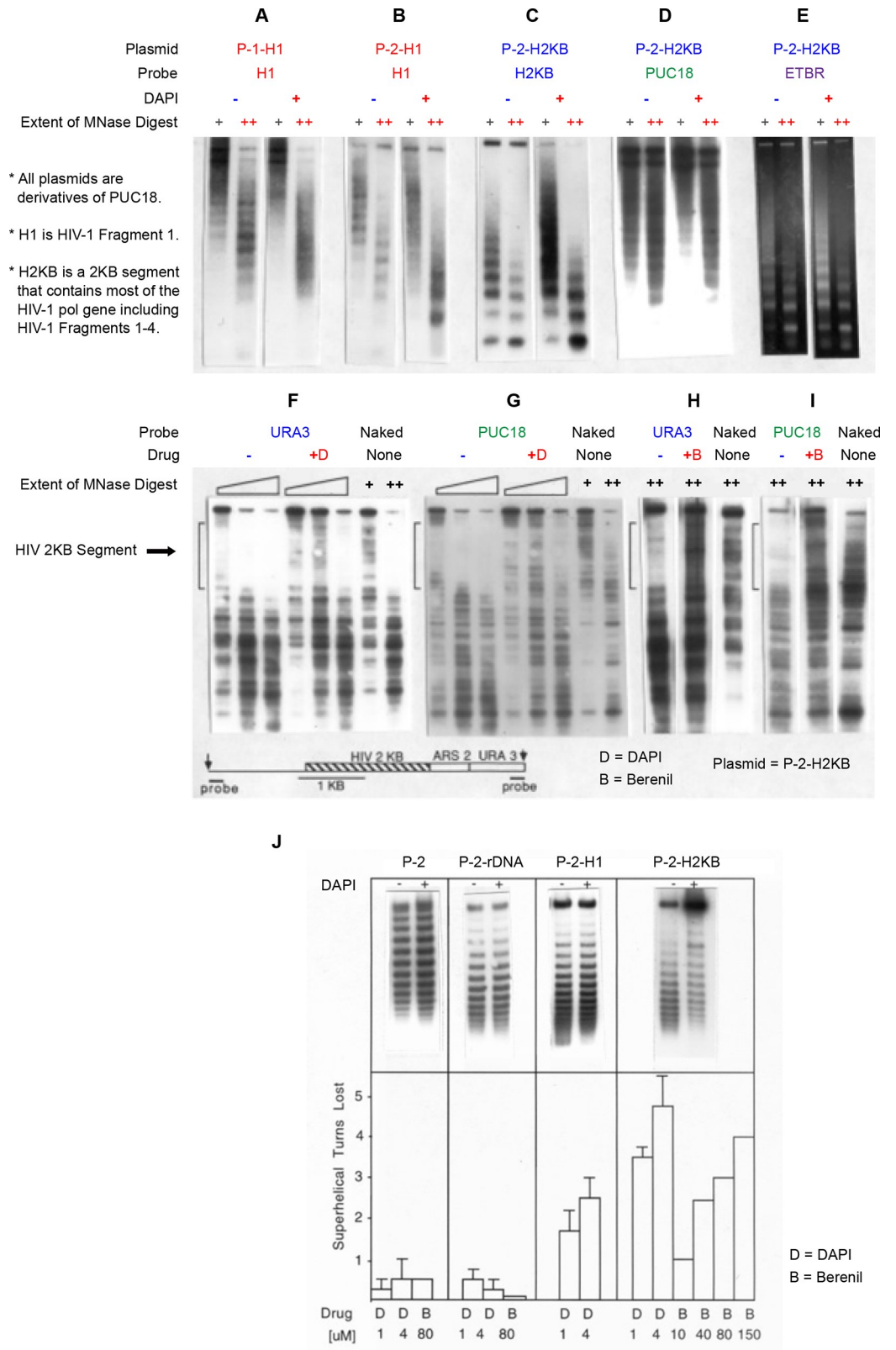


Fig 5. Effects of MGBDs on HIV chromatin in yeast. A-E) The HIV segments used in these studies are the 199 bp fragment 1 (H1) and 2 KB segment (H2KB) which contains most of the HIV *pol* gene. Yeast cells carrying plasmids P-

1-H1 (A), plasmid P-2-H1 (B) or plasmid P-2-H-2KB (C-E) were grown for about 20 hours in the absence (-) and in the presence (+) of 4 uM DAPI. Permeabilized spheroplasts were digested with MNase (100 and 350 Units/ml) and the digested DNA was electrophoresed, blotted and probed with HIV DNAs (A-C), PUC18 (D) or run on a gel stained with ethidium bromide (E). F-I) Indirect end-labeling analysis was performed on MNase digests obtained from cells that carried the longer 2 KB HIV segment on P-2-H2KB. The map of the region indicates the positions of the *Ava*II sites (arrows), hybridization probes, pUC18 DNA (single line), the 2 KB HIV segment and a portion of the *URA3* gene. HIV fragments 1–4 are contained within this 2 KB segment. Chromatin samples from control (-) and DAPI treated (+) cells were digested with increasing amounts of MNase (50, 100 and 200 U/ml). DNA was then digested with *Ava*II, electrophoresed on 1% gels and blots probed with the *URA 3* probe (F) and the pUC18 probe (G). Chromatin samples from control (-) and berenil treated (+) cells were digested with MNase (100 U/ml) and the blots probed with the *URA 3* probe (H) and the pUC18 probe (I). Naked DNAs carried through the procedure are indicated. The approximate gel positions of the 2 KB HIV segment are indicated by the vertical brackets. J) Cells that carried the indicated plasmids were grown in the presence (+) and in the absence (-) of 4 uM DAPI for 18 hours and the distributions of topoisomers were analyzed in chloroquine-agarose gels as described in [Materials and Methods](#). The DNA at the top of the gels, which is particularly prevalent in the P2-H2KB +DAPI lane, represent relaxed/nick circular plasmid [21, 32,33]. Below the gels are shown the estimate mean peak ladder positions (+S.E.M.) obtained from densitometric tracings of 2–6 independent experiments of the type shown in the top J panels were used to calculate the loss of superhelical turns induced by drug. Error bars are not given for the experiments performed only twice. The drugs (D, DAPI; B, berenil) concentrations in uM are indicated.

<https://doi.org/10.1371/journal.pone.0216515.g005>

The *pol* and *env* genes (within regions absent of overlapping reading frames) display a high density of AT sites. On average, the distances between the centers and between the starts and ends of the AT tracts (> 4 bp) are ~18.8 bp and ~12.8 bp, respectively. There are multiple nucleosome-length stretches in *pol* and *env* where more than 40% of their sequences are composed of AT oligonucleotides (Fig 1A and 1C), which corresponds to even tighter spacing between AT tracts. Nucleosomes that contain such high levels of AT oligonucleotides should be drug sensitive provided that the AT sequences have narrow minor grooves [13,14]. Taken together, these results imply that the A-bias favors a high frequency of potential drug binding sites in the HIV genome.

The occurrences of all AT tetranucleotide sequences are shown for the HIV-1 genome (Fig 1D), for the HIV *pol* region (Fig 1E) and also for five short HIV *pol* fragments (Fig 1F), which were used in biochemical studies described below. The effect of the genomic bias on the composition is clearly evident with the level of homopolymeric tracts of A being about 3-fold and 15-fold greater than those of G and C, respectively (Fig 1D–1F). The underlined sequences lack a TA step, and these AnTm tetranucleotide sequences are the most prevalent in the mixed AT sequences at each percentage of A in the three sequence sets (Fig 1D–1F). Each of these AnTm sequences is also overrepresented in HIV DNA relative to randomized HIV DNAs as shown by the numbers in Fig 1D and 1E, which represent the ratios of the occurrences in the natural to the randomized sequences.

These findings may be relevant since the TA step widens the minor groove of AT sites [25–28], and the compressed groove is required for high affinity drug binding and nucleosome disruption by the MGBDs [10,13,29,30]. In contrast, most of the sequences with TA dinucleotides are underrepresented in HIV DNA relative to the shuffled sequences, and this underrepresentation is particularly pronounced in the tetranucleotide with two TA steps and those with a

Table 1. Minor Groove Binding Drugs used in this study.

Drug Used	Synonym(s)	Empirical Formula
Berenil	Diminazene aceturate	C14H15N7 · 2C4H7NO3
DAPI	4',6-Diamidino-2-phenylindole dihydrochloride, DAPI dihydrochloride	C16H15N5 · 2HCl
Hoechst 33258	Bisbenzimidazole H 33258	C25H24N6O · 3HCl
Pentamidine	1,5-Bis(p-amidinophenoxy)pentane bis(2-hydroxyethanesulfonate salt)	CH2[CH2CH2OC6H4C(=NH)NH2]2 · 2HOCH2CH2SO3H
Distamycin	Distamycin A hydrochloride	C22H27N9O4 · HCl

<https://doi.org/10.1371/journal.pone.0216515.t001>

Table 2. Characteristics of the AT, AnTm, and GC Sites in the five HIV-1 *pol* fragments.

HIV Fragment	Basic Sequence Information				AT-Sites (> = 4 bp)				AnTm-Sites (n + m > = 4 bp)			GC-Sites (> = 4 bp)		
	Start	Stop	Len	%A+T	Freq	Avg Len	% Frag	%A *	Freq	Avg Len	% Frag	Freq	Avg Len	% Frag
1	2162	2360	199	67.3	15	5.8	43.7	67.8	12	5.3	32.2	2	4.5	4.5
2	3006	3213	208	62.5	15	5.8	41.8	69.0	11	4.5	23.6	4	4.3	8.2
3	3647	3846	200	63.5	11	6.0	33.0	63.6	9	4.6	20.5	1	5.0	2.5
4	4290	4492	203	63.5	15	5.5	40.4	63.4	10	5.1	25.1	4	5.0	9.9
5	4507	4708	202	61.9	13	4.8	30.7	62.9	6	4.0	11.9	2	4.5	4.5

* The AT-Site %Fragment values for the control 5S rDNA sequence, synthetic 61 fragment, and synthetic 67 fragment are 17%, 20%, and 20%, respectively.

<https://doi.org/10.1371/journal.pone.0216515.t002>

single TA step in the central region (Fig 1D–1F). The relative occurrences of the sequences that make up the AT sites with 50% A were AATT >> ATTA > TTAA ~ TAAT > ATAT > TATA (Fig 1D and 1E). This order is inversely related to the DNA minor groove width of these sequences [25–28] and parallels MGBD binding affinity [13,29,30] and the effectiveness of the drugs for disrupting nucleosomes on DNA that is rich in AT-sites [13,14]. The overrepresentation of AAAA, AATT and other AnTm sequences that lack TA dinucleotides also follow the occurrences of the AT sites along the viral genome while the AT tetranucleotides with TA steps display occurrence patterns that are lower and more uniform throughout the viral genome (Fig 1B). Consequently, the segments containing clustered AT sites in *pol* and *env* of HIV should be especially sensitive to MGBDs.

Table 2 gives additional characteristics for the five DNA fragments from HIV *pol* used in the biochemical studies, and these characteristics are graphically illustrated in Fig 1C. Fig 1C shows that the five HIV *pol* fragments are enriched in AT tetranucleotides and AnTm sequences relative to the entire HIV-1 genome. However, considerable variation exists in AT tetranucleotide and AnTm content amongst the five fragments (Fig 1C, Table 2). These fragments extend from the beginning of reverse transcriptase coding region into the viral integrase gene (Fig 1C), and these ~200 bp fragments have a high density of oligonucleotide AT tracts. On average, ~13.8 oligonucleotide AT tracts that are ~5.6 bp in length comprise ~37.9% of the nucleotides among the five HIV-1 *pol* fragments (Fig 1C, Table 2). Additionally, in these fragments, the frequency of the all oligonucleotide AT tracts (> = 4 bp) is about 6-fold greater than the frequency of all oligonucleotide GC tracts (> = 4 bp) as the GC tracts make up only ~5.9% of the nucleotides in the fragments on average (Table 2). The fragments also show an enrichment of AnTm tracts, AT sites that are not interrupted by TA steps. Although AnTm tetranucleotides represent only 5 of all 16 AT tetranucleotides, ~69.6% of the AT sites in the fragments consist of AnTm tracts (Fig 1C, Table 2).

In order to directly relate the DNA sequence patterns of HIV to the structure of the DNA minor groove, we performed a systematic and comprehensive bioinformatic investigation to further our understanding of the relationship between DNA sequence and minor groove width. The biophysical parameters of dinucleotides such AA, AT, and TA are influenced by the surrounding sequence context of DNA. As a result, models based on longer DNA sequences have been proposed to more accurately estimate the shape of the DNA helix in 3-dimensional space. In one particular model, a combination of computational and experimental measurements was used to predict the shapes of all 512 unique pentamer sequences in terms of roll, propeller twist, helical twist, and minor groove width (MGW) [23,24]. We utilized this pentamer model to predict the MGWs along the HIV-1 and HTLV-1 genomes and also leveraged the model's predicted MGWs across the entire human genome.

In order to illustrate how this model predicts AnTm sequences ($n + m = 5$, A+T pentamers without TA steps) to be associated with significantly narrow minor grooves, we plotted the cumulative fraction of all 512 pentamers as a function of decreasing MGW (Fig 2A). This plot shows that all 3 unique AnTm pentamer sequences are among the top 5 (or 1%) of all pentamers with the narrowest minor grooves. Notably, A+T pentamers with TA steps and G+C pentamers are characterized by much wider minor grooves (Fig 2B).

We then sought to compare the AnTm content and narrow MGW frequency between the HIV-1 and HTLV-1 genomes, which have similar sequence lengths. Fig 2C shows greater AnTm occurrence along the HIV-1 genome as compared to HTLV-1 in 1001 bp sliding windows. The most considerable increases in AnTm content in HIV-1 over HTLV-1 appear in the *gag*, *pol*, and *env* genes, especially outside of overlapping reading frames among other HIV-1 genes. In parallel, similar patterns of elevated narrow MGW frequency occur along the HIV-1 genome relative to not only HTLV-1 but also to the average narrow MGW frequency in human cDNA (Fig 2D).

In order to assess the patterns of narrow MGW frequency in HIV-1 at the scale of nucleosomal DNA, we executed the same analysis conducted in Fig 2D using 201 bp sliding windows instead of 1001 bp ones (Fig 2E). The results from this analysis revealed several hot spots of high narrow MGW density in HIV-1, which are virtually absent in HTLV-1. Moreover, several sections of HIV-1 are observed to occupy 2-8-fold increases in narrow MGW frequency at the nucleosome level over human genomic DNA (Fig 2E). Among the several spikes of AnTm and narrow MGW frequency along the HIV-1 genome, the most striking examples are observed in the HIV-1 *pol* gene with the largest spike in narrow MGW density encompassing HIV fragment 1 (Fig 2C–2E). Subsequently, we computed the fractional distributions of narrow MGW frequency and AnTm frequency in 201 bp windows for the HIV-1 *pol* gene, human genomic DNA, and human coding DNA (Fig 2F and 2G). As expected, these distributions indicated that the HIV-1 *pol* gene encompasses greater narrow MGW and AnTm density than human genomic DNA, which showed even greater concentration of these features over human coding DNA. Indeed, Kolmogorov-Smirnov tests on corresponding empirical cumulative distribution functions showed that the differences between these distributions were statistically significant.

After examining AnTm frequency and narrow MGW density in HIV-1, we investigated the spacing of AnTm sequences, and the results revealed extraordinarily organized spacing along the HIV-1 *pol* gene as compared to the bulk of the human genome (Fig 2H). Notably, resulting from an accumulation of AnTm pairwise distances, strong ~10 bp periodic patterns can be detected in AnTm spacing along the HIV-1 *pol* gene, which can be inferred by the presence of peaks near positions +/-20, +/-30, +/-40, and +/-50 in Fig 2H. In contrast, this analysis of AnTm spacing also shows that AnTm centers are often separated by a distance of 4–5 bps (Fig 2H). Upon closer inspection of the HIV-1 *pol* sequence, one can identify several examples of short AnTm sequences that are either separated by TA steps or interrupted by short C/G-containing mono and di nucleotides (data not shown).

Effects of Minor Groove Binding Drugs on HIV-1 chromatin

Since MGBDs inhibit the assembly of nucleosomes on sequences with multiple closely spaced oligo-AT sequences with narrow minor grooves which serve as drug binding sites [13,14], it was of interest to determine if the drugs exert their selective inhibitory action on HIV chromatin. To study this possibility, nucleosomes were reconstituted onto the five fragments in the presence and absence of several MGBDs. Free DNA was then separated from nucleosome DNA by electrophoresis on polyacrylamide-glycerol gels. A sample of the data is shown in Fig 3A, and the results of several similar experiments are summarized in Fig 3B. The studies

demonstrated that the MGBDs inhibited the assembly of nucleosomes onto all five of the HIV segments. The marked specificity of the drugs for HIV DNA is illustrated by the results in Fig 3A which show that these ligands had no detectable effects on the assembly of nucleosomes onto a control fragment and genomic DNA. For all HIV fragments, the relative order of inhibition was DAPI (4,6-diamidino-2-phenylindole) > distamycin > Hoechst 22358 > berenil >> pentamidine. For all drugs, the relative order of inhibition was HIV fragment 1 > 2, 3, 4 > 5. This order parallels the density of AT sites > 4 bp as can be seen from the data in Table 2. This finding is in agreement with our previous studies which demonstrated that drug sensitivity is directly related to the frequency of AT sites in nucleosome-length DNA [13,14]. Selective inhibition of nucleosome formation was seen with HIV fragment 5 which contains 29% of nucleotides in AT sites. As indicated by the dotted line in Fig 1A, approximately 65% of all of the possible 200 bp segments in HIV *pol* and 35% of all of the possible 200 bp segments in the entire viral genome have AT site densities that are equal to or greater than this level. These same sites also tend to be enriched in AT sequences without TA steps as shown by the example given in Fig 1B. These observations suggest that a relatively large fraction of the nucleosomes along HIV chromatin should be a preferred target for MGBDs.

Previous studies have shown that the MGBDs promoted the destabilization of nucleosomes that were pre-assembled onto DNA molecules rich in AT sites and this effect was most pronounced at high ionic strength [13,14]. This approach was taken in the present study to determine if MGBDs destabilize nucleosomes assembled onto the five HIV fragments. Fig 3C shows the fate of preassembled nucleosomes containing the five HIV sequences and several control DNAs when they were incubated for 1 hour at 37° C in 0.1–0.5 M NaCl in the presence and in the absence of DAPI. Nucleosomes assembled onto the HIV DNAs were preferentially destabilized by the drug and the effect was most apparent in 0.5 M NaCl where up to 95% of the histone octamers were lost from DNA. The relative order of drug-induced nucleosome loss (Fragment 1 > 2, 3, 4 > 5) was the same as the order for drug-induced inhibition of nucleosome assembly, and density of AT sites as can be seen by comparing the results in Table 2 and Fig 3B and 3C. The template selectivity is also clearly evident from the results since octamers were retained on all control fragments and genomic DNA in all salt levels.

Natural nucleosome positioning sequences are enriched in AnTm sequences which are arranged to give rise to two regions of curvature of 40–60 bp separated by a non-curved segment of 30–40 bp [31]. Similarly, HIV fragment 1 is highly enriched in these AnTm sequences and displays two curved regions, and was therefore, previously predicted to position an assembled nucleosome onto a single site [9]. This prediction of HIV fragment 1's translational positioning ability was supported by the exonuclease III digestion studies shown in Fig 4A and enabled us to study the effects of DAPI on restriction nuclease accessibility. The central position of this nucleosome along with the positions of restriction sites for HaeIII and RsaI is shown in the map at the bottom of Fig 4B. In order to see if a MGBD destabilizes this nucleosome in low ionic strength, a preassembled nucleosome containing this sequence was incubated for 40 min at 37° C in 40 mM NaCl in the presence and in the absence of DAPI. Under these conditions, no more than 5% of the DNA was rendered histone-free by drug treatment. Restriction enzymes were then added and the reaction terminated after an additional 40 min. In the absence of the drug, HaeIII site 1 which lies outside of the major positioned core was cleaved extensively while HaeIII site 2 and the RsaI sites were more protected from enzyme cleavage. The analysis demonstrated that DAPI increased the extent of cleavage of these internal nucleosome sites by 3–8-fold (Fig 4B).

The above studies established the selective inhibitory effects of the MGBDs on HIV-1 containing nucleosomes in vitro. To provide a first step to evaluate the effects of the drugs on HIV chromatin in vivo, HIV segments were cloned into yeast shuttle vectors and cells transformed

with the plasmids were grown for about 20 hours in the absence and in the presence of 4 μ M DAPI. Permeabilized spheroplasts were digested with micrococcal nuclease (MNase) and the digested DNA was electrophoresed, blotted and probed with HIV DNAs. The MNase ladder observed in the absence of drug is indicative of regularly spaced nucleosomes and the definition of the nucleosome repeat was reduced by DAPI as seen in Fig 5A–5I. Reduction in the resolution of the ladders was seen with the chromatin assembled onto the 199 bp HIV fragment 1 in two different sequence contexts (Fig 5A and 5B). Partial but not complete loss was noted in the chromatin assembled onto the longer 2 KB HIV segment, and this reduction was more clearly seen in the minimally digested samples (Fig 5C). This result is expected since some but not all of the nucleosomes assembled on this long sequence are expected to be drug sensitive. The drug had little effect on the nucleosome ladders containing bulk yeast DNA and little or no effect on the nucleosome profile of sequences from the blot probed with a segment of pUC18 (Fig 5D and 5E). The reduction in the nucleosomal repeat in Fig 5A–5C is interpreted as representative of either the preferential loss or disruption of nucleosomes along the HIV sequences.

To further characterize this effect, indirect end-labeling analysis was performed on MNase digests obtained from control and DAPI treated cells that carried the longer HIV segment (Fig 5F and 5G). In the absence of drug, the HIV sequences resided within a nuclease protected region which is nuclease sensitive in naked DNA. DAPI enhanced the nuclease cutting of the HIV sequences in chromatin rendering the digestion pattern similar, but not identical, to that seen in the protein-free samples. The cutting pattern of sequences adjacent to the HIV segment were similar in the presence and in the absence of the DAPI providing evidence for the template selectivity of this drug action. Essentially the same results were seen in experiments where 60 μ M berenil was used in place of DAPI (Fig 5H and 5I).

In order to provide an alternative strategy to study the effects of the MGBDs on the in vivo assembled chromatin, DNA topoisomers from drug-treated and control cells were resolved in agarose gels containing chloroquine (Fig 5J). DAPI promoted a topological change toward a loss of superhelical turns of the HIV-containing plasmids while having little or no effect on control plasmids that lacked HIV sequences. Such a reduction is most often interpreted as a loss of nucleosomes since each nucleosome induces a single negative superhelical turn in a closed circular DNA molecule [21,32,33]. Densitometric scans of these and other autoradiograms indicated an average loss of 1–2 nucleosomes on the plasmids with the ~200 bp HIV sequence 1 and up to 4–6 nucleosomes on the plasmid containing the larger segment of HIV DNA as seen in the histogram in Fig 5J. The value for the larger segment is significantly less than the 13 nucleosomes that should be present on this sequence in the absence of drug which implies that a fraction of the nucleosomes on this HIV segment are drug resistant. This result is consistent with the partial loss of the nucleosome ladder induced by drug treatment (Fig 5C) and with calculations of the AT-tract densities of 200 bp segments that make up this fragment which indicate that only about 2/3 of the nucleosomes in this region of HIV should display detectable drug sensitivity (Fig 1A). The results shown in Fig 5J (bottom panel) also demonstrate that DAPI was several fold more effective than berenil at inducing nucleosome loss in vivo in agreement with the relative potencies of these drugs determined by the in vitro studies shown in Fig 3B and 3C.

Since DNA that is devoid of nucleosome is hypersensitive to nucleases [34,35], it was of interest to determine if loss or disruption induced by a MGBD might render the HIV DNA nuclease hypersensitive in vivo. To test this possibility, HIV fragment 1 and the longer HIV segment on centromere-containing plasmids were introduced into yeast cells that lacked the RAD 52 gene, which is required for repair of DNA double-strand breaks [36]. DAPI and berenil had no effect on plasmid copy number or HIV sequence integrity in these cells or in wild-

type cells nor did the drugs preferentially affect the growth rate of these cells in selective media lacking uracil relative to those that were transformed with control plasmids that lacked the HIV DNA. Likewise, the drugs did not increase the ratio of relaxed circular/supercoiled DNA of the HIV-containing plasmids. Thus, it appears that nucleosome loss *in vivo* does not necessarily render the HIV DNA sensitive to cleavage by cellular nucleases under the conditions used in these studies.

Discussion

Most of the parent MGBDs were developed as trypanocides and berenil remains used for this purpose today in Africa and Asia. Pentamidine has also been used to prevent and treat opportunistic pneumonia that is caused by *Pneumocystis carinii* which is common in AIDS patients. The drugs have found additional applications in clinical medicine including use as general antitumor, antibacterial and antiviral agents [12,37–40]. However, the targets and modes of action of the drugs remain, for the most part, unknown. Although the MGBD distamycin and related analogues show anti-HIV activity in cultured cells, the MGBDs have received little attention as anti-HIV agents because many of the distamycin-related compounds are expensive and toxic to cells [41]. Likewise, to our knowledge, MGBDs have not been tested for activity against simian immunodeficiency viruses or the lentiviruses that infect non-primate mammals, though the nucleotide composition of these viral genomes are nearly identical to that of HIV-1 [8]. Much more emphasis has been devoted to the synthesis of distamycin conjugates and analogues as anti-HIV drugs with the goal of targeting the ligands to defined features of the viral life cycle [12]. For example, several distamycin-related polyanionic conjugates are inhibitors of HIV-induced cell killing and mechanistic studies suggest that the inhibition is due to the inhibition of virus attachment to CD4+ susceptible cells [42]. Another series of distamycin derivatives was designed to target the binding sites for several cellular transcription factors and a combination of these ligands inhibited HIV replication in lymphocytes [43]. However, as discussed by Tutter and Jones [44], the most significant problems associated with these analogs is the ease with which HIV-1 evolves to evade chemical inhibitors and the fact that no single sequence in the HIV genome is absolutely essential for virus replication. The studies described in this report point to an alternative global specificity of the MGBDs for the HIV provirus DNA that is based on the characteristic A-bias of the lentivirus genome. The bias dictates the unusual composition of essentially all lentiviral proteins and may also serve to enhance genetic variation and curvature of the HIV genome [8,9]. Consequently, the bias is considered to be critical for HIV phenotype making it unlikely that HIV variants that lacked this genome-wide sequence feature would appear during the course of an infection. Drugs whose anti-viral action is dependent on this global bias may thus provide a novel class of anti-lentiviral agents.

Nucleosome-length sequences with AT site densities as high as those seen in HIV are rare in the eukaryotic genome as implied from the results in Figs 3A, 3C, and 5A–5I which show that the MGBDs did not noticeably inhibit nucleosome assembly or promote nucleosome loss on genomic DNAs from human, chicken or yeast. The unusual nucleotide bias of HIV is the major driving force behind this specificity for it dictates the high frequencies of AT sites and their narrow minor groove character. All lentiviruses exhibit the strong A-bias and display little or no selection against the high levels of AT-trinucleotides [8] and tetranucleotides (Fig 1D–1F) in their coding regions. In fact, the frequency of AT sites with narrow minor grooves in HIV is higher than in random sequence DNA (Fig 1D–1F) which may be related to the propensity of the lentiviral reverse transcriptase to elongate preexisting oligo-A tracts with A at the expense of G [45,46]. In contrast, less than 0.1% of the sequences in the primate

subdivision of GenBank display an A-bias that is extreme as that seen in HIV [8]. In addition, oligonucleotide-length AT sequences are counter selected in cellular genes [47]. Thus, the unusual viral bias results in a high genomic levels of AT sites which are expected to be binding sites for the MGBDs. DNase1 footprinting studies confirmed this expectation by showing that DAPI protected nearly all of the AT sites > 2–3 bp in the HIV *pol* fragment 1. The compositional bias of the viral genome also favors AT sites with narrow minor grooves since homopolymeric A-tracts comprise a relatively large fraction of the AT sites. In addition, AT-containing oligonucleotide sequences with TA steps are underrepresented in HIV DNA as compared to random sequence and to AT sites without TA steps (Fig 1D–1F). This viral characteristic is also unusual since the AT-containing oligomeric sequences in cellular DNA are enriched in TA but not AT dinucleotides [48]. This feature contributes to the specificity of the ligands for the HIV genome since the narrow minor groove of AT sites is required for high affinity binding of most of the MGBDs [3,29,30]. In addition, our previous studies have shown that extent of drug-induced inhibition of nucleosome assembly on synthetic repetitive fragments followed the order $(CGGGAAAACC)_n > (CGGGAATTCC)_n \gg (CGGGATATCC)_n > (CGGGTTAACC)_n \sim (CGGGACAACC)_n$ [13]. This order roughly follows the relative frequencies of these AT-sequence motifs in the HIV-1 genome (Fig 1D–1F). Thus, the A-bias favors the HIV genome that is rich in clusters of AT sites that preferentially bind to MGBDs and these features are responsible for the marked specificity of the MGBDs for the disruption of chromatin [13,14].

The aromatic diamidine MGBDs including berenil, DAPI, and pentamidine are a group of related low molecular weight drugs that bind to AT-rich sequences [37–40]. As shown in Fig 3A and 3B, DAPI was the most effective drug tested at inhibiting nucleosome formation on HIV DNA. DAPI was also more effective than distamycin and Hoechst at blocking the assembly of nucleosomes and negating the effects of intrinsic DNA curvature and anisotropic bendability in studies that used other naturally occurring and synthetic DNA sequences with multiple closely spaced A-tracts with narrow minor grooves [13,14]. The high activity of DAPI as compared to distamycin and Hoechst has been attributed to the short length of this ligand and its absolute specificity for A/T sites with narrow minor grooves [13]. This specificity may ensure that only particular minor grooves that give rise to curvature and anisotropic bendability are occupied by the drug. Berenil shows the same AT-specificity as seen with DAPI although the binding affinity of this ligand for AT sites is lower than for DAPI [29]. This observation may have clinical implications since berenil was active in disrupting HIV-containing nucleosomes in vitro and in vivo (Fig 3A and 3B & Fig 5) and this drug is currently used to treat a variety of parasitic infections in humans and other mammals [38,49,50]. In addition, in contrast to many of the MGBDs, berenil and pentamidine display low toxicity and both drugs have been reported to be non-mutagenic in mammalian cells [48,51,52]. Distamycin and Hoechst 33258 have most often served as the starting point for synthesis of most MGBD conjugates and analogs [12,52]. A similar approach could be adopted for the synthesis of conjugates of the structurally simpler drugs such as berenil, pentamidine, or DAPI in order to optimize the specificity of these compounds for the clustered AT-sites in the HIV genome. In addition, berenil has previously been used to target cytotoxic platinum to AT rich regions in the genome [53]. Such conjugates targeted to the bone marrow using bone marrow targeting systems [54] might also be useful in promoting selective nucleosome disruption and lethal mutagenesis of the integrated HIV provirus. Similarly, MGBDs could conceivably be used in combination with CRISPR/Cas9-based HIV-1/AIDS therapeutic strategies, especially since recent studies have suggested that nucleosomes can strongly inhibit CRISPR-Cas9 targeting and cleavage efficiency [55–58]. The positioned nucleosome that assembles on HIV fragment 1 might represent an ideal locus for CRISPR-Cas9 targeting since this region in the reverse

transcriptase gene is among the most conserved sequences in the HIV genome [8]. In comparison to nucleosomes assembled on fragments 2–5, this nucleosome was the most sensitive to disruption by the MGBDs and has the highest AnTm occurrence as well as the greatest narrow MGW frequency (Figs 1–3, Table 2).

Dimers of distamycin and neutropsin connected by a diammine bridge displayed a higher affinity for long ($N = 8$) A/T sites in synthetic DNA than the corresponding monomers. These and similar bridged molecules show heightened anti-HIV and 2 activity in culture cells relative to the monomers [59,60]. However, such long A/T tracts are rare in the HIV genome. Rather, the A richness of the HIV genome is due principally to A-tracts that are 3–5 bp in length and these tracts are frequently arranged in a 10 bp periodicity (Fig 2H) [9], a period that is coincident with the pitch of the DNA helix. In addition, the specificity of the MGBDs for promoting nucleosome disruption requires long (>120 bp) stretches of DNA containing such phased A-tracts [13,14]. These considerations raise the interesting possibility that dimers or higher multimers of the smaller MGBDs, such as berenil, with linker lengths that correspond to the spacing of narrow minor groove sites might be more effective and specific than the monomers for HIV nucleosome disruption.

The MGBDs prevented nucleosome formation on the five segments of HIV *pol* studied in this report, and it is likely that the drugs would exert this selective inhibitory action on the bulk of HIV chromatin. It is also clear from the data in Fig 3C that the MGBDs promoted a preferential loss of histone octamers from HIV DNA in high salt since the DNA released from these particles comigrated with free DNA on PA gels and failed to bind to nitrocellulose filters which bind protein-DNA complexes. Whether the drugs induced a loss of octamers from HIV sequences in yeast is more uncertain. The results in Fig 5A–5I are consistent with such a loss but the presence of loosely bound disrupted nucleosomes at these sites cannot be formally excluded. Perhaps the most direct evidence supporting nucleosome loss is provided by the results in Fig 5J which show that the drugs reduced the number of superhelical turns of HIV-containing plasmids. However, the human SWI/SNF complex has been reported to promote a reduction in DNA supercoiling per nucleosome without causing nucleosome loss [61] so that the data in Fig 5J do not prove that the drugs cause nucleosome loss in yeast.

Nucleosome loss and/or disruption has been correlated with cell death, promoter activation, changes in higher order chromatin structure, and enhanced nuclease accessibility [34,35]. The importance of chromatin disruption in the mechanism of action of anti-cancer DNA-binding drugs has also recently been highlighted by Gurova [62]. Disruption of HIV chromatin by MGBDs is expected to be detrimental to the expression of the provirus and perhaps even to the cell that contains this integrated sequence since AT sites with narrow minor grooves are a major determinant for nucleosome stability and positioning [63,64]. In addition, the HIV genome is highly deficient in CpG dinucleotides and methylation of this dinucleotide enhances nucleosome stability [7,8,65,66].

The results in Fig 3 show that the MGBDs have no detectable effects on the assembly of nucleosomes on genomic DNA illustrating the specificity of the drugs for HIV sequences. Furthermore, the genomics studies carried in Figs 1 and 2 provide additional support for how these HIV-1 sequences are distinguished from bulk of the human genome in terms of AnTm and narrow minor groove width density. However, it is likely that there are drug sensitive genomic sequences and it might be informative to understand the nature of these sequences. We previously studied this question by first performing a computer search of all sequences in the bacterial and invertebrate subdivision of GenBank in order to identify all genes that have an unusual nucleotide composition similar to HIV [67]. The results revealed that about 80% of the identified genes coded for membrane protein antigens (virulence factors) from mammalian pathogens and the amino acid composition of these proteins were similar to those from

HIV. We also demonstrated that the MGBDs inhibited nucleosome formation on these sequences with a drug specificity similar to that displayed by the HIV sequences analyzed in Fig 3 [68]. These membrane antigen genes included those from *Pneumocystis jirovecii*, *Entamoeba histolytica*, and *Plasmodium vivax*. Whether this effect is involved in the anti-pathogenic action of the MGBDs in these organisms is a topic for future investigation.

Human fragile sites are segments of chromatin that fail to compact during mitosis and are seen as weak staining gaps and breaks in metaphase chromosomes [69,70]. A subset of these sequences is AT rich and the expandable microsatellite FRA 16B is induced MGBD [70]. This site was predicted to be packaged into MGBD sensitive nucleosomes and this prediction has been confirmed by experimental studies [14,71]. Berenil has been recommended for detection of this site as compared to the other MGBDs in metaphase chromosomes because of its lower cost, it induces higher levels of the site and has no negative morphological effect on metaphase chromosomes in optimal concentration [72]. The possibility was considered that the drugs might promote the breakage of HIV sequences in yeast as a result of the chromatin disruption effect but the data described in the Results Section suggest that this is not the case at least under the conditions used in this work. Thus, although the consequences of nucleosome disruption by the MGBDs remain to be clarified, they point to the need for future studies on the analysis of their effects on HIV infected vs non-infected CD4+human cells in terms of cell viability, proviral sequence integrity and packaging of the sequences into nucleosomes. The results of such studies might then provide an assay for the optimization of desired effects using existing MGBDs, and testing new drug derivatives and drug multimers.

Supporting information

S1 File. HIV-1 DNA minor groove widths. The DNA minor groove width values along the HIV-1 genome were determined by DNashaperR.

(XLSX)

S2 File. HTLV-1 DNA minor groove widths. The DNA minor groove width values along the HTLV-1 genome were determined by DNashaperR.

(XLSX)

S3 File. DNA Minor groove widths for all Pentamers. The DNA minor groove width values for all pentamers are given. See [Materials and Methods](#) for more details.

(XLSX)

S4 File. HIV-1 AnTm positions. The locations of non-overlapping AnTm sequences (≥ 4 bp) along the HIV-1 genome were determined. See [Materials and Methods](#) for more details.

(XLSX)

S5 File. HIV-1 Oligo AT tracts positions. The Locations of non-overlapping Oligo AT tract sequences (≥ 4 bp) along the HIV-1 genome were determined. See [Materials and Methods](#) for more details.

(XLSX)

S6 File. HIV Fragment 1 NEB cutter results for HaeIII and RsaI. The Locations of HaeIII and RsaI sites in HIV Fragment 1 were determined.

(TXT)

S7 File. HIV-1 Fragment 1 sequence. The DNA sequence of HIV Fragment 1 is given.

(FA)

Acknowledgments

The author wishes to thank Irwin Tessman, P.T. Gilham and L.Weith for reviewing the manuscript and for helpful discussions during the course of these studies.

Author Contributions

Conceptualization: John N. Anderson.

Data curation: John N. Anderson.

Formal analysis: Clayton K. Collings, John N. Anderson.

Investigation: Donald W. Little, III, Samuel J. Schafer, John N. Anderson.

Methodology: John N. Anderson.

Project administration: John N. Anderson.

Software: Clayton K. Collings.

Supervision: John N. Anderson.

Validation: John N. Anderson.

Visualization: Clayton K. Collings.

Writing – original draft: John N. Anderson.

Writing – review & editing: Clayton K. Collings.

References

1. Andino R, Domingo E. Viral quasispecies. *Virology*. 2015; 279: 46–51.
2. Nomaguchi M, Doi N, Koma T, Adachi A. HIV-1 mutates to adapt in fluxing environments. *Microbes Infect*. 2018; 20(9–10): 610–4. <https://doi.org/10.1016/j.micinf.2017.08.003> PMID: 28859896
3. Gunthard HF, Aberg JA, Eron JJ, Hoy JF, Telenti A, Benson CA, et al. Antiretroviral treatment of adult HIV infection: 2014 recommendations of the International Antiviral Society-USA Panel. *JAMA*. 2014; 312(4): 410–25. <https://doi.org/10.1001/jama.2014.8722> PMID: 25038359
4. Zuo XF, Huo ZP, Kang DW, Wu GC, Zhou ZX, Liu XY, et al. Current insights into anti-HIV drug discovery and development: a review of recent patent literature (2014–2017). *Expert Opinion on Therapeutic Patents*. 2018; 28(4): 299–316. <https://doi.org/10.1080/13543776.2018.1438410> PMID: 29411697
5. Bashiri K, Rezaei N, Nasi M, Cossarizza A. The role of latency reversal agents in the cure of HIV: A review of current data. *Immunol Lett*. 2018; 196: 135–9. <https://doi.org/10.1016/j.imlet.2018.02.004> PMID: 29427743
6. Margolis DM, Garcia JV, Hazuda DJ, Haynes BF. Latency reversal and viral clearance to cure HIV-1. *Science*. 2016; 353(6297): aaf6517. <https://doi.org/10.1126/science.aaf6517> PMID: 27463679
7. Kypr J, Mrázek J, Reich J. Nucleotide composition bias and CpG dinucleotide content in the genomes of HIV and HTLV 1/2. *Biochim Biophys Acta*. 1989; 1009(3): 280–2. [https://doi.org/10.1016/0167-4781\(89\)90114-0](https://doi.org/10.1016/0167-4781(89)90114-0) PMID: 2597678
8. Bronson EC, Anderson JN. Nucleotide Composition as a Driving-Force in the Evolution of Retroviruses. *J Mol Evol*. 1994; 38(5): 506–32. <https://doi.org/10.1007/bf00178851> PMID: 8028030
9. Albert FG, Bronson EC, Fitzgerald DJ, Anderson JN. Circular structures in retroviral and cellular genomes. *J Biol Chem*. 1995; 270(40): 23570–81. <https://doi.org/10.1074/jbc.270.40.23570> PMID: 7559522
10. Khan GS, Shah A, Zia ur R, Barker D. Chemistry of DNA minor groove binding agents. *J Photochem Photobiol B*. 2012; 115: 105–18. <https://doi.org/10.1016/j.jphotobiol.2012.07.003> PMID: 22857824
11. Neidle S. DNA minor-groove recognition by small molecules. *Nat Prod Rep*. 2001; 18(3): 291–309. <https://doi.org/10.1039/a705982e> PMID: 11476483
12. Rahman A, O'Sullivan P, Rozas I. Recent developments in compounds acting in the DNA minor groove. *Medchemcomm*. 2019; 10(1): 26–40. <https://doi.org/10.1039/c8md00425k> PMID: 30774852

13. Albert FG, Eckdahl TT, Fitzgerald DJ, Anderson JN. Heterogeneity in the actions of drugs that bind in the DNA minor groove. *Biochemistry*. 1999; 38(31): 10135–46. <https://doi.org/10.1021/bi990382p> PMID: 10433722
14. Fitzgerald DJ, Anderson JN. Selective nucleosome disruption by drugs that bind in the minor groove of DNA. *J Biol Chem*. 1999; 274(38): 27128–38. <https://doi.org/10.1074/jbc.274.38.27128> PMID: 10480928
15. Chiu TP, Comoglio F, Zhou T, Yang L, Paro R, and Rohs R. DNASHapeR: an R/Bioconductor package for DNA shape prediction and feature encoding. *Bioinformatics* 2016; 32(8): 1211–1213. <https://doi.org/10.1093/bioinformatics/btv735> PMID: 26668005
16. Chiu TP, Yang L, Zhou T, Main BJ, Parker SC, Nuzhdin SV, Tullius TD, and Rohs R. GBshape: a genome browser database for DNA shape annotations. *Nucleic Acids Res*. 2015; 43: D103–9. <https://doi.org/10.1093/nar/gku977> PMID: 25326329
17. Tanese N, Sodroski J, Haseltine WA, Goff SP. Expression of reverse transcriptase activity of human T-lymphotropic virus type III (HTLV-III/LAV) in *Escherichia coli*. *J Virol*. 1986; 59(3): 743–5. PMID: 2426471
18. Williams JS, Eckdahl TT, Anderson JN. Bent DNA Functions as a Replication Enhancer in *Saccharomyces-Cerevisiae*. *Mol Cellular Biol*. 1988; 8(7): 2763–9.
19. Fitzgerald DJ, Anderson JN. Unique translational positioning of nucleosomes on synthetic DNAs. *Nucleic Acids Res*. 1998; 26(11): 2526–35. <https://doi.org/10.1093/nar/26.11.2526> PMID: 9592133
20. Brown JW, Anderson JA. The binding of the chromosomal protein HMG-2a to DNA regions of reduced stabilities. *J Biol Chem*. 1986; 261(3): 1349–54. PMID: 3003066
21. Brock JA, Bloom K. A chromosome breakage assay to monitor mitotic forces in budding yeast. *J Cell Sci*. 1994; 107(4): 891–902.
22. Kent NA, Bird LE, Mellor J. Chromatin analysis in yeast using NP-40 permeabilised sphaeroplasts. *Nucleic Acids Res*. 1993; 21(19): 4653–4.
23. Shure M, Pulleyblank DE, Vinograd J. The problems of eukaryotic and prokaryotic DNA packaging and in vivo conformation posed by superhelix density heterogeneity. *Nucleic Acids Res*. 1977; 4(5): 1183–205. <https://doi.org/10.1093/nar/4.5.1183> PMID: 197488
24. Robyr D, Gegonne A, Wolffe AP, Wahli W. Determinants of vitellogenin B1 promoter architecture—HNF3 and estrogen responsive transcription within chromatin. *J Biol Chem*. 2000; 275(36): 28291–300. <https://doi.org/10.1074/jbc.M002726200> PMID: 10854430
25. Burkhoff AM, Tullius TD. The unusual conformation adopted by the adenine tracts in kinetoplast DNA. *Cell*. 1987; 48(6): 935–43. [https://doi.org/10.1016/0092-8674\(87\)90702-1](https://doi.org/10.1016/0092-8674(87)90702-1) PMID: 3030560
26. Goodsell DS, Kaczor-Grzeskowiak M, Dickerson RE. The crystal structure of CCATTAATGG: implications for bending of B-DNA at TA steps. *J Mol Biol*. 1994; 239(1): 79–96. <https://doi.org/10.1006/jmbi.1994.1352> PMID: 8196049
27. Fernandez AG, Anderson JN. Nucleosome positioning determinants. *J Mol Biol*. 2007; 371(3): 649–68. <https://doi.org/10.1016/j.jmb.2007.05.090> PMID: 17586522
28. Vasudevan D, Chua EYD, Davey CA. Crystal Structures of Nucleosome Core Particles Containing the '601' Strong Positioning Sequence. *J Mol Biol*. 2010; 403(1): 1–10. <https://doi.org/10.1016/j.jmb.2010.08.039> PMID: 20800598
29. Abu-Daya A, Brown PM, Fox KR. DNA sequence preferences of several AT-selective minor groove binding ligands. *Nucleic Acids Res*. 1995; 23(17): 3385–92. <https://doi.org/10.1093/nar/23.17.3385> PMID: 7567447
30. Bailly C, Waring MJ. Comparison of different footprinting methodologies for detecting binding sites for a small ligand on DNA. *J Biomol Struct Dyn*. 1995; 12(4): 869–98. <https://doi.org/10.1080/07391102.1995.10508782> PMID: 7779305
31. Fitzgerald DJ, Dryden GL, Bronson EC, Williams JS, Anderson JN. Conserved patterns of bending in satellite and nucleosome positioning DNA. *J Biol Chem*. 1994; 269(33): 21303–14. PMID: 8063755
32. Germond JE, Hirt B, Oudet P, Gross-Bellark M, Chambon P. Folding of the DNA double helix in chromatin-like structures from simian virus 40. *Proc Natl Acad Sci U S A*. 1975; 72(5): 1843–7. <https://doi.org/10.1073/pnas.72.5.1843> PMID: 168578
33. Keller W. Determination of the number of superhelical turns in simian virus 40 DNA by gel electrophoresis. *Proc Natl Acad Sci U S A*. 1975; 72(12): 4876–80. <https://doi.org/10.1073/pnas.72.12.4876> PMID: 174079
34. Lanzer M, Wertheimer SP, de Bruin D, Ravetch JV. Chromatin structure determines the sites of chromosome breakages in *Plasmodium falciparum*. *Nucleic Acids Res*. 1994; 22(15): 3099–103. <https://doi.org/10.1093/nar/22.15.3099> PMID: 8065922

35. Legault J, Tremblay A, Ramotar D, Mirault ME. Clusters of S1 nuclease-hypersensitive sites induced in vivo by DNA damage. *Mol Cell Biol*. 1997; 17(9): 5437–52. <https://doi.org/10.1128/mcb.17.9.5437> PMID: 9271420
36. Sung P, Trujillo KM, Van Komen S. Recombination factors of *Saccharomyces cerevisiae*. *Mutat Res*. 2000; 451(1–2): 257–75. [https://doi.org/10.1016/s0027-5107\(00\)00054-3](https://doi.org/10.1016/s0027-5107(00)00054-3) PMID: 10915877
37. Baraldi PG, Bovero A, Fruttarolo F, Preti D, Tabrizi MA, Pavani MG, Romagnoli R. DNA minor groove binders as potential antitumor and antimicrobial agents. *Med Res Rev*. 2004; 24(4): 475–528. <https://doi.org/10.1002/med.20000> PMID: 15170593
38. Dardonville C. Recent advances in antitrypanosomal chemotherapy: patent literature 2002–2004. *Expert Opin Ther Pat*. 2005; 15(9): 1241–57.
39. Vanden Eynde JJ, Mayence A, Huang TL, Collins MS, Rebholz S, Walzer PD, Cushion MT. Novel bis-benzamidines as potential drug candidates for the treatment of *Pneumocystis carinii* pneumonia. *Bioorg Med Chem Lett*. 2004; 14(17): 4545–8. <https://doi.org/10.1016/j.bmcl.2004.06.034> PMID: 15357989
40. Huang TL, Tao B, Quarshie Y, Queener SF, Donkor IO. N, N'-Bis [4-(N-alkylamidino) phenyl] homopiperazines as anti-*Pneumocystis carinii* agents. *Bioorg Med Chem Lett*. 2001; 11(20): 2679–8. [https://doi.org/10.1016/s0960-894x\(01\)00541-8](https://doi.org/10.1016/s0960-894x(01)00541-8) PMID: 11591500
41. Cortesi R, Esposito E. Distamycins: Strategies for possible enhancement of activity and specificity. *Mini Rev Med Chem*. 2010; 10(3): 217–30. <https://doi.org/10.2174/138955710791185055> PMID: 20408803
42. Howard OM, Oppenheim JJ, Hollingshead MG, Covey JM, Bigelow J, McCormack JJ, Buckheit RW Jr, Clanton DJ, Turpin JA, Rice WG. Inhibition of in vitro and in vivo HIV replication by a distamycin analogue that interferes with chemokine receptor function: A candidate for chemotherapeutic and microbicidal application. *J Med Chem*. 1998; 41(13): 2184–93. <https://doi.org/10.1021/jm9801253> PMID: 9632350
43. Dickinson LA, Gulizia RJ, Trauger JW, Baird EE, Mosier DE, Gottesfeld JM, et al. Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proc Natl Acad Sci U S A*. 1998; 95(22): 12890–5. <https://doi.org/10.1073/pnas.95.22.12890> PMID: 9789010
44. Tutter A, Jones KA. Chemicals that footprint DNA: Hitting HIV-1 in the minor groove. *Proc Natl Acad Sci U S A*. 1998; 95(22): 12739–41.
45. Fitzgibbon JE, Mazar S, Dubin DT. A new type of G→A hypermutation affecting human immunodeficiency virus. *AIDS Res Hum Retroviruses*. 1993; 9(9): 833–8. <https://doi.org/10.1089/aid.1993.9.833> PMID: 7504935
46. Vartanian JP, Plikat U, Henry M, Mahieux R, Guillemot L, Meyerhans A, Wain-Hobson S. HIV genetic variation is directed and restricted by DNA precursor availability. *J Mol Biol*. 1997; 270(2): 139–51. <https://doi.org/10.1006/jmbi.1997.1104> PMID: 9236117
47. Nussinov R. Compositional variations in DNA sequences. *Comput Appl Biosci*. 1991; 7(3): 287–93. <https://doi.org/10.1093/bioinformatics/7.3.287> PMID: 1913208
48. Phillips MD, Nascimbeni B, Tice RR, Shelby MD. Induction of micronuclei in mouse bone marrow cells: an evaluation of nucleoside analogues used in the treatment of AIDS. *Environ Mol Mutagen*. 1991; 18(3): 168–83. <https://doi.org/10.1002/em.2850180305> PMID: 1915312
49. Rogers DJ. Trypanosomiasis' risk'or'challenge': a review. *Acta Trop*. 1985; 42(1): 5–23. PMID: 2859750
50. Tidwell RR, Jones SK, Geratz JD, Ohemeng KA, Cory M, Hall JE. Analogs of 1, 5-bis (4-amidinophenoxy) pentane (pentamidine) in the treatment of experimental *Pneumocystis carinii* pneumonia. *J Med Chem*. 1990; 33(4): 1252–7. <https://doi.org/10.1021/jm00166a026> PMID: 2319567
51. Connor TH, Trizna ZJTI. Pentamidine isethionate is negative in tests for microbial mutagenicity and chromosomal breakage in vitro. *Toxicol Lett*. 1992; 63(1): 69–74. [https://doi.org/10.1016/0378-4274\(92\)90108-v](https://doi.org/10.1016/0378-4274(92)90108-v) PMID: 1412524
52. Boos G, Stopper H. Genotoxicity of several clinically used topoisomerase II inhibitors. *Toxicol Lett*. 2000; 116(1–2): 7–16. [https://doi.org/10.1016/s0378-4274\(00\)00192-2](https://doi.org/10.1016/s0378-4274(00)00192-2) PMID: 10906417
53. Bielawski K, Bielawska A, Poplawska B, Bolkun-Skornicka U. Synthesis, DNA-binding affinity and cytotoxicity of the dinuclear platinum(II) complexes with berenil and amines ligands. *Acta Pol Pharm*. 2008; 65(3): 363–70. PMID: 18646556
54. Sou K, Goins B, Oyajobi BO, Travi BL, Phillips WT. Bone marrow-targeted liposomal carriers, *Expert Opin Drug Deliv*. 2011; 8(3): 317–328. <https://doi.org/10.1517/17425247.2011.553218> PMID: 21275831
55. Xiao Q, Guo D, Chen S. Application of CRISPR/Cas9-Based Gene Editing in HIV-1/AIDS Therapy. *Front Cell Infect Microbiol*. 2019; 9: 69. <https://doi.org/10.3389/fcimb.2019.00069> PMID: 30968001

56. Yarrington RM, Verma S, Schwartz S, Trautman JK, and Carroll D. Nucleosomes inhibit target cleavage by CRISPR-Cas9 in vivo. *Proc Natl Acad Sci U S A*. 2018; 115(38): 9351–9358. <https://doi.org/10.1073/pnas.1810062115> PMID: 30201707
57. Jensen KT, Fløe L, Petersen TS, Huang J, Xu F, Bolund L, Luo Y, Lin L. Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. *FEBS Lett*. 2017; 591: 1892–1901. <https://doi.org/10.1002/1873-3468.12707> PMID: 28580607
58. Uusi-Mäkelä MIE, Barker HR, Bäuerlein CA, Häkkinen T, Nykter M, Rämetsä M. Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (*Danio rerio*). *PLoS One*. 2018; 13(4): e0196238. <https://doi.org/10.1371/journal.pone.0196238> PMID: 29684067
59. Neamati N, Mazumder A, Sunder S, Owen JM, Tandon M, Lown JW, Pommier Y. Highly potent synthetic polyamides, bisdistamycins, and lexitropsins as inhibitors of human immunodeficiency virus type 1 integrase. *Mol Pharmacol*. 1998; 54(2): 280–90. <https://doi.org/10.1124/mol.54.2.280> PMID: 9687569
60. Reddy BS, Sondhi SM, Lown JW. Synthetic DNA minor groove-binding drugs. *Pharmacol Ther*. 1999; 84(1): 1–111. [https://doi.org/10.1016/s0163-7258\(99\)00021-2](https://doi.org/10.1016/s0163-7258(99)00021-2) PMID: 10580832
61. Sudarsanam P, Winston F. The Swi/Snf family nucleosome-remodeling complexes and transcriptional control. *Trends Genet*. 2000; 16(8): 345–51. [https://doi.org/10.1016/s0168-9525\(00\)02060-6](https://doi.org/10.1016/s0168-9525(00)02060-6) PMID: 10904263
62. Gurova KV, Chromatin Stability as a Target for Cancer Treatment. *Bioessays*. 2019; 41(1): e18000141.
63. Satchwell SC, Drew HR, Travers AA. Sequence periodicities in chicken nucleosome core DNA. *J Mol Biol* 1986; 191(4): 659–675. [https://doi.org/10.1016/0022-2836\(86\)90452-3](https://doi.org/10.1016/0022-2836(86)90452-3) PMID: 3806678
64. Collings CK, Fernandez AG, Pitschka CG, Hawkins TB, Anderson JN. Oligonucleotide Sequence Motifs as Nucleosome Positioning Signals. *PLoS One*. 2010; 5(6): e10933. <https://doi.org/10.1371/journal.pone.0010933> PMID: 20532171
65. Collings CK, Waddell PJ, Anderson JN. Effects of DNA methylation on nucleosome stability. *Nucleic Acids Res*. 2013; 41(5): 2918–31. <https://doi.org/10.1093/nar/gks893> PMID: 23355616
66. Collings CK, Anderson JN. Links between DNA methylation and nucleosome occupancy in the human genome. *Epigenetics & Chromatin*. 2017; 10(1): 18.
67. Fitzgerald DJ, Bronson EC, Anderson JN. Compositional similarities between the human immunodeficiency virus and surface antigens of pathogens. *AIDS Res Hum Retroviruses* 1996; 12: 99–106. <https://doi.org/10.1089/aid.1996.12.99> PMID: 8834459
68. Tan VY. Effects of minor groove binding drugs on nucleosome assembly disruption and positioning of surface antigen genes from parasites. Unpublished Master's thesis. 2007; Purdue University, West Lafayette, Indiana 47906.
69. Schmid M, Feichtinger W, Jessberger A, Köhler J, Lange R. The fragile site (16) (q22): Induction by AT-specific DNA-ligands and population frequency. *Hum Genet*. 1986; 74(1): 67–73. <https://doi.org/10.1007/bf00278788> PMID: 3759087
70. Sarni D, Kerem B. The complex nature of fragile site plasticity and its importance in cancer. *Curr Opin Cell Biol*. 2016; 40: 131–6. <https://doi.org/10.1016/j.ceb.2016.03.017> PMID: 27062332
71. Hsu YY, Wang YH. Human fragile site FRA16B DNA excludes nucleosomes in the presence of distamycin. *J Biol Chem*. 2002; 277(19): 17315–9. <https://doi.org/10.1074/jbc.M200901200> PMID: 11880377
72. Yu S, Mangelsdorf M, Hewett D, Hobson L, Baker E, Eyre HJ, Lapsys N, Le Paslier D, Doggett NA, Sutherland GR, Richards RI. Human chromosomal fragile site FRA16B is an amplified AT-rich minisatellite repeat. *Cell*. 1997; 88(3): 367–74. [https://doi.org/10.1016/s0092-8674\(00\)81875-9](https://doi.org/10.1016/s0092-8674(00)81875-9) PMID: 9039263