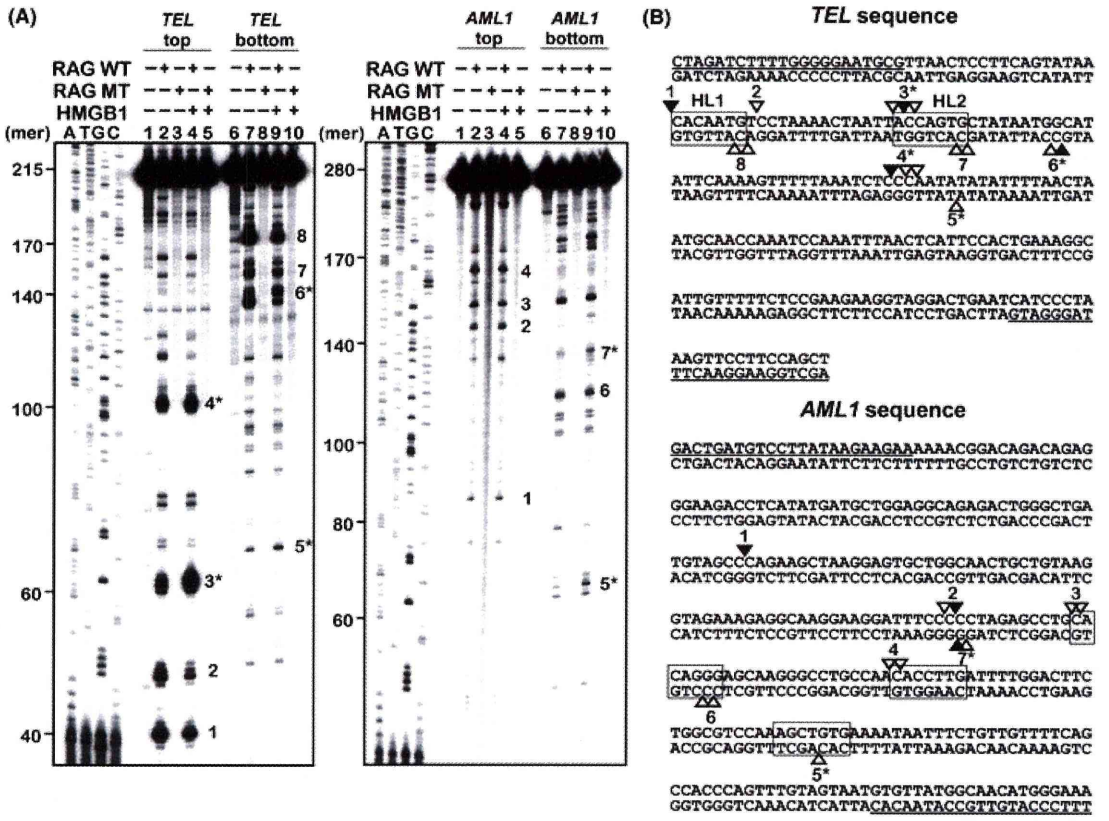

Figure 3 Recombination activating gene (RAG) proteins introduce nicks into the *TEL* and *AML1* translocation regions in the presence or absence of HMGB1. (A) The 215-bp-long *TEL* DNA fragment and the 280-bp-long *AML1* DNA fragment, which are [^{32}P]-labeled at the 5′ end of either top or bottom strand, were incubated with purified GST–RAG1ΔN and GST–RAG2ΔC (RAG WT) or GST–RAG1ΔND711A and GST–RAG2ΔC (RAG MT) in the absence or presence of HMGB1 as indicated at the top of the gel. DNA samples were separated through a 6% sequencing gel along with DNA ladder markers. The major nicked products by RAG proteins are numbered (*TEL*; from 1 to 8, *AML1*; from 1 to 7) and also summarized in panel B. The HMGB1-dependent enhanced nicks were denoted by asterisks with the numbers. The DNA fragments for the cleavage assays were amplified by PCR using sets of primers, TEL-F82 and TEL-R296, and AML1-54F3 and AML1-R54, respectively. (B) The break points in the *TEL* and *AML1* DNA substrates are summarized. The positions of major RAG-mediated nicks are indicated by numbered arrowheads. Closed arrowheads show the positions of RAG-mediated nicks generated in *in vitro* cleavage assays, which exactly correspond to the break points determined in the extrachromosomal recombination assays (Numata *et al.* 2010). Open arrowheads show break points detected only in *in vitro* cleavage assays. The primer sequences used for PCR amplification were underlined, and heptamer-like sequences (HL1 and HL2) within these DNA fragments are surrounded by open squares.

RAG1ΔN). Wild-type RAG proteins introduced some nicks not only in the vicinity of heptamer-like elements but also into other sites within the *TEL* and *AML1* DNA fragments (Fig. 3A, lanes 2, 4, 7, and 9). However, only a trace level of cleaved products was observed in the presence of D711A RAG1ΔN and RAG2 (Fig. 3A, lanes 3, 5, 8, and 10, respectively). The precise positions of major nicks are mapped in Fig. 3B. Some of nicks generated by RAG proteins *in vitro* were completely crossovered with those determined in extrachromosomal recombination assays (Fig. 3B, closed arrowheads), while some others were generated only in the *in vitro* system (Fig. 3B, open arrowheads). We noted that there are two heptamer-like (HL) nucleotide sequences, HL1 (5′-CACAATG-3′) and HL2 (5′-ACCAGTG-3′), within the *TEL* fragment (Fig. 3B, open boxes). Even though a putative cRSS contains at least the first three residues (5′-CAC-3′) of the consensus heptamer (5′-CACAGTG-3′) (Marculescu *et al.* 2006), recently it has been shown the CAC motif of the heptamer is not always needed for RAG-mediated aberrant nicking (Zhang & Swanson 2009).

Thus, we thought that HL2 site lacking the CAC motif also might act as a functional heptamer-like element. It has been reported that an aberrant RAG-mediated DSB is followed by cryptic nicks leading to hairpin formation, known as standard nick/hairpin mechanism (Zhang & Swanson 2008, 2009). To examine whether the nicks at HL1 and HL2 transit to the formation of hairpin, we carefully investigated RAG-mediated cleaved products (Fig. 3A). However, we could not detect the plausible cleaved products derived from expected hairpin structure of HL1 and HL2 nicks. We performed primer extension assays using the [^{32}P] end-labeled primers (Fig. S2 in Supporting Information). Based on the length of primer extension products, we concluded that the cleaved products are nicked ones but not hairpins. However, we found that RAG proteins introduced the nicks into both top and bottom strands within HL1 and HL2 regions (Fig. S2 in Supporting Information). This result suggests that the staggered nicks might be converted into DSBs by the nick–nick mechanism (Raghavan *et al.* 2004). Indeed, we detected many cleaved double-stranded DNA fragments near HL1 and HL2 sites in the *TEL* and *AML1* translocation regions by separating DNA fragments in a native PAGE (Fig. S3 in Supporting Information). Several RAG-mediated DSBs were enhanced in the presence of HMGB1.

RAG proteins recognize the heptamer-like sequence as a functional heptamer

We identified two remarkable RAG-mediated nicks at HL1 and HL2 sites within the *TEL* recombination region. To examine that these HL sequences are functional in RAG-mediated nicking, we prepared *TEL* DNA fragments containing nucleotide mutations within HL1 and HL2 sites (Fig. 4A). We constructed two kinds of mutated sequences, designated 'Consensus' and 'Divergent' sequences, respectively. In 'Consensus' and 'Divergent' sequences, wild-type HL1 and HL2 contain mutations so as to become authentic and more diverse sequences from the consensus sequence. The nicking level on 'Divergent' substrates of HL1 and HL2 was strongly diminished (Fig. 4B, lanes 2 compared with 4 and 6). On the other hand, replacement of the cryptic heptamers with consensus heptamers restored the level of its RAG-mediated nicking a little more than those of wild-type substrates (Fig. 4B, lanes 2 compared with 8 and 10). In addition, the major nicked position in the 'Consensus' substrate of HL2 was changed by one nucleotide to become adjacent to the 5′ direction, namely a nick

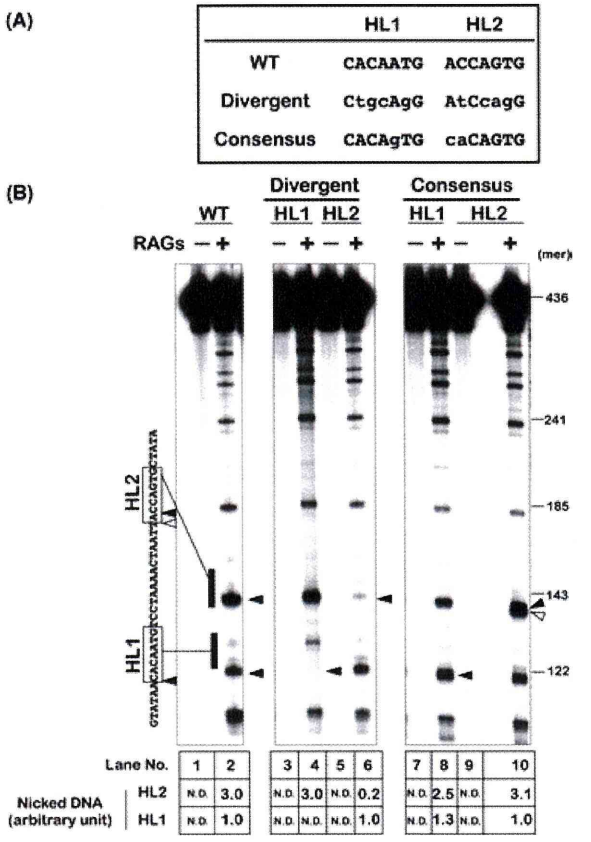

Figure 4 Recombination activating gene (RAG) recognizes and cleaves at heptamer-like DNA regions within the translocation region of *TEL*. (A) The sequence summary of mutants. Mutations introduced to HL sequences (HL1 and HL2) are indicated by lower-case letters. The positions of HL1 and HL2 within *TEL* are shown on the left of panel B. The two kind of mutations, 'Divergent' and 'Consensus', contain completely divergent form and exactly the same with the consensus heptamer (5′-CACAGTG-3′), respectively. These nucleotide mutations in 'HL1-Divergent', 'HL1-Consensus', 'HL2-Divergent', and 'HL2-Consensus' were introduced by PCR amplification using sets of primers, TEL-F and TEL436-R2Mutant, TEL-F and TEL436-R2WT, TEL-F and TEL436-R3Mutant, and TEL-F and TEL436-R3WT, respectively. (B) The 444-bp-long *TEL* DNA fragment, which is [^{32}P]-labeled at the 5′ end of top strand, was used for the cleavage assays. The nicked products in the cleavage assay using mutant *TEL* DNA substrates ('HL1-Divergent', 'HL2-Divergent', 'HL1-Consensus', and 'HL2-Consensus') were separated on a 6% sequence gel. The nicked positions of HL1 and HL2 are indicated by closed arrowheads. The open arrowhead shows the precise nick site of the consensus heptamer on the 'HL2-Consensus'. 'RAGs' contains both GST-core RAG1, GST-core RAG2, and His-HMGB1 proteins. Relative amounts of the nicked DNA at HL1 and HL2 were shown below the gel, where the amount of HL1 in lane 2 is set to be 1. 'N.D.' means 'Not Detected'.

exactly at the 5′ end of the consensus heptamer element (Fig. 4B, lane 10). Taken together, it is quite likely that RAG proteins recognize HL1 and HL2 sites within the *TEL* translocation region as functional heptamers and introduce nicks into these elements in a sequence-specific manner.

RAG-mediated nicking was enhanced by both HMGB1 and an additional NL sequence

We found that the level of RAG-mediated nicking in several sites is enhanced by HMGB1 (asterisks in Fig. 3A; 3★ (3.5-fold), 4★ (1.4-fold), 5★ (2.3-fold), and 6★ (1.2-fold) on the *TEL* and 5★ (3.1-fold) and 7★ (1.8-fold) on the *AML1*). As we found that HMGB1 enhances RAG-dependent recombination between *TEL* and *AML1 in vivo* (Fig. 1) and the enhancement of the aberrant RAG-mediated nicking by HMGB1 could be important with regard to the genome instability leading to lymphoid malignancies (Lee *et al.* 2004; Zhang & Swanson 2009), we decided to focus our study on the HL2 site, where the RAG-mediated nicking was most enhanced (3.5-fold) by HMGB1, to reveal the HMGB1-enhanced nicking mechanism. RAG-mediated nicking occurs by recognizing a heptamer alone, although nicking is most efficient when both heptamer and nonamer are present and separated by either 12 or 23 bp. Based on the fact that HL2 site is recognized by RAG proteins as heptamer-like element, we next tried to determine which NL element plays a functional role as nonamer for HL2 and then show more clearly whether HL and NL elements cooperate or function independently in this DNA fragment. We found three putative NL sequences, 12NL, 23NL, and 56NL, which are 12, 23, and 56 nucleotides away from HL2, respectively (Fig. 5A). We substituted three adenosines with cytocines at the 4–6th or the 5–7th nucleotide positions in NL (Fig. 5B), because it is indicated that the fifth and the sixth adenosines within the nonamer sequence are critical for efficient binding of RAG proteins (Hesse *et al.* 1989). The amount of nicks at HL2 site shown in Fig. 5C was quantitatively determined (Fig. 5D). Mutations in 12NL diminished the nicking activity of RAG proteins in the both absence and presence of HMGB1 (Fig. 5C, lanes 5 and 6). Substitution of triple adenosines with triple cytocines in 23NL, 56NL, and both 23NL and 56NL gave no effect on the level of RAG-mediated nicking at HL2 in the absence of HMGB1. In contrast, it is noted that the nicking enhancement by HMGB1 at HL2 site of 23NLm and 56NLm substrates was significantly impaired compared with that of wild type, respectively (Fig. 5C, lanes 3 compared with 9 and 12). Mutations in both 23NL and 56NL resulted in completely loss of the HMGB1-dependent enhancement of RAG-mediated nicking (Fig. 5C, lanes 14 and 15). This result indicates that 23NL and 56NL independently function as a NL element, and these elements additively enhance the RAG-mediated and HMGB1-enhaned nicking at HL2 site. Then, we performed EMSAs with the wild type and 56NL-mutated *TEL* DNA fragments (Fig. S4A in Supporting Information). In the both absence and presence of HMGB1, the amount of RAG-bound DNA was reduced in the 56NL-mutated substrate, compared with that of wild type (Fig. S4A in Supporting Information). Collectively, these results suggest that a set of HL2 and 12NL is recognized as a cryptic 12-RSS by RAG proteins and indispensable for the aberrant RAG-mediated nicking at HL2 site. Moreover, 23NL and 56NL orphan NL sequences are required for the enhancement of RAG-mediated nicking at HL2 by HMGB1.

DNA binding activity of HMGB1 is required for the HMGB1-enhanced RAG-mediated nicking

HMGB protein consists of two tandem HMG boxes, called box A and box B, and an acidic C-terminal tail separated by a basic linker residues (Thomas & Travers 2001). Box A binds to distorted DNA selectively, whereas box B does not have such a DNA binding specificity but can itself induce bending in linear DNA. Based on *in vitro* analyses using an extensive panel of mutant HMGB1 proteins, it is reported that HMGB1 promotes RAG-mediated RSS binding and cleavage largely through the activity of box B, and a functional box A tethered in the correct orientation is required for the maximal stimulation (Bergeron *et al.* 2005). Thus, we tried to examine which HMGB1 region is required to support the HMGB1-dependent enhancement of aberrant RAG nicking at a cryptic site. To address this, we prepared four mutant HMGB1 proteins (Bergeron *et al.* 2005; Dai *et al.* 2005) (Fig. 6A). As the essential hydrophobic residues in HMG boxes that anchor the protein to the DNA minor grooves are well characterized (Thomas & Travers 2001; Dai *et al.* 2005), we disrupted the DNA binding activity of HMGB1 by substituting three intercalating residues with alanines (Fig. 6A, 'HMGB1ΔDNA'). We obtained highly purified mutant HMGB1 proteins (Fig. 6B).

Using purified proteins, we performed an EMSA with the [^{32}P] end-labeled wild-type and

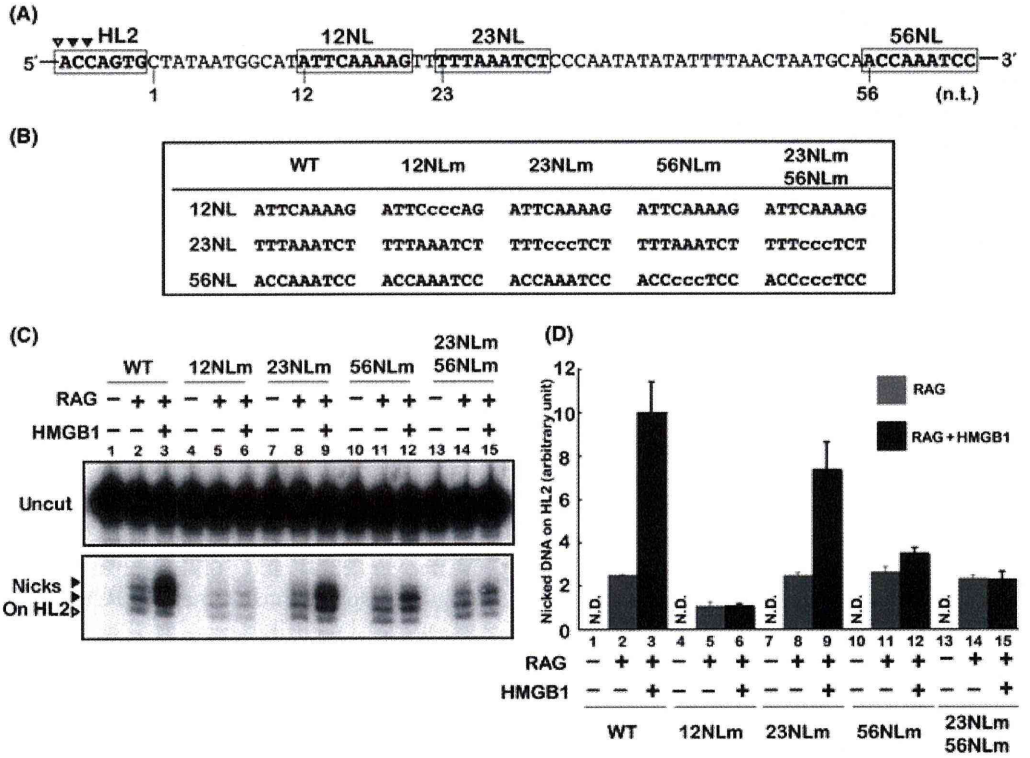

Figure 5 Effect of orphan nonamer-like sequences on the recombination activating gene (RAG)-mediated nicking at HL2 site. (A) A part of the top-strand sequence of the 134-bp-long *TEL* DNA used for the cleavage assay. HL2 site and NL sequences are indicated by both bold font and open box. These NL sites are 12, 23, and 56 bp away from HL2 site as depicted. The RAG-mediated nicking sites are indicated by open and closed arrowheads. The open arrowhead indicates the 5′ side of RAG-mediated nicks at HL2. (B) Replaced nucleotides in 12NL, 23NL, and 56NL. 'WT', '12NLm', '23NLm', '56NLm', and '23NLm56NLm' are substrates for cleavage assays. Nucleotide mutations introduced to the three possible NL sequences are indicated by lower-case letters. '23NLm56NLm' substrate includes both mutated 23NL and 56NL sequences. These nucleotide mutations in '12NLm', '23NLm', and '56NLm' were introduced by PCR amplification using sets of primers, TEL-F82 and TEL436-R6MT, TEL-F82 and TEL436-R7MT, and TEL-F82 and TEL-R215MT, respectively. Each DNA fragment for the cleavage assay was amplified by PCR using sets of primers, TEL-F82 and TEL-R215, or TEL-F82 and TEL-R215MT, respectively. (C) Cleavage assays were performed with GST-core RAG proteins and substrates shown in panel B in the presence or absence of HMGB1. Nicked products at HL2 site were separated through a 6% sequence gel. The upper panel shows uncut DNA, whereas the bottom panel shows nicked DNAs at HL2 site. The positions of open and closed arrowheads are the same with those in panel A. (D) The relative amounts of the RAG-mediated nicked products at HL2 site. The abundance of nicks shown in panel C was quantitatively analyzed by IMAGE GAUGE analysis software (Fujifilm, Tokyo, Japan). The results are represented as mean values with SD of independent three to four experiments. 'N.D.' means not detected.

56NL-mutated *TEL* DNA fragments (Fig. 6C). We found that wild-type HMGB1 and C-terminal acidic region-deleted mutant (HMGB1ΔC) supershift a RAG-cRSS complex, and the amount of supershifted band by HMGB1ΔC is slightly higher than wild-type HMGB1 (Fig. 6C, lanes 3, 5, 9, and 11). HMGB1ΔDNA completely diminished the ternary complex formation (Fig. 6C, lanes 4 and 10). Each HMG box apparently supershifted a RAG–cRSS complex, but its effect was relatively low (Fig. 6C, lanes 6, 7, 12, and 13). Taken together, both the DNA binding activity of HMGB1 and two HMG boxes are required for maximal formation of RAG–HMGB1–RSS complex. Moreover, it is worthwhile to note that the pattern of supershifted RAG-56NLm DNA bands in the presence of each HMGB1 mutant protein was quite similar to that of wild-type DNA. Thus, it is presumed that the stabilization of RAG–DNA complex by HMGB1 is not a major cause for HMGB1- and 56NL-dependent enhancement of

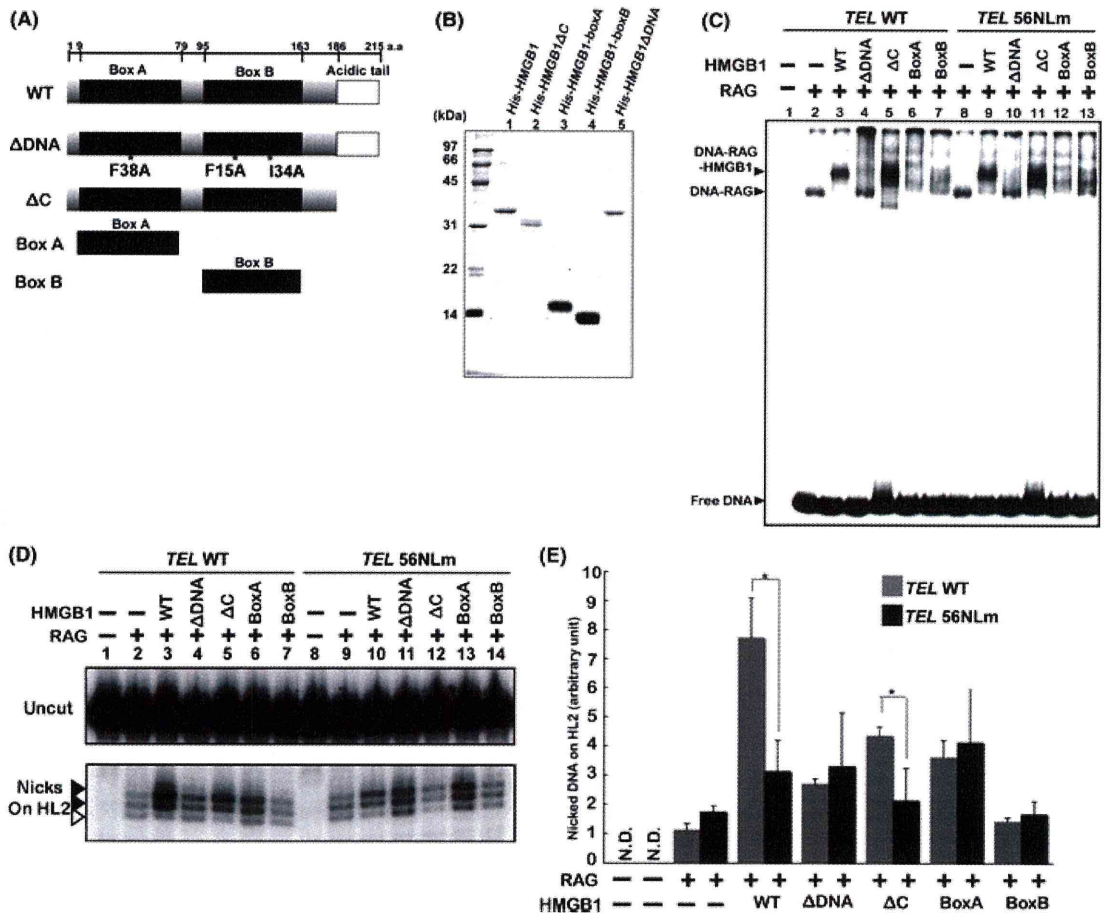

Figure 6 HMGB1 domains required for the formation of recombination activating gene (RAG)–HMGB1–cRSS complex and enhancement of RAG-mediated nicking at HL2 site. (A) Diagrams of full-length, a DNA binding mutant, truncated forms of HMGB1 used in this study. The high mobility group boxes and the acidic tail are schematically indicated by closed and open squares, respectively. The amino acid residues encompassing box A (Box A, residues 9–79), box B (Box B, residues 95–163), and a mutant lacking acidic tail (HMGB1ΔC, residues 1–186) are indicated at the top of panel. The positions of substituted amino acid residues in HMGB1ΔDNA are indicated by asterisks (F38A, F15A, and I34A). (B) Purified HMGB1 proteins were separated through a 15% SDS-PAGE and stained with CBB. (C) GST-core RAG proteins were incubated at 30 °C for 30 min with [^{32}P]-end labeled 134-bp-long *TEL* DNA fragment used in Fig. 5A in the absence or presence of various forms of HMGB1 proteins. After incubation, protein–DNA complexes were separated by electrophoresis mobility shift assays and visualized by autoradiography. (D) Cleavage assays were performed with GST-core RAG proteins in the presence or absence of HMGB1 using wild-type and 56NLm *TEL* DNA fragments shown in Fig. 5B. Nicked products at HL2 site were separated through a 6% sequence gel. The upper panel shows uncut DNA, whereas the bottom panel shows nicked DNAs at HL2 site. The positions of open and closed arrowheads are the same as those in panel A. (E) The relative amounts of RAG-mediated nicked products at HL2 site. The abundance of nicks shown in panel D was quantitatively analyzed by Image Gauge analysis software. The results are represented as mean values with SD of independent three to four experiments. 'N.D.' means not detected. Statistical *P*-values were calculated by *t*-tests and indicated with ⋆ for $P < 0.05$. cRSS, cryptic recombination signal sequence.

RAG-mediated nicking on HL2 and rather that a step of RAG-mediated nicking enhancement on HL2 by HMGB1 could be the nicking reaction.

Next to examine which domain or function is important for RAG-mediated nicking at HL2 site that is enhanced by HMGB1 and dependent on 56NL, we performed cleavage assays using HMGB1 mutant proteins. In case of wild-type DNA, nicking at HL2 was stimulated by wild-type HMGB1, but HMGB1ΔDNA and HMGB1ΔC could not fully stimulate the RAG-

mediated nicking (Fig. 6D,E, lanes 3–5). However, when 56NLm DNA was used, the nicking enhancement was significantly decreased using wild-type HMGB1 and HMGB1ΔC (Fig. 6D,E, lanes 10 and 12). Unexpectedly, box A retained slight enhancing effect on RAG-mediated nicking, whereas box B did not stimulate nicking at all (Fig. 6D,E, lanes 6, 7, 13, and 14). Nevertheless, these effects seemed independent of 56NL, and the level of nicking enhancement by box A is less than that by wild-type HMGB1. The results of DNA binding and cleavage assays suggest that the DNA binding activity and both of HMG boxes are needed for the full HMGB1-enhanced RAG-mediated nicking at the HL2 site together with the orphan 56 NL element.

Stimulation of DNA bending is responsible for the HMGB1-enhanced RAG-mediated nicking at HL2 site

As HMGB1 is known as a DNA bending protein, we examined the DNA bending activity of HMGB1 on the *TEL* DNA fragment containing HL2, 12NL, 23NL, and 56NL elements. First to identify a DNA bending center within the *TEL* DNA fragment, we performed a circular permutation assay (See Data S1 and Fig. S5 in Supporting Information). We found that the AT-rich element (5′-AATATA-3′) could be a plausible RAG-dependent DNA bending center with or without HMGB1 (Fig. S5 in Supporting Information). To evaluate the functional role of the bending center, we carried out ligation-mediated circularization (LMDC) assays with the *TEL* DNA fragments (Fig. 7A,B). To disrupt the bending capacity of the DNA fragment, we introduced GC-rich mutations into the AT-rich bending center (Fig. 7A, 'BCm'). In the presence of either RAG proteins or HMGB1, we could not detect any circularized DNA (Fig. 7B, lanes 2, 3, 8, 9, 14, and 15). On the other hand, we found that the increasing amounts of HMGB1 in the presence of RAG proteins dramatically enhance the formation of circularized DNA in a dose-dependent manner when wild-type and 56NLm substrates are used (Fig. 7B, lanes 4–6 and 10–12). However, the HMGB1-dependent stimulation of DNA circularization on BCm substrate was threefold less than that on the wild-type substrate (Fig. 7B, lanes 6 compared with 18). We found that there is no difference in the amount of DNA bound by RAG proteins between the wild-type and BCm substrates by EMSA, confirming that RAG proteins do not bind to the bending center in a sequence-specific manner (data not shown). These results suggest that the mutation in the bending center moderately impairs the DNA bending enhancement by HMGB1.

Next, to test the assumption that the HMGB1-mediated enhancement of DNA bending plays a role in the HMGB1-enhanced RAG-mediated nicking, we performed the cleave assay with the substrates shown in Fig. 7A. The levels of RAG-mediated nicking at HL2 site on three substrates in the absence of HMGB1 were similar to each other (Fig. 7C,D, lanes 2, 5, and 8). In contrast, the HMGB1-dependent enhancement of the RAG-mediated nicking on 56NLm and BCm DNA fragments was repressed (Fig. 7C,D, lanes 3, 6, and 9). These results suggest that the moderate enhancement of the DNA distortion by HMGB1 plays an important role in the HMGB1- and orphan nonamer-dependent enhancement of RAG-mediated nicking at HL2 site. We performed the cleavage assay using a purified HMGA1 protein, which is also known as a DNA bending (Falvo *et al.* 1995). However, we could not detect HMGA1-dependent enhancement of the RAG-mediated nicking (data not shown), suggesting that enhanced DNA bending is not enough for the RAG-mediated nicking enhancement, and thus a specific interaction between RAG proteins and HMGB1 could be required for the enhancement process.

HMGB1 and 56NL are critical for RAG-mediated nicking at cryptic 12-RSS but not canonical 12-RSS

It is well characterized that HMGB1 enhances the nicking activity of RAG proteins on authentic RSSs *in vitro*, in particular 23-RSS (van Gent *et al.* 1997; Yoshida *et al.* 2000). Then, we wondered whether the 56 NL-dependent and HMGB1-enhanced RAG-mediated nicking is specific for the cryptic 12-RSS or also an authentic 12-RSS. To address this, we prepared substrates, in which HL2 (5′-ACCAGTG-3′) and 12NL (5′-ATTCAAAAG-3′) were replaced with consensus heptamer (5′-CACAGTG-3′) and nonamer (5′-ACAAAAACC-3′), respectively (Fig. 8A,B). The level of nicks at HL2 in the wild-type substrate was increased as a function of the increasing amounts of HMGB1 (Fig. 8C, lanes 2–5), while the enhancing effect was much less in the 'HL2-12-RSS' substrate (Fig. 8C, lanes 12–15). Moreover, to examine the 56NL-dependent effect on the authentic 12-RSS, we prepared HL2-12-RSS substrate containing mutated 56NL, designed 'HL2-12-RSS-56NLm'. The HMGB1-dependent effect on the RAG-mediated nicking at HL2 site in the context of HL2-12-RSS-

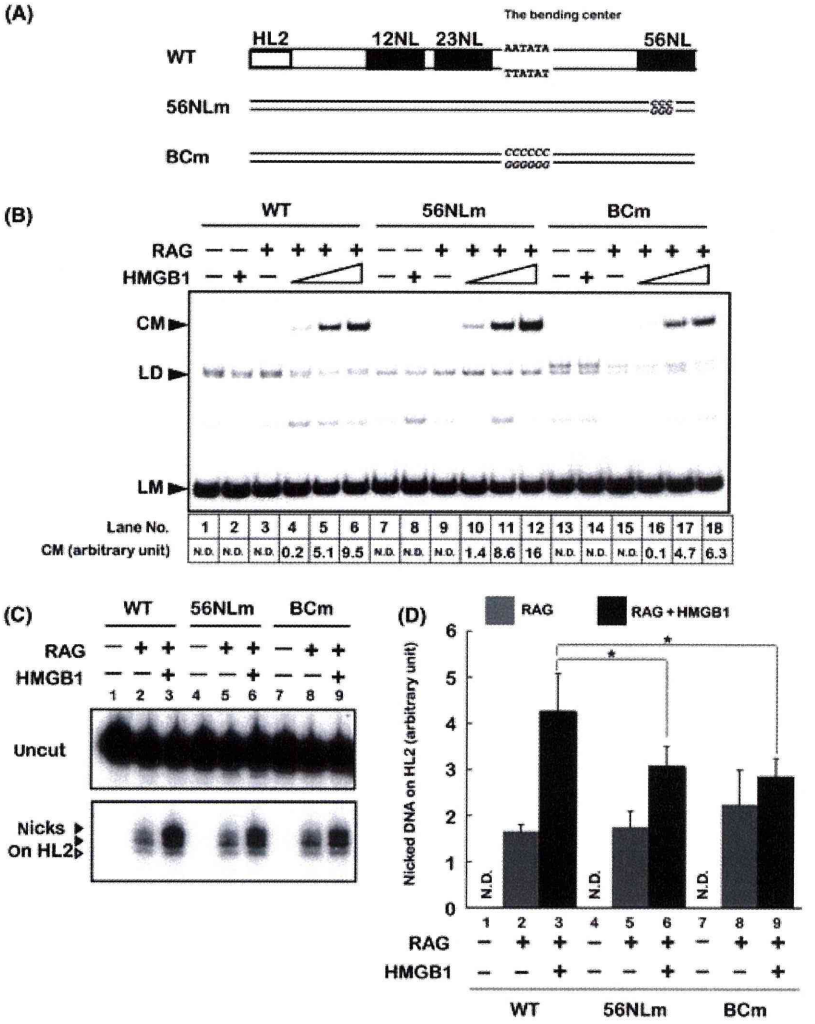

Figure 7 Involvement of DNA bending in HMGB1-dependent enhancement of recombination activating gene (RAG)-mediated nicking at HL2 site. (A) The diagram of a part of 134-bp-long *TEL* DNA fragment used for LMDC and cleavage assays. The positions of HL2, 12NL, 23NL, and 56NL are indicated by open or closed squares only in the 'WT' DNA fragment. The putative bending center (5′-AATATA-3′) determined by a circular permutation assay (Fig. S5 in Supporting Information) is located between 23NL and 56NL elements. 'BCm' substrate contains a mutated bending center (5′-CCCCCC-3′). (B) LMDC assays were performed with RAG proteins in the presence or absence of HMGB1 using the [^{32}P] end-labeled linear 134-bp-long *TEL* DNA fragments as shown in panel A. The ligation products were separated through a 5% native PAGE. Digestion with exonuclease III (exoIII) was carried out to distinguish linear DNA from circular DNA (data not shown). CM, circular monomer; LM and LD, linear monomer and dimer, respectively. (C) Cleavage assays were performed with GST-core RAG proteins and substrates shown in panel A in the presence or absence of HMGB1. Nicked product at HL2 site was separated through a 6% sequence gel. The upper panel shows uncut DNA, whereas the bottom panel shows nicked DNAs at HL2 site. Open and closed arrowheads are the same as those used in Fig. 5A. (D) The relative amounts of the RAG-mediated nicked products at HL2 site. The abundance of nicks shown in panel C was quantitatively analyzed by Image Gauge analysis software. The results are represented as mean values with SD of independent three to four experiments. 'N.D.' means not detected.

56NLm was similar to that in HL2-12-RSS (Fig. 8C, lanes 11–15 compared with lanes 16–20). The amount of nicks on HL2 site shown in Fig. 8C was quantitatively determined (Fig. 8D). Thus, it is concluded that HMGB1 and the additional orphan 56 NL element synergistically enhance the RAG-mediated nicking at the cryptic 12-RSS but little at the authentic 12-RSS.

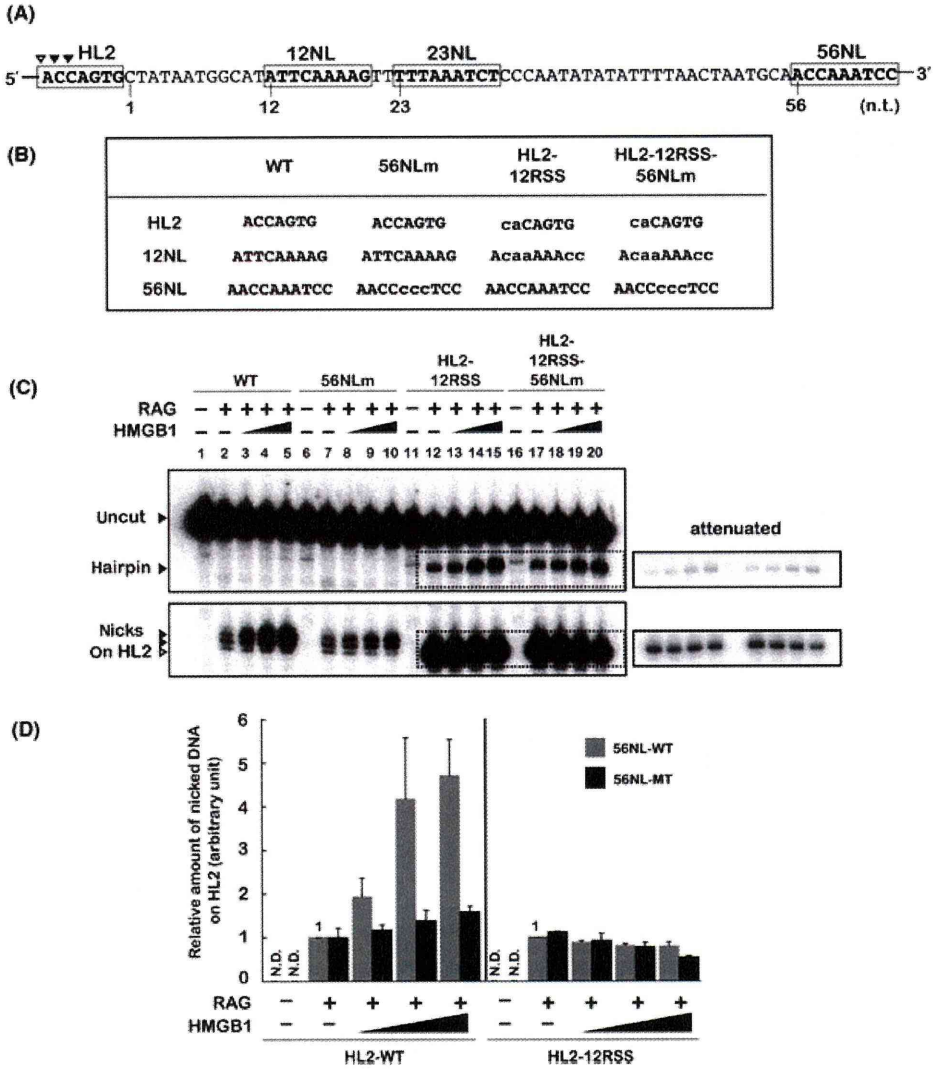

Figure 8 HMGB1- and 56NL-dependent enhancement of recombination activating gene (RAG)-mediated nicking is evident at cryptic 12-recombination signal sequence (RSS) but not at canonical 12-RSS in the *TEL* locus. (A) A part of the top-strand sequence of the 134-bp-long *TEL* DNA used for the cleavage assay is depicted as similar to Fig. 5A. (B) Replaced nucleotides in 12NL, 23NL, and 56NL. 'WT', '56NLm', 'HL2-12RSS', and 'HL2-12RSS-56NLm' are substrates for cleavage assays. Nucleotide mutations introduced to the three possible NL sequences are indicated by lower-case letters. 'HL2-12RSS' contains a consensus heptamer and a consensus nonamer in place of HL2 and NL12 sites, respectively. 'HL2-12RSS-56NLm' is the same as 'HL2-12RSS' except for possessing mutated 56NL. 'WT' and '56NLm' are the same as those shown in Fig. 5B. Nucleotide mutations in 'HL2-12RSS' and 'HL2-12RSS-56NLm' were introduced by PCR amplification using sets of primers, HL2-12RSS and TEL-R215, and HL2-12RSS and TEL-R215MT, respectively. (C) Cleavage assays were performed with GST-core RAG proteins and substrates shown in panel B in the presence or absence of HMGB1. Nicked products at HL2 site were separated through a 6% sequence gel. The upper panel shows uncut DNA, whereas the bottom panel shows nicked DNAs at HL2 site. Open and closed arrowheads are the same as those used in panel A. The positions of 'uncut', 'hairpin', and 'nicks at HL2' are indicated by arrowheads, respectively. The short-exposed figures (dotted squares) are also shown. (D) The relative amounts of the RAG-mediated nicked products at HL2 site. The abundance of nicks shown in panel C was quantitatively analyzed by Image Gauge analysis software. The amount of the RAG-mediated nicks at HL2 site is set to be 1 (lanes 2 and 12 in panel C), and the relative amount of nicked products in each lane is summarized in panel D. The results are represented as mean values with SD of independent three to four experiments. 'N.D.' means not detected.

Discussion

Illegitimate V(D)J recombination is one of the contributing factors to recurrent chromosomal translocations or deletions, major hallmarks in certain classes of leukemia and lymphoma. Recently, we have shown that RAG proteins induce the recombination in the *TEL–AML1* t(12;21)(p13;q22) chromosomal translocation regions DNA *in vivo* (Numata *et al.* 2010). Because there are several heptamer or NL sequences in the vicinity of RAG-dependent recombination sites determined by the plasmid-based recombination assays (Numata *et al.* 2010), we suspected that the aberrant RAG-dependent recombination could be because of the mistargeting of RAG proteins to cryptic elements in the *TEL* and *AML1* translocation regions. Here, we have shown that RAG proteins directly bind to and introduce nicks into the *TEL* and *AML1* translocation regions *in vitro*, which contains several heptamer and NL sequences. Moreover, we found that the cryptic nicking site within the *TEL* fragment was cleaved by RAG proteins functionally depending on a 12-RSS framework (Fig. 5). Some of RAG-mediated nicking sites detected *in vitro* were coincident with the RAG-dependent recombination sites *in vivo* (Fig. 3B), while some others were generated only the *in vitro* system. It is well known that the recombination activity of RAG is also regulated by the accessibility of RAG proteins to nucleosomal DNA. Thus, the difference between RAG-dependent break points within *TEL* locus determined by extrachromosomal recombination assays and those by *in vitro* cleavage assays might be because of the difference between nucleosomal DNA and naked DNA.

The RAG-mediated nicking level at cryptic sites was <1% compared with that at authentic 12-RSS, but the nicking level was much higher than the background level seen for DNA cleavage-defective RAG1 (Fig. 3A). In addition to RAG-mediated nicks at cryptic sites, we also detected RAG-mediated DSBs on the *TEL* and *AML1* translocation regions (Fig. S3 in Supporting Information). Recent studies have suggested that RAG-mediated nicks at cRSSs in proto-oncogenes (*LMO2*, *TAL1*, *Ttg-1*, and *SIL*) are efficiently converted into hairpin products *in vitro* (Zhang & Swanson 2008). We could not detect any plausible hairpin structure at HL1 and HL2 sites but found that RAG proteins generate nicks into top and bottom strands at HL1 and HL2 sites, these nicks of which are close to each other but seem mutually independent (Fig. 3; Fig. S2B in Supporting Information). Thus, it is likely that RAG proteins introduce staggered nicks at HL1 and HL2 sites in a way different from the standard nick/hairpin mechanism, resulting in DSBs with overhanged DNA ends. It is evident that RAG proteins introduce DSBs into the non-B DNA structure region by the nick–nick mechanism (Raghavan *et al.* 2004, 2005). We investigated the DNA structure of the *TEL* region used in our study by the same experiments to examine non-B DNA structure as shown in the previous report (Raghavan *et al.* 2005) but could not detect any distinct structural features in the *TEL* translocation region (data not shown). Therefore, the detailed characteristics of RAG-mediated nicking at HL2 site on the *TEL* translocation region is likely to be different from those on the non-B DNA substrate analyzed by Raghavan *et al.* (2004). Previous studies on aberrant RAG-mediated DNA cleavages were summarized in Table S2 in Supporting Information.

It has been suggested based on *in vitro* studies that HMGB1 facilitates the normal V(D)J recombination through the enhancement of correct DNA binding and cleavage activities of RAG proteins at RSSs followed by coordinated DSB depending on the 12/23 rule (Swanson 2002a,b). However, it has been largely unknown whether HMGB1 enhances V(D)J recombination *in vivo*. Here, we found that HMGB1 enhances authentic V(D)J recombination in extrachromosomal recombination assays. Furthermore, an aberrant RAG-dependent *TEL–AML1* recombination was enhanced significantly more than that of the authentic RSSs in the same assay (Fig. 1). We found that HMGB1 enhances the DNA binding activity of the RAG complex to the *TEL* and *AML1* DNA fragments containing cryptic sites *in vitro* (Fig. 2). In addition, the several RAG-mediated nicks and DSBs in the *TEL* and *AML1* translocation regions were found to be enhanced by HMGB1, whereas most of nicks and DSBs were unchanged or slightly reduced in the presence of HMGB1 (Fig. 3; Fig. S4 in Supporting Information). In the course of RAG-dependent recombination, the RAG-mediated nicks might not necessary be converted to DSBs, because it was shown that unstable RAG postcleavage complexes allow DNA ends to participate in both homologous recombination (HR) and the error-prone alternative NHEJ pathway, and furthermore, RAG-mediated nicks efficiently stimulate HR pathway (Curry *et al.* 2005). Based on this concept, it is assumed that RAG proteins at cryptic RSSs are unstable because of less DNA binding affinity and might lead RAG-mediated aberrant nicks to aberrant HR pathway, resulting in generation of *TEL–AML1* chromosomal translocation. Taken together, it is postu-

lated that both of HMGB1-enhanced RAG-mediated nicks and DSBs might contribute to the stimulation of RAG-dependent *TEL–AML1* chromosomal translocation in acute lymphoblastic leukemia.

We found a functional bending center in the *TEL* DNA fragment by circular permutation assays (Fig. S5 in Supporting Information) but could not detect an evident difference between the bending angle with HMGB1 and that without HMGB1 in circular permutation assays. Instead, we figured out that the addition of HMGB1 to a DNA–RAG complex efficiently enhances the DNA bending activity by LMDC assays. The position of the bending center was slightly changed in the presence of HMGB1 compared with that in the absence of HMGB1 (Fig. S5C in Supporting Information). Although the exact reason for this change is not known presently, it is possible that the change of the bending position may contribute to the HMGB1-dependent enhancement of RAG nicking at HL2 site.

We showed that two possible NL elements, 23NL and 56NL, enhance the RAG-mediated nicking at HL2 in the presence of HMGB1. The function of 23NL could be considered in the framework of the 12/23 rule, whereas that of 56NL should be different from that of the authentic NL separated by the regular spacer sequence. Recently, it is suggested that the sequence preference of HMGB1 is important for HMGB1-dependent enhancement of the nicking activity of RAG on a cRSS (Zhang & Swanson 2009). However, the DNA binding affinity of HMGB1 to the *TEL* fragment containing HL2 was not changed by nucleotide mutations in 56NL substrate (data not shown). The HMGB1- and 56NL-dependent nicking enhancement was shown to be valid for a cryptic 12-RSS in the *TEL* recombination region, while this was not observed in a substrate containing the authentic 12-RSS (Fig. 8). The formation of synaptic complex consisting of 12-RSS, 23-RSS and RAG tetramer enhances the cleavage efficiency by RAG proteins in the process of V(D)J recombination. Based on the fact that RAG molecule forms oligomers and the results obtained in this study, we would propose a hypothetical model as shown in Fig. 9 (right panel). We postulate that the interaction of an RAG complex

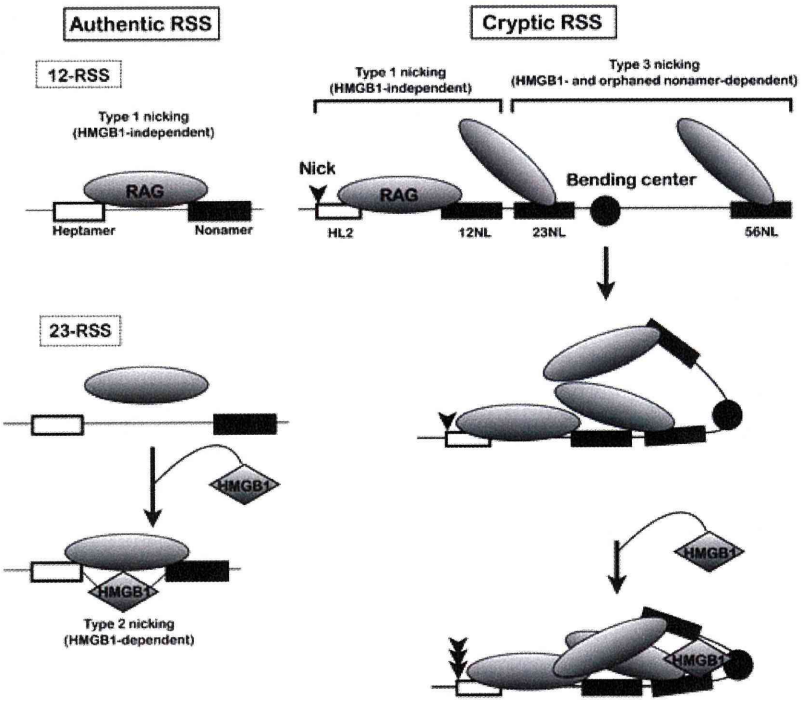

Figure 9 A hypothetical model of HMGB1- and orphan nonamer-like element-dependent enhancement of recombination activating gene (RAG)-mediated nicking activity at a cryptic recombination signal sequence (RSS). In the authentic RSSs, RAG complex efficiently introduces nicks into a 12-RSS without HMGB1 (Left upper panel, 'Type 1 nicking'), whereas HMGB1 stimulates RAG-mediated nicking at a 23-RSS by bending the DNA (Left bottom panel, 'Type 2 nicking'). In the cryptic 12-RSS within the *TEL* translocation region, orphan nonamer-like sequences (23NL and 56NL) and DNA bending enhanced by HMGB1 synergistically enhance the nicking activity of RAG complex at the cryptic 12-RSS (Right panel, 'Type 3 nicking').

containing a single 56NL with the other RAG complex on the 12-RSS-like sequence (HL2 and 12NL) *in cis* is brought by HMGB1-dependent stimulation of the DNA bending, and the RAG–RAG interaction might eventually enhance the nicking activity of RAG proteins (Fig. 9, 'Type 3'). However, we also considered a possibility that the interaction between RAG1 and HMGB1 might be involved in regulation of the enhancement of the RAG-mediated nicking through an allosteric effect of HMGB1 on the RAG-nicking complex, besides the DNA bending enhancement by HMGB1. It is also possible that HMGB1 enhances the formation of a synaptic complex *in trans*, consisting of RAG proteins on the cryptic 12-RSS and another RAG proteins on 23NL or 56NL element. In conclusion, we have found *in vitro* that the RAG-mediated nicking at the cryptic 12-RSS within the *TEL* translocation region is synergistically enhanced by HMGB1 and orphan NL sequences, through a moderate stimulation of the DNA bending angle by HMGB1. Based on our *in vitro* results, we speculate that the combination of structural features of DNA and the position of RAG-binding target sites could be a potential factor that contributes to an aberrant RAG-mediated nicking activity leading to chromosomal translocation in lymphoblastic leukemia and lymphoma.

Experimental procedures

Extrachromosomal recombination assay

The extrachromosomal recombination assay using pAT-12-RSS-23-RSS and pAT-*TEL–AML1* plasmid substrates was performed as described previously (Numata *et al.* 2010). 293T cells were transfected with a pAT substrate with HA–RAG expression vectors in the presence or absence of HA-HMGB1 expression vectors. Plasmid DNA was purified after 48 h post-transfection by rapid alkaline lysis method. Purified plasmid DNA was used for the transformation of *Escherichia coli*, and transformants were plated on agar containing 100 μg/mL ampicillin alone and one containing 100 μg/mL ampicillin and 6 or 14 μg/mL tetracycline. The Western blotting for the detection of HA–RAG proteins and FLAG–HMGB1 was carried out using rat anti-HA (sigma) and mouse anti-HMGB1 (MBL) antibodies, respectively.

Electrophoresis mobility shift assay

DNA probes containing 12-RSS wild-type and 12-RSS mutant with nucleotides changes in nonamer and heptamer sequences were prepared for EMSA as described (Rodgers *et al.* 1999). Each double-stranded DNA substrate was prepared by annealing a set of top and bottom oligonucleotides or PCR amplifications. Sequences of all oligonucleotides used in this study are summarized in Table S1 in Supporting Information. Only the top-strand oligonucleotides were labeled at the 5′ end using [γ-^{32}P]ATP and T4 polynucleotide kinase (TOYOBO). For EMSA, 50 fmol of DNA substrate was incubated with 45 ng of GST–RAG1ΔN and/or 20 ng of GST–RAG2ΔC in the presence or absence of 10 ng of His-HMGB1 at 30 °C for 30 min in 10 μL of a binding buffer (10 mM HEPES-KOH (pH 7.4), 10 mM $MgSO_4$, 100 mM KCl, 0.02 mg/mL BSA, 6% Glycerol, and 2 mM DTT). After incubation, DNA–protein complexes were separated through 6% native polyacrylamide gels in 45 mM Tris-borate, and the [^{32}P] end-labeled DNA was visualized by autoradiography. All [^{32}P]-labeled DNAs were gel-purified.

Cleavage assay

Cleavage reactions were performed essentially according to the previous studies (McBlane *et al.* 1995; Hiom & Gellert 1997). Substrate DNA (50 fmol), which was [^{32}P]-labeled at the 5′ end of the top or bottom strand, was incubated at 37 °C for 1 h with 60 ng of GST–RAG1ΔN, 30 ng of GST–RAG2ΔC, 15 ng of His-HMGB1, in 20 μL of reaction buffer (5 mM HEPES-KH (pH 7.4), 2 mM $MgSO_4$, 45 mM KCl, 4% Glycerol, and 2.2 mM DTT). After incubation, the reaction mixture was extracted with phenol/chloroform (1 : 1), and DNA was precipitated with ethanol. DNAs were separated through 6% denaturing polyacrylamide gels or 6% native polyacrylamide gels, and the [^{32}P] end-labeled DNA was visualized by autoradiography. All DNA substrates for the cleavage assay were prepared by PCR amplification followed by gel-purification.

LMDC assay

The ring closure assay is performed with the radiolabeled 134 bp fragments as described previously (Pasheva *et al.* 2004; Ugrinova *et al.* 2007). Briefly, 25 fmol of [^{32}P] end-labeled DNA fragment was preincubated at 30 °C for 30 min with 30 ng of core GST–RAG proteins in the presence or absence of 3, 6, and 9 ng of His-HMGB1 in 50 μL of reaction buffer (25 mM HEPES-KOH (pH 7.4), 10 mM $MgCl_2$, 2 mM dithiothreitol, and 2 mM ATP). After preincubation for 30 min, 2 U of T4 DNA ligase (TOYOBO) was added to the mixture and then subjected to an additional incubation at 25 °C for 30 min. The samples were extracted with phenol/chloroform (1 : 1), and DNAs were precipitated with ethanol. DNAs were separated through 5% native polyacrylamide gels, and the [^{32}P] end-labeled DNAs were visualized by autoradiography.

Acknowledgements

We thank Drs S. Saito (University of Tsukuba) and H. Haruki (University of Tsukuba and presently Ecole Polytechnique Federale de Lausanne) for the grateful discussion and suggestion and help for purification of His-HMGB1, respectively.

References

Aidinis, V., Bonaldi, T., Beltrame, M., Santagata, S., Bianchi, M.E. & Spanopoulou, E. (1999) The RAG1 homeodomain recruits HMG1 and HMG2 to facilitate recombination signal sequence binding and to enhance the intrinsic DNA-bending activity of RAG1-RAG2. *Mol. Cell. Biol.* **19**, 6532–6542.

Bergeron, S., Madathiparambil, T. & Swanson, P.C. (2005) Both high mobility group (HMG)-boxes and the acidic tail of HMGB1 regulate recombination-activating gene (RAG)-mediated recombination signal synapsis and cleavage *in vitro*. *J. Biol. Chem.* **280**, 31314–31324.

Cuomo, C.A., Mundy, C.L. & Oettinger, M.A. (1996) DNA sequence and structure requirements for cleavage of V(D)J recombination signal sequences. *Mol. Cell. Biol.* **16**, 5683–5690.

Curry, J.D., Geier, J.K. & Schlissel, M.S. (2005) Single-strand recombination signal sequence nicks *in vivo*: evidence for a capture model of synapsis. *Nat. Immunol.* **6**, 1272–1279.

Dai, Y., Wong, B., Yen, Y.M., Oettinger, M.A., Kwon, J. & Johnson, R.C. (2005) Determinants of HMGB proteins required to promote RAG1/2-recombination signal sequence complex assembly and catalysis during V(D)J recombination. *Mol. Cell. Biol.* **25**, 4413–4425.

Falvo, J.V., Thanos, D. & Maniatis, T. (1995) Reversal of intrinsic DNA bends in the IFN beta gene enhancer by transcription factors and the architectural protein HMG I(Y). *Cell* **83**, 1101–1111.

Gellert, M. (2002) V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu. Rev. Biochem.* **71**, 101–132.

van Gent, D.C., Hiom, K., Paull, T.T. & Gellert, M. (1997) Stimulation of V(D)J cleavage by high mobility group proteins. *EMBO J.* **16**, 2665–2670.

Hesse, J.E., Lieber, M.R., Mizuuchi, K. & Gellert, M. (1989) V(D)J recombination: a functional definition of the joining signals. *Genes Dev.* **3**, 1053–1061.

Hiom, K. & Gellert, M. (1997) A stable RAG1-RAG2-DNA complex that is active in V(D)J cleavage. *Cell* **88**, 65–72.

Kitagawa, Y., Inoue, K., Sasaki, S., Hayashi, Y., Matsuo, Y., Lieber, M.R., Mizoguchi, H., Yokota, J. & Kohno, T. (2002) Prevalent involvement of illegitimate V(D)J recombination in chromosome 9p21 deletions in lymphoid leukemia. *J. Biol. Chem.* **277**, 46289–46297.

Lee, G.S., Neiditch, M.B., Salus, S.S. & Roth, D. (2004) RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. *Cell* **117**, 171–184.

Marculescu, R., Le, T., Simon, P., Jaeger, U. & Nadel, B. (2002) V(D)J-mediated translocations in lymphoid neoplasms: a functional assessment of genomic instability by cryptic sites. *J. Exp. Med.* **195**, 85–98.

Marculescu, R., Vanura, K., Montpellier, B., Roulland, S., Le, T., Navarro, J.M., Jager, U., McBlane, F. & Nadel, B. (2006) Recombinase, chromosomal translocations and lymphoid neoplasia: targeting mistakes and repair failures. *DNA Repair (Amst)* **5**, 1246–1258.

McBlane, J.F., van Gent, D.C., Ramsden, D.A., Romeo, C., Cuomo, C.A., Gellert, M. & Oettinger, M.A. (1995) Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell* **83**, 387–395.

Nagawa, F., Ishiguro, K., Tsuboi, A., Yoshida, T., Ishikawa, A., Takemori, T., Otsuka, A.J. & Sakano, H. (1998) Footprint analysis of the RAG protein recombination signal sequence complex for V(D)J type recombination. *Mol. Cell. Biol.* **18**, 655.

Numata, M., Saito, S. & Nagata, K. (2010) RAG-dependent recombination at cryptic RSSs within TEL-AML1 t(12;21)(p13;q22) chromosomal translocation region. *Biochem. Biophys. Res. Commun.* **402**, 718–724.

Pasheva, E., Sarov, M., Bidjekov, K., Ugrinova, I., Sarg, B., Lindner, H. & Pashev, I.G. (2004) *In vitro* acetylation of HMGB-1 and -2 proteins by CBP: the role of the acidic tail. *Biochemistry* **43**, 2935–2940.

Raghavan, S.C., Kirsch, I.R. & Lieber, M.R. (2001) Analysis of the V(D)J recombination efficiency at lymphoid chromosomal translocation breakpoints. *J. Biol. Chem.* **276**, 29126–29133.

Raghavan, S.C., Swanson, P.C., Ma, Y. & Lieber, M.R. (2005) Double-strand break formation by the RAG complex at the bcl-2 major breakpoint region and at other non-B DNA structures *in vitro*. *Mol. Cell. Biol.* **25**, 5904–5919.

Raghavan, S.C., Swanson, P.C., Wu, X., Hsieh, C.L. & Lieber, M.R. (2004) A non-B-DNA structure at the Bcl-2 major breakpoint region is cleaved by the RAG complex. *Nature* **428**, 88–93.

Ramsden, D.A., Baetz, K. & Wu, G.E. (1994) Conservation of sequence in recombination signal sequence spacers. *Nucleic Acids Res.* **22**, 1785–1796.

Ramsden, D.A., McBlane, J.F., van Gent, D.C. & Gellert, M. (1996) Distinct DNA sequence and structure requirements for the two steps of V(D)J recombination signal cleavage. *EMBO J.* **15**, 3197–3206.

Rodgers, K.K., Villey, I.J., Ptaszek, L., Corbett, E., Schatz, D.G. & Coleman, J.E. (1999) A dimer of the lymphoid protein RAG1 recognizes the recombination signal sequence and the complex stably incorporates the high mobility group protein HMG2. *Nucleic Acids Res.* **27**, 2938–2946.

Sawchuk, D.J., Weis-Garcia, F., Malik, S., Besmer, E., Bustin, M., Nussenzweig, M.C. & Cortes, P. (1997) V(D)J recombination: modulation of RAG1 and RAG2 cleavage activity on 12/23 substrates by whole cell extract and DNA-bending proteins. *J. Exp. Med.* **185**, 2025–2032.

Swanson, P.C. (2002a) A RAG-1/RAG-2 tetramer supports 12/23-regulated synapsis, cleavage, and transposition of V(D)J recombination signals. *Mol. Cell. Biol.* **22**, 7790–7801.

Swanson, P.C. (2002b) Fine structure and activity of discrete RAG-HMG complexes on V(D)J recombination signals. *Mol. Cell. Biol.* **22**, 1340–1351.

Thomas, J.O. & Travers, A.A. (2001) HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends Biochem. Sci.* **26**, 167–174.

Tonegawa, S. (1983) Somatic generation of antibody diversity. *Nature* **302**, 575–581.

Ugrinova, I., Mitkova, E., Moskalenko, C., Pashev, I. & Pasheva, E. (2007) DNA bending versus DNA end joining activity of HMGB1 protein is modulated *in vitro* by acetylation. *Biochemistry* **46**, 2111–2117.

Yoshida, T., Tsuboi, A., Ishiguro, K., Nagawa, F. & Sakano, H. (2000) The DNA-bending protein, HMG1, is required for correct cleavage of 23 bp recombination signal sequences by recombination activating gene proteins *in vitro*. *Int. Immunol.* **12**, 721–729.

Zhang, M. & Swanson, P.C. (2008) V(D)J recombinase binding and cleavage of cryptic recombination signal sequences identified from lymphoid malignancies. *J. Biol. Chem.* **283**, 6717–6727.

Zhang, M. & Swanson, P.C. (2009) HMGB1/2 can target DNA for illegitimate cleavage by the RAG1/2 complex. *BMC Mol. Biol.* **10**, 24.

Received: 18 February 2011
Accepted: 9 May 2011

Supporting Information/Supplementary material

The following Supporting Information can be found in the online version of the article:

Figure S1 Purification of recombinant proteins and their recombination signal sequence binding activity.

Figure S2 Detection of recombination activating gene nicking sites on top and bottom strands of the *TEL* fragment by primer extension.

Figure S3 HMGB1 enhances recombination activating gene-mediated double-strand breaks in the *TEL* and *AML1* translocation regions.

Figure S4 Recombination activating gene proteins specifically bind to 56NL.

Figure S5 Identification of the bending center on the *TEL* DNA fragment between 23NL and 56NL.

Table S1 Primers and oligonucleotides used in this study

Table S2 Recombination activating gene-associated chromosomal translocations

Data S1 Supporting text.

Additional Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

JOURNAL OF VIROLOGY, Aug. 2011, p. 7504–7512 Vol. 85, No. 15
0022-538X/11/$12.00 doi:10.1128/JVI.02375-10

Recognition of Cap Structure by Influenza B Virus RNA Polymerase Is Less Dependent on the Methyl Residue than Recognition by Influenza A Virus Polymerase▽†

Chitose Wakai,[1,2] Minako Iwama,[2] Kiyohisa Mizumoto,[2,3] and Kyosuke Nagata[1]*

Department of Infection Biology, Graduate School of Comprehensive Human Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba 305-8575, Japan[1]; Department of Biochemistry, School of Pharmaceutical Sciences, Kitasato University, 5-9-1 Shirokane, Minato-ku, Tokyo 108-8641, Japan[2]; and Microbial Chemistry Research Center, 3-14-23 Kamiosaki, Shinagawa-ku, Tokyo 141-0021, Japan[3]

Received 12 November 2010/Accepted 10 May 2011

The cap-dependent endonuclease activity of the influenza virus RNA-dependent RNA polymerase cleaves host mRNAs to produce capped RNA fragments for primers to initiate viral mRNA synthesis. The influenza A virus (FluA) cap-dependent endonuclease preferentially recognizes the cap1 structure (m^7GpppNm). However, little is known about the substrate specificity of the influenza B virus (FluB) endonuclease. Here, we determined the substrate specificity of the FluB polymerase using purified viral RNPs and ^{32}P-labeled polyribonucleotides containing a variety of cap structures (m^7GpppGm, m^7GpppG, and GpppG). We found that the FluA polymerase cleaves m^7G-capped RNAs preferentially. In contrast, the FluB polymerase could efficiently cleave not only m^7G-capped RNAs but also unmethylated GpppG-RNAs. To identify a key amino acid(s) related to the cap recognition specificity of the PB2 subunit, the transcription activity of FluB polymerases containing mutated cap-binding domains was examined by use of a minireplicon assay system. In the case of FluA PB2, Phe323, His357, and Phe404, which stack the m^7GTP, and Glu361 and Lys376, which make hydrogen bonds with a guanine base, were essential for the transcription activity. In contrast, in the case of FluB PB2, the stacking interaction of Trp359 with a guanine base and putative hydrogen bonds using Gln325 and Glu363 were enough for the transcription activity. Taking these results together with the result for the cap-binding activity, we propose that the cap recognition pocket of FluB PB2 does not have the specificity for m^7G-cap structures and thus is more flexible to accept various cap structures than FluA PB2.

Influenza A virus (FluA) and influenza B virus (FluB) belong to the family of *Orthomyxoviridae*. The genomes of FluA and FluB are composed of a set of eight segments of RNA (vRNA) of negative polarity. vRNA is complexed with nucleoprotein (NP) and associated with the RNA polymerase to form viral ribonucleoprotein (vRNP) complexes. vRNP is an essential unit for both transcription and replication (9). In transcription, the RNA polymerase catalyzes not only RNA polymerization and polyadenylation of mRNA but also cleavage of host mRNAs to generate capped RNA fragments. The RNA polymerase is composed of one molecule each of three viral proteins, PB1, PB2, and PA. PB1 plays central roles in both RNA polymerase assembly (27, 31) and RNA polymerization (6). It contains the conserved motifs characteristic of RNA-dependent RNA polymerases and is directly involved in RNA chain elongation (1, 2). It binds to 5′- and 3′-terminal sequences of vRNA and cRNA (cRNA to vRNA), which are conserved in all segments and act as *cis*-acting elements for the viral RNA synthesis. PB2 is required for transcription and binds to the cap structures of host mRNAs. Recently, the structural features of the cap-binding site in FluA PB2 and the FluA PB1-PB2 contact site have been determined by functional studies and crystallography (12, 31). PA is involved in not only virus genome replication but also transcription as an endonuclease for generation of primers for RNA synthesis (8, 10, 13, 19, 36). It is also reported that PA is important for the polymerase assembly (19). The structure of the PB1-PA contact site has also been determined crystallographically (14, 27).

The FluA polymerase exhibits a cap-dependent endonuclease activity, which cleaves host mRNAs to produce capped RNA fragments with lengths of 11 to 13 nucleotides (nt). The resulting capped RNA fragment serves as a primer to initiate viral mRNA synthesis. It is well known that in the case of the FluA polymerase, eukaryotic mRNAs containing m^7G(5′)ppp(5′)Nm (cap1) and m^7G(5′)ppp(5′)NmN′m (cap2) structures stimulate *in vitro* viral RNA transcription strongly (4, 5, 29). Removal of m^7G of the cap from mRNA eliminates the priming activity, and naturally occurring uncapped mRNAs do not prime transcription (5, 29). In addition, the presence of methyl groups in the cap is required for the priming activity; reovirus mRNAs with 5′-terminal GpppG are inactive as primers (3). It has also been demonstrated that each of the two methyl groups in the cap1 structure, the 7-methyl residue of guanine and the 2′-*O*-methyl on the ribose of guanosine, strongly influences the capped RNA-primed transcription activity (4).

Biochemical and structural studies revealed the functional structures of the cap-binding proteins, including FluA PB2 (12), human translation initiation factor 4E (eIF4E) (33, 34),

* Corresponding author. Mailing address: Department of Infection Biology, Graduate School of Comprehensive Human Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba 305-8575, Japan. Phone and fax: 81-29-853-3233. E-mail: knagata@md.tsukuba.ac.jp.

† Supplemental material for this article may be found at http://jvi.asm.org/.

▽ Published ahead of print on 18 May 2011.

human nuclear cap-binding protein 20 (CBP20) (23), and vaccinia virus (nucleoside-2′-*O*-)-methyltransferase (VP39) (16). The overall structures of these four cap-binding proteins differ widely due to their evolutionarily unrelated origins, but the cap-binding pockets form a common structure and preferentially bind to the 7-methylated cap structure. These cap-binding proteins hardly bind to the unmethylated cap structure.

Most of our knowledge on the transcription mechanism of the influenza virus genome has been derived from studies on the FluA polymerase, whereas little is known about the FluB polymerase. It is reported that α-amanitin, a potent inhibitor for the host cell RNA polymerase II, inhibits influenza virus transcription, suggesting that eukaryotic mRNAs containing the cap structure are essential for influenza virus transcription (21). Using α-amanitin, we found that the growth of FluB is more sensitive to the amount of cellular mRNA than that of FluA (data not shown). To elucidate the transcription initiation mechanism of the FluB polymerase, we tried to determine the specificity of cap recognition by the FluB polymerase. First, we compared the substrate specificities of FluA and FluB polymerases using purified vRNPs and various capped RNA substrates (m^7GpppGm-, m^7GpppG-, and GpppG-RNA) and found that the FluB polymerase efficiently cleaves not only m^7G-capped RNAs but also unmethylated GpppG-RNA, whereas the FluA polymerase cleaves m^7G-capped RNAs specifically. We then tried to identify key amino acids related to the cap recognition of FluB PB2. In order to examine the transcription activity using mutated PB2 proteins, we utilized FluA and FluB minireplicon assay systems using a virus polymerase-dependent reporter gene (17, 35). The minireplicon system has been utilized for a number of functional analyses of *cis*-acting elements with the viral genome and *trans*-acting viral factors (10, 35). The reporter gene contains a coding region flanked by each viral 5′ and 3′ untranslated region (UTR), which function as promoters, and therefore mimics an influenza virus genomic segment. Using this assay system, we identified the important amino acids required for the cap recognition by the FluB polymerase by referencing functionally important amino acids in the FluA polymerase (12).

Based on the findings using the assay systems, we propose that the FluB polymerase possesses a novel cap recognition mechanism, which is different not only from the FluA polymerase but also from well-known cap-binding proteins. These findings could be important to develop novel anti-influenza virus drugs targeting the cap recognition and cleavage reaction.

MATERIALS AND METHODS

Biological materials. Monolayer cultures of 293T and MDCK cells were maintained at 37°C in Dulbecco's modified Eagle medium (DMEM) and minimal essential medium (MEM) (Nissui), respectively, supplemented with 10% fetal calf serum (Cell Culture Technologies). Influenza virus A/Panama/2007/99 (A/PA/99) and B/Shanghai/361/2002 (B/SH/02) were kindly supplied by Y. Suzuki and T. Gotanda (Kitasato Institute, Research Center for Biologicals, Saitama, Japan). Vaccinia virus capping enzyme and recombinant human mRNA (guanine-7-)methyltransferase (rhMTase) were prepared according to a previously described procedure (28).

Cloning of cDNAs for viral RNA polymerase subunits and nucleoprotein cDNA. For construction of mammalian expression vectors for influenza virus polymerase subunits (PB1, PB2, and PA) and nucleoprotein (NP), cDNAs corresponding to the full-length PB1, PB2 with a FLAG tag at its C terminus (PB2cFLAG), PA, and NP were amplified by reverse transcription-PCR (RT-PCR) from vRNAs of influenza virus A/PA/99 and B/SH/02 as templates using the following sets of phosphorylated primers (see Table S1 in the supplemental material): A-PB1-FOR and A-PB1-REV for FluA-PB1, A-PB2-FOR and A-PB2-cFLAG-REV for FluA-PB2cFLAG, A-PA-FOR and A-PA-REV for FluA-PA, A-NP-FOR and A-NP-REV for FluA-NP, B-PB1-FOR and B-PB1-REV for FluB-PB1, B-PB2-FOR and B-PB2-cFLAG-REV for FluB-PB2cFLAG, B-PA-FOR and B-PA-REV for FluB-PA, and B-NP-FOR and B-NP-REV for FluB-NP. The PCR products were then cloned into the EcoRV site of pCAGGS-P7 (7), resulting in construction of pCAGGS-Panama-PB1, pCAGGS-Panama-PB2-cFLAG, pCAGGS-Panama-PA, pCAGGS-Panama-NP, pCAGGS-Shanghai-PB1, pCAGGS-Shanghai-PB2-cFLAG, pCAGGS-Shanghai-PA, and pCAGGS-Shanghai-NP. cDNAs for PB2 mutants were prepared by site-directed mutagenesis using the primer sets for FluA-PB2-cFLAG and FluB-PB2-cFLAG and mutant primer sets (see Table S2 in the supplemental material). The PB2 mutant genes have been fully sequenced by standard methods (35).

Preparation of influenza virus vRNP. To prepare vRNP, we first treated purified influenza virus virions at 30°C for 60 min with a disruption buffer consisting of 50 mM Tris-HCl (pH 8.0), 100 mM KCl, 5 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 5% glycerol, 2% Triton X-100, and 2% lysolecithin according to a method described previously (32). The sample was then directly subjected to centrifugation on a 30 to 60% (wt/vol) linear gradient of glycerol on a 70% (wt/vol) glycerol cushion in 50 mM Tris-HCl (pH 8.0) and 150 mM NaCl in a Beckman MLS-50 rotor with adapters at 163,000 × g_{AV} for 3 h at 4°C. Fractionation was carried out from the top of the gradient. Fractions containing vRNP were pooled and then used for *in vitro* endonuclease and elongation assays.

Preparation of various RNA substrates. Triphosphate-ended RNA with the 33-nucleotide sequence 5′-GAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA UAAA-3′, designated pppG-RNA, was synthesized by using T7 RNA polymerase (Amersham Biosciences) and a synthetic DNA template. The protocol was previously described (30). Briefly, to prepare the template for the T7 RNA polymerase, the oligonucleotide T7P (5′-TAATACGACTCACTATA-3′), corresponding to the T7 promoter (−17 to −1), was annealed to the template oligonucleotide T7-polyA-R1 (5′-TTTATTTTTTTTTTTTTTTTTTTTTTTTT TTCTATAGTGAGTCGTATTA-3′, where the underlined sequence is complementary to the T7 promoter [−17 to −1]). After the transcription reaction, the transcription mixture was treated with DNase I (Roche Applied Science). RNA was then extracted with phenol-chloroform, ethanol precipitated, and used as a capping substrate. To synthesize m^7G[^{32}P]pppGm-RNA and G[^{32}P]pppG-RNA, 50 pmol of pppG-RNA was incubated at 37°C for 2 h in the presence of 8 μM [α-^{32}P]GTP (800 cpm/fmol) and an appropriate amount of purified vaccinia virus capping enzyme, which has guanylyltransferase, guanine-7-methyltransferase, and ribose-2′-*O*-methyl-transferase activities, in a reaction mixture (50 μl) containing 50 mM Tris-HCl (pH 7.9), 2 mM $MgCl_2$, 40 mM NaCl, and 20 mM DTT in the presence or absence of 150 μM *S*-adenosyl-L-methionine (AdoMet). After the reaction, capped RNA was extracted with phenol-chloroform, ethanol precipitated, and dissolved in H_2O. To synthesize m^7G[^{32}P]pppG-RNA, 0.4 pmol of G[^{32}P]pppG-RNA was incubated at 30°C for 20 min with 15 ng/μl of rhMTase in a reaction mixture (20 μl) containing 25 mM Tris-HCl (pH 7.9), 0.5 mM DTT, 0.1 mg/ml bovine serum albumin (BSA), and 50 μM AdoMet. The RNA was extracted with phenol-chloroform, ethanol precipitated, and dissolved in H_2O. To confirm the cap structure on the synthesized RNA, the cap structure of the synthesized ^{32}P-capped RNA was liberated by digestion with nuclease P_1 (Wako) (28). The reaction product was analyzed by thin-layer chromatography (TLC) on a polyethyleneimine (PEI)-cellulose plate (PEI-CEL UV_{254}; Macherey-Nagel) with 0.65 M LiCl and visualized by autoradiography.

***In vitro* capped RNA cleavage and RNA elongation reactions.** The determination of Flu cap-dependent endonuclease activity and the subsequent RNA elongation reaction were carried out in a reaction mixture (25 μl) containing 50 mM Tris-HCl (pH 7.9), 0.1 mM ammonium acetate, 5 mM $MgCl_2$, 2.5 mM DTT, 0.1% Nonidet P-40, 8 U of RNasin, 3 to 5 fmol of each ^{32}P-capped RNA (800 cpm/fmol), and an appropriate amount of purified vRNPs without or with ATP, UTP, GTP, or CTP. The reaction mixture was incubated at 30°C for 2 h, and then RNA products were extracted with phenol-chloroform and ethanol precipitated. The RNA products denatured with formamide were electrophoresed in a 20% acrylamide gel containing 8 M urea. After electrophoresis, the gel was dried, and RNAs were visualized by autoradiography. The amount of synthesized RNA was measured with a liquid scintillation counter (LS6000IC; Beckman). The endonuclease activity was represented as a ratio of the amount of cleaved RNAs to that of total capped RNAs, and the RNA elongation efficiency was represented as a ratio of the amount of transcripts to that of total capped RNAs.

Cap-binding assay. UV cross-linking was carried out to measure the cap-binding activity of viral RNA polymerases. A reaction mixture (12 μl) containing

50 mM Tris-HCl (pH 7.9), 0.1 mM ammonium acetate, 5 mM $MgCl_2$, 2.5 mM DTT, 250 fmol of uncapped RNA substrate, 50 fmol of each ^{32}P-capped RNA (~800 cpm/fmol), and an appropriate amount of purified vRNPs was incubated for 30 min on ice and then irradiated on ice for 10 min with 254-nm UV light (FUNA-UV-Linker FS-1500 [Funakoshi, Japan]) with 0.2 mg/ml of heparin. The ^{32}P-labeled products were digested with nuclease P_1, analyzed by 6% SDS-PAGE, and detected by autoradiography.

Minireplicon assay. Two plasmid vectors carrying a reporter gene (an artificial influenza virus genome containing the firefly luciferase gene of negative polarity, which is synthesized in cells by the human DNA-dependent RNA polymerase I [Pol I]), were constructed as described previously (35). A fragment containing the luciferase gene sandwiched by 5′- and 3′-terminal sequences of FluA/PA/99 and FluB/SH/02 segment 8 was amplified by PCR with specific primers 5′-GTAGTAGAAACAAGGGTGTTTTTTACTCGAGATCTTACAATTTGGACTTTCCGCCCTT-3′ and 5′-GATCCGTCTCCGGGAGCAAAAGCAGGGTGACAAAGACATAATGCATATGGAAGACGCCAAAAACATAAAGAAAGG-3′ for FluA/PA/99 and 5′-TATTCGTCTCAGGGAGCAGAAGCAGAGGATTTGTTTAGTCACTGGCAAACGGAAAAAAATGGAAGACGCCAAAAACATAAAG-3′ and 5′-ATATCGTCTCGTATTAGTAGTAACAAGAGGATTTTATTTTAAATTTACAATTTGGACTTTCCGCC-3′ for FluB/SH/02, using pGV-B (the promoterless luciferase reporter vector; Toyo Inc.) as a template. The amplified PCR products were digested with BsmBI and cloned into pHH21 containing the promoter region of the human rRNA gene (24, 25), which had been digested with BsmBI. The constructed plasmids were designated pHH-A-vNS-Luc and pHH-B-vNS-Luc, in which the luciferase gene in reverse orientation sandwiched with 23- and 26-nucleotide 5′- and 3′-terminal sequences of the FluA/PA/99 segment 8 or 30 and 44-nucleotide 5′- and 3′-terminal sequences of the FluB/SH/02 segment 8, respectively, is placed under the control of the human Pol I promoter. 293T cells were transfected with plasmids for the expression of the FluA minireplicon (pCAGGS-Panama-PB1, pCAGGS-Panama-PB2-cFLAG, pCAGGS-Panama-PA, pCAGGS-Panama-NP, and pHH-A-vNS-Luc) or FluB minireplicon (pCAGGS-Shanghai-PB1, pCAGGS-Shanghai-PB2-cFLAG, pCAGGS-Shanghai-PA, pCAGGS-Shanghai-NP, and pHH-B-vNS-Luc). A plasmid for the expression of *Renilla* luciferase driven by the simian virus 40 (SV40) promoter was used as an internal control for the dual-luciferase assay. As a negative control, 293T cells were transfected with the same plasmids, except for the omission of the PB2 expression plasmid. After transfection, the cells were incubated at 37°C for 24 h, and then the luciferase activity was determined using commercially available reagents (Promega) according to the manufacturer's protocol. The relative luminescence intensity was measured with a luminometer for 20 s. To measure the levels of accumulation of viral mRNA, cRNA, and vRNA, quantitative RT-PCR was performed. Total RNA was extracted from transfected cells and then reverse transcribed with either (i) oligo$(dT)_{20}$ for synthesizing cDNA from viral mRNA, (ii) 5′-ATATCGTCTCGTATTAGTAGTAACAAGAGCATT-3′, which is complementary to the 3′ portion of cRNA of the reporter gene, for synthesizing cDNA from cRNA, or (iii) 5′-TCCATCACGGTTTTGGAATGTTTACTACAC-3′, which is complementary to vRNA, for synthesizing cDNA from vRNA of the reporter gene. These single-stranded cDNAs were subjected to real-time quantitative PCR analyses (Thermal Cycler Dice real-time system TP800; TaKaRa) with SYBR Premix Ex *Taq* (TaKaRa) and two specific primers, 5′-TCCATCACGGTTTTGGAATGTTTACTACAC-3′, corresponding to the firefly luciferase mRNA between nucleotide sequence positions 728 and 757, and 5′-GTGCGCCCCCAGAAGCAATTTC-3′, which is complementary to the firefly luciferase mRNA between nucleotide sequence positions 931 and 952. *Renilla* luciferase mRNA was also amplified with two specific primers, 5′-GCAGCATATCTTGAACCATTC-3′, corresponding to the *Renilla* luciferase mRNA between nucleotide sequence positions 598 and 618, and 5′-CATCACTTGCACGTAGATAAG-3′, which is complementary to the *Renilla* luciferase mRNA between nucleotide sequence positions 725 and 745. The relative amounts of mRNA, cRNA, and vRNA were calculated by using the second-derivative maximum method and normalized to the amount of *Renilla* luciferase mRNA. The ratio of the amounts of mRNA and cRNA relative to vRNA is shown.

Detection of capped RNA coprecipitated with the viral RNA polymerase. 293T cells were transfected with plasmids for the expression of the FluB viral proteins, PB1, FLAG-tagged PB2 (wild-type or mutated PB2), and PA. At 24 h posttransfection, cells were resuspended in a lysis buffer (20 mM Tris-HCl [pH 7.9], 100 mM NaCl, 30 mM KCl, and 0.1% Nonidet P-40). The RNA polymerase complex composed of PB1, FLAG-tagged PB2, and PA was purified by incubation with anti-FLAG M2 agarose (Sigma) at 4°C for 3 h and eluted with an elution buffer (50 mM Tris-HCl [pH 7.9], 100 mM ammonium acetate, 5 mM $MgCl_2$, and 10% [vol/vol] glycerol) containing 0.1 mg/ml FLAG peptide (Sigma). RNAs which interact with the viral RNA polymerase was extracted from recombinant RNA polymerase complexes (100 ng of PB1 equivalents) with phenol-chloroform and ethanol precipitated with 20 μg of carrier tRNA. After treatment with calf intestinal alkaline phosphatase (CIAP), which removes free phosphate groups, periodate oxidation under mild conditions followed by β-elimination with aniline was carried out to remove 5′-terminal m^7G from capped RNA, generating RNA with 5′-triphosphate, which is the substrate for vaccinia virus capping enzyme, as described previously (4, 11). The RNA was then recapped using vaccinia virus capping enzyme with [α-^{32}P]GTP as described in the previous section. To measure the amount of ^{32}P-labeled capped RNA, the RNA was digested with tobacco acid pyrophosphatase (TAP) (Sigma) at 37°C for 1 h in a buffer containing 50 mM sodium acetate (pH 5.5), 5 mM EDTA, and 10 mM 2-mercaptoethanol. The reaction product was analyzed by thin-layer chromatography on a PEI-cellulose plate as described above, and the amount of [^{32}P]m^7Gp was measured with a liquid scintillation counter.

RESULTS

***In vitro* capped RNA cleavage reaction and subsequent RNA elongation reaction.** The FluA polymerase requires the cap1 structure (m^7GpppNm) stringently for transcription (4). In contrast, little is known about the requirement for the cap structure of the FluB polymerase. Thus, we first examined the efficiency of the capped RNA cleavage reaction and subsequent RNA elongation reaction by FluA and FluB polymerases using cap1-RNA (m^7GpppGm-RNA). The cap1-RNA labeled with ^{32}P in the cap structure was incubated with purified vRNP (see Fig. S1A in the supplemental material) in the absence or presence of nucleoside triphosphates (NTPs) (Fig. 1A). RNA products were analyzed by 15% PAGE containing 8 M urea. FluA and FluB polymerases cleaved the cap1-RNA and produced 11- to 13-nucleotide and 11- to 12-nucleotide RNAs, respectively, in the absence of NTPs (Fig. 1A, lanes 2 to 7), indicating that the endonuclease activity of FluB is different from that of FluA in the distance of cleavage site from the cap structure. This cleavage pattern was observed commonly among FluA strains and among FluB strains (see Fig. S1B, lanes 2 to 6, in the supplemental material). The cleaved RNA products were elongated in the presence of NTPs in a dose-dependent manner (Fig. 1A, lanes 8 to 13), but the elongation efficiency of the FluB polymerase was lower than that of the FluA polymerase. We also confirmed that these elongated products contain full-length transcripts from 8 segments and are partially polyadenylated (see Fig. S2 in the supplemental material). To investigate the cap-binding activity of the polymerases, UV cross-linking assays were carried out (Fig. 1B). Cap1-RNA specifically bound to PB2 in both the FluA and FluB polymerases, although the cap-binding activity of FluB PB2 is less (~25%) than that of FluA PB2. These results suggest that the FluA and FluB polymerases are different in their binding to RNA containing the cap1 structure and in their cleavage modes.

Specificity of recognition of cap structures by Flu polymerases. To investigate the specificity of recognition of cap structures by FluA and FluB polymerases, we carried out similar experiments using RNA primers containing various cap structures. To this end, we prepared ^{32}P-labeled RNAs containing differently methylated cap structures, such as m^7GpppGm, m^7GpppG, and GpppG, as described in Materials and Methods. After preparation, we analyzed the terminal cap structure using nuclease-digested samples (see Materials and Methods) and thin-layer chromatography on a PEI-cellulose plate. As shown in Fig. 2A, we confirmed that each RNA

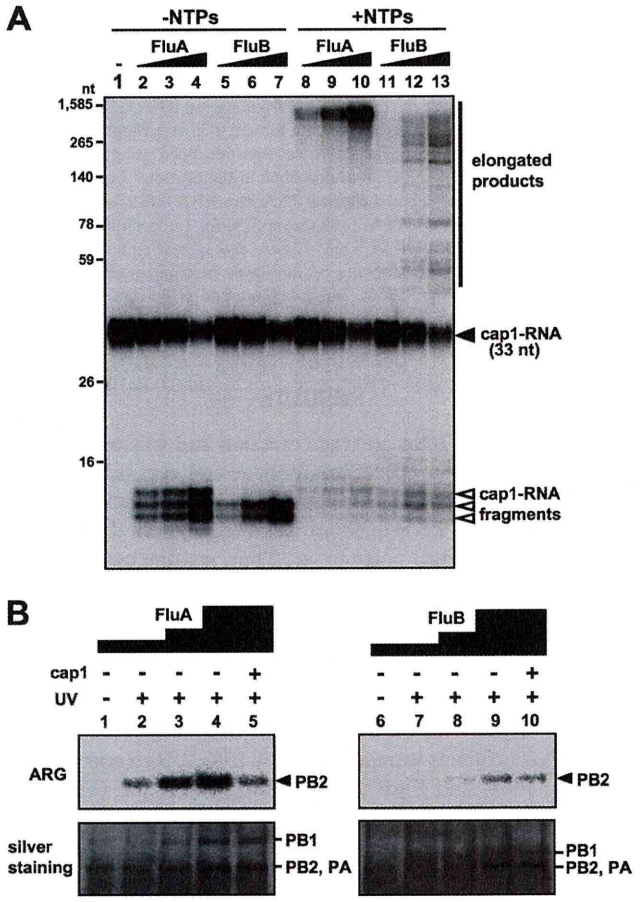

FIG. 1. *In vitro* capped RNA cleavage, RNA elongation and cap-binding reactions. (A) Dose dependency of *in vitro* capped RNA cleavage and subsequent RNA elongation by vRNP. *In vitro* capped RNA cleavage and RNA elongation reactions were performed with 20, 40, and 80 ng of FluA (lanes 2 to 4 and 8 to 10) and FluB (lanes 5 to 7 and 11 to 13) vRNP using 2 fmol of m^7GpppGm-RNA. Capped RNA cleavage was performed in the absence of NTPs (lanes 2 to 7), while RNA elongation was performed in the presence of NTPs (lanes 8 to 13). Synthesized RNA products were analyzed by 15% PAGE containing 8 M urea. (B) Interaction of vRNP with the cap1 structure. UV cross-linking was performed using 50, 100, and 200 ng of FluA (lanes 1 to 5) and FluB (lanes 6 to 10) vRNPs with (lanes 2 to 5 and 7 to 10) or without (lanes 1 and 6) UV irradiation at 254 nm. Competition experiments were performed in the presence of 100 pmol of m^7GpppGm analogue (lanes 5 and 10). Upper panels show autoradiography (ARG), while lower panels show silver staining patterns.

had the expected cap structure. Using these RNAs as substrates, we carried out *in vitro* capped RNA cleavage and subsequent RNA elongation reactions with FluA or FluB vRNPs. As expected, FluA vRNP specifically cleaved both m^7GpppGm-RNA and m^7GpppG-RNA, although the latter was less efficiently cleaved (Fig. 2B, lanes 2, 5, and 8, and D). The m^7GpppGm-RNA fragments were most successfully elongated into viral mRNAs (Fig. 2C, lane 2, and E). In contrast, FluB vRNP could cleave GpppG-RNA efficiently in addition to the m^7GpppGm-RNA and m^7GpppG-RNA (Fig. 2B, lanes 3, 6, and 9, and D). It is noteworthy that m^7GpppGm-RNA fragments also served as an efficient primer for chain elongation, as is the case for the FluA polymerase (Fig. 2C, lane 3, and E). Moreover, we carried out UV cross-linking assays using RNA primers containing various cap structures (Fig. 2F). Interestingly, the cap-binding activity was detected just using m^7GpppGm-RNA with both FluA and FluB vRNPs. These results indicate that the guanine-7-methyl residue is a key for stable cap binding of both FluA and FluB polymerases. It is also indicated that the cap-binding activity is strictly related to the elongation efficiency but not to the cleavage reaction. It is presently unknown why the binding of GpppG and m^7GpppG was not detected under the conditions employed, while m^7GpppG-RNA was recognized and cleaved by both FluA and FluB polymerases and GpppG-RNA was by the FluB polymerase. Since m^7GpppG-RNA and GpppG-RNA were not effective for elongation, the cleavage of these cap structures would be abortive for transcription, possibly due to improper recognition.

Identification of key amino acids involved in the cap recognition specificity of the PB2 subunit of the FluB polymerase. To clarify the cap recognition mechanism, we focused our structure-related functional studies on the interaction between the cap1 structure and the PB2 subunit, which has the cap-binding domain. It is quite likely that amino acid residues essential for cap binding are conserved between FluA and FluB (Fig. 3A). Three-dimensional (3D) structural studies (12) revealed that in the FluA PB2 cap-binding domain (Fig. 3B), Phe404 and His357 sandwich the methylated guanine and Phe323 stacks on the ribose of m^7GTP. Glu361 makes hydrogen bonds with the N1 and N2 positions of guanine, and Lys376 also makes a hydrogen bond with position O6 of guanine. Computer-associated modeling could make the FluB PB2 cap-binding domain fit on the FluA PB2 cap-binding domain (Fig. 3C). In the model of the FluB cap-binding domain, 2 amino acids, Gln325 and Trp359, are different from Phe323 and His357 of the FluA cap-binding domain, respectively.

To determine key amino acids related to the cap recognition specificity, the transcription activity was measured using a minireplicon assay system. In this assay system, we have used a transient-transfection system with a viral genome, in which the coding region for a viral gene is replaced with a luciferase reporter gene while *cis*-acting regulatory regions (24) remain intact (35). The cellular RNA polymerase I produces a negative-sense luciferase RNA sandwiched with viral terminal sequences. Luciferase mRNA is synthesized by transcription of the negative-sense RNA with the viral RNA polymerase and NP and subjected to translation. This system has been used to measure the transcription activity of the Flu polymerase (20, 22).

In the case of FluA PB2, His357, with which methylated guanine is stacked, could be replaced by other aromatic residues such as Trp and Phe, while Phe404, which is also involved in stacking methylated guanine, could not be (Fig. 4A). Leu could not substitute for either His357 and Phe404. On the other hand, in the case of FluB PB2, Trp359 could be replaced with other aromatic residues (but with less efficiency than for the FluA polymerase), but Phe406 could be replaced with hydrophobic residues such as Tyr and Leu (Fig. 4B). To confirm the importance of the hydrogen bonds with methylated guanine, Glu361 and Lys376 in FluA PB2 and Glu363 and Lys378 in FluB PB2 were replaced with alanine (Ala). Ala substitutions in FluA PB2 abolished the transcription activity,

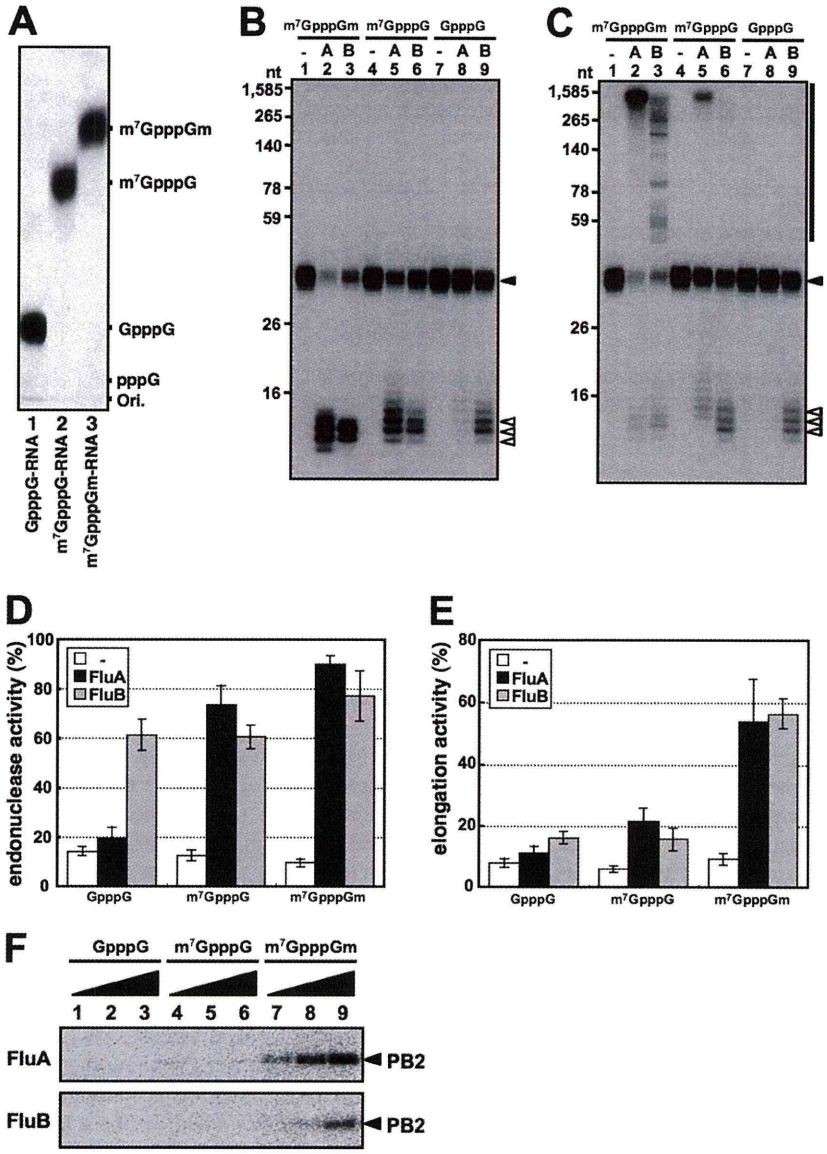

FIG. 2. Specificity of recognition of cap structures by Flu polymerases. (A) Analysis of 5′-terminal cap structures of RNAs. T7 RNA polymerase-synthesized RNAs were treated with nuclease P_1 and analyzed by TLC (PEI-CEL, 0.65 M LiCl), and radioactive nucleotides were detected by autoradiography. (B and C) *In vitro* capped RNA cleavage (B) and RNA elongation (C) reactions were performed with 600 ng of FluA (lanes 2, 5, and 8) or FluB (lanes 3, 6, and 9) vRNP using 2 fmol of variously methylated capped RNAs (m^7GpppGm-RNA, lanes 1 to 3; m^7GpppG-RNA, lanes 4 to 6; GpppG-RNA, lanes 7 to 9). RNA products were analyzed by 15% PAGE containing 8 M urea. The input capped RNAs (33 nt), the cleaved capped RNA products, and the elongated products are indicated as a closed triangle, open triangles, and a black bar, respectively, at the right. (D and E) Ratios of cleaved RNA products (D) and RNA transcripts (E) to total input primer RNAs. (F) Cap-binding activity for variously methylated capped RNAs. UV cross-linking was performed using 50, 100, and 200 ng of FluA (upper panel) and FluB (lower panel) vRNP and 50 fmol of variously methylated capped RNAs (GpppG-RNA, lanes 1 to 3; m^7GpppG-RNA, lanes 4 to 6; m^7GpppGm-RNA, lanes 7 to 9).

while Ala substitution for Lys378 of FluB PB2 caused only a small decrease in the transcription activity (Fig. 4C and D). These results suggest that the stacking interaction of His357 and Phe404 and the hydrogen bonds of Glu361 and Lys376 with methylated guanine are essential for cap recognition by the FluA polymerase. This is in good agreement with a previous report (12). In contrast, it is suggested that the stacking interaction of Trp359 and the hydrogen bonds of Glu363 with methylated guanine are sufficient for cap recognition by the FluB polymerase. These results indicate that the mechanism for recognition of methylated guanine by the FluB polymerase could be different from that for the FluA polymerase. It is also speculated that the cap-binding pocket of the FluB polymerase may be more flexible or less stringent than that of the FluA polymerase in recognition of various cap structures, since Phe406 of FluB PB2 is changeable with other amino acids.

Phe323 in FluA PB2 stacks on the ribose of m^7GTP and was essential for cap recognition (see Fig. S3 in the supplemental

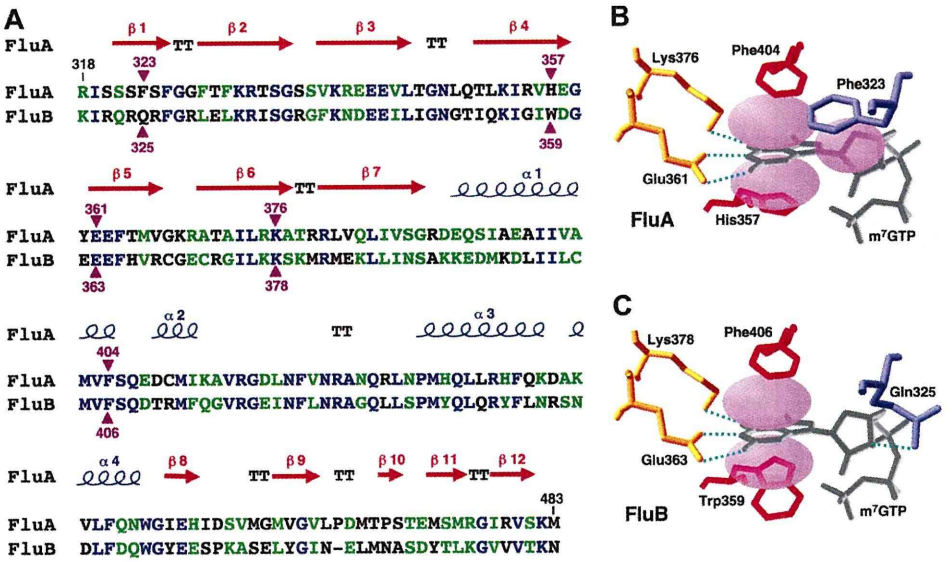

FIG. 3. Structure of the PB2 cap-binding domain. (A) Sequence alignment of the PB2 cap-binding domains of FluA (A/Panama/2007/99) and FluB (B/Shanghai/361/2002). The secondary structure of FluA is displayed over the sequence alignment. Blue letters and green letters show identical residues and similar residues, respectively. Purple triangles indicate the residues in contact with the cap analogue m^7GTP. (B and C) Model of m^7GTP interaction with the cap-binding domains of FluA PB2 (B) (10) and FluB PB2 (C) drawn by computer-associated calculation, with putative hydrogen bonds shown as green dotted lines.

material) (12). However, it is likely that Gln325 in FluB PB2, which is located in the same position of Phe323 in FluA PB2, makes a hydrogen bond with the ribose of m^7GTP. We speculated that FluB PB2 recognizes the cap structure in a flexible pocket as discussed above, so that the hydrogen bonds made by Gln325 and Glu363 could be more crucial for cap binding than those in FluA PB2. In addition, there could be an appropriate amino acid in the amino acid combination between amino acid positions 325 and 363 in FluB PB2 in order to keep the flexible pocket. To confirm this prediction, the transcription activities of mutants with substitutions at position 325 were examined in the presence of the Asp363 mutant (Fig. 4E). The transcription activity of the Asp363 single mutant was reduced to 20% of the wild-type level, possibly because of a longer distance between Asp363 and guanine residues for hydrogen bonds (Fig. 4E; see Fig. S4B in the supplemental material). Interestingly, Lys and Arg mutations but not Ala and Asn mutations at position 325 could rescue the transcription activity of Asp363 (Fig. 4E). We also examined the effect of an Asp363 single mutation and an Arg325-Asp363 double mutation on the transcription and replication processes and the cap-binding activity (Fig. 5). According to the levels of accumulation of mRNA (Fig. 5A) and cRNA (Fig. 5B), the level of reporter expression (Fig. 4E) is well correlated with the transcription but not the replication activities. To examine the cap-binding activity *in vivo*, capped RNAs that could interact with the viral RNA polymerase were coprecipitated from cells expressing the recombinant RNA polymerase, and the cap structure was detected by recapping of RNA which had been CIAP treated and then decapped (β-eliminated) (Fig. 5C). We could detect the [^{32}P]m^7Gp labeled by [α-^{32}P]GTP and vaccinia virus capping enzyme, depending on TAP digestion. In contrast, uncapped RNA treated with CIAP was poorly labeled by this protocol. These results indicate that this recapping method is suitable for the detection of capped RNA specifically. Using this method, we found that the cap-binding activities of these mutants (Fig. 5D) are well correlated with these transcription activities (Fig. 4E) and the mRNA accumulation levels (Fig. 5A). These results indicate that the Arg at position 325 in FluB PB2 supports cap recognition when Glu363 is replaced with Asp363.

DISCUSSION

Most of our knowledge on the transcription mechanism of the influenza virus genome has been derived from studies on FluA, while little has been demonstrated for FluB. This is also the case for studies on the enzymatic aspects of these viral RNA polymerases. Each of the two methyl groups in the cap1 structure, the 7-methyl residue of the guanine base and the 2′-*O*-methyl residue on the ribose of the penultimate base, strongly influences the transcription activity of the FluA polymerase (4). Recently, the structure of the PB2 cap-binding domain of the FluA polymerase with m^7GTP has been clarified (12). Based on these reports, we tried to identify the specificity of cap recognition and characterize key amino acids for cap recognition of the FluB polymerase.

First, we compared the efficiencies of capped RNA cleavage and subsequent RNA elongation reactions of the FluA polymerase with those of the FluB polymerase using cap1-RNA. As expected, the FluA polymerase exhibited efficient endonuclease activity, elongation activity, and cap-binding affinity. The pattern of cleavage of cap1-RNA by the FluB polymerase was different from that by the FluA polymerase (Fig. 1A; see Fig. S1B in the supplemental material), and the RNA elongation and cap-binding activities of the FluB polymerase were lower than those of the FluA polymerase (Fig. 1A and B). These results indicate that the cap binding and cleavage mechanism

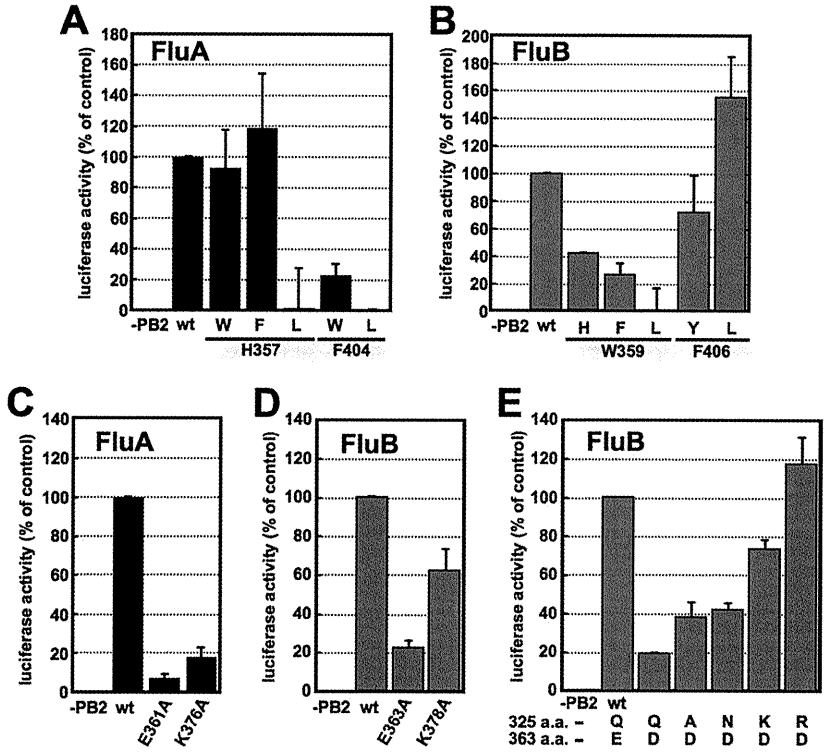

FIG. 4. Transcription activities of PB2 mutants in a minireplicon system. (A and B) Effects of mutations of m^7GTP stacking residues in FluA (A) and FluB (B) PB2 on transcription activity. (C and D) Effects of mutations in residues involved in hydrogen bonds with the guanine residue of m^7GTP in FluA (C) and FluB (D) PB2 on transcription activity. (E) Effect of mutations in Gln325 with an Asp mutation at position 363 in FluB PB2 on transcription activity. The firefly luciferase activity was normalized to *Renilla* luciferase activity. The results are averages and standard deviations (SD) from four independent experiments.

of the FluB polymerase are different from those of the FluA polymerase.

We then examined the specificity of recognition of cap structures by the FluB polymerase compared with that by the FluA polymerase. Using various methylated capped RNAs, we performed capped RNA cleavage and RNA elongation assays (Fig. 2). The FluA polymerase cleaved RNA containing m^7G specifically, while the FluB polymerase could cleave GpppG-RNA as well as RNA containing m^7G. Both the FluA and FluB polymerases elongated and bound to the cap structure efficiently only in the case of $m^7GpppGm$-RNA compared with other capped RNAs (Fig. 2C, 2E, and 2F). Based on these results, we propose that the FluA polymerase recognizes strictly the guanine-7-methyl residue in the cleavage reaction and that the FluB polymerase recognizes only the cap core structure (GpppX), which may result in its weak cap1-binding activity. In addition, these results suggest that the ribose 2′-*O*-methyl residue and/or the guanine-7-methyl residue may be responsible for the elongation reaction by both FluA and FluB polymerases, because cap binding and efficient elongation could not be observed except for $m^7GpppGm$-RNA.

To elucidate the mechanism of cap recognition by the FluB polymerase, we studied the PB2 subunit, which has the cap-binding domain. Recently, the 3D structure of the FluA PB2 cap-binding domain was revealed (12). Amino acid residues essential for cap binding were identified and found to be conserved between FluA and FluB polymerases (Fig. 3A). In the FluA PB2 cap-binding domain (Fig. 3B), the methylated guanine base is sandwiched with His357 and Phe404, and Phe323 stacks on the ribose of m^7GTP. Glu361 makes hydrogen bonds with the N1 and N2 positions of guanine, and Lys376 also makes hydrogen bonds with the O6 position of guanine. Based on the structure of the FluA PB2 cap-binding domain, a model of the FluB PB2 cap-binding domain was postulated (Fig. 3C). Five amino acids which contact the guanine-7-methyl residue are highlighted. Minireplicon assays showed that Trp359 in FluB PB2 is crucial for possible stacking interaction with a methylated guanine base without sandwiching with Phe406 (Fig. 4B). Moreover, the hydrogen bond made by Lys378 to the O6 position of guanine seemed not to be essential for cap recognition (Fig. 4D). These results suggest that the FluB polymerase recognizes the cap structure in a manner different from the FluA polymerase. We illustrated a new proposed computer-associated model for cap recognition by FluB PB2 (see Fig. S4A in the supplemental material), although the 3D structure of the FluB PB2 cap-binding domain has not been determined. The overall structures of four cap-binding proteins, FluA PB2 (12), eIF4E (33, 34), CBP20 (23), and VP39 (16), differ each other widely due to their evolutionarily unrelated origins, but the cap-binding pockets are essentially quite similar (see Fig. S5 in the supplemental material), although there are some differences in details. In addition to the two aromatic amino acids, an acidic residue is directed toward the pocket to accommodate the positively charged π-ring system of