Molecular Characterization of Gap Region in 28S rRNA Molecules in Brine Shrimp *Artemia parthenogenetica* and Planarian *Dugesia japonica*

Shuhong Sun*, Hui Xie*, Yan Sun*, Jing Song, and Zhi Li*

College of Life Sciences, Shaanxi Normal University, Xi'an 710062, Shaanxi, P.R. China; E-mail: sunyan@snnu.edu.cn; mmbiolab@gmail.com; izhi@snnu.edu.cn; mmbiolab@163.com

Received November 25, 2011 Revision received February 20, 2012

Abstract—In most insects and some other protostomes, a small stretch of nucleotides can be removed from mature 28S rRNA molecules, which could create two 28S rRNA subunits (28Sα and 28Sβ). Thus, during electrophoresis, the rRNA profiles of these organisms may differ significantly from the standard benchmark since the two subunits co-migrate with the 18S rRNA. To understand the structure and mechanism of the atypical 28S rRNA molecule, partial fragments of 28Sα and 28Sβ in brine shrimp *Artemia parthenogenetica* and planarian *Dugesia japonica* were cloned using a modified technology based on terminal transferase. Alignment with the corresponding sequences of 28S rDNAs indicates that there are 41 nucleotides in *A. parthenogenetica* and 42 nucleotides in *D. japonica* absent from the mature rRNAs. The AU content of the gap sequences of *D. japonica* and *A. parthenogenetica* is high. Both the gaps may form stem-loop structure. In *D. japonica* a UAAU cleavage signal is identified in the loop, but it is absent in *A. parthenogenetica*. Thus, it is proposed that the gap processing of 28S rRNA was a late enzyme-dependent cleavage event in the rRNA maturational process based on the AU rich gap sequence and the formation of the stem-loop structure to expose the processing segment, while the deletion of the gap region would not affect the structure and function of the 28S rRNA molecule.

DOI: 10.1134/S000629791204013X

Key words: 28S rRNA, gap region, fragmentation, Artemia parthenogenetica, Dugesia japonica

During the maturation of eukaryotic rRNAs (18, 5.8, and 28S), external and internal transcribed spacer (ETS and ITS) sequences are removed from 45S pre-rRNA to become functional molecules. Furthermore, additional excision events of rRNAs have also been observed. For example, the 28S rRNA from most protostomes is processed and dissociated into two subunits, α and β , which can be detected under denaturing conditions [1-3]. This specific cleavage was once known as a "hidden break" [4]; however, further studies on insect rDNA structure have shown that the "break" was generated by the removal of a short stretch, also known as a gap region, from the rRNA precursors [5]. Thus, this phenomenon might be better termed as "gap deletion" rather than "hidden break".

Fragmented large subunit ribosomal RNAs have been observed in a variety of organisms [6, 7]. It was first found in silk moth pupae [3]. In *Galleria mellonella*, the fragmented rRNA was presumed to have break(s) in its primary structure rather than a change in conformation

[4]. Subsequently, the discontinuity position of 26S rRNA has been located in Acanthamoeba castellani [8]. From the late 20th century more organisms were reported to have this specific "gap deletion" phenomenon, particularly in protostomes such as Tetrahymena pyriformis [9], and insects such as Bombyx mori [10], Aedes aegypti [7], Musca carnaria [11], Drosophila species [12], Actinobacillus actimomycetemcomitans [13], and Sciara coprophila [5]. Although many organisms contain gaps within the 28S rRNA, the size of the gap region differs among them, for example, 19 nt in S. coprophila [5], 54 nt in Schistosoma mansoni [14], and 86 nt in Trichinella spiralis [15]. Although the size varied between organisms, the gap sequence resided in the same region within the 28S rRNA, locating in the D7a expansion segment of eukaryotes [16-19]. Some researchers tried to explain the molecular mechanism of the excision of the gap sequence from mature 28S rRNA, but after investigation for nearly 40 years its mechanism remains unclear [14, 16, 20-22].

When using a denaturing agarose gel to analyze the RNA samples of *Artemia parthenogenetica* and *Dugesia japonica*, the band corresponding to the 18S rRNA was

[#] These authors contributed equally to this work.

^{*} To whom correspondence should be addressed.

denser than the 28S band, unlike the typical rRNA profile. For *A. parthenogenetica*, this phenomenon had been observed in 1984, and it was considered that there was a "hidden break" in the 28S rRNA molecule [23]. Now we know that it is a "gap deletion" rather than a "hidden break" that may result in the breakage of 28S rRNA. Therefore, in this study, terminal transfer and PCR methods were employed to verify the gap regions of 28S rRNA in *A. parthenogenetica* and *D. japonica*. Compared to traditional and classical S1 nuclease protection assay, which was used to determine the gap regions in the previous studies, these strategies are simpler and more efficient. Finally, the two gap regions are confirmed, and the possible processing mechanisms of 28S rRNA in *A. parthenogenetica* and *D. japonica* are proposed.

MATERIALS AND METHODS

Primers. Primers (28SαF1, 5'-AAGCAGAACTG-GCGCTGTGGGATG-3'; 28SαF2, 5'-AGACAKCA-GGACGGTGGCCATGG-3'; 28SβR1, 5'-GTCATA-GTTACTCCGCCGTTTACCCGCGC-3'; 28SβR2, 5'-YCYTYAGAGCCAATCCTTATCCCGAAG-3') were designed according to the highly conserved regions of 28S rRNAs after alignment with the 28S rRNA sequences of *Artemia* sp. (AY210805), *Schmidtea mediterranea* (http://smedgd.neuro.utah.edu/cgi-bin/gbrowse/smed_mysql/#search), *Xenopus laevis* (X02995), and *S. mansoni* [14].

Cloning the partial sequences of A. parthenogenetica and D. japonica 28S rDNA. Whole planarians, starved for one week, were homogenized in lysis buffer containing 100 mM Tris HCl, pH 8.0, 20 mM EDTA, 1.4 M NaCl, 20 mg/ml CTAB, 2% β-mercaptoethanol. The brine shrimp were also homogenized using the same lysis buffer. After incubation at 65°C for 1 h, the genome DNA was extracted with phenol/chloroform, and chloroform, ethanol-precipitated in the presence of 0.3 M NaAc, dissolved in pure water. DNA was further treated with RNase A for RNA degradation. Purified genome DNA was then detected on an agarose gel. The first run of PCR amplification was carried out with primers $28S\alpha F1$ and $28S\beta R1$. $28S\alpha F2$ and $28S\beta R2$ were applied to amplify the partial A. parthenogenetica and D. japonica 28S rDNA sequences within the first run product. The PCR products were purified and cloned into the pMD-20T vector (TaKaRa, China). After colony PCR screening, positive colonies were sequenced.

Cloning the 28S β partial fragments of *A. parthenogenetica* and *D. japonica* 28S rRNA. Total RNA was prepared from *A. parthenogenetica* and *D. japonica* by the Trizol method and reverse-transcribed using 28S β R1 as primer. The addition of polyG tail to the 3' terminus of the first-strand cDNA was catalyzed by terminal deoxynucleotidyl transferase (TdT; TaKaRa). The Oligo dC₁₈ was employed as a forward primer and 28S β R1 as a reverse primer to amplify the target fragment. To increase the specificity, the obtained PCR products were subject-

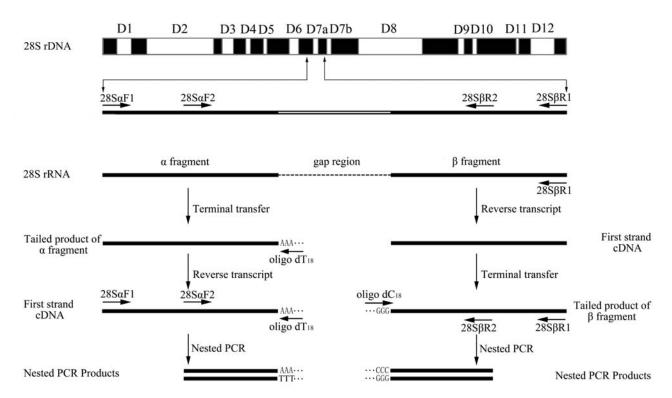


Fig. 1. Strategies used for determining the 5' and 3' boundaries of the gap region in A. parthenogenetica and D. japonica. The primers used in cloning the partial $28S\alpha$ and $28S\beta$ fragments of A. parthenogenetica and D. japonica are indicated as arrows.

ed to nested PCR using oligo dC_{18} and $28S\beta R2$ as primers (Fig. 1). The target fragment was cloned, and after colony PCR screening the positive colonies were sequenced.

Cloning the 28S α partial fragments of A. parthenogenetica and D. japonica 28S rRNA. The same batch of total RNA was used in the next experiments. First, a tail of polyA was added directly to the 3'-OH of the A. parthenogenetica and D. japonica 28S rRNA using TdT enzyme. The tailed RNA served as template in the reverse transcript reaction with oligo dT₁₈ as primer. Then oligo dT₁₈ and 28S α F1 were applied to amplify the target fragment. The nested PCR was then carried out using 28S α F2 and oligo dT₁₈ (Fig. 1). Finally, the target fragments were cloned and sequenced.

S1 nuclease protection assay. A modified S1 nuclease protection assay was carried out to verify the results obtained from TdT methods and is briefly described as follows. The 28S rDNA fragments covering the assumed gap regions of A. parthenogenetica and D. japonica 28S rRNAs were amplified with primers $28S\alpha F1$ and $28S\beta R1$, respectively. After the rDNA fragments were denatured and hybridized with total RNA, the mixture was incubated with S1 nuclease. The S1 protected rDNA fragments were purified and concentrated through ethanol precipitation. For the 28S\alpha subunit, the corresponding rDNA fragment was subjected to a primer extension experiment using 28SαF1 as a primer. A polyG tail was then added to the 3' terminus of the primer extension products by TdT. PCR reaction was performed with 28SαF1 and oligo dC₁₈. Subsequently, the amplified PCR product was cloned and sequenced. For the 28Sβ subunit, a polyG tail was added to the 3' terminus of the rDNA fragment directly. The PCR reaction was performed with oligo dC_{18} and 28SβR1. Finally, the amplified PCR product was cloned and sequenced.

RESULTS AND DISCUSSION

Identification of the gaps in A. parthenogenetica and **D.** japonica. The 5' end of the 28Sβ fragment and the 3' end of the 28Sa fragment of A. parthenogenetica were cloned and sequenced. After comparing with the 28S rDNA of A. parthenogenetica (accession number JN009670, cloned in our laboratory), the 5' and 3' boundaries of the A. parthenogenetica 28S rRNA gap sequence were identified and a 41 nt gap sequence was obtained (Fig. 2). Nevertheless, when compared with the former research results, which reported the 3' end sequence of Artemia salina 28Sa fragment, several nucleotides are different and one more "G" was present in our sequence [23]. This may have resulted from the inaccuracy of the sequencing method utilized at that time or the difference between different species. In spite of this, the cleavage sites were identified to be located almost in the same position. Meanwhile, the 5' and 3' boundaries of D. japonica 28S rRNA gap sequence were confirmed and a 42 nt gap was absent from the D. japonica mature 28S rRNA, compared with the corresponding rDNA sequence (accession number JF827607, cloned in our laboratory).

The nucleotides of and around the D7a expansion segments of *A. parthenogenetica* and *D. japonica* 28S rDNA were compared to those of *Artemia* sp., *S. mansoni*, *S. coprophila*, *T. spiralis*, and *X. laevis* (Fig. 2). The gaps of *A. parthenogenetica* and *D. japonica* are located in the same region, the D7a expansion segment of eukaryotic 28S rRNA, just as other species reported previously. The AU content of the *A. parthenogenetica* D7a expansion segment was 46.5%, lower than that of *D. japonica* (67.9%) and most of the other species studied (table). So far, the D7a expansion segments in most of the organisms demonstrating gap processing are characterized by their

Putative processing features in different organisms

Organism	Gap processing	CGAAAGGG	UAAU	AU content in D7a, %	Gap sequence, nt
D. japonica	+	+	+	67.90	42
S. coprophila	+	+	+	74.60	19
S. mansoni	+	+	+	47.50	54
T. spiralis	+	+	+	59.10	86
B. mori	+	+	+	68.80	30
A. parthenogenetica	+	+	_	46.50	41
T. thermophila	+	+	_	63.80	4
A. albopictus	+	+	_	65.30	350
X. laevis	_	+	_	27.60	NA
A. pisum	_	+	_	34.50	NA

Note: NA, information not available; +, presence; -, absence.



Fig. 2. Alignment of the partial 28Sα, 28Sβ sequences of *A. parthenogenetica* and *D. japonica* with the counterparts in *S. mansoni*, *S. coprophila*, *T. spiralis*, and *X. laevis*, also with the corresponding sequences of 28S rDNA of *A. parthenogenetica*, *Artemia* sp. and *D. japonica*, which indicated as "*A. parthenogenetica* (G), *Artemia* sp. (G) and *D. japonica* (G)". The 28Sα and 28Sβ fragments are indicated. The "G" in *A. parthenogenetica*, which is inconsistent with a previous report, is marked with an asterisk. The conserved motif CGAAAGGG is boxed. The CGAAAG conserved sequences located near both termini of the gap are boxed with a rounded rectangle. The D7a expansion segment is underlined.

extremely high AU content, especially the gap region. So, there is the opinion that the high AU content in the D7a expansion segment may be a necessary condition for the excising of the gap [21]. And in *X. laevis* and *A. pisum*, organisms not having gap processing, the AU contents are only 27.6 and 34.5%, respectively. So, it was suggested that a relatively higher AU content of the gap sequence could be one gap processing symbol.

A conserved motif CGAAAGGG proximal to the 3'-end of the gap was observed in *A. parthenogenetica* and *D. japonica*. This conserved element was once proposed as a possible gap processing signal even though it was found in organisms such as *X. laevis* and *C. elegans*, whose 28S rRNA was intact [15, 24]. Another sequence CGAAAG, which had been proposed as a possible recognition signal [15], was located near both termini of the gap in *T. spiralis* and *S. coprophila*. But these flanking sequences were not

found in *A. parthenogenetica*, *D. japonica*, *S. mansoni*, and some other organisms demonstrating gap processing [14]. Thus, any consensus sequence functions as specific recognition signal or processing site had not been identified within the D7a expansion segment through the sequence comparisons.

Analysis of the 28S rRNA secondary structure in A. parthenogenetica and D. japonica. Secondary structure as well as primary sequence may also play a role in the 28S rRNA gap processing. The possible secondary structures were constructed for the D7a expansion segments of A. parthenogenetica and D. japonica (Fig. 3). The secondary structures of the corresponding sequences of 28S rRNAs in S. mansoni, S. coprophila, T. spiralis, and X. laevis were predicted at the same time. The D7a expansion segments in all gap processing species analyzed were folded into the stemloop structure or similar, with the cleavage sites locating in

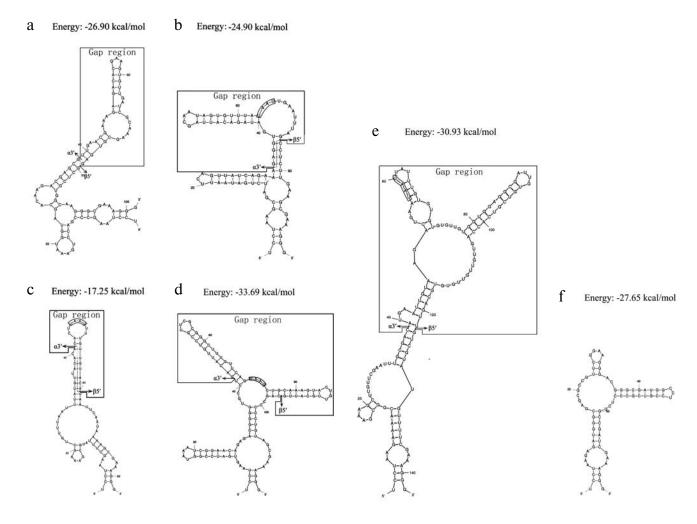


Fig. 3. Secondary structure models for the D7a expansion segment of the 28S rRNAs of *A. parthenogenetica* (a), *D. japonica* (b), *S. coprophila* (c), *S. mansoni* (d), *T. spiralis* (e), and *X. laevis* (f). The calculated free energy of each structure is shown on the top of each model. The putative processing sites are indicated by arrows. The 3' end of the 28S α fragment and the 5' end of the 28S β fragment are expressed using the symbols α 3' and β 5'. The conserved UAAU sequence is boxed. The gap sequences that are excised from the mature 28S rRNAs are boxed.

the stem portion of the structure and adjacent to an interior loop that is consistent with previous study findings.

It was noteworthy that one conserved motif, the UAAU tract resides within the hairpin loop had been found in D. japonica, also in S. coprophila, S. mansoni, and some other gap processing organisms. The UAAU tract had been suggested to be a signal for excision of the gap region to create the fragmented 28S rRNA [14, 21, 22]. This assumption was supported by the fact that the UAAU sequence had also been identified in the 5.8S rRNA of Drosophila melanogaster in which an AU rich stretch of nucleotides was removed [25]. However, there were some exceptions. For example, the lacking of the UAAU motif did not affect the fragmentation of the 28S rRNA molecule in *Aedes albopictus* [26]. In this study, the UAAU motif was also not found in A. parthenogenetica, while in T. spiralis the sequence was located in the stem structure [5, 14, 15]. So, whether the UAAU motif was one of the gap processing symbols was controversial.

However, one consistent event was that the removing of the gap sequences did not affect the overall structures of the rRNA (Fig. 3), since sufficient hydrogen bonds existed between the α and β fragments of 28S rRNA molecules. That was the reason why the atypical electrophoretic rRNA profiles were observed under denaturing conditions, since the hydrogen bonds were broken under heat or formaldehyde treated, while under nondenaturing conditions, the fragmentation of 28S rRNA molecules could not be detected. Thus, combined with the analyses of the primary sequences of the gaps, we postulated that the cleavage of 28S rRNA molecules did not depend on the conserved motifs, but might associated with the recognition of the stem-loop secondary structures. As the gap processing would not affect the function of 28S rRNA molecules in vivo, the cleavage could be a late event in rRNA maturational process, which meant that the stem-loop secondary structure of 28S rRNA molecules formed prior to the gap processing.

Recently, an increasing number of three-dimensional (3D) structures of eukaryotic ribosomes are being constructed. For example, the crystal structure of *Saccharomyces cerevisiae* 80S ribosome at a resolution of 3.0 Å was reported [27]. Based on the 3D model, the spatial distribution of the ribosome elements is clear. So that it was identified that ES19, which corresponding to D7a expansion segment including the gap region in RNA molecular is distributed on the surface of the *Saccharomyces cerevisiae* ribosome [27, 28]. This may help to expose the processing sites of the gap region on the ribosome surface and make the recognition of the sites easier. Also, this might be evidence that the specific rRNA conformation was required for the excision of the gap.

However, Lava-Sanchez and Puppo [11] predicted that nonspecific endonucleases might participate in the phenomenon, because it was observed *in vivo* that the fragmentation pattern in *Musca carnaria* 28S rRNA could be reproduced by treating newly assembled ribosomes with pancreatic RNase. In conjunction with the above analyses, we proposed that the gap processing of 28S rRNA was a late enzyme-dependent cleavage event in rRNA maturational process based on: 1) the AU rich gap sequence; 2) the formation of the specific stem-loop conformation to expose the processing segment.

Techniques used to determine the terminal sequences. The techniques of isotope labeling and S1 nuclease protection assay were employed to identify the terminals of the 28S rRNA gap region in most studies. In this study, terminal transfer (relying on the activity of TdT enzyme) was applied to determine the exact boundaries of the gap sequences. Usually, the TdT was applied to add the nucleotides to the 3'-OH end of DNA molecule. However, in this study the TdT also helped the addition of nucleotides directly to the 3'-OH end of RNA. Combined with reverse transcription and the polymerase chain reaction, the gap regions of 28S rRNA in A. parthenogenetica and D. japonica were determined successfully and efficiently. To verify the accuracy of the results from the TdT method, a modified S1 nuclease protection experiment was carried out in parallel. The agreement between techniques indicates that the TdT method was believable. Compared to the S1 nuclease protection assay, the TdT method has several advantages: 1) safe from radiation; 2) fewer steps renders it simpler; 3) more efficient, and 4) economical. Therefore, the TdT method could be easily applied in other laboratories to determine the terminal sequences of any kind of nucleic acid molecules.

Several similar methods were employed to map 5' and 3' ends of RNA molecules. For example, TdT was used to identify the 5' end of *Schizosaccharomyces pombe* TER1 telomerase RNA [29], and the method of tailing RNA with poly(A) polymerase was applied to determine the 3' ends of human telomerase RNA [30]. In our study, the TdT method was utilized to map the boundaries of the

gap regions in RNA molecules for the first time, and the TdT enzyme adds dATP to the 3'-OH end of RNA directly and successfully. In addition, TdT could add any type of the dNTPs to the 3'-OH end of RNA molecules, which make the TdT method more flexible in different research fields.

In summary, the gap deletion of 28S rRNA in A. parthenogenetica and D. japonica was confirmed. According to the analyses of the primary and secondary structures, we presumed that the gap processing of 28S rRNA is a late enzyme-dependent cleavage event in the rRNA maturational process based on the AU rich gap sequence and the formation of the stem-loop structure to expose the processing segment. The gap sequence removed from 28S rRNAs could be considered as special internal transcribed spacers, which are similar to the internal transcribed spacers flanking the 5.8S rRNA gene. As the excision of the small gap does not affect the functions of the mature 28S rRNA in vivo, there remains many questions in this area of research. For example, as it is well known that the excision of external and internal transcribed spacers is necessary for rRNA maturation, is it an essential issue to create a mature rRNA molecular? It is unclear why most excision is limited to insects and some protostomes, and therefore further study is recommended.

The work described in this paper was supported by the Natural Science Basic Research Plan in Shaanxi Province of China (No. 2009JQ3002), the Fundamental Research Funds for the Central Universities (No. GK200902030, GK200902031), the Innovative Experiment Projects of Educational Ministry of China for Undergraduate (No. 091071842) and LOFCS (No. 09057, 09063, 2010023) from Shaanxi Normal University.

REFERENCES

- Ishikawa, H. (1973) Comp. Biochem. Physiol. B, 46, 217-227.
- 2. Greenberg, J. R. (1969) J. Mol. Biol., 46, 85-98.
- 3. Applebaum, S. W., Ebstein, R. P., and Wyatt, G. R. (1966) *J. Mol. Biol.*, **21**, 29-41.
- Ishikawa, H., and Newburgh, R. W. (1972) J. Mol. Biol., 64, 135-144.
- Ware, V. C., Renkawitz, R., and Gerbi, S. A. (1985) Nucleic Acids Res., 13, 3581-3597.
- Winnebeck, E. C., Millar, C. D., and Warman, G. R. (2010) J. Insect. Sci, 10, 159.
- 7. Shine, J., and Dalgarno, L. (1973) J. Mol. Biol., 75, 57-72.
- Stevens, A. R., and Pachler, P. F. (1972) J. Mol. Biol., 66, 225-237.
- Eckert, W. A., Kaffenberger, W., Krohne, G., and Franke, W. W. (1978) Eur. J. Biochem., 87, 607-616.
- Fujiwara, H., Ogura, T., Takada, N., Miyajima, N., Ishikawa, H., and Maekawa, H. (1984) *Nucleic Acids Res.*, 12, 6861-6869.

- Lava-Sanchez, P. A., and Puppo, S. (1975) J. Mol. Biol., 95, 9-20.
- 12. Delanversin, G., and Jacq, B. (1983) CR Seances Acad. Sci. III, 296, 1041-1044.
- Haraszthy, V. I., Sunday, G. J., Bobek, L. A., Motley, T. S., Preus, H., and Zambon, J. J. (1992) *J. Dental Res.*, 71, 1561-1568.
- Van Keulen, H., Mertz, P. M., LoVerde, P. T., Shi, H., and Rekosh, D. M. (1991) Mol. Biochem. Parasitol., 45, 205-214.
- 15. Zarlenga, D. S., and Dame, J. B. (1992) *Mol. Biochem. Parasitol.*, **51**, 281-289.
- Basile-Borgia, A. E., Dunbar, D. A., and Ware, V. C. (2005) Insect. Mol. Biol., 14, 523-536.
- De Rijk, P., van de Peer, Y., Chapelle, S., and de Wachter,
 R. (1994) Nucleic Acids Res., 22, 3495-3501.
- Clark, C. G., Tague, B. W., Ware, V. C., and Gerbi, S. A. (1984) Nucleic Acids Res., 12, 6197-6220.
- 19. Ishikawa, H. (1977) Comp. Biochem. Physiol. B, 58, 1-7.
- Melen, G. J., Pesce, C. G., Rossi, M. S., and Kornblihtt,
 A. R. (1999) *EMBO J.*, 18, 3107-3118.
- Ogino, K., Eda-Fujiwara, H., Fujiwara, H., and Ishikawa, H. (1990) J. Mol. Evol., 30, 509-513.

- Fujiwara, H., and Ishikawa, H. (1986) Nucleic Acids Res., 14, 6393-6401.
- Nelles, L., van Broeckhoven, C., de Wachter, R., and Vandenberghe, A. (1984) Naturwissenschaften, 71, 634-635
- Ware, V. C., Tague, B. W., Clark, C. G., Gourse, R. L., Brand, R. C., and Gerbi, S. A. (1983) *Nucleic Acids Res.*, 11, 7795-7817.
- Pavlakis, G. N., Jordan, B. R., Wurst, R. M., and Vournakis, J. N. (1979) *Nucleic Acids Res.*, 7, 2213-2238.
- Park, Y. J., and Fallon, A. M. (1990) *Insect. Biochem.*, 20, 1-11.
- Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G., and Yusupov, M. (2011) *Science*, 334, 1524-1529.
- Taylor, D. J., Devkota, B., Huang, A. D., Topf, M., Narayanan, E., Sali, A., Harvey, S. C., and Frank, J. (2009) Structure, 17, 1591-1604.
- 29. Webb, C. J., and Zakian, V. A. (2008) *Nat. Struct. Mol. Biol.*, **15**, 34-42.
- Zaug, A. J., Linger, J., and Cech, T. R. (1996) Nucleic Acids Res., 24, 532-533.