Acta Biochimica Polonica

Vol. 41 No. 1/1994

QUARTERLY

Evolutionary conservation of the transcribed spacer sequences of the rDNA repeat unit in three species of the genus *Aspergillus**

Piotr Borsuk, Marek Gniadkowski, Robert Kucharski, Mariola Bisko, Magdalena Kanabus, Piotr P. Stępień and Ewa Bartnik**

Department of Genetics, University of Warsaw, Al. Ujazdowskie 4, 00–478 Warsaw, Poland

Received 20 November, 1993

Key words: *Aspergillus*, rDNA, intervening transcribed sequences

We have cloned and sequenced the two intervening transcribed spacers in the rDNA repeat unit of three *Aspergillus* species — *A. nidulans*, *A. awamori* and *A. wentii*. The *A. wentii* and *A. awamori* spacers are almost identical and share a high degree of homology with the *A. nidulans* spacers. All spacers have a high G-C content (66%–76%) and the potential of forming complex secondary structures, which may indicate that they play a role in the maturation of pre-rRNA molecules.

In most eucaryotes the genes coding for the three larger classes of ribosomal RNA (rRNA)[1] form a tandemly repeated transcriptional unit with the following organization:

NTS*ETS*18SrRNA *ITS1*5.8S rRNA *ITS2*26SrRNA

where NTS is the non-transcribed spacer, ETS – the external transcribed spacer and ITS1 and ITS2 are the internal transcribed spacers [1]. The primary transcript is processed to the mature rRNAs. The signals recognized by the processing mechanisms are poorly known, and transcription has been extensively analyzed only in mammals [2] where U3 snRNA (small nuclear RNA) is known to participate in primary transcript processing [3]. Recent data have implicated other snRNAs (U8, U14) in pre-rRNA processing in yeast, *Xenopus* and mammals [4, 5].

We have cloned the rDNA repeat unit from the filamentous fungus *Aspergillus nidulans* [6]. With a size of 7.8 kb it is one of the smallest rDNA repeat units in eucaryotes. Its nucleotide sequence has been established (Borsuk, unpublished). In order to determine what sequences are important for processing we decided to examine rDNA intervening sequences in two other closely related species — *A. awamori* and *A. wentii*. As spacer sequences are in general not conserved between species, similarities in sequence among the three species could be an indication of some function. We have been using this approach in analyzing the evolution of dispersed 5S rRNA genes in the genus *Aspergillus* [7]. This paper presents the analysis of ITS1 and ITS2 sequences from the three *Aspergillus* species: *A. nidulans*, *A. awamori* and *A. wentii*.

MATERIALS AND METHODS

Appropriate fragments of the plasmid pMN1 [6] which contains the rDNA repeat unit from *A. nidulans* were subcloned into mp18 and

*This work was partly supported by Grant 103 and 304 from the State Committee for Scientific Research.

**All correspondence to Dr. Ewa Bartnik

[1] Abbreviations: ETS, external transcribed spacer; ITS, internal transcribed spacer; NTS, non-transcribed spacer; rRNA, ribosomal RNA; snRNA, small nuclear RNA.

mp19 and sequenced as described [8]. Gene libraries of *A. wentii* and *A. awamori* were constructed after digestion with *Eco*RI and ligation with pUC19. Fragments of the pMN1 insert corresponding to ITS1 and ITS2 were used to screen these libraries. The restriction maps of the plasmids isolated from these libraries were established and subfragments containing ITS1 and ITS2 were cloned in mp18 and mp19 for sequence analysis.

Secondary structures of ITS1 and ITS2 were analyzed using the PC Gene program.

RESULTS AND DISCUSSION

The sequences of ITS1 and ITS2 from the three *Aspergillus* species are shown in Fig. 1 and 2, respectively. The ITS1 sequences from *A. wentii* and *A. awamori* are identical, 185 bp long and contain 65.9% GC. The *A. nidulans* sequence is considerably shorter, 150 bp in length and even more GC rich – 76.7%. The ITS2 sequences of *A. wentii* and *A. awamori* differ by a deletion of two nucleotides in the former and are thus 155 and 157 bp in length, with GC content of 69.7 and 70.1%, respectively. The *A. nidulans* ITS2 sequence differs in many places from the *A. awamori* and *A. wentii* sequences, is 158 bp in length and contains 74.1% GC.

The presented findings show an unexpected extensive homology among the ITS sequences of different species of *Aspergillus*. The *A. wentii* and *A. awamori* sequences are identical (ITS1) or almost identical (ITS2), the *A. nidulans* sequence shares a high degree of homology (approx. 58% and 80% for ITS1 and ITS2, respectively) with the two other species.

All the ITS show a high GC content and a number of long GC tracts. All the sequences were analyzed for the ability to produce secondary structures using the PC gene program. The data for the six analyzed ITS sequences are presented in Figs. 3 and 4. In all cases long stretches of double stranded structures formed by complementary sequences are visible. The ITS2 structures are very similar for all three species. In ITS1 there are differences in second-

```
A. awamori   aacctgcggaaggatcattaCCGAGTGCGGGT-CCTTTGGG--CCCAACCTCCCATCCGTGTCTATTGTACCCTGTTGCTT
A. wentii
A. nidulans                                  C G   CC   CG          C    A   CCTA A

A. awamori   CGGCGGGCCCGCCGCTTGTCG-GCCGCCGGGGGGGCGCCTCTGCCCCCCGGGCCCGTGCCCGCCGGAGACCCCAACACGAA
A. wentii
A. nidulans        GG C  C AG G  A          ----A  A   AA TT AT----------   T    ------------

A. awamori   CACTGTCTGAAAGCGTGCAGTCTGAG-TTGATTGAATGCA-ATCAGTTAaaactttcaacaatggatctcttggttcoggc
A. wentii
A. nidulans  ----  --  -----          CC ----   A A    C   g --
```

Fig. 1. Sequences of ITS1 from Aspergillus awamori, A. wentii and A. nidulans.
The 18S rRNA and 5.8S rRNA sequences are given in small letters. Only differences in nucleotide sequences are shown.

```
A. awamori   gggggggcatgcctgtccgagc-gtcattGCTGCCCTCAAGCCCGGCTTGTGTGTTGGGTCGCCGTCCCCCTCTCCGGG
A. wentii                          ct
A. nidulans                                                                T       C ----

A. awamori   GGGAC-GGCCCGAAAGGCAGCGGCGGCACCG-GTCCGATCCTCGAGCGTATGGGGCTTTGTCACATGCTCTGTAGGAT
A. wentii
A. nidulans      G                          T    G                              CC    -----

A. awamori   TGGCC-------GGCGCCTGCCGACGTTTTCCAACC--ATTCTTTCCAGGttgacctcggatcaggtagggatacccg
A. wentii                          --
A. nidulans  A GGCCGGCCG      A   G    C-      TT   T  CT
```

Fig. 2. Sequences of ITS2 from A. awamori, A. wentii and A. nidulans.
The 5.8S and 26S rRNA sequences are shown. – Denotes a deletion.

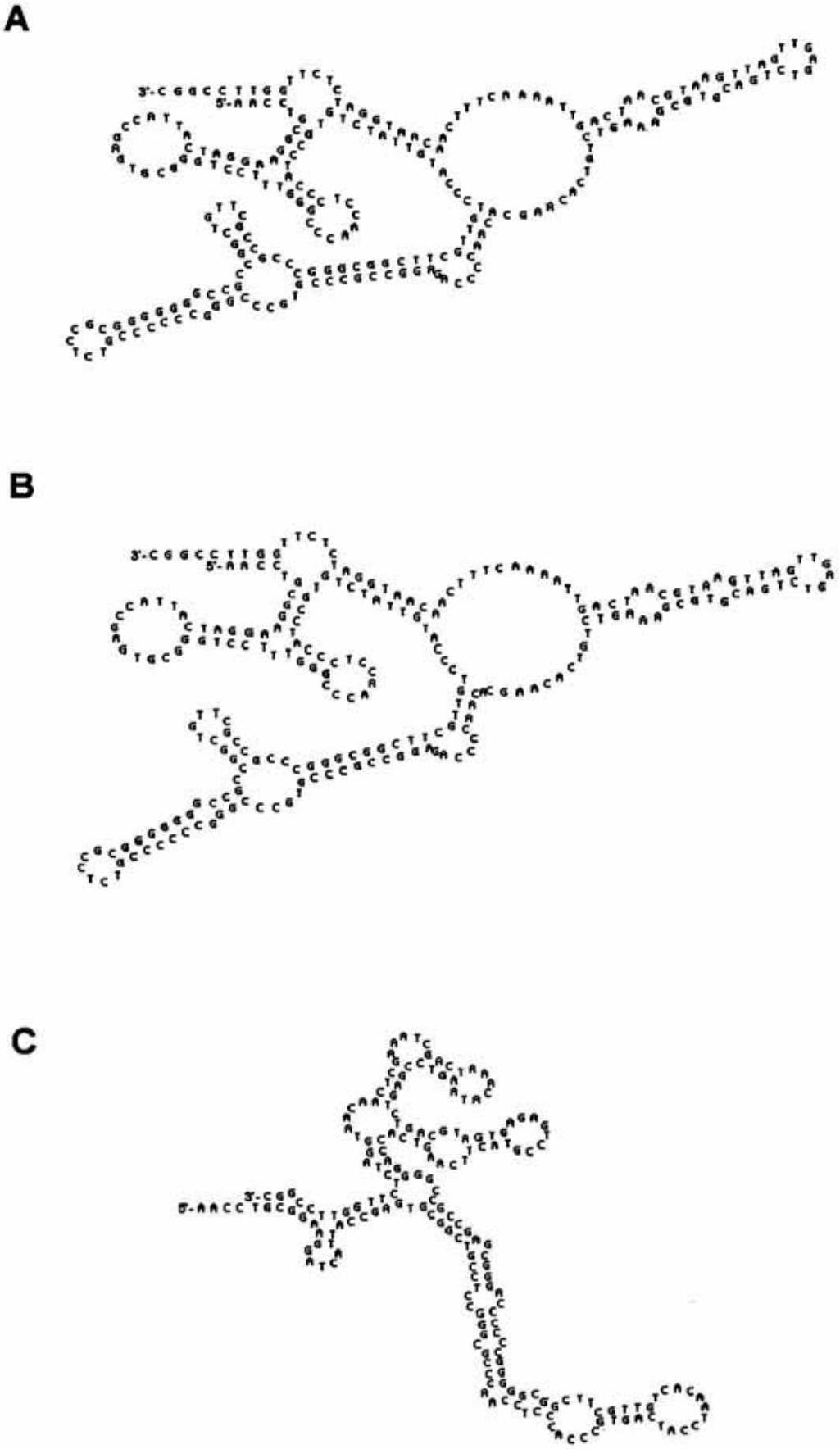

Fig. 3. ITS1 secondary structures for A. awamori (A), A. wentii (B) and A. nidulans (C).

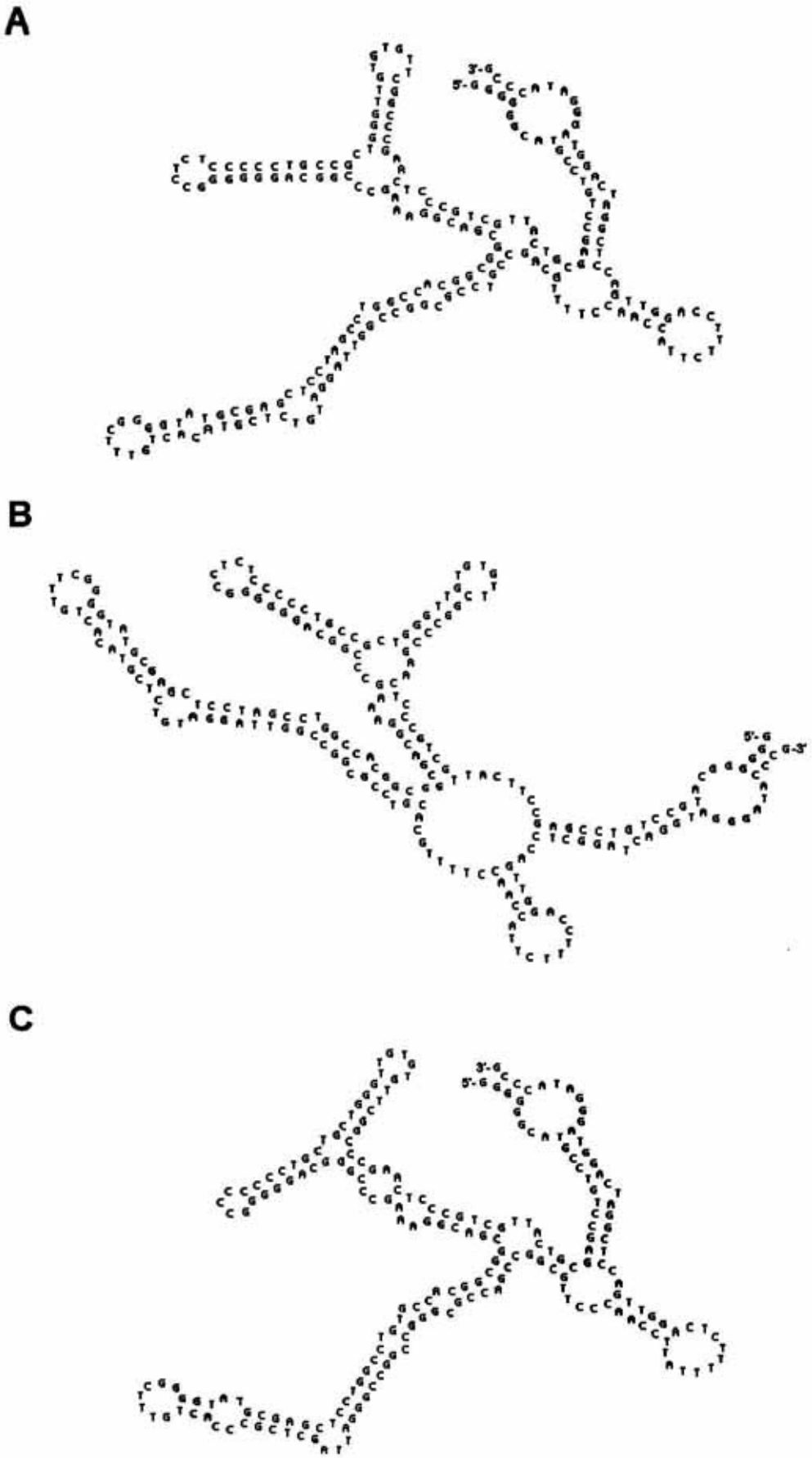

Fig. 4. ITS2 secondary structures for A. awamori (A), A. wentii (B) and A. nidulans (C).

ary structure, as the shorter *A. nidulans* ITS1 could only form two secondary structures, whereas *A. awamori* ITS1 sequences could form three. Extensive secondary structures in the ITS regions have been found in other species: in the yeast ITS2 [9] and in human ITS1 and ITS2 [10].

The existence of high degrees of sequence homology between the rRNA spacers in three species of *Aspergillus* indicates that these intergenic regions contain the necessary information for the processing of the rRNA precursor. Other studies of non-coding regions in the genus *Aspergillus* by us and by others: of the upstream region of the ornithine carbamoyltransferase gene [11] and the flanking regions of 5S rRNA genes [7] show very little or no homology.

The sequences studied here contain four different processing sites, needed for the generation of mature rRNA molecules. However, we did not discover any consensus sequences residing at the junctions of 18S, 5.8S and 26S rRNAs with their respective flanking sequences. Therefore we may assume that their processing must be dependent on signals other than junction sequences. These may, of course, be in the rRNA sequences themselves [4]. However, the detection of long stretches of sequence homologies among the different species of *Aspergillus* plus the existence of thermodynamically stable secondary structures within the ITS regions strongly suggests that the processing of pre-rRNA is guided by the three dimensional structure of the ITS sequences. Some enzymes, such as bacterial RNAse P, are known to recognize RNA structure rather than sequence, and this phenomenon may be quite common.

REFERENCES

1. Sollner-Webb, B. & Tower, J. (1986) Transcription of cloned eukaryotic ribosomal RNA genes. *Annu. Rev. Biochem.* **55**, 801–830.
2. Grummt, I. (1989) Mammalian ribosomal gene transcription; in *Nucleic Acids and Molecular Biology*, pp. 148–163, Springer Verlag, Berlin.
3. Kass, S., Tyc, K., Steitz, J. & Sollner-Webb, B. (1990) The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing. *Cell* **60**, 897–907.
4. Fournier, M.J. & Maxwell, E.S. (1993) The nucleolar snRNAs: catching up with the spliceosomal snRNAs. *Trends Biochem. Sci.* **18**, 131–135.
5. Peculis, B.A. & Steitz, J. (1993) Disruption of U8 nucleolar snRNA inhibits 5.8S and 28S rRNA processing in the *Xenopus* oocyte. *Cell* **73**, 1233–1245.
6. Borsuk, P., Nagięć, M., Stępień, P.P. & Bartnik, E. (1982) Organization of the rRNA gene cluster in *Aspergillus nidulans*. *Gene* **17**, 147–152.
7. Gniadkowski, M., Fiett, J., Borsuk, P., Hoffman-Zacharska, D., Stępień, P.P. & Bartnik, E. (1991) Structure and evolution of 5S rRNA genes and pseudogenes in the genus *Aspergillus*. *J. Mol. Evol.* **33**, 175–178.
8. Bartnik, E., Bartoszewski, S., Borsuk, P. & Empel, J. (1986) *Aspergillus nidulans* 5S rRNA genes and pseudogenes. *Curr. Genet.* **10**, 453–457.
9. Yeh, L.C.C. & Lee, J.C. (1990) Structural analysis of the internal transcribed spacer 2 of the precursor ribosomal RNA from *Saccharomyces cerevisiae*. *J. Mol. Biol.* **211**, 699–712.
10. Gonzales, I.L., Chambers, C., Gorski, J.L., Stambolian, D., Schmickel, R.D. & Sylvester, J.E. (1990) Sequence and structure correlation of human ribosomal transcribed spacers. *J. Mol. Biol.* **212**, 27–35.
11. Buxton, F.P., Gwynne, D.I., Garven, S., Sibley, S. & Davies, R.W. (1987) Cloning and molecular analysis of the ornithine carbamoyltransferase gene of *Aspergillus niger*. *Gene* **60**, 255–266.