Published in final edited form as:

Biochem Biophys Res Commun. 2008 May 9; 369(3): 818–823. doi:10.1016/j.bbrc.2008.02.097.

Molluscan mobile elements similar to the vertebrate recombination-activating genes

Yuri Panchin¹ and Leonid L. Moroz^{2,3,*}

¹ Institute of Problems of Information Transmission, Russian Academy of Science, 127994, Moscow, Russia

² The Whitney Laboratory for Marine Bioscience, University of Florida, FL 32080, USA

³ Evelyn F. & William McKnight Brain Institute of the University of Florida and Department of Neuroscience, FL 32611, USA

Abstract

Animal genomes contain ~20,000 genes. Additionally millions of genes for antigen receptors are generated in cells of the immune system from the sets of separate gene segments by a mechanism known as the V(D)J somatic recombination. The components of the V(D)J recombination system, Recombination-Activating Gene proteins (RAG1 and RAG2) and recombination signal sequence (RSS), are thought to have “entered” the vertebrate genome as a hypothetical “RAG transposon”. Recently discovered mobile elements have terminal inverted repeats (TIRs) similar to RSS and may encode proteins with a different degree of similarity to RAG1. We describe a novel *N-RAG-TP* transposon identified from the sea slug *Aplysia californica* that encodes a protein similar to the N-terminal part of RAG1 in vertebrates. This refines the “RAG transposon” hypothesis and allows us to propose a scenario for V(D)J recombination machinery evolution from a relic transposon related to the existing mobile elements *N-RAG-TP*, *Chapaev* and *Transib*.

Keywords

Aplysia; Mollusca; Recombination-activating gene; RAG-1; Evolution; Transposase; V(D)J recombination; Immune system; Transposon; Somatic recombination

Although only vertebrates possess the mechanism of V(D)J somatic recombination [1], another type of DNA rearrangement called transposition, by which sequences of DNA (transposons) can move around to different positions within the genome of a single cell, is common for different organisms.

Even before RAG proteins were identified and sequenced, the similarity of these two mechanisms of DNA rearrangements had been recognized and the “RAG transposon” hypothesis was proposed [1–6]. It was suggested that the components of the V(D)J recombination system, Recombination-Activating Gene proteins (RAG1 and RAG2) and recombination signal sequence (RSS), have “entered” the vertebrate genome by horizontal transfer as components of a transposable element (a hypothetical “RAG transposon”, Fig. 1)

*Corresponding author. Fax: +1 904 461 4052; Email address: E-mail: moroz@whitney.ufl.edu (L. L. Moroz).

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.

[1–6]. As a result, it contributed to the enormous adaptive radiation of jawed vertebrates (also dubbed the “Immunological Big Bang” [2,7]). However, due to the lack of comparative data evolutionary events leading to the origin of RAG genes (and adaptive immunity in general) have been obscure.

Here, in the opisthobranch mollusc, *Aplysia californica* we have found transcripts representing the N-terminal part of a vertebrate RAG1 gene. Cloning and sequencing of the related full-length genomic DNA and cDNA fragments revealed that in *Aplysia* they belong to new mobile elements that we named as N-RAG-transposons (*N-RAG-TP*). *NRAG-TP* is distinct from *Transib*, the transposon that was recently identified in several invertebrate species [8] and encodes a putative transposase homologous to the C-terminal part of RAG1 protein. It is also related to *Chapaev* transposon family that were currently independently described by Kapitonov V.V. and Jurka J. bioinformatical studies [9].

Materials and methods

Standard molecular biology procedures were used for cDNA and genomic DNA cloning and sequencing as described by Moroz *et al.* [10]. RACE procedure is also described elsewhere [11]. For microarray experiments two custom 44,000 oligonucleotide *Aplysia* arrays were constructed in collaboration with Agilent Technologies using 60-mer oligonucleotide sequences designed from each nonredundant sequence in the *Aplysia* EST database [10]. The data have been deposited in NCBI’s Gene Expression Omnibus (GEO) and are accessible through the GEO Series accession number GSE4628 [10]. Some experimental modifications are discussed in the text and supplements 1–7.

Results and discussion

The sea slug *Aplysia californica* is a well-established model in neuroscience and, recently, genomics [10] providing unique insights into molecular and cellular mechanisms of intercellular communications, long-term plasticity and behavior [12]. Our genome-wide screening of *Aplysia* transcriptome and molecular experiments led to discovery of a novel class of RAG1-like transcripts.

N-RAG-TP encodes a protein similar to the N-terminal region of the human RAG1 protein

N-RAG-TP is expressed in adult *Aplysia* tissues as polyadenylated non-spliced mRNA (See Fig. 2 and supplements 1 and 6). The absence of introns is also characteristic for ORFs encoding transposases in other mobile elements and to RAG1 and RAG2 genes.

N-RAG-TP consensus sequence and most of the directly cloned sequences’ (NRAG1 to NRAG6, **EF042298**, **EF042299**, **EF042300**, **EF042301**, **EF042302**, and **EF042303**) conceptual translation revealed ORFs of about 250 aa of putative N-RAG protein and one sequence revealed an ORF 523 aa long (NRAG8, **EF042305**). These *Aplysia* proteins have similarity to vertebrate RAG1 protein (Fig. 3B) and to the proteins of unknown function from several invertebrate species such as *Hydra*, sponges and sea urchins (see alignment in supplement 4).

RAG1 proteins are the closest homologs for *Aplysia* N-RAG proteins in the NCBI database. A PSI-BLAST search of the NR database for homology matches to the new mollusc *N-RAG-TP* sequences revealed similarity to vertebrate RAG1, proteins with *E*-value 2e-15 to 2e-10 at the first iteration and with *E* values of 9e-153 to 1e-113 in the second iteration. (*E*-value inclusion threshold 0.005). It is important to note that all the top 1000 hits in this blast search represented RAG1 proteins with other protein scores substantially lower. (For instance human

RAG1 protein shows BLASTP *E*-value $3\text{e-}10$ while the next closest match to human protein shows *E*-value of only 0.6).

In a reciprocal BLASTP search initiated by the human homolog, the *Aplysia* N-RAG proteins were reliably detected with *E* values of $8\text{e-}14$ to $1\text{e-}10$.

The regions of similarity revealed by PSI-BLAST (second iteration, *E*-value inclusion threshold 0.005) between RAG1 and N-RAG8 protein spans from 87 to 408 amino acids in the human sequence and from 12 to 325 amino acids in the *Aplysia* sequence (Fig. 3A, B). Several conserved domains are found in this area (Fig. 3A). Once the primary amino acid sequence of RAG1 was known, it was found that a sequence within the N-terminal third of the protein was homologous to a zinc-binding motif referred to as the C3HC4 or the zinc RING finger domain [13]. RING finger domain in human RAG1 is recognized by SMART program (<http://smart.embl-heidelberg.de/>) with *E*-value $2.82\text{e-}04$. It entirely complies with RING motif (Fig. 3C). In *Aplysia* RAG1 homolog long ORF (NRAG8) equally well corresponds to RING motif (SMART *E*-value $9.17\text{e-}04$) (Fig. 3C). The shorter ORF of NRAGTP1-6 has distorted RING domain sequence although it is still recognized by SMART program with *E*-value $6.69\text{e-}02$. Its sequence lacks the first two conserved cysteins and it is not clear if they are substituted by other neighboring cysteine or histidine residues.

RING domains are rather common in different proteins so it is possible that the similarity of *Aplysia* protein to RAG-1 is based only on this sequence. However, in the case of N-RAG sequences, because of the presence of a relatively well-conserved region N-terminal to the RING domain sequence, there is sufficient similarity outside the RING domain region to indicate homology: a PSI-BLAST search seeded by the NRAG8 fragment containing the first 194 amino acid residues without RING domain sequence with the *E*-value inclusion threshold of 0.005 revealed similarity to vertebrate RAG-1 proteins with *E*-values of $6\text{e-}09$ at the first iteration and with *E* values of $5\text{e-}72$ in the second iteration. PSI-BLAST search seeded by the NRAG1- 6 fragments containing the first 180 amino acid residues with deleted RING domain revealed similarity to vertebrate RAG-1 with *E*-values of $2\text{e-}06$ at the first iteration and with *E* values of $4\text{e-}65$ in the second iteration.

In humans and other vertebrates RAG1 also contains a C2H2 zinc finger sequence, referred to as ZFA, that is adjacent to the C3HC4 motif within the primary sequence [14,15]. Short CDS of N-RAG1-6 terminates shortly after the putative C3HC4 motif region and does not include the ZFA region. Longer CDS of NRAG8 extends its similarity to this region, although the last histidine in the zinc finger sequence is substituted by cysteine, as frequently happens with C2H2 sequences (Fig. 3A, B).

Rodgers et al [14] determined the part of RAG1 corresponding to human protein aa 269–389 as a zinc-binding dimerization domain of RAG1. The end of this domain delaminates the border between the N-terminal RAG1 region and the C-terminal “core domain” of RAG1. Only these carboxyl-terminal two-thirds of RAG1 are considered to be necessary for recombination [16, 17]. This region, referred to as core RAG1, consists of amino acid residues 384 to 1008. SSEARCH3, BLAST and FASTA programs detect no direct similarity of N-RAG8 protein to core RAG1 region, yet comparison of different RAG1 related sequences as *Transib* and *Chapaev* imply that they may have a common origin.

Genes encoding proteins similar to the N-terminal RAG1 portion were previously described in *Branchiostoma*, *Hydra*, and *Nematostella* [8]. Sea urchins appear to have a longer gene including a C-terminal part “core RAG” but no RING domain [4,18]. We also found and reconstructed similar sequences from the sponge *Reniera* WGS project (supplement 4). Initially it was suggested that in contrast to *Aplysia* N-RAG-TP, sequences from the other animals encoding homologs to the RAG1 N-terminal part have no features of mobile elements

[8], yet new data from the same authors suggest that these sequences may also derive from other mobile elements [9]. The regions of similarity between RAG1 protein sequence and protein sequences encoded by different mobile elements are presented in Fig. 3A.

Our finding that *Aplysia* mobile element encodes just one protein that is similar to the N-terminal region of the human RAG1 is surprising because the human RAG1 N-terminal region was shown to be dispensable for recombination [16,17,19]. This part of RAG1 has no similarity to known transposases or integrases.

It is possible that the *N-RAG-TP* is a non-autonomous mobile element and its transposition is mediated by a protein encoded by some other gene outside of the *N-RAG-TP* sequence. It is also possible that only the form of *N-RAG-TP* with a longer ORF (*N-RAG-TP8*) represents the autonomous mobile element. Although even longer *N-RAG-TP* isoforms have no direct similarity to known transposases or integrases or to the core RAG-1 domain, they contain a number of carboxylate residues that are believed to be responsible for coordinating metal ions needed for catalysis in endonucleases of the DDE superfamily.

N-RAG-TP sequences show hallmarks of similarly structured transposable elements

The *N-RAG-TP* family is composed of multiple copies (estimated >50; see supplement 2) dispersed in the *Aplysia* genome and characterized by >90% sequence identity with a 3425 bp long consensus sequence. In some cases it was found that *N-RAG-TP* sequences are inserted into the introns of protein coding genes (see supplement 1). *N-RAG-TP* sequences show hallmarks of similarly structured transposable elements: terminal inverted repeats (TIRs) and target-site duplications (TSDs) flanking each individual copy (Fig. 2). TSDs are usually formed by transposon insertion because transposase makes a staggered cut at the target site and the gaps are filled by DNA polymerase with identical sequence on both flanks of the inserted transposon. In all studied cases where the bordering regions were identified on both sides of the transposon (n=12, see supplement 3), *N-RAG-TP* copies inserted in the *Aplysia* genome were flanked by direct repeats 4 bp long. Human RAG proteins are also able to catalyze a transposition reaction and produce target-site duplication immediately flanking the recombination signals [20]. They are normally 5 bp long although some 4 bp and 3 bp TSDs were also found. Insertion sites of *Transib* transposon similar to the C-terminal portion of RAG1 protein are flanked by 5 bp TDSs [8].

Similarly to other DNA-only transposons, the putative *Aplysia N-RAG-TP* transposons contain TIRs at their ends. Transposon TIRs serve as recognition signals for a transposase enzyme associated with the element. The signals for V(D)J recombination are also formed by inverted repeats and are called RSS.

As shown in Fig. 2b *N-RAG-TP* and *Transib* TIRs and recombination signal sequences from vertebrate antigen receptor gene segments have sequence similarity. They all share the first 4 identical nucleotides (CACA) at their ends. It is important to mention that the first three nucleotides of RSS (CAC) were shown to be the most crucial for V(D)J recombination, whereas mutations at the other RSS positions still allow recombination [21].

Evidence for ribosomal frame shifting in N-RAG-TP translation

The ORF sequences from different *N-RAG-TP* copies in the *Aplysia* genome are very conservative with most sequences encoding consensus N-RAG protein without interrupting stop codons or frame shifts. The only exception is a consistent frame shift in some *N-RAG-TP* copies. Sequences with and without a frame shift differ in the number of adenosine nucleotides in a row of 7 or 8 “A” residues at position 1058. Four out of eight non-redundant *N-RAG-TP* clones with thoroughly sequenced and verified ORFs have this frameshift. We

hypothesize that ribosomal frame shifting [22,23] takes place during translation of this protein because “slippery” sites, where the ribosome changes reading frames, and stimulatory stem-loop RNA structures downstream are identified in their sequences (see supplement 5).

“RAG transposon” hypothesis

Even before RAG proteins were identified and sequenced, the similarity of V(D)J somatic recombination and transposition had been recognized and the “RAG transposon” hypothesis was proposed [1–6]. It was suggested that the components of the V(D)J recombination system, Recombination-Activating Gene proteins (RAG1 and RAG2) and recombination signal sequence (RSS), have “entered” the vertebrate genome by horizontal transfer as components of a transposable element (a hypothetical “RAG transposon”, Fig. 1) [1–6]. For a long time no direct evidence in favor of this hypothesis was available until *N-RAG-TP* (this paper) and *Transib* [8] transposons were discovered.

While this paper was in preparation a new mobile element with resemblance to RAG1 named *Chapaev* was discovered [9] providing more information on relations of V(D)J somatic recombination to mobile elements. We have included one *Chapaev* transposase sequence (*Chapaev_2ACa*) in Fig 3A to demonstrate that it has some similarity to different parts of RAG1 protein. It is important to note that different *Chapaev* proteins display different degrees of such similarity.

We may assume that the relic “RAG transposon” was a complex mobile element equipped with TIR/RSS-like signal sequences, with ancestors of coding sequences corresponding to both the N- and C- portions of RAG1 protein and with the RAG2 prototype. This, now extinct, mobile element was acquired from the common ancestor or has been horizontally transferred to different taxonomic groups. Eventually multiple gene loss events gave rise to discrete sets of “RAG transposon” derivatives in different animal lineages or individual species. These derivatives include two types of transposons (*N-RAG-TP* and *Transib*), non-mobile N-RAG related genes of unknown function and the components of the somatic recombination system in vertebrate’s adaptive immune system- RAG1, RAG2 and RSS. It is assumed that in jawed vertebrates the RAG transposon was split so that its coding sequence lost TIRS and became RAG1 and RAG2 genes whereas TIR sequences became RSS in the segments of antigen receptor genes (Fig. 1).

As an alternative it could be that V(D)J recombination machinery emerged not from one complex transposon but from the simpler transposons that still exist as separate mobile elements in some invertebrates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by grants from the National Institutes of Health (P50 HG002806), the McKnight Brain Research Foundation, and the University of Florida Opportunity Funds (LLM) as well as Russian Foundation for Basic Research (grant: 08-04-00478 to Y.P.). We would like to thank James Netherton for his comments on earlier versions of the manuscript

References

1. Sakano H, Huppi K, Heinrich G, Tonegawa S. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature* 1979;280:288–294. [PubMed: 111144]

2. Schatz DG. Antigen receptor genes and the evolution of a recombinase. *Semin Immunol* 2004;16:245–256. [PubMed: 15522623]
3. Fugmann SD, Lee AI, Shockett PE, Villey IJ, Schatz DG. The RAG proteins and V(D)J recombination: complexes, ends, and transposition. *Annu Rev Immunol* 2000;18:495–527. [PubMed: 10837067]
4. Fugmann SD, Messier C, Novack LA, Cameron RA, Rast JP. An ancient evolutionary origin of the Rag1/2 gene locus. *Proc Natl Acad Sci U S A* 2006;103:3728–3733. [PubMed: 16505374]
5. Thompson CB. New insights into V(D)J recombination and its role in the evolution of the immune system. *Immunity* 1995;3:531–539. [PubMed: 7584143]
6. Oettinger MA, Schatz DG, Gorka C, Baltimore D. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 1990;248:1517–1523. [PubMed: 2360047]
7. Bernstein RM, Schluter SF, Bernstein H, Marchalonis JJ. Primordial emergence of the recombination activating gene 1 (RAG1): sequence of the complete shark gene indicates homology to microbial integrases. *Proc Natl Acad Sci U S A* 1996;93:9454–9459. [PubMed: 8790351]
8. Kapitonov VV, Jurka J. RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. *PLoS Biol* 2005;3:e181. [PubMed: 15898832]
9. Kapitonov VV, Jurka J. Chapaev - a novel superfamily of DNA transposons. *Rebase Reports* 2007;7:777–777.
10. Moroz LL, Edwards JR, Puthanveetil SV, Kohn AB, Ha T, Heyland A, Knudsen B, Sahni A, Yu F, Liu L, Jezzini S, Lovell P, Iannuccilli W, Chen M, Nguyen T, Sheng H, Shaw R, Kalachikov S, Panchin YV, Farmerie W, Russo JJ, Ju J, Kandel ER. Neuronal transcriptome of aplysia: neuronal compartments and circuitry. *Cell* 2006;127:1453–1467. [PubMed: 17190607]
11. Chenchik A, Diachenko L, Moqadam F, Tarabykin V, Lukyanov S, Siebert PD. Full-length cDNA cloning and determination of mRNA 5' and 3' ends by amplification of adaptor-ligated cDNA. *Biotechniques* 1996;21:526–534. [PubMed: 8879595]
12. Kandel ER. The molecular biology of memory storage: a dialogue between genes and synapses. *Science* 2001;294:1030–1038. [PubMed: 11691980]
13. Schatz DG, Oettinger MA, Baltimore D. The V(D)J recombination activating gene, RAG-1. *Cell* 1989;59:1035–1048. [PubMed: 2598259]
14. Rodgers KK, Bu Z, Fleming KG, Schatz DG, Engelman DM, Coleman JE. A zinc-binding domain involved in the dimerization of RAG1. *J Molecular Biology* 1996;260:70–84.
15. De P, Rodgers KK. Putting the pieces together: identification and characterization of structural domains in the V(D)J recombination protein RAG1. *Immunological Review* 2004;200
16. Sadofsky MJ, Hesse JE, McBlane JF, Gellert M. Expression and V(D)J recombination activity of mutated RAG-1 proteins. *Nucleic Acids Res* 1993;21:5644–5650. [PubMed: 8284210]
17. Silver DP, Spanopoulou E, Mulligan RC, Baltimore D. Dispensable sequence motifs in the RAG-1 and RAG-2 genes for plasmid V(D)J recombination. *Proc Natl Acad Sci U S A* 1993;90:6100–6104. [PubMed: 8327489]
18. Hibino T, Loza-Coll M, Messier C, Majeske AJ, Cohen AH, Terwilliger DP, Buckley KM, Brockton V, Nair SV, Berney K, Fugmann SD, Anderson MK, Pancer Z, Cameron RA, Smith LC, Rast JP. The immune gene repertoire encoded in the purple sea urchin genome. *Dev Biol* 2006;300:349–365. [PubMed: 17027739]
19. van Gent DC, McBlane JF, Ramsden DA, Sadofsky MJ, Hesse JE, Gellert M. Initiation of V(D)J recombination in a cell-free system. *Cell* 1995;81:925–934. [PubMed: 7781069]
20. Agrawal A, Eastman QM, Schatz DG. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* 1998;394:744–751. [PubMed: 9723614]
21. Hesse JE, Lieber MR, Gellert M, Mizuuchi K. Extrachromosomal DNA substrates in pre-B cells undergo inversion or deletion at immunoglobulin V-(D)-J joining signals. *Cell* 1987;49:775–783. [PubMed: 3495343]
22. Hammell AB, Taylor RC, Peltz SW, Dinman JD. Identification of putative programmed –1 ribosomal frameshift signals in large DNA databases. *Genome Res* 1999;9:417–427. [PubMed: 10330121]
23. Moon S, Byun Y, Kim HJ, Jeong S, Han K. Predicting genes expressed via –1 and +1 frameshifts. *Nucleic Acids Res* 2004;32:4884–4892. [PubMed: 15371551]

Fig. 1.

Schematic representations of V(D)J somatic recombination. (A) and transposition (B) illustrate that these two processes are similar. (A), V(D)J somatic recombination serves to generate numerous antigen receptor genes from sets of separate gene segments in cells of the immune system. Recombination activating genes-1 and -2 (*RAG-1* and *RAG-2*) were identified based on their ability to synergistically confer this recombination. Vertebrate RAG1 protein consists of an N-terminal region (NRAG), a RING domain and a C-terminal region (CRAG). In the course of V(D)J somatic recombination RSS sequences serve as signals for the RAG recombinase complex. Recombinase thus cuts out the spacers flanked by RSS and joins the segments (V and J on this figure A) of the antigen receptor gene. This recombination reaction generates diverse molecules for antigen receptors in the adaptive immune system. The fragment excised by the recombinase complex is normally discarded, although in special conditions it is inserted into a new position within the genome, revealing the transposase activity of RAG proteins. (B), Transposons, or mobile genetic elements, are sequences of DNA that can move around to different positions within the genome of a single cell. DNA transposons move directly from one position to another using a transposase enzyme to “cut and paste” them within the genome. Specific sequences at the ends of a transposon, terminal inverted repeats (TIRs), serve as recognition signals for the transposase encoded by that element. For this reason, RAG protein function (A) is akin to that of transposase whereas RSS serves the same role as transposon TIRs. Specific transposase enzyme coding sequence is usually contained inside a DNA transposon and jumps as its part within the genome. (C), Hypothetical “RAG transposon”. In the case of V(D)J recombination enzyme, coding sequences of RAG1 and RAG2 have fixed positions in the genome and proteins that they encode control the excision and recombination of RSS containing sequences in remote DNA locations. Consequently the key elements of V (D)J somatic recombination RAG proteins and RSS (A) were suggested to derive from a hypothetical “RAG transposon” (C). It was assumed that both RAG1 and RAG2 genes might have evolved from transposase proteins that catalyzed transpositions of ancient transposons flanked by TIRs that were precursors of RSSs. Such a transposon can also be an ancestor to simplified *N-RAG-TP* and *Transib* transposons that have lost the sequences encoding RAG2 protein and N- or C- parts of RAG1 protein correspondingly. Alternatively, it could be proposed that vertebrate RAG1 proteins have emerged by fusion of *N-RAG-TP* and *Transib* transposons while the RAG2 gene was acquired independently. See text for details.

Fig. 2.

The organization of the *Aplysia N-RAG-TP* gene. (A). Mobile element *N-RAG-TP* (red) inserted into genomic DNA (in this example into the intron of a protein coding gene [see Supplement Figure 1 for the actual sequence]) is flanked by terminal inverted repeats (TIR) and target site duplications (TSD) characteristic for many types of mobile elements (yellow). Transcribed mRNA is shown in black. According to RACE experiments (see Supplements 1 and 6) the *N-RAG-TP* transcription start is located at position 128 of the *N-RAG-TP* consensus. The TGA stop codon is located at position 1535 of the consensus sequence, 3 nucleotides from an AATAAA polyadenylation signal followed by several more AATAAA signals further downstream. This is in agreement with 3' RACE products of several different lengths observed in our experiment. (B). *N-RAG-TP* TIRs consist of two tridecamer sequences separated by 39 bp on the 5' end and by 13 bp on the 3' end. *N-RAG-TP* TIRs, TIRs of TRANSIB1 and TRANSIB5 and recombination signal sequences from vertebrate antigen receptors segments RSS have some sequence similarity (see text).

Fig. 3.

Proteins encoded by different mobile elements (*Chapaev*, N-RAG and *Transib*) show some degree of similarity to RAG1 protein in primary amino acid sequence and domain composition. (A), Schematic, yet proportional representation of RAG1 protein and transposon encoded proteins. Areas in transposon encoded proteins aligned by PSIBLAST (second iteration) to human RAG1 are painted in black color. RING and zinc finger sequences are shown in red. Carboxylate residues needed for catalysis in endonucleases of the DDE superfamily are indicated (proteins Aca_N-RAG1 and Aca_N-RAG8, N-RAG-TP are from *Aplysia*). (C), RING domain motif and corresponding sequences from human RAG1 and N-RAG1 and N-RAG8 *Aplysia* proteins found by SMART program.