Role of the silkworm argonaute2 homolog gene in double-strand break repair of extrachromosomal DNA

Haruna Tsukioka, Masateru Takahashi, Hiroaki Mon, Kazuhiro Okano1, Kazuei Mita2, Toru Shimada3, Jae Man Lee, Yutaka Kawaguchi, Katsumi Koga and Takahiro Kusakabe*

Laboratory of Silkworm Sciences, Kyushu University Graduate School of Bioresource and Bioenvironmental Sciences, 6-10-1 Hakozaki, Fukuoka 812-8581, Japan, 1Laboratory of Molecular Entomology and Baculovirology, The Institute of Physical and Chemical Research (RIKEN), Hirosawa 2-1, Wako, Saitama 351-0198, Japan, 2Laboratory of Insect Genome, National Institute of Agrobiological Sciences, Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan and 3Department of Agricultural and Environmental Biology, University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-8657, Japan

Received as resubmission January 25, 2006; Accepted January 27, 2006

ABSTRACT

The argonaute protein family provides central components for RNA interference (RNAi) and related phenomena in a wide variety of organisms. Here, we isolated, from a Bombyx mori cell, a cDNA clone named BmAGO2, which is homologous to Drosophila ARGONAUTE2, the gene encoding a repressive factor for the recombination repair of extrachromosomal double-strand breaks (DSBs). RNAi-mediated silencing of the BmAGO2 sequence markedly increased homologous recombination (HR) repair of DSBs in episomal DNA, but had no effect on that in chromosomes. Moreover, we found that RNAi for BmAGO2 enhanced the integration of linearized DNA into a silkworm chromosome via HR. These results suggested that BmAgo2 protein plays an indispensable role in the repression of extrachromosomal DSB repair.

INTRODUCTION

DNA double-strand breaks (DSBs), caused by a variety of endogenous and exogenous DNA-damaging agents, are one of the most serious forms of DNA damage that can occur in a cell’s genome, so that the preservation of genomic integrity in all organisms is an indispensable process. DNA replication in cells having DSBs, or incorrect repair which follows, may result in chromosomal fragmentation, translocation or deletion, and in mammals may lead to diseases including cancer (1,2). There are several DSB repair pathways, homologous recombination (HR), non-homologous end-joining (NHEJ) and single-strand annealing (SSA), in all organisms from bacteria to mammals (3–5), and which of these provides the major pathway for DSB repair depends upon the cell state (6,7); HR frequently contributes in early development and the G2 phase of the cell cycle (2,6), while NHEJ primarily takes place in the G1 phase of the cell cycle (8).

The DSB repair process in extrachromosomes is said to include recombination between transfected DNA substrates, between a transfected substrate and chromosomal DNA or both. Extrachromosomal HR along close direct repeats is predominately non-conservative with a resulting loss of sequence between the direct repeats (9). This can be explained by the SSA model. Previous researchers using mammalian cells reported that extrachromosomal DSBs are repaired predominantly through the SSA pathway. When a DSB is situated within one of the two repeats, it can be repaired either by SSA or by gene conversion. Increasing the distance between the two repeats lowers the efficiency of SSA in competition with gene conversion, consistent with the need for more time for 5′–3′ resection to expose complementary homologies (10,11).

In the silkworm, Bombyx mori, DSBs in extrachromosomal DNA with two homologous direct repeats are repaired mainly...
by SSA (12), as is the case for mammalian cells. In the previous report (12), we used an intramolecular HR substrate consisting of truncated but overlapping luciferase genes that could be reconstructed into a functional luciferase by cellular HR activities. It seemed highly improbable that synthesis-dependent strand annealing happened on the substrate. The insertion of a long intervening sequence, over 1 kb, between the tandem repeat sequences of the former substrate promoted recombination repair by gene conversion, although DSB was still mainly repaired through SSA.

In this study, we tried to identify a gene involved in DSB repair using a convenient assay system that detects repair activities in cultured silkworm cells (12,13). After screening several hundred clones in a database of B. mori expressed sequence tag (EST) by RNA interference (RNAi), we focused on one clone which augmented the repair of extrachromosomal DSBs >2-fold when silenced by RNAi. Isolation of full-length cDNA and nucleotide sequence determination indicated that this clone had the signature domains of Argonaute proteins, PAZ and PIWI.

**MATERIALS AND METHODS**

**Construction of assay systems for DSB repair activities**

The substrates used in the assay of the DSB repair activities were as described previously (12,13). These included the following plasmids: pSK8Fb-Luc, pLuc5 were as described previously (12,13). These included the substrates used in the assay of the DSB repair activities.

**Materials and methods**

**Construction of assay systems for DSB repair activities**

The substrates used in the assay of the DSB repair activities were as described previously (12,13). These included the following plasmids: pSK8Fb-Luc, pLuc5 were as described previously (12,13). These included the substrates used in the assay of the DSB repair activities.

**Expression plasmid**

Full-length cDNA for clone No.128 was obtained by PCR using the primers, 5'-GCTAGAGGAAAAACAAAGGTG-GTAAG-3' and 5'-CCCTCGAGATCATCTTGTTAGACGA-3'. The amplified cDNA was fused to 3x HA-tag or green fluorescent protein (GFP) coding sequences, respectively. The resulting fragments were inserted into a KpnI–XbaI site of pIZT (Invitrogen).

**Preparation of double-stranded RNA**

Double-stranded RNA (dsRNA) for *in vivo* RNAi reactions was produced by the following procedures. (i) GFPdsRNA for control: templates were synthesized by two-step PCR amplification of pET32 (inserted with GFP sequence) using the primers pET5' (5'-CCGGAGAATATACGACTCAGTACAT-AGGG-3') and pET/T7-INTER-1 (5'-TAATACGACTCACTATAGGC-3'). The amplified cDNA was fused to 3x HA-tag or green fluorescent protein (GFP) coding sequences, respectively. The resulting fragments were inserted into a KpnI–XbaI site of pIZT (Invitrogen).

Preparation of double-stranded RNA

Full-length cDNA for clone No.128 was obtained by PCR using the primers, 5'-GCTAGAGGAAAAACAAAGGTGGTAAG-3' and 5'-CCCTCGAGATCATCTTGTTAGACGA-3'. The amplified cDNA was fused to 3x HA-tag or green fluorescent protein (GFP) coding sequences, respectively. The resulting fragments were inserted into a KpnI–XbaI site of pIZT (Invitrogen).

**Preparation of double-stranded RNA**

Double-stranded RNA (dsRNA) for *in vivo* RNAi reactions was produced by the following procedures. (i) GFPdsRNA for control: templates were synthesized by two-step PCR amplification of pET32 (inserted with GFP sequence) using the primers pET5' (5'-CCGGAGAATATACGACTCAGTACAT-AGGG-3') and pET/T7-INTER-1 (5'-TAATACGACTCACTATAGGC-3'). The amplified cDNA was fused to 3x HA-tag or green fluorescent protein (GFP) coding sequences, respectively. The resulting fragments were inserted into a KpnI–XbaI site of pIZT (Invitrogen).

**RESULTS**

**Screening of genes that affect DSB repair activities from the silkworm EST database**

In order to find the genes responsible for DSB repair in insect cells, we searched, in a *B. mori* EST database, for clones that showed any homologies to well-known genes participating in recombination, repair, replication, chromatin regulation or cell cycle, and attempted to knock down the genes by RNAi. A total of 241 clones were selected and subjected to the synthesis of dsRNAs, of which 124 were successfully obtained and transfected into silkworm BmN4 cells together with the uncut or BamHI-cut pLuc5' substrates (Figure 1A), DSB repair via both SSA and HR (the latter includes the crossing over type and gene-conversion type) possibly takes place on these substrates (Figure 1B). We then measured luciferase activity 72 h after transfection (Figure 1C). Clone no. 128 was found to increase DSB repair activity >2-fold with respect to both the uncut and cut substrates, and this clone was solely investigated in subsequent experiments. To confirm the above result, the effect of silencing clone no. 128 in DSB repair was analyzed using siRNA (Figure 1D).
Deduced repair pathways of pLuc5 transfection of pLuc5 full-length luciferase gene but with deletions at the region (C box) where the BamHI site is inserted. Luc3 cut) with siRNAs for BmRAD51 and clone no. 128. (c) Relative luciferase activities (normalized for β-galactosidase) in BmN4 cells measured 72 h after co-transfection of pLuc5‘ (BamHI cut) with siRNAs for BmRAD51 and clone no. 128. (E) Structures of the substrates pLuc5‘DR and pLuc5‘IR (illustrated by A/B/C boxes) for DSB repair assay through intramolecular HR, and deduced repair pathways following induced DSB. Both substrates consist of Luc5‘ΔC and Luc5‘ΔA moieties. Luc5‘ΔC has almost a full-length luciferase gene but with deletions at the region (C box) where the BamHI site is inserted. Luc5‘ΔA lacks a 5’ region (A box) of the luciferase gene. pLuc5‘DR contains Luc5‘ΔC and Luc5‘ΔA as direct repeats, whereas pLuc5‘IR contains them as inverted repeats. RluC is an intervening sequence between the repeats. (F) The expression plasmid for HA-clone no. 128 was co-transfected with or without dsRNA for clone no. 128 into BmN4 cells. HA-tagged protein was analyzed by western blot using anti-HA antibody (Sc-7392; Santa Cruz Biotechnology). (G) The expression plasmid for GFP-clone no. 128 was co-transfected with or without dsRNA for clone no. 128. GFP fluorescence was observed by microscopy 72 h after transfection. (H) Relative luciferase activities (normalized for β-galactosidase) in BmN4 cells measured 72 h after co-transfection of pLuc5‘IR or pLuc5‘DR with dsRNA constructed from EST no.128 sequence. See the legend to (C) for other details.

Molecular characterization of the clone no. 128

Since EST clone no. 128 contained a partial cDNA sequence lacking 5’ end, we isolated a full-length cDNA clone by 5’-RACE and determined its nucleotide sequence. It has an open reading frame of 3087 bp, which encodes a putative protein of 1029 amino acids. Homology search using the BLAST program revealed that this clone showed high homology with Argonaute family proteins and had the signature domains of proteins PAZ and PIWI in the C-terminus (Figure 2A). We then made a phylogenetic tree based on the amino acid sequence data of this clone and representative Argonaute family proteins from other organisms (Figure 2B). The current protein was proved to be the most similar to Drosophila melanogaster Argonaute2 (Dm Ago2) and was named BmArgonaute2 (BmAgo2). This protein had highly conserved sequence regions in the C-terminus, although without a glutamine-rich domain in the N-terminus, which is present in DmAgo2 [cf. Ref. (14)].

Alignment of the BmAgo2 amino acid sequence with those of D melanogaster, Rattus norvegicus, Mus musculus and Homo sapiens counterparts (Figure 2C) showed that identical residues are rich in the C-terminus region.

Expression analysis of BmAgo2 mRNA

The Argonaute protein family constitutes the largest group of proteins specifically participating in dsRNA-triggered gene silencing process.
silencing such as PTGS (posttranscriptional gene silencing) and RNAi (15). One could argue that RNAi intended for the targeted knockdown brings about sequence non-specific down regulation of a transgene. This possibility could be excluded by the observation that there were no effects of RNAi induction on luciferase activity when a luciferase expression vector was co-transfected with dsRNA for *BmAGO2* into BmN4 cells (Figure 3A). We also analyzed

Figure 2. Cloning and sequencing of full-length EST no. 128 (*BmAgo2*). (A) Deduced amino acid sequence (the *BmAgo2* protein) analyzed by the Pfam program. PAZ domain is shaded in black and PIWI domain is shaded in gray. (B) Phylogenetic tree for Argonaute family protein sequences (*Ago2* and so on). The CLUSTALW program was used for the calculation. Dm, *D.melanogaster*; Ce, *Caeorhabditis elegans*; At, *Arabidopsis italiana*; Mm, *M.musculus*; Hs, *H.sapiens*; Rn, *R.norvegicus*; Oc, *Oryctolagus cuniculus*; Sp, *Schizosaccharomyces pombe*; Nc, *Neurospora crassa*; Bm, *B.mori*. (C) Alignment of *BmAgo2* sequence with those of four related proteins, *DmAgo2*, *RnAgo3*, *MmAgo2* and *HsPIWI*. The CLUSTALW program and BOXSHADE were used for arrangement. Identities and similarities of amino acid residues are shaded in black and gray, respectively. Biochemically characterized critical residues in PAZ domain (37) and PIWI domain (DDH motif) (38,39) are indicated by closed triangles and circles, respectively. A portion of the N-terminal region of *BmAgo2* shows no significant similarity and is not shown in this figure.
the expression profiles of BmAgo2 mRNA by RT–PCR analysis. In agreement with previous reports that argonaute2 family proteins were rather ubiquitously expressed in various tissues (16), this mRNA was detected in all tissues examined (Figure 3B).

Effects of BmAgo2 on intrachromosomal recombination

As described above, BmAgo2 may function in the HR pathway without undergoing SSA. This prompted us to examine the effects of BmAGO2 RNAi on intrachromosomal recombination, using the cell line BmN4-DR11, which bears the functional sequence of pLuc5’3’ (BamHI cut) (Figure 4A) and dsRNA constructed from BmAGO2 sequence. As controls, dsRNA for GFP and pBluescript II SK (–) were transfected as mock substrates.

Figure 3. Some characterization of BmAgo2. (A) Relative luciferase activities (normalized for β-galactosidase) in BmN4 cells measured 12, 24, 72 and 144 h after co-transfection with ‘pLuc5’3’ (BamHI cut) (Figure 1A) and dsRNA constructed from BmAgo2 sequence. As controls, dsRNA for GFP and pBluescript II SK (–) were transfected as mock substrates. (B) Expression analysis of BmAgo2 messenger in various tissues from normal B. mori larvae (the strain r06) on day 3 of the fifth instar. RT–PCR was performed using the specific primers for BmAgo2. TE, testis; OV, ovary; SG, silk glands; MG, mid-gut; FB, fat body; BC, blood cells (males and females).

Figure 4. Effects of BmAgo2 on intrachromosomal recombination. (A) Schematic views of co-transfection strategy in the intrachromosomal recombination assay. The efficiency of the assay was measured by luciferase activity in BmN4-DR11 cells after transfection with dsRNA for the specific gene, with or without pSK8I-SceI, which contains an I-SceI endonuclease site under the control of SK8 promoter and a polyadenylation signal. (B) Relative luciferase activities (normalized for β-galactosidase) in cells measured 72 h after the transfection of dsRNA for partial BmAgo2 (5’AGO2 and 3’AGO2), constructed from the regions of its 5’ and 3’ ends, respectively. As controls, dsRNA for GFP and pBluescript II SK (–) were transfected as mock substrates.
nearly identical or rather tended to be lower than those of controls. The assay system itself was valid, as luciferase activity was significantly reduced in a positive control, RNAi for BmRAD51. This suggested that BmAgo2 exerted almost no effect on intrachromosomal recombination.

BmAgo2 represses HR repair in extrachromosomal DSBs

We then investigated the extrachromosomal HR process, exploiting the previously described gene targeting system with two targeting vectors [Figure 5A; cf. Ref. (13)].

**Figure 5.** Effects of BmAgo2 on extrachromosomal recombination as assayed after gene targeting. (A) Schematic views of co-transfection strategy in the extrachromosomal recombination assay. The efficiency of the assay was measured by the luciferase activity in BmN4-Luc5 D cells after co-transfection with the targeting vector and dsRNA for BmAgo2, together with pSK8I-SceI. Two kinds of targeting vectors were used, pZErOI-SceILuc3 and pZErOLuc3, which contain a truncated 3′ fragment of the luciferase gene and a polyadenylation signal. pZErOI-SceILuc3 has the I-SceI recognition site, but pZErOLuc3 does not. (B) Relative luciferase activities (normalized for β-galactosidase) in the cells measured 72 h after co-transfection with pZErOI-SceILuc3 and dsRNA for BmAgo2, together with pSK8I-SceI. As controls, dsRNA for GFP (GFP) and pBlueScript II SK (−) were transfected as mock substrates. (C) Relative luciferase activities (normalized for β-galactosidase) in the cells measured 72 h after co-transfection with pZErOLuc3 and dsRNA for BmAgo2, together with pSK8I-SceI. See the legend to (B) for controls.
These plasmids contain a truncated 3' fragment of the luciferase gene, with or without an I-SceI recognition site 5' upstream of the luciferase fragment. The efficacy of extrachromosomal recombination was measured as relative luciferase activity using BmN4-Luc5ΔC cells, which contains the 5' fragment of luciferase sequence as a genomic target on its chromosome. First, we carried out a targeting assay using targeting vector with an I-SceI recognition site (pZErOI-SceLuc3). When dsRNA for BmAgo2 was co-transfected with the I-SceI expression vector (pSK&I-SceI) and pZErOI-SceILuc3 into BmN4-Luc5ΔC cells, luciferase activity was ~1.5-fold higher than controls (Figure 5B). In negative control experiments using the plasmid without the I-SceI recognition site (pZErOLuc3'), there was almost no effect on luciferase activity even when BmAgo2 was knocked down (Figure 5C). These results strongly indicated that BmAgo2 recognizes extrachromosomal DSBS and represses their HR repair.

**DISCUSSION**

Using a database searching technique, we isolated cDNA encoding a silkworm homolog of *D.melanogaster* Argonaute2 as a repressive factor for extrachromosomal HR repair. Argonaute2 is a protein with highly conserved sequence motifs, referred to as Argonaute family proteins (5). These proteins are divided into two groups: Argonaute1 (Argo1) and Argonaute2 (Argo2). The Argonaute1 family proteins are involved in RNAi and transcriptional gene silencing (TGS), while the Argonaute2 family proteins are involved in post-transcriptional gene silencing (PTGS). The crystal structure of the PIWI domain of the PIWI family proteins is similar to that of ribonuclease H (15), which implies that Argonaute is an endonuclease that cleaves target mRNAs. The current study showed that the presently characterized BmAgo2 contains the conserved amino acid residues D965, D1037 and H1173, which are critical for Ago2 nuclease activity and identified in the DmAgo2 PIWI domain, but do not possess, in the PAZ domain, any of the residues F292, Y309, Y314 and L337, which have been identified in HseIF2c1 and pyrococcus furiosus (15). According to a recent study on *Pyrococcus furiosus* (17), the crystal structure of the PIWI domain is similar to that of ribonuclease H, implying that Argonaute is an endonuclease that cleaves target mRNA. It was confirmed that the presently characterized BmAgo2 contains the conserved amino acid residues D965, D1037 and H1173, which are critical for Ago2 nuclease activity and identified in the DmAgo2 PIWI domain, but does not possess, in the PAZ domain, any of the residues F292, Y309, Y314 and L337, which have been identified in HseIF2c1 and are said to be necessary for the binding to 5'-phosphate of the guide RNA, siRNA (18–20; see Figure 2C and the legend). These findings imply that Argonaute protein family members, found in diverse species of organisms, are crucial factors with special reference to small RNAs which are of significance beyond what is already known.

RNAi factors have been discussed in relation to germline development (21) and stem cell maintenance (22). Several studies have shown that Argonaute family proteins are important in the development in diverse species. Argonaute itself is involved in leaf and flower formation (23). Some family members, including Aubergine/Sting, are required for normal germline development (22, 24–26) seemingly via an RNAi-like mechanism (27). These results showed a clear relationship between recombination and RNAi or Argonaute protein. To our knowledge, such presentations are very scarce except for two previous articles reporting that Argonaute family proteins are implicated in programmed genome rearrangements in tetrahymena (28) and in the developmentally programmed elimination of genome sequences in ciliates (29).

Our observation that knockdown of the silkworm Argonaute2 homolog gene, *BmAgo2*, augments the rate of extrachromosomal HR appeared to be explicable by two alternative assumptions. First, BmAgo2 may act as an inhibitor of HR. This inference comes from the notion that Argonaute protein possibly inhibits the binding of HR-related proteins to the strands via a PAZ domain, which can directly bind to substrate DNA. Second, BmAgo2 is needed to discriminate extrachromosomal DNA from chromosomal DNA and represses extrachromosomal HR repair indirectly. We favor the latter possibility, because Argonaute family proteins have been reported to participate in heterochromatin formation. Argonaute1 is one of the components of the RNA-induced initiation of the transcriptional gene silencing (RTS) complex, which was shown to be necessary for heterochromatin assembly (30). Recent studies indicate that RNAi is involved in heterochromatin formation at the centromere and therefore in chromosome segregation (31–34). If BmAgo2 plays a role in the extrachromosomal DNA-specific assembly of heterochromatin, in which HR repair of DSBS are repressed, the decrease in BmAgo2 expression would not affect intrachromosomal HR repair, and this was in fact the case as described above.

It is reasonable to predict that cells have defense systems for exogenous DNA, e.g. viral DNA. If such exogenous DNA has a homologous sequence to that of a host genome region, HR between these may frequently cause a partial loss of genomic information. Therefore, the cells would acquire mechanisms to repress the HR repair of extrachromosomal DSBS by using an Argonaute protein. Indeed, baculoviruses, DNA viruses widely isolated from lepidopteran insects, often carry DNA transposable elements, such as piggybac and Tc1-like elements, which apparently originate from the cellular genomes and are inserted into infecting baculovirus genomes (35,36). These viruses are to be excluded from the HR-related integration pathway leading to the modification of host genomes.

**ACKNOWLEDGEMENTS**

This work was supported in part by grants Nos. 14704010 and 14656024, and the National Bioresource Project, from the Ministry of Education, Science and Culture of Japan. We thank Ms. Junko Nohata and Ms. Motoe Sasanuma, Insect Genome Lab., National Institute of Agrobiological Sciences, for preparation of EST clones. Funding to pay the Open Access publication charges for this article was provided by the grant No. 1470410.

Conflict of interest statement. None declared.

**REFERENCES**

maintenance of chromosomal integrity in vertebrate cells. EMBO J., 17, 5497–5508.