Nucleosome stability dramatically impacts the targeting of somatic hypermutation

Prashant Kodgire¹, Priyanka Mukkawar¹, Justin A. North³, Michael G. Poirier³,⁴,⁵ and Ursula Storb¹,²,#

¹ Department of Molecular Genetics and Cell Biology, and ² Committee on Immunology, University of Chicago, 920 East 58th Street, Chicago, IL 60637, United States
³ Department of Physics, ⁴ Department of Biochemistry, ⁵ Department of Molecular Virology, Immunology and Medical Genetics, Ohio State University, Columbus, OH, 43210, United States

# Correspondence should be addressed to Dr. Ursula Storb
Email : stor@uchicago.edu
Phone : 773-702-4440

Running title: Nucleosome stability impacts the targeting of SHM

Word count for the introduction, Results, and Discussion sections: 3877
Word count for the Materials and Methods section: 1243
<table>
<thead>
<tr>
<th>Abbreviations used</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AID</td>
<td>activation-induced cytidine deaminase</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>EMSA</td>
<td>electromobility shift analysis</td>
</tr>
<tr>
<td>Ig</td>
<td>immunoglobulin</td>
</tr>
<tr>
<td>MNase</td>
<td>micrococcal nuclease</td>
</tr>
<tr>
<td>NPS</td>
<td>nucleosome positioning sequence</td>
</tr>
<tr>
<td>SHM</td>
<td>somatic hypermutation</td>
</tr>
</tbody>
</table>
ABSTRACT

Somatic hypermutation (SHM) of immunoglobulin (Ig) genes is initiated by the activation-induced cytidine deaminase (AID). However, the influence of chromatin on SHM remains enigmatic. Our previous cell-free studies indicate that AID cannot access nucleosomal DNA in the absence of transcription. We have now investigated the influence of nucleosome stability on mutability in vivo. We introduced two copies of a high affinity nucleosome positioning sequences (MP2) into a variable Ig gene region to assess its impact on SHM in vivo. The MP2 sequence significantly reduces the mutation frequency throughout the nucleosome and especially near its center, despite similar proportions of AID hotspots as in Ig genes. A weak positioning sequence (M5) was designed based on rules deduced from published whole genome analyses. Replacement of MP2 with M5 resulted in much higher mutations throughout the nucleosome. This indicates that both nucleosome stability and positioning significantly influence the SHM pattern. We postulate that, unlike RNA polymerase, AID has reduced access to stable nucleosomes. This study outlines the limits of nucleosome positioning for SHM of Ig genes and suggests that stable nucleosomes may need to be disassembled for access of AID. Possibly the variable regions of Ig genes have evolved for low nucleosome stability to enhance access to AID, DNA repair factors and error-prone polymerases and hence, maximize variability.
INTRODUCTION

The somatic hypermutation of antibody genes is initiated by the activation-induced cytidine deaminase (AID) that creates cytosine (C) to uracil (U) mutations, starting after ~100-200 bp from the promoter and extending for about 2 kb. During SHM these 'U's are repaired in error-prone fashion via translesion DNA polymerases leading to mutations at and near the 'U', reviewed in (38).

Absence of AID, results in a variety of immunodeficiencies (6), but on the other hand, AID is a dangerous oncogenic mutator, reviewed in (24). Additionally, DNA demethylation via AID may be essential for normal early development and perhaps some aspects of DNA methylation in general (9, 12, 21, 30, 31). Thus, the study of the molecular mechanisms of AID action is essential for understanding the roles of AID in immunity and oncogenesis, as well as development. The process of SHM requires transcription without requiring a specific promoter (3, 4) but is linked to transcription initiation (25). We have postulated that AID is crucially associated with the transcription complex and may target negative supercoils, as they arise in the wake of the transcription complex during transcript elongation (34). Transcription occurs in the context of chromatin, which sterically occludes DNA binding complexes (15, 18, 28, 29). To begin to understand the role of chromatin in SHM we previously investigated the effect of a strong nucleosome positioning sequence (MP2) on the function of AID in a cell-free system (33). Nucleosomes positioned within a circular plasmid that was susceptible to AID-induced cytosine deamination when the DNA was naked, inhibited AID access specifically to the sequences associated with histone
octamers. However, when the nucleosomal region was transcribed by the phage RNA polymerase T7, it underwent efficient cytosine deaminations suggesting that AID, unlike RNA polymerase, cannot access tightly wrapped DNA. Since transcription in these cell-free assays was by the small T7 polymerase it was possible that the eukaryotic polymerase pol II was less able to unwrap tight nucleosomes sufficiently. Indeed, the effect of tightly positioned nucleosomes and chromatin on the targeting of AID \textit{in vivo} remained unknown. We report here a study in which the same MP2 sequence was introduced into the variable (V) region of an immunoglobulin (Ig) gene by homologous integration in cells that normally undergo SHM in culture. Surprisingly, the presence of the MP2 sequence affected the efficiency of SHM \textit{in vivo}.

\section*{MATERIALS AND METHODS}

\subsection*{Cell culture and transgenic clones}

DT40 \psi V knock-out cells derived from avian leukosis virus–induced chicken bursal B cells were a gift of H. Arakawa and J.M. Buerstedde (Institute of Molecular Radiology, Neuherberg, Germany) \cite{2}. The cells were cultured in RPMI 1640 with 1\% penicillin/streptomycin, 1\% l-glutamine (Invitrogen), 1\% chicken serum, and \(\beta\)-mercaptoethanol (Sigma-Aldrich) at 39.5\(^\circ\)C with 5\% CO\(_2\). MP2-MP2 (M5-MP2, MP2-M5 and M5-M5) knock-in constructs were linearized with Sall and transfected as previously described \cite{2}. After 12 h, transfected cells were treated with 20 \(\mu\)g/ml Blasticidin for selection of blasticidin resistance and single clones were isolated by subsequent limiting dilutions. Single clones from
limiting dilutions were expanded to perform FACS analysis for surface-IgM negative clones and subsequently collect their genomic DNA for Southern blotting. Genomic DNA was digested with SbfI and MluI, respectively and a radioactive probe for the MP2 region was used for Southern blot analysis.

**Flow-cytometric analysis and cell sorting**

DT40 knock-in clones for MP2 / M5 sequence were stained with PE-conjugated anti–chicken IgM antibody (Santa Cruz Biotechnology, Inc.) and were analyzed for loss of surface-IgM and presence of AID-IRES-GFP expression of 50,000 live cells on an LSR II (BD) using DT40 CL18 cells and GFP+ ψV KO (a gift of H. Arakawa and J.M. Buerstedde (2)) as gating controls. DT40 knock-in clones treated with tamoxifen were sorted for AID-IRES-GFP+ single cells on a cell sorter (FACSAria; BD) at the University of Chicago Flow Cytometry Facility.

**Q-PCR analysis**

Real-time PCRs were run and analyzed on a MYiQ system with SYBR Green SuperMix (both from Bio-Rad Laboratories). Primers used were pk58 and pk59 (Table S1) for the MP2 region, and pk61 and pk62 (Table S1) for the spacer between the MP2 regions. For the analysis of the M5 mono-nucleosomes, the primers used were pk71 and pk72 (Table S1) for the M5 region and pk156 and pk157 (Table S1) for the spacer between the M5 regions. PCR conditions were 95°C for 30 s, 64°C for 45 s, and 72°C for 60 s for 40 cycles. The values are
normalized for the copy number and primer efficiencies using the Pfaffl method (26).

RT-PCR analysis of transcripts in MP2/M5 knock-in clones

Total RNA was made from DT40 cells with RNA STAT-60 (Tel-Test Inc.), recovered in 50 µl, and stored at −80°C. Equal amounts of RNA were used for making cDNA by the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen). Real-time PCRs were run and analyzed on a MYiQ system with SYBR Green SuperMix (both from Bio-Rad Laboratories). Primers used were pk142, pk24 (Table S1) for the IgL V region, and gg-actin1, gg-actin2 (Table S1) for the chicken actin region. PCR conditions were 95°C for 30 s, 64°C for 45 s, and 72°C for 60 s for 40 cycles. The data from chicken actin were used as a reference for the relative quantification of IgL V region levels using the Pfaffl method (26).

Identification of somatic mutations

Mutations in knock-in clones were detected by PCR cloning using Pfu polymerase (Agilent Technologies), and primers PK64 and PK65 (Table S1) were used for PCR cloning with Pfu polymerase at 95°C for 30 s, 67°C for 30 s, and 72°C for 160 s for 25 cycles and cloned in a PCR cloning kit (Zero Blunt TOPO; Invitrogen). DNA sequencing was performed by the University of Chicago Cancer Research Center DNA Sequencing Facility.
DNA Synthesis

The 147bp MP2, M5, or *L. variegates* 5S (35) nucleosome positioning sequences were cloned into pUC19 such that the nucleotide sequences flanking MP2, M5, or 5S were homologous. MP2-247, M5-247, and 5S-247 (5S ribosomal RNA sequence) were amplified by PCR to contain 50bp of DNA flanking each side of the positioning sequence, Cy5 on the 5’ end of the forward strand and Cy3 on the 5’ end of the reverse strand. 5’-amine-labeled primers (Sigma) were conjugated with Cy3-NHS or Cy5-NHS (GE Healthcare) and purified by reverse phase HPLC (Vydac C18). The forward primer was mgp1 and the reverse primer was mgp2 (Table S1). Amplified DNA was purified by HPCL on a Gen-Pak FAX ion exchange column (Waters).

Nucleosome Preparation

Nucleosomes for exonuclease III mapping were reconstituted by salt double dialysis as previously reported (19) with 1 μg of cy3/cy5 labeled MP2-247 or M5-247 DNA, 3 μg of lambda DNA (Invitrogen) and 1.5 μg of purified HO (histone octamer). The DNA and HO were mixed in 50 μl of 0.5x TE (pH 8.0) with 2 M NaCl and 1 mM BZA (benzamidine). The sample was loaded into an engineered 50 μl dialysis chamber which was placed in a large dialysis tube with 80 ml of 0.5x TE (pH 8.0) with 2 M NaCl and 1 mM BZA. The large dialysis tube was extensively dialyzed against 0.5x TE with 1 mM BZA. The 50 μl sample was extracted from the dialysis button and purified by sucrose gradient centrifugation.
**Competitive Reconstitutions**

Competitive reconstitutions were performed as previously described (19). Reconstitutions were prepared in 2 M NaCl, 0.5x TE, 1 mM BZA with 6 ng/μl labeled MP2-247, M5-247, or 5S-247 DNA (5S ribosomal RNA sequence), 50 ng/μl buffer DNA, and 10 ng/μl of HO in a volume of 50 μl. To minimize variation in DNA and HO concentrations, we first prepared a HO and buffer DNA master mix that was split and combined with each DNA stock. Each DNA sample was then split into thirds and dialyzed separately. Each sample was dialyzed against the same reservoir containing 0.2 liters of 2 M NaCl, 0.5x TE, and 1 mM BZA. The concentration of salt in the dialysis reservoir was slowly reduced to 200 mM over 24 h; the samples were then dialyzed overnight against 0.5x TE and 1 mM BZA to reduce the final NaCl concentration to 1 mM NaCl. The reconstitution products were examined by PAGE, scanned with a Typhoon 8600 variable mode imager (GE Healthcare), and analyzed with ImageQuant (GE Healthcare).

**Electrophoresis Mobility Shift Assay**

The population of positioned and depositioned nucleosomes on the MP2 and M5 sequences was resolved by 5% native acrylamide gel with 0.3x TBE at 300V for 1 hr and imaged by Cy5 using a Typhoon 8600 variable mode imaged (GE Healthcare). The fraction of centrally positioned nucleosomes was calculated by measuring the image intensity within a box drawn around the positioned nucleosome band and dividing it by the intensity within a box drawn around all...
nucleosome bands using Image Quant software (Invitrogen) with local median background subtraction enabled.

Exonuclease III Mapping

The nucleosome positions within the MP2-247 or M5-247 DNA molecules were determined with ExoIII mapping as previously reported (19). Reactions were carried out with 10nM nucleosomes in 30U/ml of ExoIII (NEB) and Buffer 1 (NEB) at 37°C. At each time point, 7ul of the reaction was quenched with a final concentration of 20mM EDTA. A final concentration of 1 mg/ml of proteinase K and 0.02% of SDS was added to each time point to remove the histone octamer from the DNA. Samples were separated by 8% denaturing PAGE in 7 M Urea and 1x TBE. The sequence markers were prepared with a SequiTherm Excel II DNA sequencing kit (Epicentre) using the Cy5 or Cy3 labeled primers, MP2-247 or M5-247 DNA template and either ddATP or ddTTP. Results were imaged by a Typhoon 8600 variable mode imager (GE Healthcare), which detects cy3 and cy5 separately in the same gel. The cy3 and cy5 ladders could be loaded in the same lanes to increase accuracy of the mapping gel readout.

RESULTS

Controlling nucleosome positioning and stability within the IgL locus.

To explore how SHM is influenced by nucleosomes within chromatin, we placed two copies of the strong nucleosome positioning sequence (NPS), MP2, that inhibited AID access in vitro (33) into the active lambda gene of mutating B cells,
DT40, by homologous recombination (Fig. 1). The cell line we used is a variant of DT40 cells that is an AID knock-out and expresses AID as a transgene (AID-IRES-GFP) (2); all 25 ψV IgL genes are deleted (2) to make sure that these cells do not undergo IgL gene conversion. We chose a 147 bp MP2 sequence as a NPS for our study since it is reported to be a strong nucleosome positioning sequence (28). It has a significant number of AID hotspots (14.3 %), very similar to the IgL gene (11.6 %). We integrated two MP2 sequences, with a spacer of 76 bp between them, into the IgL locus in such a way that the MP2 sequences will be at 235 bp and 458bp, respectively, from the transcription start site, which is the peak region of SHM in Ig genes.

Nucleosome occupancy and positioning appear to be partially regulated by the underlying DNA sequence (16, 22, 32, 43). AA/TT and/or TA di-nucleotides spaced every 10 base pairs (bp) and out of phase with GC di-nucleotides that are also spaced by 10 bp appear to have the highest preference for forming nucleosomes (17). The MP2 sequence is a variant of the strong 601 positioning sequence, which was determined by SELEX experiments (17). The MP2 sequence has a significant number of AA, TT and TA di-nucleotides at around 10 bp intervals that are out of phase with GC di-nucleotides (Fig. 2A).

To validate our results with the stable nucleosome positioning sequence, MP2, we created a control sequence that is less favored to interact with histone proteins, lacking all the AA/TT/TA repeats that are a signature of a strong NPS, eventually reducing its affinity to the histone octamer, which is expected to make it a weaker positioning sequence, and test the influence of nucleosome stability.
and positioning on AID accessibility. During the design of the 147 bp control sequence we replaced all AA/TT/TA repeats, however, kept the number of GCs and AID hotspots the same as that of the MP2 sequence (Fig. 2B). Aligning this control sequence M5 with MP2, there is no periodicity of either of AA/TT/TA di-nucleotides as well as GC di-nucleotides (Fig. 2A, B).

To assess the influence of the DNA base changes, which convert MP2 to M5, on the DNA-histone binding affinity, we carried out competitive nucleosome reconstitutions (40). Nucleosomes were reconstituted with histone octamer, an excess of low affinity competitor DNA, and either the MP2-247, M5-247 or 5S-247 DNA molecules that were fluorophore labeled at the 5-prime ends. The 5S positioning sequence was used to allow for comparison to previous competitive reconstitution studies (40). A dynamic equilibrium between free DNA and DNA wrapped around a histone octamer is established and EMSA (Fig. 3A) is used to determine the equilibrium constant, $K_{eq}$, between these DNA states. We determined $K_{eq}$ relative to the 5S sequence for MP2 (11 ± 4) and M5 (0.6 ± 0.2) (Fig. 3B) from the relative intensities of the nucleosome bands to the DNA band (Fig. 3C). The ratio of the relative $K_{eq}$ for MP2 and M5 indicates that MP2 is 18 (11/0.6) times more probable to form a nucleosome than M5. We also determined the relative free energy of nucleosome formation with MP2-247 (-1.4 ± 0.2 kcal/mol) and M5-247 (0.4 ± 0.3 kcal/mol) relative to the 5S-247 DNA molecule from $\Delta\Delta G = -k_B T \ln(K_{eq}/K_{eq-5S})$, where $k_B T = 0.6$ kcal/mol (Fig. 3C). This demonstrates that MP2 has a 1.8 kcal/mol lower free energy than M5 and that...
nucleosomes containing MP2 are significantly more stable than nucleosomes containing M5.

To determine the influence of this reduction in DNA-histone binding on nucleosome positioning, we quantified nucleosome positions by electromobility shift analysis (EMSA) of nucleosomes reconstituted with the MP2 sequence and the M5 sequence (Fig. 3D). Each sequence was centrally located within a 247 bp DNA test molecule (Fig. 3F). In the EMSA assay the slowest mobility band contains nucleosomes that are centrally located within the DNA molecule and the highest mobility band contains nucleosomes that are located at either end of the 247 bp DNA molecule (19) (Fig. 3D). Fig. 3D clearly shows that nucleosomes within the MP2 sequence are largely centrally located with the only other position being at the ends of the DNA molecule. The M5 control sequence has a reduced fraction of nucleosomes at the central position, additional shifted positions and is largely positioned at the ends of the DNA molecule (Fig. 3D). Quantification of the band intensities finds that 70% of nucleosomes are centrally positioned within the MP2 DNA molecule, whereas less than 35% of the nucleosomes within M5 are centrally positioned (Fig. 3E).

To confirm that the findings in the gel shift analysis are due to changes in nucleosome position, we performed exonuclease III mapping of both the MP2 positioning sequence as well as the M5 sequence (Fig. 3F-H). In this experiment a 247 bp linear DNA was used in which the central 147 bp DNA is either the MP2 or the M5 sequence (Fig. 3F). Purified nucleosomes are reconstituted with the same 247 bp DNA molecules that are labeled at one 5’ end with Cy5 and at the
other end with Cy3. The nucleosomes are treated with exonuclease III (Fig. 3F) and subsequently analyzed by denaturing polyacrylamide gel electrophoresis. Nucleosomes containing the MP2 sequence had a single stall position about 50 bp into the DNA molecule for both strands (Fig. 3G). This confirms that the nucleosomes are centrally located with a 147 bp footprint. However, exonuclease III mapping of nucleosomes containing the M5 positioning sequence showed a number of stall positions (Fig. 3H), that are consistent with nucleosome positions observed by EMSA (Fig. 3D, E), which confirms that positioning within DNA molecules containing M5 is significantly reduced relative to MP2.

After confirming that the M5 sequence has low affinity to histones we replaced either the first or the second or both copies of the MP2 positioning sequence with the less efficient M5 positioning sequence (Fig. 4A). Similar to the MP2-MP2 knock-in construct (Fig. 1), we created three control knock-in constructs, namely, M5-MP2, MP2-M5 and M5-M5, respectively.

The knock-in plasmids were constructed by cloning genomic sequences that flank the integration site in the IgL gene. A surface IgM positive clone of DT40 ψV knock-out cells was used for transfection so that targeted integration into the IgL locus can be easily detected by the loss of surface IgM expression. The targeted integrations were confirmed by Southern blotting with the MP2/M5 region and V region of the IgL gene as probes (data not shown). The blasticidin (Bsr) drug marker gene present between two LoxP sites (Fig. 1) was excised by treating the cells with tamoxifen. Incidentally, the AID-IRES-GFP transgene is also flanked by two LoxP sites and is likely to be excised after treatment with
tamoxifen. We used a very low concentration of tamoxifen (25 nM) for 24 hours so that a fraction of cells retained the AID-IRES-GFP transgene. Cell clones that were GFP positive and had excised the blasticidin drug marker gene were selected and the blasticidin drug marker gene excision was confirmed by PCR and Southern blotting (data not shown).

To test whether the presence of a strong or weak NPS influenced transcription through the IgL gene, we performed RTQ-PCR in the V region of the IgL gene in MP2-MP2, M5-MP2, MP2-M5 and M5-M5 knock-in clones, respectively and observed that knocking in a pair of strong positioning sequences (MP2) does not decrease transcription through the IgL gene (Fig. 4B).

To confirm whether nucleosomes are assembled at the MP2 sequence in the DT40 cells, we performed a Micrococcal Nuclease (MNase) assay for the MP2-MP2 knock-in clones. We treated nuclei from these cells with MNase, gel eluted mono-nucleosomes (i.e. a 150 bp band), and then performed PCR amplification for the positioning sequence as well as the spacer region. In Fig. 5A, lanes 1 & 2 are PCR bands with MP2 sequence specific primers (pk58 and pk59, Table S1), whereas lanes 3 & 4 are specific for the spacer region (pk61 and pk62, Table S1). Lanes 6 and 7 are control PCR reactions with the genomic DNA from the MP2-MP2 containing cell clones as templates. We observed a very strong band for the nucleosome region compared with the spacer region, suggesting that indeed nucleosomes were assembled at the MP2 sequence in the DT40 cells. Quantitation by Q-PCR analysis showed that the MP2 region was ~7 fold more abundant compared to the spacer region (Fig. 5B) confirming
nucleosome assembly at the MP2 sequence. Similarly, to compare the relative
stability of nucleosomes at the M5 sequence, Q-PCR analysis was performed
with the gel eluted mono-nucleosomes from the M5-M5 knock-in clones using
primers specific to the M5 sequence (pk71 and pk72, Table S1) as well as the
spacer region (pk156 and pk157, Table S1) and observed that the M5 region
was only ~2 fold more abundant compared to the spacer region (Fig. 5B). These
results combined with our observation that MP2 is 7 times more abundant than
the spacer region suggest that nucleosome occupancy at the M5 sequence is
about four times less than at the MP2 sequence in the DT40 cells. This is
consistent with our in vitro results that MP2 is significantly more stable than M5
(Fig. 3).

Nucleosomes significantly influence the mutation pattern of the IgL locus
The MP2-MP2 knock-in clones were cultured for 5 weeks to acquire mutations in
the IgL gene. A 1.2 kb PCR product of genomic DNA was amplified and
sequenced encompassing MP2 as well as the VJ region (Fig. 6A). Fig. 6B
shows the somatic hypermutation pattern of the MP2-MP2 knock-in clones. In the
strong nucleosome positioning sequences (MP2) we observed reduction in
mutation frequencies compared with the flanks (Fig. 6C). The number of
mutations was 2-4 times reduced in both copies of the MP2 nucleosome
positioning sequence compared with the spacer and neighboring IgL V region.
Moreover, mutation frequencies per AID hotspot in the two nucleosome
positioning sequences were also considerably lower, despite a similar proportion
of AID hotspots (Fig. 6C). We observed a predominance of single nucleotide substitutions with few insertions and deletions (Fig. S1A). The clones show very few mutations in A/T bases and a preference for transversion mutations (Figs. S1B, S2) as was found by others in DT40 (2). Mutations at G bases were more than at C bases in the regions surrounding MP2 (Fig. 7B); this relationship was however reversed in MP2 (Fig. 7A). Finally, we find mutations are most significantly suppressed in the central region of the MP2 nucleosome, while mutations near the entry/exit regions of the nucleosome are the least suppressed (Fig. 7C, E). These results indicate that the presence of stably positioned nucleosomes in the immunoglobulin gene significantly affects the accessibility of AID to and the mutation patterns within Ig genes.

Reduction in nucleosome stability alters the mutation pattern within the IgL locus

After 5 weeks of culturing we analyzed the mutation profiles of the M5-MP2, MP2-M5 and M5-M5 knock-in clones and compared them with the SHM profile of MP2-MP2 knock-in clones. With the M5-MP2 construct the percent mutations in the M5 region were around 3 times higher than in the corresponding MP2 region of the MP2-MP2 construct (P value: 0.0026), whereas mutations in the second MP2 region were almost the same (Fig. 6D). We also observed significant increases in the percent of mutations in the neighboring regions of M5 (Fig. 6E). Similarly, with MP2-M5, the percent mutations in the M5 region were around 3 times higher than in the corresponding MP2 region in the MP2-MP2 construct (P
value: 0.0001), whereas the mutation frequency in the first MP2 region was
almost the same (Fig. 6F, G). Finally, when we replaced both copies of the MP2
sequence with M5, we observed a significant increase in the percent of mutations
in both copies of M5 (Fig. 6H). Very similar to M5-MP2 and MP2-M5 constructs,
the first M5 again showed an around three times increase in the percent of
mutations compared to MP2 (P value: 0.0089), and the second M5 showed an
around two times increase in the percent of mutations (P value: 0.0447), but not
in the region 3' of the second NPS (Fig. 6I). Replacing the MP2 sequence with
the M5 sequence also changed the mutation pattern within the nucleosome. The
mutations within the MP2 sequence are suppressed in the center of the
nucleosome relative to the DNA entry/exit region of the nucleosome (Fig. 7C). In
contrast, the distribution of mutations within the M5 sequence remained relatively
constant (Fig. 7D). Furthermore, these patterns were not influenced by whether
the adjacent NPS was M5 or MP2.

We observed a predominance of single nucleotide substitutions with few
insertions and deletions in MP2-MP2, M5-MP2, MP2-M5 and M5-M5 knock-in
clones (Figs. S1A, S2). All four types of knock-in clones show very few
mutations in A/T bases and a preference for transversion mutations (Fig. S1B);
two independent cell clones for each of the four types of nucleosome
combinations were very similar (data not shown). In the total 1.1 kb sequenced
for each cell type, mutations at G bases were more than at C bases (Fig. S1B);
this reflects the SHM pattern outside the NPSs (Fig. 7B), but within the NPSs,
the C, G frequencies were reversed (Fig. 7A). Considering the three 49 bp
regions of the NPSs separately, the central region is less mutated in the MP2 sequence compared with the entry/exit points i.e. sequences on the left and right (Fig. 7C, D). However, the M5 sequence has more mutation events in the central 49bp region, although both the MP2 and the M5 sequence have the same number of seven AID hotspots in this region (Fig. 7E, F). Thus we conclude that, when we replaced a strong nucleosome positioning sequence (MP2) with a weak positioning sequence (M5), mutations in M5 were much higher than in MP2.

DISCUSSION

The rules for nucleosome assembly deduced from total genome analyses (22, 32) enabled us to change the MP2 sequence with high affinity for histone cores to the low affinity M5 sequence. While regulatory mechanisms also play a major role in chromatin structure (16, 43), the striking difference in the biophysical properties of MP2 and M5 nucleosomes validates previous conclusions that the primary DNA sequence can considerably affect the propensity for its assembly into nucleosomes (22, 32).

The findings show that both the presence and stability of the nucleosome strongly influence mutation patterns during SHM: the number of mutations was significantly reduced in both copies of the nucleosome positioning sequence (NPS) MP2 and mutations per AID hotspot in the MP2 sequence were considerably lower as compared to the IgL gene, despite a similar proportion of hotspots. Moreover, replacement of a stable NPS (MP2) with a less stable sequence (M5) resulted in higher mutations. We conclude that the stability of
nucleosomes in the IgL gene significantly affects the outcome of the somatic
hypermutation process.

There are two mechanisms by which AID could gain access to
nucleosomal DNA. One possible mechanism is that the DNA must be
nucleosome-free for AID to access DNA (Fig. 8A). Nucleosomes could be
disassembled by RNA transcription, which is important for deamination by AID
(33). Histone chaperones (7) and/or chromatin remodeling (5, 36) could further
enhance nucleosome disassembly. In this model, the mutation rate should
remain constant through the nucleosome positioning sequence since mutations
only occur when the DNA is nucleosome free. Alternatively, the nucleosomes
could be retained during transcription and AID gains access to DNA through
partial DNA unwrapping (Fig. 8B) and/or nucleosome repositioning (Fig. 8C).

These nucleosome alterations expose DNA that is originally wrapped into a
nucleosome. Transcription through a nucleosome (13) and nucleosome
remodeling could induce nucleosome repositioning (5) and enhance nucleosomal
DNA unwrapping fluctuations, which occur rapidly many times a second (14). In
this model, the mutation rates are expected to be the highest near the entry/exit
regions and lowest near the nucleosome center. DNA site exposure by
unwrapping is greatest near the entry/exit regions and exponentially reduced for
DNA sites further into the nucleosome (1, 29). Also, since nucleosomes are
spaced by 30 to 60 base pairs (42), nucleosome repositioning is restricted, so
again sites near the DNA entry/exit regions will be the most accessible. The
unwrapping/repositioning models are not mutually exclusive, so both could occur in vivo. Our studies of the mutation levels within both MP2 and M5 and our biophysical characterization of MP2 and M5 nucleosomes indicate that both the disassembly and unwrapping/repositioning mechanisms occur in vivo. M5 has both a reduced affinity to the histone octamer and reduced nucleosome positioning strength relative to MP2. Therefore, M5 has the ability to enhance mutations by both nucleosome repositioning and disassembly. Furthermore, the mutation patterns within MP2 and M5 indicate that both mechanisms occur. The mutation frequency within the MP2 sequence is greatest near the entry/exit regions, which is consistent with the unwrapping/repositioning models (Fig. 8B, C). This mutation pattern suggests that the dominant mechanism by which AID gains access to highly stable nucleosomes is by either the unwrapping (Fig. 8B) or the repositioning model (Fig. 8C). Repositioning would likely be affected by the neighboring nucleosomes. Since there is no increase in the mutability of MP2 when the other NPS is M5 rather than MP2 (Fig. 6), the current study supports unwrapping (Fig. 8B). However, the mutation frequency within the M5 sequence is relatively constant in relative concordance with the AID hotspot distributions across M5 (Fig. 7D, F); this finding supports the disassembly mode for M5 (Fig. 8A). Since the M5 sequence has a lower affinity to the histone octamer, this DNA sequence could reduce nucleosome occupancy by both enhancing the rate of nucleosome disassembly and reducing re-assembly. The ~four-fold reduction in nucleosome occupancy produced by replacing MP2 with
M5 (Figs. 3, 5) also resulted in a 3-fold increase in mutation frequency (Fig. 6). The combination of these results strongly suggests that AID accesses less stable nucleosomes largely by the disassembly model.

These *in vivo* findings are interesting given our previous results with *in vitro* assays of the mutability by AID of MP2 embedded in a supercoiled circular plasmid, pKMP2 (33). In contrast to naked pKMP2 DNA, MP2 containing nucleosomes in the plasmid were not mutated by AID alone. However, there were ample mutations in the MP2 nucleosome sequences when the plasmids were transcribed. The arrangement and sequences of the two MP2 and spacer elements were the same as used in this paper. The conclusion was that AID cannot access nucleosomes unless they are transcribed (33). Clearly the Ig lambda gene in the DT40 cells used here is continuously transcribed and MP2 is mutated, however at a reduced frequency compared with the flanking DNAs. It is not simple to make a direct comparison between MP2 and flanks in the *in vitro* experiments; there some regions without a defined nucleosome were slightly more mutable than MP2, others had very few mutations [Fig. 5K in reference (33)]. We do not know whether and where nucleosomes were randomly placed in the ~3.9 kb plasmid outside of the MP2 regions. Interestingly, in the *in vitro* experiments the pKMP2 plasmid was transcribed by T7 RNA polymerase that is considerably smaller than the pol II operating in vertebrate cells. It had been shown that nucleosomes containing a T7 promoter are completely displaced by T7 pol (39). The results reported here show that RNA polymerase PolII can deal
with nucleosomes more efficiently than AID and suggest that subtle epigenetic
events may be best investigated in vivo.

AID requires single-stranded DNA for access (8) and operates
processively (27). Our previous data support the idea that negative supercoils
behind the RNA polymerases (pol) extrude single-stranded Cs as AID targets
(34). The propagation of negative supercoils is probably inhibited at the next
nucleosome. This is supported by comparing the processivity of AID in cell free
assays with that in vivo (37). In vitro up to 16 consecutive Cs are deaminated by
AID in stretches of up to ~60 total NTs (37). In vivo maximally 4-5 consecutive
Cs, but mostly only 2, are mutated in up to 11 total NTs (37), but nevertheless
the AID processitivity is significantly greater than expected also in vivo (p<0.01).
Thus apparently the average length of the spacers between nucleosomes allows
sufficiently large stretches of negative supercoils to develop and become
accessible to AID.

We find that while nucleosome occupancy influences SHM within the Ig
locus, the level of transcription is not significantly influenced by a strong NPS.
Only a 2 fold change in transcription was also observed in budding yeast when
the high affinity NPS, 603, was inserted at the beginning of the CUP1 gene (10).
Interestingly, Gaykalova et al. (10) found that 603 did not position nucleosomes
well in vivo, while our studies find that the MP2 sequence significantly positions
nucleosomes. In addition, a high throughput sequencing study (11) found that a
601 NPS inserted in the EF1a promoter of the human Factor IX gene contained
well-positioned nucleosomes that then became depositioned as the gene was silenced.

An important difference between these previous studies and ours is the location of the NPS with respect to the transcription start site. We inserted the two NPSs into the transcribed region of the gene (235 and 458 base pairs respectively, from the transcription start site). The 603 sequence in the studies by Gaykalova et al. was at the first nucleosome within the transcribed region of the gene (56 base pairs from the transcription start site) and the 601 sequence in the Gracey et al. studies was inserted in the promoter upstream from the transcription start site. The combinations of these results are consistent with the idea that chromatin remodeling may selectively influence nucleosome position near the promoter region of genes. Interestingly, *in vitro* measurements by Gaykolova et al. suggested that chromatin remodeling near the promoter could be responsible for their observed depositioning at the 603 NPS.

SHM experiments in mice showed a certain periodicity of the mutation patterns in a highly mutable Ig transgene, RS (20). The results were consistent with the conclusion that the Ig gene was organized into nucleosomes, but that different cells had different nucleosome phasings and that the nucleosome pattern was relatively stable for a given cell for several generations throughout the hypermutation process.

Clearly, in the current study the inserted MP2 must have caused rather stable nucleosome phasing: In MP2-MP2 on average seven times more of the MP2 sequences are in nucleosomes rather than of the spacers (*Fig. 5B*). Even
the M5 sequence is slightly more nucleosomal than the spacer (Fig. 5B) and about 30% of the M5 sequences are stably associated with histone cores (Fig. 3E). We find that the MP2 sequence, which is a variant of the 601 sequence, is 1.4 kcal/mol lower in free energy than the 5S sequence, while the M5 sequence is 0.4 kcal/mol higher than the 5S sequence. The MP2 sequence is about 1 kcal/mol (23, 40) higher in free energy than the original 601 sequence and is similar to one of the highest in vivo NPSs (40, 41). This indicates that the MP2 sequence is at the extreme of high affinity nucleosome positioning sequences in vivo and that it is not representative of the typical nucleosomal DNA in vivo. The M5 sequence is lower affinity than the well studied 5S NPS, but similar to mouse minor satellite DNA (40). These sequences are about 0.6 kcal/mol lower than average affinity of mouse DNA to histone octamers. This indicates that the M5 sequence is representative of typical nucleosomal DNA in vivo and that our observations of the influence of nucleosomes on SHM at the M5 sequence may apply to nucleosomes in the native Ig locus. Thus this study suggests that the limits of nucleosome positioning for Ig genes may be below MP2 stability and around and below that of M5. It will be interesting to investigate the propensity for nucleosome positioning of endogenous Ig genes in mice and human. It seems possible that the variable regions of Ig genes have evolved for low nucleosome stability to enhance the chance for increased access to AID, DNA repair factors and error-prone DNA polymerases and hence creation of maximal variability by somatic hypermutation.
ACKNOWLEDGEMENTS

We are greatly indebted to J. Widom (Northwestern University, Evanston, Illinois) for advice on the design of the M5 sequence. We thank H. Arakawa and J.M. Buerstedde (Institute of Molecular Radiology, Neuherberg, Germany) for the DT40 CL18 and ψV knock-out cells, W. Buikema and C. Hall for DNA sequencing, and R. Duggan for flow cytometric cell sorting. We are grateful to B. Kee for technical advice and use of her instrument for real-time PCR. We thank T. E. Martin for constructive discussions of these experiments and critical reading of the paper. P.K. thanks the Lady Tata Memorial Trust, UK and the Cancer Research Institute, USA for postdoctoral fellowships. J.A.N. acknowledges support from an Ohio State Comprehensive Cancer Center Pelotonia predoctoral fellowship. This work is supported by the NIH (AI047380 and AI053130 to U.S., and GM083055 to M.G.P.) and a Burroughs-Wellcome Career award to M.G.P.

The authors have no conflicting financial interests.

This article contains supporting information. Primers used in this study are listed in Table S1.
REFERENCES


32


FIGURE LEGENDS

FIG 1 Map of the rearranged Ig light chain locus in the chicken B-cell line DT40. The locus contains a leader, V, J and C region of IgL gene. The strategy of knocking in 2 MP2s (440bp) by the targeted integration is shown.

FIG 2 M5 control sequence (A) Alignment of the MP2 and M5 sequences showing periodicity of AA/TG/TA and GC di-nucleotides. Stretches of AA, TT and TA are underlined, GC are in bold letters. The numbers 0 and 5 relate to the dyad center. AA/TG/TA di-nucleotides are shown in solid arrows on the top, and GC di-nucleotides are shown in dashed arrows on the bottom of the sequence. (B) Sequences of MP2 and M5. Stretches of AA, TT and TA are underlined, GC are in bold letters, c and g in the AID hotspots (WRC and GYW) are in small letters, 47 changes in the MP2 sequence to create M5 are marked with *.

FIG 3 Properties of MP2 and M5 NPSs. (A, B, C) Competitive reconstitution. The probability and differences in the relative free energy of nucleosome formation were determined by competitive reconstitution. (A) Electromobility shift assay was done to quantify the fraction of nucleosomes that formed within MP2-247 (lanes 2-4), M5-247 (lanes 5-7) and 5S-247 in the presence of low affinity and unlabeled carrier DNA. Competitive reconstitution was done in triplicates for the MP2, M5 as well as 5S sequence and each lane represents a separate reconstitution. (B) The equilibrium constants, $K_{eq}$, for the formation of
nucleosomes with MP2 ($K_{eq} = 11 \pm 4$) and M5 ($K_{eq} = 0.6 \pm 0.2$) relative to the 5S positioning sequence. The error bars represent the variation in between the three reconstitutions. (C) The difference in the free energy for nucleosome formation between 5S-247, and either MP2 ($\Delta \Delta G = -1.4 \pm .2$) or M5 ($\Delta \Delta G = 0.4 \pm 0.3$). (D, E) Electrophoretic Mobility Shift Assay. (D) 5% Native PAGE gel of nucleosomes reconstituted on MP2-247 or M5-247 DNA. Mobility of a nucleosome through the gel is dependent upon its position on the DNA. (E) Fraction of centrally positioned nucleosomes with respect to all nucleosome positions from MP2-247 and M5-247. (F, G, H) Exonuclease III mapping of MP2 and M5 nucleosome positioning sequences. (F) Map of the nucleosome and flanking DNA. (G, H) Exonuclease assays, details in Methods (33). The Cy5 or Cy3 labeled nucleosomes were treated with exonuclease III (G: lanes 1-8, MP2 and H: lanes 1-7, M5 are increasing incubation times from 0 to 30 mins). Sequencing ladders (two right lanes in each group) were prepared with ddATP and ddTTP.

FIG 4 Replacing either the first or second or both the MP2 with M5. (A) Four combinations of MP2 and M5 inserted in the DT40 Ig lambda locus. (B) Transcription levels at the IgL V region in the four knock-in clones. Histograms show relative mRNA levels to the MP2-MP2 knock-in clone; the values are normalized with chicken $\beta$-actin levels. The data represent means and SDs of three independent experiments.
FIG 5 Nucleosomes are assembled at the MP2 and M5 sequence in vivo. (A) MP2-MP2 cells: Lanes 1 and 2, a 147 bp amplification band with MP2 specific primers using the mono-nucleosomes as the template; Lanes 3 and 4, a 144 bp amplification band with spacer specific primers using of the mono-nucleosomes as the template; Lane 5, 100-bp DNA ladder (100 bp to 1000 bp); Lane 6, 147 bp and 370 bp amplification bands with MP2 specific primers using the genomic DNA from the MP2-MP2 knock-in clones; Lane 7, a 144 bp amplification band with spacer specific primers using the genomic DNA from the MP2-MP2 knock-in clones as the template. (B) Fold stability of the MP2 and the M5 positioning sequences in the MP2-MP2 and M5-M5 DT40 knock-in clones. The histograms represent relative abundance of either the MP2 or the M5 positioning sequence compared to the respective spacer regions, as analyzed by Q-PCR. The values are normalized for the copy number and primer efficiencies. The data represent means and SDs of two independent experiments.

FIG 6 Ig light chain sequence analysis of the nucleosome positioning sequence knock-in clones. (A) Map of Ig gene with 2 MP2s (not to scale); the triangle represents the two recombined loxP sites. (B), (D), (F) and (H) Mutations in 1.1 kb from the start of transcription (=1); nos. on Y-axis: point mutations at the indicated positions in MP2-MP2 (B), M5-MP2 (D), MP2-M5 (F), and M5-M5 (H). 1 to 165, IgL gene containing the leader region; 235 – 382, first NPS (MP2 / M5); 383 – 457, spacer between two NPS; 458 – 605, second NPS; 606 – 695, loxP site generated from the Bsr marker excision; 695 – 1100, IgL gene containing V
and J regions; (C), (E), (G) and (I), summary of mutations; CDR, complementarity determining region.

**FIG 7** Mutation events in the MP2 and the M5 sequences. (A, B) Patterns of nucleotide substitutions within 2x 147 bp of MP2-MP2 and M5-M5 nucleosome positioning knock-in clones (A) and 0.8 kb of flanking DNA, that includes 5' of the first NPS, spacer between the two NPSs and 3' of the second NPS (B); the ratios of transitions (ts) to transversions (tv) are also shown. (C, D) Histograms show distribution of mutations across the 147bp region of either MP2 or M5 sequences. The 147bp positioning sequence was divided into three 49bp regions and total mutation events are shown in each region. The bottom panel shows positions and the number of AID hotspots (WRC and GYW) in the three regions. (E, F) Summary of mutations in the 147bp region of either the MP2 or M5 sequences. MP2-I (MP2-MP2) = the first of the two MP2 inserts; etc. The top panel shows all mutation events and the bottom panel shows mutations per AID hotspots in the three 49bp regions of the MP2 and M5 sequence, respectively.

**FIG 8** Models of nucleosomal DNA exposure for SHM. The nucleosome can be (A) disassembled and reassembled, which requires all of the DNA-histone contacts to be broken. Alternatively, the DNA can (B) partially unwrap from or (C) reposition with respect to the histone octamer. Both of these models maintain DNA-histone contacts.
\[ \text{FIG 2A} \]

\[ \text{CTCGAAGGCTTGGTGCATTCTGGATCCGTTTTAAAACAATGTCTCCAGTCTCCCTAGCACGTGTCAGATATATAATCCTGT} \]

\[ \text{dyad} \]

\[ \text{MP2} \]

\[ \text{M5} \]

\[ \uparrow : \text{AA/TT/TA di-nucleotides} \]

\[ \uparrow : \text{GC di-nucleotides} \]
FIG 2B

<table>
<thead>
<tr>
<th></th>
<th>MP2</th>
<th>M5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AID hotspots</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>AA/TT/TAs</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>GCs</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>GC %</td>
<td>55.8</td>
<td>54.4</td>
</tr>
</tbody>
</table>
FIG 3

A

MP2  M5  5S

Nuc

DNA

B

K_{eq} (Relative to 5S)

MP2  M5

C

ΔΔ_G° (kcal/mo)

MP2  M5
FIG 3

D

E

DNA M5-nuc MP2-nuc

Fraction centrally positioned

MP2 M5
FIG 3

DNA arm (50bp)
MP2 / MS (147 bp)
DNA arm (50bp)

Cy5 labeled top strand
Cy3 labeled bottom strand

G  Cy5 labeled top strand
0  1  2  5  10  20  30  A  T

H  Cy3 labeled bottom strand
0  1  2  5  10  20  30  A  T

(247 bp total)
FIG 4

A

MP2
MP2

M5
MP2

MP2
M5

M5
M5

B

Relative m-R

knock-in clones

MP2-MP2
M5-MP2
MP2-M5
M5-M5
FIG 5

A  Template for PCR
Mono-nucleosome  gDNA
1  2  3  4  5  6  7

g
10
12
MP2
M5
4
6
8
Fold stability
0 2

B

Fold stability

MP2
M5
0 2 4 6 8 10 12
### FIG 7

#### A

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>G</th>
<th>T</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both MP2</td>
<td>A</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>2</td>
<td>7</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>G</th>
<th>T</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both M5</td>
<td>A</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>4</td>
<td>12</td>
<td>12</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

**Notes:**
- Cell line (ts:tv)
- MP2-MP2 (8:16)
- M5-M5 (23:24)

#### B

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>G</th>
<th>T</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flanks</td>
<td>A</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>3</td>
<td>49</td>
<td>14</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>39</td>
<td>31</td>
<td>25</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>G</th>
<th>T</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flanks</td>
<td>A</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>8</td>
<td>26</td>
<td>23</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>39</td>
<td>32</td>
<td>15</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

**Notes:**
- Cell line (ts:tv)
- MP2-MP2 (55:112)
- M5-M5 (65:84)
FIG 7

C  MP2

D  M5

<table>
<thead>
<tr>
<th></th>
<th>MP2-MP2 clone (MP2-II)</th>
<th>MP2-M5 clone (M5-I)</th>
<th>MP2-M5 clone (M5-II)</th>
<th>MP2 AID hotspots</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>10 1 5 17 7 6</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>2</td>
<td>6</td>
<td>11 7 3 10 7 4</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>17 7 6</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1</td>
<td>5</td>
<td>10 7 4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>M5-M5 clone (M5-I)</th>
<th>M5-MP2 clone (MP2-I)</th>
<th>M5-MP2 clone (MP2-II)</th>
<th>M5 AID hotspots</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12</td>
<td>9</td>
<td>6</td>
<td>01</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>24</td>
<td>2</td>
<td>15 0 9 9</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>10</td>
<td>4</td>
<td>01</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>7</td>
<td>6</td>
<td>15 0 9 9</td>
</tr>
</tbody>
</table>
### FIG 7

#### MP2 mutation events

<table>
<thead>
<tr>
<th>Event</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP2-I (MP2-MP2)</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>MP2-II (MP2-MP2)</td>
<td>10</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>MP2-I (MP2-M5)</td>
<td>9</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>MP2-II (M5-MP2)</td>
<td>10</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td><strong>AID hotspots</strong></td>
<td>11</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

#### M5 mutation events

<table>
<thead>
<tr>
<th>Event</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>M5-I (M5-MP2)</td>
<td>12</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>M5-II (MP2-M5)</td>
<td>20</td>
<td>24</td>
<td>2</td>
</tr>
<tr>
<td>M5-I (M5-M5)</td>
<td>8</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>M5-II (M5-M5)</td>
<td>17</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td><strong>AID hotspots</strong></td>
<td>10</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>

#### MP2 mutation events / AID hotspot

<table>
<thead>
<tr>
<th>Event</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP2-I (MP2-MP2)</td>
<td>0.5</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>MP2-II (MP2-MP2)</td>
<td>0.9</td>
<td>0.3</td>
<td>2.0</td>
</tr>
<tr>
<td>MP2-I (MP2-M5)</td>
<td>0.8</td>
<td>0.1</td>
<td>0.7</td>
</tr>
<tr>
<td>MP2-II (M5-MP2)</td>
<td>0.9</td>
<td>0.1</td>
<td>1.7</td>
</tr>
</tbody>
</table>

#### M5 mutation events / AID hotspot

<table>
<thead>
<tr>
<th>Event</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>M5-I (M5-MP2)</td>
<td>1.2</td>
<td>1.3</td>
<td>1.5</td>
</tr>
<tr>
<td>M5-II (MP2-M5)</td>
<td>2.0</td>
<td>3.4</td>
<td>0.5</td>
</tr>
<tr>
<td>M5-I (M5-M5)</td>
<td>0.8</td>
<td>1.4</td>
<td>1.0</td>
</tr>
<tr>
<td>M5-II (M5-M5)</td>
<td>1.7</td>
<td>1.0</td>
<td>1.5</td>
</tr>
</tbody>
</table>
FIG 8

A  $k_{\text{disassemble}}$  
    $\leftrightarrow$  
    $k_{\text{assemble}}$

B  $k_{\text{unwrap}}$  
    $\leftrightarrow$  
    $k_{\text{rewrap}}$

C  $k_{\text{slide right}}$  
    $\leftrightarrow$  
    $k_{\text{slide left}}$