Published in final edited form as:

DNA Repair (Amst). 2007 June 1; 6(6): 751–759. doi:10.1016/j.dnarep.2006.12.016.

The Structure-Specific Nicking of Small Heteroduplexes by the RAG Complex: Implications for Lymphoid Chromosomal Translocations

Sathees C. Raghavan¹, Jiafeng Gu², Patrick C. Swanson³, and Michael R. Lieber^{2,4}

¹ Department of Biochemistry, Indian Institute of Science, Bangalore 560 012, India

² USC Norris Comprehensive Cancer Ctr., Rm. 5428, Departments of Pathology, of Biochemistry & Molecular Biology, of Biological Sciences, and of Molecular Microbiology & Immunology, 1441 Eastlake Ave., MC 9176, Los Angeles, CA 90089-9176

³ Department of Medical Microbiology & Immunology, Creighton University Medical Center, 2500 California Plaza, Omaha, NE 68164

Abstract

During V(D)J recombination, the RAG complex binds at recombination signal sequences and creates double-strand breaks. In addition to this sequence-specific recognition of the RSS, the RAG complex has been shown to be a structure-specific nuclease, cleaving 3' overhangs and 3' flaps, and, more recently, 10 nucleotides (nt) bubble (heteroduplex) structures. Here, we assess the smallest size heteroduplex that core and full-length RAGs can cleave. We also test whether bubbles adjacent to a partial RSS are nicked any differently or any more efficiently than bubbles that are surrounded by random sequence. These points are important in considering what types and what size of non-B DNA structure that the RAG complex can nick, and this helps assess the role of the RAG complex in mediating lymphoid chromosomal translocations. We find that the smallest bubble nicked by the RAG complex is 3 nt, and proximity to a partial or full RSS sequence does not affect the nicking by RAGs. RAG nicking efficiency increases with the size of the heteroduplex and is only about two-fold less efficient than an RSS when the bubble is 6 nt. We consider these findings in the context of RAG nicking at non-B DNA structures in lymphoid chromosomal translocations.

Keywords

structure-specific nuclease; chromosomal translocation; RAG; V(D)J recombination; non-B DNA

1. INTRODUCTION

V(D)J recombination breaks are created by a protein complex of RAG1, RAG2 and HMGB1 at antigen receptor recombination signal sequences (RSS). Each RSS consists of a heptamer and a nonamer, separated by a 12 or a 23 bp nonconserved spacer sequence, and an RSS is located directly adjacent to each V, D or J coding element that undergoes V(D)J recombination. A single recombination event requires two RSSs, one with a 12 bp RSS and the other with a

⁴Corresponding author. E-mail: lieber@usc.edu, ph. 323-865-0568.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

23 bp RSS; this is the 12/23 rule [1–6]. The RAG complex makes an initial nick adjacent to each heptamer and then uses the 3'OH of the nick as a nucleophile to attack the opposite strand. This generates a hairpin at each V, D, or J coding end, leaving each signal end blunt. After this, the RAG complex remains tightly bound to the two RSSs and less tightly bound to the two coding ends in a post-cleavage complex [4,7–9]. The two coding ends are then released and joined by nonhomologous DNA end joining (NHEJ) to create the new VJ or DJ or VDJ variable domain exon [10,11]. The RSSs are also joined by NHEJ to form a signal joint [12–16].

When the RAG complex converts the nicked intermediate to a DNA hairpin, some distortion of the DNA is presumably required [3]. For this reason, the RAG complex active site is thought to be able to accommodate such DNA distortions during the course of its normal function. In addition to sequence-specific recognition of the RSS, the RAG complex has been shown to have some features of a structure-specific nuclease [17]. Specifically, the RAG complex can cleave 3' overhangs and 3' flaps [17,18]. The types of deviations from duplex DNA which can be cleaved by RAGs in a structure-specific manner was significantly broadened by studies showing that bubble or heteroduplex structures can also be cleaved [19]. All of these studies were done using Mg^{2+} as the divalent cation. With Mn^{2+} as the divalent cation, earlier studies had shown that the RAG complex can open DNA hairpins, and this was observed at much lower levels in Mg^{2+} [20,21].

In order to assess how well the RAG complex can nick small discontinuities in the duplex, we have begun to examine RAG nicking at small bubble or heteroduplex structures (Fig. 1). Such characterization is important as we consider the contribution of RAG nicking to the process of lymphoid chromosomal translocation, where small deviations from B-form DNA may account for some of the major neoplastic chromosomal translocation fragile sites [22]. We find here that full-length and core RAGs cleave symmetrical heteroduplex structures as small as 3 nt. An adjacent RSS or heptamer sequence does not alter the action of the core RAG complex at the bubble structure. The efficiency of nicking approaches that of an RSS when the bubble size is 6 nt or greater. These findings are relevant to considerations of RAG action at chromosomal translocation fragile sites.

2. MATERIALS AND METHODS

2.1 Enzymes, chemicals, and reagents

Chemical reagents were from Sigma Chemical Co. (St. Louis, MO). Restriction enzymes and DNA modifying enzymes were from New England Biolabs (Beverly, MA). Radioisotope-labeled nucleotides were purchased from NEN (Boston, MA).

2.2 Oligomers

Oligomers were from Qiagen/Operon (Richmond, CA) or from the USC Norris Cancer Center Microchemical Core Facility. The following oligomers were used in this study. SCR245, 5'-GACCTGAGGGCGAGCCCGTTGCACAGTGCTTAACAG-3'; SCR246, 5'-CTGTAAAGCACTGTGTTTTTGTCTCGCCCTCAGGTC-3'; SCR247, 5'-GACCTGAGGGCGAGCCGTCACAGTGCTTAACAG-3'; SCR248, 5'-CTGTAAAGCACTGTGTTTGTCTCGCCCTCAGGTC-3'; SCR249, 5'-CTGTAAAGCACTGTGCAACGGGCTCGCCCTCAGGTC-3'; SCR250, 5'-CTGTAAAGCACTGTGACGGCTCGCCCTCAGGTC-3'; SCR251, 5'-GACCTGAGGGCGAGCCCGTTGCGAGTAACTTAACAG-3'; SCR252, 5'-CTGTAAAGTTACTCGTTTTTGTCTCGCCCTCAGGTC-3'; SCR253, 5'-GACCTGAGGGCGAGCCGTCGAGTAACTTAACAG-3'; SCR254, 5'-CTGTAAAGTTACTCGTTTTGTCTCGCCCTCAGGTC-3'; SCR255, 5'-

GACCTGAGGGCGAGCTCGAGTAACTTAACAG-3'; SCR256, 5'-CTGTAAAGTTACTCGGGCTCGCCCTCAGGTC-3'; SCR257, 5'-GACCTGAGGGCGAGCTCACAGTGCTTAACAG-3'; SCR258, 5'-CTGTAAAGCACTGTGGGCTCGCCCTCAGGTC-3'; SCR259, 5'-CTGTAAAGTTACTCGCAACGGGCTCGCCCTCAGGTC-3'; KY28, 5'-GATCAGCTGATAGCTACCACAGTGCTACAGACTGGAACAAAAACCCTGCT-3'; KY29, 5'-TAGCAGGGTTTTTGTTCAGTCTGTAGCACTGTGGTAGCTATCAGCTGAT-3'; YM21, TTGTCGACCTGAGGGCGAGCCCGATGAATTCCAGATACTTAACACAGCCT-3'; YM231, 5'-AGGCTGTGTAAAGTATCTGGTTTTTTTTTTGCTCGCCCTCAGGTCGACAA-3'; SCR268, 5'-GATCAGCTGATAGCTACCGTCACAGTGCTACAGACTGGAACAAAAACCCT-3'; SCR269, 5'-GCAGGGTTTTTGTTCAGTCTGTAGCACTGTGTTTGTAGCTATCAGCTGATC-3'; SCR270, 5'-GATCAGCTGATAGCTACTCACAGTGCTACAGACTGGAACAAAAACCCT-3'; SCR271, 5'-AGCAGGGTTTTTGTTCAGTCTGTAGCACTGTGGGTAGCTATCAGCTGATC-3'; SCR274, 5'-GATCAGCTGATAGCTACCGCACAGTGCTACAGACTGGAACAAAAACCCT-3'; SCR275, 5'-AGCAGGGTTTTTGTTCAGTCTGTAGCACTGTGTTGTAGCTATCAGCTGATC-3'; SCR276, 5'-GATCAGCTGATAGCTACCCGTTGCACAGTGCTACAGACTGGAACAAAAACCCT-3'; SCR277, 5'-AGCAGGGTTTTTGTTCAGTCTGTAGCACTGTGTTTTTTGTAGCTATCAGCTGATC-3'.

The oligomers are purified using 8–15% denaturing polyacrylamide gel electrophoresis. The complementary oligomers were annealed in 100 mM NaCl and 1 mM EDTA by heating in a beaker of boiling water for 10 min, followed by slow cooling.

2.3 Preparation of RAG proteins

Core murine GST- RAG-1 (aa 330-1040) and GST-RAG-2 (aa 1-383) or core MBP-RAG-1 and full-length MBP-RAG-2 proteins were overexpressed in the human 293T cells and purified as previously described [19,23,24].

2.4 Preparation of oligonucleotide DNA substrates

The oligomeric DNA containing 1 nt heteroduplex was created by annealing a [γ - 32 P] ATP end-labeled 31 nt oligomer SCR255 with unlabeled 31 nt complementary oligomer SCR256 (1:1 ratio) with slow annealing in 10 mM Tris (pH8), 100 mM NaCl. Similarly, a 3 nt, 6 nt, or 10 nt bubble structure was prepared by annealing a [γ - 32 P] ATP end-labeled top strand (SCR247, SCR 245 and YM21, respectively) with the corresponding cold bottom strand (SCR248, SCR246 and YM231, respectively). Control duplex DNA of 36 bp length was created by annealing SCR245 and SCR249 (Fig. 1A). A standard 12-signal is prepared by annealing [γ - 32 P] ATP end-labeled KY28 and KY29 (cold) (Fig. 1C). In order to study the bottom strand, in each of the above cases, the bottom strand was [γ - 32 P] ATP end-labeled and annealed with the unlabeled complementary top strand.

The oligomeric DNA containing a 1 nt bubble and a V(D)J recombination heptamer (CACAGTG) was created by annealing a [γ - 32 P] ATP end-labeled 31 nt oligomer SCR257

with unlabeled a 31 nt complementary oligomer SCR258 (1:1 ratio), as described above. Similarly a 3 nt or 6 nt bubble structure with an adjacent heptamer or a control duplex containing a heptamer was prepared by mixing a [γ - 32 P] ATP end-labeled top strand (SCR253, SCR251 and SCR251, respectively) and an unlabeled bottom strand (SCR254, SCR252 and SCR259, respectively) (Fig. 1B). The oligomeric DNA containing a 1 nt, 2 nt, 3 nt or 6 nt bubble with an RSS or an RSS alone were created by annealing a [γ - 32 P] ATP end-labeled top strand oligomers (SCR270, SCR273, SCR268, SCR276 and KY28, respectively) with unlabeled complementary oligomers (SCR271, SCR274, SCR269, SCR277 and KY29, respectively) (Fig. 1C). In order to study the bottom strand, in each of the above cases, the bottom strand was also [γ - 32 P] ATP end-labeled and paired with a complementary unlabeled top strand.

2.5 In vitro RAG nicking of DNA structures

The substrate DNA containing a bubble structure alone, a bubble structure with an adjacent heptamer, or a bubble structure with an adjacent RSS (10 nM) was incubated with core GST-RAG1 and RAG2 proteins (or core MBP-RAG1 and full-length MBP-RAG2), as indicated in the figure legend) in the presence of HMGB1 for 1 h at 37°C (or 30°C, as specified) in a buffer containing 25 mM MOPS, pH 7.0, 30 mM KCl, 30 mM potassium glutamate, and 5 mM MgCl₂ with 100 nM of HMGB1. In control reactions, RAG reaction buffer plus HMGB1 was used. Reactions were terminated by adding loading dye, containing formaldehyde. The reaction products were then resolved on a 15 % denaturing polyacrylamide gel. The gel was dried and exposed to a PhosphorImager screen, and the signal was detected using a Molecular Dynamics PhosphorImager 445SI (Molecular Dynamics, Sunnyvale, CA) and analyzed with ImageQuant software (v5.0). Each experiment described in the present study was done a minimum of two independent times (independent reaction incubations) with very good agreement. The experiments shown in Figures 2 and 3 were done more than three independent times with very good agreement, and representative gels are shown.

3. RESULTS

3.1 The core RAG complex nicking at an RSS and at a heteroduplex site in DNA are comparable

Previously we have seen that the RAG complex is capable of cleaving symmetrical bubbles (heteroduplexes) and at heterologous loops. In the present study, we compare the core RAG nicking activity at an RSS with that at a 10 nt bubble structure (Fig. 2A, compare lanes 1–4 to 21–24 and Fig. 2B). In order to perform this, one of the two strands (either top or bottom strands) was [γ - 32 P] ATP end-labeled and annealed with the complementary strand to create the DNA substrate with a bubble, an RSS, or both (Fig. 1A). The end-labeled DNA was then incubated with core RAGs at 37°C for 1 hr. The products were resolved using 15% denaturing PAGE.

The results show that new bands due to core RAG nicking are visible in the presence of the RAGs on DNA substrates containing an RSS or a bubble structure (Fig. 2A, compare lanes 1–4 to 21–24). In the case of the RSS, as expected, core RAG nicking resulted in a prominent band exactly at the 5' end of the heptamer. Minor bands are present both upstream and downstream of the main nicking products. The prominent band just below the primary band may be due to core RAG nicking of the 1 nt flap that may arise from breathing of the nick.

For the 10 nt bubble DNA substrate, core RAG nicking observed was indistinguishable from that which was previously described. The RAG nicking occurs primarily at multiple locations adjacent to the heteroduplex within the double-stranded portion (Fig. 2A, lanes 21–24). Among the various RAG sensitive sites, we find that the nucleotides at the junction of the single- and double-stranded DNA are the most sensitive to core RAG nicking on a 10 nt bubble structure

(Fig. 2A). The observed core RAG nicking efficiency at the RSS and 10 nt bubble are comparable (Fig. 2B). Previously, we noted that there is no nicking on a 10 nt bubble structure when an active site point mutant of RAG is used as a control (Figure 10, lane 10 of [19]).

We also examined the core RAG sensitivity on the anti-parallel strand of the 10 nucleotide bubble substrate. In order to perform such an experiment, we [γ - 32 P] ATP end-labeled either the top or bottom strands of the symmetrical bubble (or RSS, for comparison) and annealed it with the unlabeled antiparallel strand. Results show that upon incubation with the core RAG complex, we detect RAG nicking on both the top and bottom strands in the bubble structures, though the efficiency of the nicking is weaker on the bottom strand (Fig. 2A, lane 24). (In the case of the RSS substrate, we do not expect to see RAG nicking on the bottom strand because the hairpin formation step is dependent on the presence of a partner RSS (Fig. 2A, lane 4).) These results show that core RAGs nick on both the top and bottom strands of the symmetrical bubble with different efficiencies.

3.2 RAG nicking efficiency at bubble structures is dependent on the size of the bubble

The minimum length bubble required for core RAG nicking was tested by generating a series of substrates that contain bubble structures. In addition to the 10 nt bubble described above, DNA containing 6 nt, 3 nt, and 1 nt bubbles were created by annealing oligomers (see Materials and Methods) (Fig. 1A). In each case, either the top or bottom strand was labeled. Results show that core RAGs nick DNA containing a 6 nt and 3 nt bubble, but not a 1 nt bubble (Fig. 2A, lanes 5–20 and Fig. 2B). The efficiency of nicking is lower when the bubble length is 3 (Fig. 2, lanes 13–16 and Fig. 2B), consistent with the lack of nicking seen with the 1 nt bubble or the no-bubble control (Fig. 2, lanes 5–12). In all cases (3 nt bubble, 6 nt bubble, and 10 nt bubble), RAG nicking occurs on both the top strand and bottom strands, but the efficiency of the nicking is not symmetrical (Fig. 2B). (This is likely due to minor sequence preferences at the RAG nicking step, similar to the coding end sequence effects on nicking at an RSS [25–28].) The RAG nicking predominantly occurs at the single-/double-strand DNA transitions and into the adjacent double-stranded region, probably due to breathing of the bubble into the adjacent duplex DNA (see below). These results suggest that a minimum of 3 nt single-strandedness is required for RAG nicking at non-B DNA structures. It appears that at a bubble length close to 6 nt or 10 nt, RAG nicking is as efficient as at an optimal RSS (Fig. 2B).

3.3 Presence of a heptamer adjacent to a heteroduplex structure does not affect RAG nicking at the heteroduplex

In light of the length-dependent RAG nicking of bubbles in the above experiments, we wondered whether or not the presence of a heptamer near a bubble would affect RAG nicking efficiency. In order to perform such an experiment, we designed oligomers containing a V(D) J recombination heptamer, CACAGTG, adjacent to the bubble region (Fig. 1B). Then the oligomer was end-labeled ([γ - 32 P] ATP end-labeled) either on the top or bottom strands of the bubble and annealed with the respective unlabelled anti-parallel strands (Fig. 1B). Results show that upon incubation with the core RAG complex, we detect RAG nicking on both the top and bottom strands at bubbles of 6 nt length or 3 nt length (Fig. 3A, lanes 1–8 and Fig. 3B). The pattern of nicking and its efficiency is indistinguishable to that observed earlier when the heptamer was not present. Similar to bubble DNA without heptamers, even when the heptamer is present, we do not find any nicking when the bubble length is 1 nt (Fig. 3A, lanes 9–12) or 0 nt (two independent top strands were used in this case, Fig. 3A, lanes 13–16). For the 3 and 6 bubbles, the nicking is limited to the adjacent duplex region at the single- to double-strand transitions (Fig. 3A). Hence, the overall nicking efficiency is comparable in the presence or absence of a heptamer (compare Fig. 2B and 3B).

3.4 Full-length RAG2 and core RAG1 complex nick the heteroduplex structures

Since the experiments thus far are done using GST-core RAG1 and 2, we decided to test and compare the RAG nicking efficiency using full-length MBP-RAG2 and core-MBP RAG1 on heteroduplex substrates. We used DNA substrates with the RSS or the 6 nt bubble, 3 nt bubble, 1 nt bubble, or no-bubble control (Fig. 1A). We find a similar pattern of nicking when GST-core RAGs are replaced with full-length RAG2 and core RAG1, under otherwise similar reaction conditions (Fig. 4A). The full-length RAG containing complex nicked at an RSS and at a 6 nt bubble substrate with comparable efficiency (Fig. 4A, compare lanes 12 to 16 and Fig. 4B).

We wondered if breathing directly adjacent to the bubble structure was important for RAG nicking. To assess this, we compared RAG nicking activity on RSS and bubble substrates at 30°C instead of the usual 37°C, under otherwise identical conditions. In these experiments, we chose to examine only the strand in each case where the maximum RAG nicking occurs. The strand of interest was [γ - 32 P] ATP end-labeled and annealed with the complementary strand to create the DNA with bubble or RSS (see Materials and Methods) (Fig. 1A). The end-labeled DNA was incubated with core RAGs at 30°C for 1 hr. The products were then resolved by 15% denaturing PAGE.

Results showed that there is no significant difference in the efficiency of RAG nicking at an RSS and at a 6 nt bubble when the incubation is done at 30°C versus 37°C (Fig. 5, compare lanes 3 and 4 to 7 and 8). As noted earlier, the 3 nt bubble was cleaved less efficiently than the RSS or the 6 nt bubble (Fig. 5, lanes 1 and 2). These results suggest that the observed RAG nicking does not require melting of the DNA around the bubble site.

3.5 Comparison of RAG nicking efficiency at an RSS and at a heteroduplex on the same DNA molecule

In cases where there is an RSS adjacent to a bubble, we wondered whether the bubble would affect the core RAG nicking at the RSS or vice versa. In order to test this, we generated oligomers containing an RSS without a bubble, an RSS with a 1 nt bubble, an RSS with a 2 nt bubble, an RSS with a 3 nt bubble, and an RSS with a 6 nt bubble (Fig. 1C). Then the oligomer was end-labeled ([γ - 32 P] ATP end-labeled) either on the top or on the bottom strands of the bubble and annealed with the respective unlabelled anti-parallel strand. The end labeled DNA was incubated with 100 ng of core RAGs at 37°C for 1 hr. The products were then resolved by denaturing PAGE.

The results show that in the case of an RSS with no bubble, as expected, core RAGs nick exactly at the 5' end of the heptamer, whereas on the other strand, there is no nicking at all (Fig. 6A, lanes 1–4). In the case of an RSS with a 1 nt bubble, again the nicking occurs exactly at the 5' end of the heptamer. However, the band position is 1 nt higher on the gel because the overall length of the oligomer is longer by 1 nt due to the bubble, and there is no nicking on the bottom strand (Fig. 6A, lanes 5–8). For a substrate with an RSS adjacent to a 2 nt bubble, the predominant nicking position is at the 5' end of the heptamer. However, bands are present at or near the opposite side of the bubble junction also, and there is no nicking on the bottom strand (Fig. 6A, lanes 9–12). For an RSS with a 3 nt bubble, the RAG nicking resulted in two similar intensity bands, one due to a nick at the 5' end of the heptamer and the other one exactly at the single-strand/double-strand transition (Fig. 6B). Interestingly, even in this case, there is no bottom strand nicking (Fig. 6A, lanes 13–16). When the RSS with an adjacent 6 nucleotide bubble is the substrate, we note that almost all of the nicking now occurs at the single-strand/double-strand transition of the bubble (Fig. 6A, lanes 17 and 18). In contrast to our previous results with bubble structures that lack an adjacent heptamer, we do not find any nicking on the bottom strand (Fig. 6A, lanes 19 and 20). The absence of apparent nicking at the heptamer

is, in part, for technical reasons. Because the radioactive label is at the 5' end, once the bubble is nicked, nicking at the RSS on the top strand can not be visualized. Nevertheless, these results indicate that an RSS and a bubble structure are roughly equally preferred sites for RAG nicking (see Fig. 2B also), and among the bubble structures, a single-stranded region of 6 nt is the minimal length for optimal nicking, with lower efficiencies of nicking at 3 nt bubbles.

4. DISCUSSION

4.1 The core RAG complex can nick small heteroduplexes in a manner that is entirely unaffected by adjacent heptamer or heptamer/nonamer sequences

The RAG complex has well-known *sequence*-specific endonuclease activity [1–4]. However, the RAG complex also has documented *structure*-specific endonuclease activity [17–21]. This structure-specific endonuclease activity was thought to be limited to 3' overhangs from a duplex [17,18], with a much less efficient action at hairpins [20,21]. However, more recently, heteroduplex regions were documented to be substrates of RAG endonuclease action, independent of any specific sequences [19]. We initiated the current studies in order to discern how small of a heteroduplex deviation that the RAG complex would nick in a structure-specific manner. It was surprising to us that the RAG complex could nick bubbles as small as 3 nt. We have previously shown that the RAG complex can bind to small heteroduplex structures nearly as well as it can bind to an RSS [19].

Binding to and nicking at a 3 nt (less efficient) or 6 nt (as efficient as an RSS) fixed bubble raises the possibility that the RAG complex may stabilize a transient site of breathing that might be encountered at any of a wide range of sequences throughout the genome. This may account for observations of apparent nicking at non-RSS sites by the RAG complex in the mammalian genome.

Though we did not observe nicking at a 1 nt heteroduplex, the observation that a 3 nt heteroduplex is nicked by core RAGs raises the possibility that single base pair mismatches might occasionally breath into the larger, 3 nt bubble, which could then be nicked by the RAG complex. Earlier studies had shown that RAG mediated transposition is favored when single nucleotide mismatches are present within GC rich sequences [29]. This difference could be due to the difference in the chemistry of the reactions. The RAG nicking at a heteroduplex results from a hydrolytic reaction at the phosphodiester bond, whereas RAG transposition occurs by a transesterification reaction.

The presence of an adjacent heptamer had no effect on the RAG nicking of the bubble structures. Hence, the structure-specific nicking dominated any potential nicking at the heptamer. However, an RSS (heptamer/nonamer) adjacent to a bubble was nicked readily, and this was independent of the nicking at the bubble. This could even be seen at a 2 bp heteroduplex adjacent to an RSS, where the structure-specific nicking was throughout the heteroduplex, but predominantly at the distal edge of the bubble, 1 nt into the duplex portion. Previous work with heteroduplex regions adjacent to an RSS were limited to studies in Mn^{2+} , and did not draw any inferences about heteroduplexes nicked in a structure-specific manner [27,28].

4.2 Implications of nicking at short heteroduplexes for lymphoid chromosomal translocations

Some lymphoid chromosomal translocations occur at RSS-like sequences, which are sometimes called pseudo-RSS sites or cryptic RSS sites [30–32]. However, a much larger number of lymphoid neoplasm arises from translocations within broader zones of DNA. For example, the t(14;18) translocation occurs predominantly at the bcl-2 major breakpoint region (Mbr), which is a 150 bp region that contains no RSS-like sequences [33,34]. Moreover, this

region fails to function in an RSS-like manner when paired with a partner 12-RSS or 23-RSS [22,32]. We have shown that this region is highly reactive with bisulfite in vitro and with KMnO_4 in vivo. These nucleophiles attack C and T, respectively, but only if these bases are unstacked. Hence, the bcl-2 Mbr region frequently adopts an unstacked or non-B DNA conformation [22]. We do not know with certainty the precise conformation of this non-B structure or its longevity. Some data is consistent with a possible triplex conformation [35]. However, given the large number of degrees of freedom at non-B DNA conformations, we are open to consideration of a broader range of conformations, including simple slipped configurations that would have foci of heteroduplex bubbles. The studies here are important in showing that small slippage events that generate 3 nt heteroduplexes can be targets of the RAG complex.

Results from the current study indicate that, like core RAGs, full-length RAGs also nick the non-B DNA structures in vitro. Since full-length RAGs (native RAGs) are present inside the cell, it raises the question of how often such heteroduplex structures are encountered and then nicked by RAGs in the genome? Though it is almost impossible to predict the frequency of such RAG nicking in vivo, several considerations limit the genetic instability that might otherwise occur. First, RAGs are only expressed in early B and T cells, and therefore only these cells are at risk. Second, even in these pre-B and pre-T cells, the RAG2 protein is only stable during G0/G1 of the cell cycle [36,37]; hence, there is no RAG activity during DNA replication, when junctions of single-/double-stranded DNA might frequently arise. Third, during G1, when RAGs are active, any nicks created by RAGs may be religated by DNA ligase I. Hence, only when there are nearby nicks on both strands, as in the case of the bcl-2 Mbr, would such RAG activity lead to DSBs and chromosomal abnormality [22].

Acknowledgments

This work was supported by NIH grants to MRL. The authors thank C.-L. Hsieh, K. Yu, Y. Ma, A. Tsai, D. Roy, F. Huang and H. Lu for help and suggestions.

References

1. Fugmann SD, Lee AI, Shockett PE, Villey IJ, Schatz DG. The RAG proteins and V(D)J recombination: complexes, ends, and transposition. *Ann Rev Immunol* 2000;18:495–527. [PubMed: 10837067]
2. Gellert M. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu Rev Biochem* 2002;71:101–132. [PubMed: 12045092]
3. Swanson PC. The bounty of RAGs: recombination signal complexes and reaction outcomes. *Immunol Rev* 2004;200:90–114.
4. Sadofsky MJ. Recombination-activating gene proteins: more regulation, please. *Immunol Rev* 2004;200:83–89. [PubMed: 15242398]
5. Hiom K, Gellert M. Assembly of a 12/23 paired signal complex: a critical control point in V(D)J recombination. *Mol Cell* 1998;1:1011–1019. [PubMed: 9651584]
6. Jones JM, Gellert M. Intermediates in V(D)J recombination: a stable RAG1/2 complex sequesters cleaved RSS ends. *Proc Natl Acad Sci* 2001;98:12926–12931. [PubMed: 11606753]
7. Agrawal A, Schatz DG. RAG1 and RAG2 form a stable postcleavage synaptic complex with DNA containing signal ends in V(D)J recombination. *Cell* 1997;89:43–53. [PubMed: 9094713]
8. Tsai CL, Drejer AH, Schatz DG. Evidence of critical architectural function for the RAG proteins in end processing, protection, and joining in V(D)J recombination. *Genes Dev* 2002;16:1934–1949. [PubMed: 12154124]
9. Mo X, Bailin T, Sadofsky MJ. A C-terminal region of RAG1 contacts the coding DNA during V(D)J recombination. *Mol Cell Biol* 2001;21:2038–2047. [PubMed: 11238939]
10. Ma Y, Lu H, Tiffin B, Goodman MF, Shimazaki N, Koiwai O, Hsieh CL, Schwarz K, Lieber MR. A Biochemically Defined System for Mammalian Nonhomologous DNA End Joining. *Molecular Cell* 2004;16:701–713. [PubMed: 15574326]

11. Ma Y, Lu H, Schwarz K, Lieber MR. Repair of double-strand DNA breaks by the human nonhomologous DNA end joining pathway: the iterative processing model. *Cell Cycle* 2005;4:1193–2000. [PubMed: 16082219]
12. Grawunder U, Wilm M, Wu X, Kulesza P, Wilson TE, Mann M, Lieber MR. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* 1997;388:492–495. [PubMed: 9242410]
13. Li Z, Otevrel T, Gao Y, Cheng HL, Seed B, Stamato TD, Taccioli GE, Alt FW. The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination. *Cell* 1995;83:1079–1089. [PubMed: 8548796]
14. Frank KM, Sekiguchi JM, Seidl KJ, Swat W, Rathbun GA, Cheng HL, Davidson L, Kangaloo L, Alt FW. Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. *Nature* 1998;396:173–177. [PubMed: 9823897]
15. Ferguson DO, Sekiguchi JM, Chang S, Frank KM, Gao Y, DePinho RA, Alt FW. The nonhomologous end-joining pathway of DNA repair is required for genomic stability and the suppression of translocations. *Proc Natl Acad Sci* 2000;97:6630–6633. [PubMed: 10823907]
16. Jung D, Alt FW. Unraveling V(D)J recombination; insights into gene regulation. *Cell* 2004;116:299–311. [PubMed: 14744439]
17. Santagata S, Besmer E, Villa A, Bozzi F, Allingham JS, Sobacchi C, Haniford DB, Vezzoni P, Nussenzweig MC, Pan ZQ, Cortes P. The RAG1/RAG2 complex constitutes a 3' flap endonuclease: implications for junctional diversity in V(D)J and transpositional recombination. *Mol Cell* 1999;4:935–947. [PubMed: 10635319]
18. Grawunder U, Lieber MR. A complex of RAG-1 and RAG-2 persists on the DNA after single-strand cleavages at V(D)J recombination signal sequences. *Nucl Acids Res* 1997;25:1375–1382. [PubMed: 9060432]
19. Raghavan SC, Swanson PC, Ma Y, Lieber MR. Double-strand break formation by the RAG complex at the bcl-2 Mbr and at other non-B DNA structures in vitro. *Mol Cell Biol* 2005a;25:5904–5919. [PubMed: 15988007]
20. Besmer E, Mansilla-Soto J, Cassard S, Sawchuk DJ, Brown G, Sadofsky M, Lewis SM, Nussenzweig MC, Cortes P. Hairpin coding end opening is mediated by the recombination activating genes RAG1 and RAG2. *Molecular Cell* 1998;2:817–828. [PubMed: 9885569]
21. Shockett PE, Schatz DG. DNA hairpin opening mediated by the RAG1 and RAG2 proteins. *Mol Cell Biol* 1999;19:4159–4166. [PubMed: 10330156]
22. Raghavan SC, Swanson PC, Wu X, Hsieh CL, Lieber MR. A non-B-DNA structure at the bcl-2 major break point region is cleaved by the RAG complex. *Nature* 2004a;428:88–93. [PubMed: 14999286]
23. Swanson PC. The DDE motif in RAG-1 is contributed in trans to a single active site that catalyzes the nicking and transesterification steps of V(D)J recombination. *Mol Cell Biol* 2001;21:449–458. [PubMed: 11134333]
24. Swanson PC, Volkmer D, Wang L. Full-length RAG-2, and not full-length RAG1, specifically suppress RAG-mediated transposition but not hybrid joint formation or disintegration. *J Biol Chem* 2004;279:4034–4044. [PubMed: 14612436]
25. Yu K, Lieber MR. The mechanistic basis for coding end sequence effects in the initiation of V(D)J recombination. *Mol Cell Biol* 1999;19:8094–8102. [PubMed: 10567535]
26. Yu K, Lieber MR. The Nicking Step of V(D)J Recombination is Independent of Synapsis: Implications for the Immune Repertoire. *Mol Cell Biol* 2000;20:7914–7921. [PubMed: 11027262]
27. Ramsden DA, McBlane JF, Gent DCv, Gellert M. Distinct DNA sequence and structure requirements for the two steps of V(D)J recombination signal cleavage. *EMBO J* 1996;15:3197–3206. [PubMed: 8670820]
28. Cuomo CA, Mundy CL, Oettinger MA. DNA sequence and structure requirements for cleavage of V(D)J recombination signal sequences. *Mol Cell Biol* 1996;16:5683–5690. [PubMed: 8816481]
29. Tsai CL, Chatterji M, Schatz DG. DNA mismatches and GC-rich motifs target transposition by the RAG1/RAG2 transposase. *Nucleic Acids Res* 2003;31:6180–6190. [PubMed: 14576304]
30. Lewis SM, Agard E, Suh S, Czyzyk L. Cryptic signals and the fidelity of V(D)J Joining. *Mol Cell Biol* 1997;17:3125–3136. [PubMed: 9154811]

31. Raghavan SC, Kirsch IR, Lieber MR. Analysis of the V(D)J recombination efficiency at lymphoid chromosomal translocation breakpoints. *J Biol Chem* 2001;276:29126–29133. [PubMed: 11390401]
32. Marculescu R, Le T, Simon P, Jaeger U, Nadel B. V(D)J-mediated translocations in lymphoid neoplasms: a functional assessment of genomic instability by cryptic sites. *J Exp Med* 2002;195:85–98. [PubMed: 11781368]
33. Wyatt RT, Rudders RA, Zelenetz A, Delellis RA, Krontiris TG. BCL2 oncogene translocation is mediated by a c-like consensus. *J Exp Med* 1992;175:1575–1588. [PubMed: 1588282]
34. Jager U, Bocskor S, Le T, Mitterbauer G, Bolz I, Chott A, Kneba A, Mannhalter C, Nadel B. Follicular lymphomas BCL-2/IgH junctions contain templated nucleotide insertions: novel insights into the mechanism of t(14;18) translocation. *Blood* 2000;95:3520–3529. [PubMed: 10828038]
35. Raghavan SC, Chastain P, Lee JS, Hegde BG, Houston S, Langen R, Hsieh CL, Haworth I, Lieber MR. Evidence for a triplex DNA conformation at the bcl-2 Major breakpoint region of the t(14;18) translocation. *J Biol Chem* 2005c;280:22749–22760. [PubMed: 15840562]
36. Lin WC, Desiderio S. Regulation of V(D)J recombination activator protein RAG-2 by phosphorylation. *Science* 1993;260:953–959. [PubMed: 8493533]
37. Lin WC, Desiderio S. Cell cycle regulation of RAG-2 V(D)J recombinase. *Proc Natl Acad Sci USA* 1994;91:2733–2737. [PubMed: 8146183]

A

SCR245 5' ★ _____ 3' double-stranded DNA
 SCR249 _____

SCR255 5' ★ _____ 3' DNA with 1 nt bubble
 SCR256 _____

SCR247 5' ★ _____ 3' DNA with 3 nt bubble
 SCR248 _____

SCR245 5' ★ _____ 3' DNA with 6 nt bubble
 SCR246 _____

YM21 5' ★ _____ 3' DNA with 10 nt bubble
 YM231 _____

B

SCR251 5' ★ _____ 3' double-stranded DNA
 SCR259 _____ with heptamer

SCR257 5' ★ _____ 3' DNA with 1 nt bubble
 SCR258 _____ and heptamer

SCR253 5' ★ _____ 3' DNA with 3 nt bubble
 SCR254 _____ and heptamer

SCR251 5' ★ _____ 3' DNA with 6 nt bubble
 SCR252 _____ and heptamer

Figure 1. Diagram of RAG nicking substrates used in this study

Positions of bubble and RSS (open triangle) are indicated. Names of oligomers used for making each substrate are also indicated. **A.** Oligomeric double-stranded DNA with a bubble in the middle. The length of the bubble is indicated at the right-hand side. In the case of 1 nt, 3 nt and 6 nt bubbles, the length of the DNA duplex upstream and downstream is 15 bp. In the case of the 10 nt bubble, the length of duplex DNA upstream and downstream is 20 bp. Sequences of the bubbles are, G/T (1 nt bubble), CGT/TTT (3 nt bubble), CCGTTG/TTTTTT (6 nt bubble), CCGATGAATT/TTTTTTTTTT (10 nt bubble). **B.** Oligomeric bubble DNA substrate with V (D)J recombination heptamers (CACAGTG). In these cases, a heptamer is placed just downstream of the bubble and indicated with a bold red line. The length of the bubble is indicated. In the case of 1 nt, 3 nt and 6 nt bubbles, the length of DNA duplex upstream and downstream is 15 bp. Sequence at the bubbles are, G/T (1 nt bubble), CGT/TTT (3 nt bubble), CCGTTG/TTTTTT (6 nt bubble). **C.** Oligomeric DNA substrates with bubbles and RSS (recombination signal sequence). The RSS used here has a 12 bp spacer between the heptamer and nonamer. The length of the bubble is indicated on the right side. In the case of 1 nt, 2nt, 3 nt and 6 nt bubbles, the length of DNA duplex upstream is 17 bp and downstream is 33 bp. The sequences of the bubbles are, G/T (1 nt bubble), CG/TT (2 nt bubble), CGT/TTT (3 nt bubble), CCGTTG/TTTTTT (6 nt bubble).

Figure 2. RAG nicking of heteroduplex regions of various lengths

Core RAG complexes (100 ng) were incubated with a [γ - ^{32}P] ATP end-labeled oligomeric double-stranded DNA (30 to 50 bp) containing either varying lengths of heteroduplex (bubble) regions or a 50 bp duplex containing an RSS for 1 h at 37°C. RAG nicking was done in a buffer containing 25 mM MOPS, pH 7.0, 30 mM KCl, 30 mM potassium glutamate, and 5 mM MgCl_2 with 100 nM of HMG1. Afterwards, the nicking products were resolved on a 15% denaturing polyacrylamide gel. The RSS and heteroduplex-containing duplexes were prepared by annealing respective oligomers as described in the Materials and Methods. The length of the heteroduplex is indicated. The RAG nicking positions on the heteroduplex or RSS are

indicated by the bracket. In all cases, core GST-RAG proteins were used. In order to detect the RAG nicking on the top (indicated as “top”) or bottom strand (indicated as “bot”), the respective strand of the bubble DNA was [γ - ^{32}P] ATP end-labeled and used for RAG nicking. **A.** The gel profile of core RAG nicking products. M stands for molecular weight marker, and it is a 36 nt oligomer digested with Klenow, to generate a ladder. For further details see Fig. 1A legend.

B. A histogram derived from part A showing comparison of the RAG nicking efficiencies of various heteroduplexes relative to a standard RSS. Substrate DNA alone (when no RAG is present) is designated as 100%. RAG nicking efficiency is calculated for each substrate as a percentage, with respect to the total substrate DNA of that particular category.

Figure 3. Comparison of RAG nicking efficiency on heteroduplex DNA substrates containing a V(D)J recombination heptamer (CACAGTG) adjacent to the bubble

DNA substrates containing heteroduplex regions directly adjacent to heptamers are shown in Fig. 1B. The length of the bubble in the DNA is indicated. The RAG nicking positions within the bubble are indicated by the bracket alongside the gel. In all cases, core GST-RAG proteins were used. “Top” indicates that the [γ - 32 P] ATP end-label is on the top strand. “Bot” indicates that [γ - 32 P] ATP end-label is on the bottom strand. The RAG nicking is visualized on the strand that is labeled. **A.** The gel profile of core RAG nicking products. M indicates the molecular weight marker, consisting of a 36 nt or 20 nt oligomer digested with the Klenow fragment of *E. coli* DNA polymerase I (hereafter referred to as Klenow fragment) for various times, to generate a ladder. **B.** A histogram derived from the gel in part A showing a comparison of the core RAG nicking efficiencies of various heteroduplexes with a neighboring heptamer. For quantitation, the substrate DNA alone (when no RAG protein is present) is set to 100%. RAG nicking efficiency is calculated for each substrate in percentage, with respect to the total substrate DNA.

Figure 4. RAG nicking efficiency with respect to the length of the bubble region when full-length RAGs are used

Full-length RAG (core MBP-RAG1/full-length MBP-RAG-2) complexes (100 ng) were incubated with a [γ - 32 P] ATP end-labeled oligomeric double-stranded DNA (30 to 36 bp) containing varying lengths of bubble regions or a 50 bp RSS for 1 h at 37°C. For RAG nicking, a buffer containing 25 mM MOPS, pH 7.0, 30 mM KCl, 30 mM potassium glutamate, and 5 mM MgCl₂ with 100 nM of HMG1 was used. Afterwards, the nicking products were resolved on a 15% denaturing polyacrylamide gel. In order to detect the RAG nicking on the top (indicated as “top”) or bottom strand (indicated as “bot”), the respective strand was [γ - 32 P] ATP end-labeled. For other details, see Figure 2 legend. **A.** The gel profile of full-length RAG nicking products. The RAG nicking positions of the bubble or RSS are indicated by the arrowhead. M stands for molecular weight marker, and it is a 36 nt or 20 nt oligomer digested with the Klenow fragment for various times, to generate a ladder. **B.** Histogram derived from the results in the gel in part A comparing the full-length RAG nicking efficiencies on various heteroduplexes and a standard RSS. Substrate DNA alone (when no RAG protein is present) is set to 100%. The RAG nicking efficiency is calculated for each substrate as a percentage, relative to the total substrate DNA.

Figure 5. RAG nicking on heteroduplex DNA substrates at 30°C

Core RAG complexes (100 ng) were incubated with a [γ - 32 P] ATP end-labeled duplex DNA (30 to 50 bp) containing a 3 nt or a 6 nt bubble region, or a 50 bp duplex containing an RSS for 1 h at 30°C in nicking buffer containing 25 mM MOPS, pH 7.0, 30 mM KCl, 30 mM potassium glutamate, and 5 mM MgCl₂ with 100 nM of HMG1. Afterwards, the nicking products were resolved on a 15% denaturing polyacrylamide gel. The RAG nicking positions of the bubble or RSS are indicated by the arrowhead. For further details, see Fig. 2 legend.

Figure 6. Comparison of RAG nicking at an RSS and an adjacent heteroduplex

The DNA substrates containing an RSS and an adjacent bubble are shown in Figure 1C. The core RAG complexes (100 ng) were incubated with a [γ - 32 P] ATP end-labeled oligomeric duplex DNA (50 to 56 bp) containing varying lengths of a bubble region for 1h at 37°C. For RAG nicking, a buffer containing 25 mM MOPS, pH 7.0, 30 mM KCl, 30 mM potassium glutamate, and 5 mM MgCl₂ with 100 nM of HMG1 was used. Afterwards, the nicking products were resolved on a 15% denaturing polyacrylamide gel. In all cases, core GST-RAG proteins were used. In order to detect the RAG nicking at the top (indicated as “top”) or bottom strand (indicated as “bot”), the respective strand of the bubble DNA was [γ - 32 P] ATP end-labeled and used for RAG nicking. **A.** The gel profile of core RAG nicking products. The length of the bubble is indicated. The RAG nicking positions of the bubble or RSS are indicated by the bracket. M stands for molecular weight marker, and this is a 36 nt or 20 nt oligomer digested with Klenow fragment for different times, to generate a ladder. **B.** A histogram derived from the gel in part A (only from the top strand data) showing comparison of core RAG nicking at an RSS and an adjacent heteroduplex. Substrate DNA alone (when no RAG protein is present) is set to 100%. In each case, RAG nicking efficiency at an RSS is compared with an adjacent bubble and is calculated and presented as a percentage relative to total substrate.