Schizosaccharomyces U6 genes have a sequence within their introns that matches the B box consensus of tRNA internal promoters

David Frendewey*, Imre Barta, Marc Gillespie and Judith Potashkin+
Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, NY 11724, USA

Received January 16, 1990; Revised and Accepted March 1, 1990

ABSTRACT

The gene for the U6 small nuclear RNA (snRNA) in the fission yeast *Schizosaccharomyces pombe* is interrupted by an intron whose structure is similar to those found in messenger RNA precursors (pre-mRNAs) (1). This is the only known example of a split snRNA gene from any organism—animal, plant, or yeast. To address the uniqueness of the *S. pombe* U6 gene, we have investigated the structures of the U6 genes from five *Schizosaccharomyces* strains and three other fungi. A fragment of the U6 coding sequence was amplified from the genomic DNA of each strain by the polymerase chain reaction (PCR). The sizes of the PCR products indicated that all of the fission yeast strains possess intron-containing U6 genes; whereas, the U6 genes from the other fungi appeared to be uninterrupted. The sequences of the *Schizosaccharomyces* U6 gene fragments revealed that each had an intron of approximately 50 base pairs in precisely the same position. In addition to the splice sites and putative branch point regions, a sequence immediately upstream of the branch point consensus was found to be conserved in all of the *Schizosaccharomyces* U6 genes. This sequence matches the consensus for the B box of eukaryotic tRNA promoters. These results raise the interesting possibility that synthesis of U6 RNA in fission yeast might involve the use of internal promoter elements similar to those found in other genes transcribed by RNA polymerase III.

INTRODUCTION

The small nuclear ribonucleoproteins (snRNPs) are ubiquitous components of eukaryotic cells (for reviews see ref. 2). These particles are usually composed of a single small nuclear RNA (snRNA) and a set of associated proteins; however, the U4/6 snRNP contains two snRNAs in a base-paired interaction (3–6). Members of the most thoroughly studied class of snRNPs are known as U-snRNPs. Approximately one dozen U-snRNPs have been characterized in vertebrate cells (7–10) and perhaps 20–30 snRNAs of this class have been identified in fungi and plants (11–13). Some of the proteins in the particles are shared among a group of snRNPs. For example, the U1, U2, U4/6, U5, U7, U11 and U12 snRNPs all possess a set of small proteins that together comprise the antigen recognized by the Sm autoantibody often expressed by patients suffering from connective tissue disorders (reviewed in ref. 14). The nucleolar snRNPs U3, U8 and U13 have at least one common protein (10), and the RNase P and RNase MRP snRNPs appear to share protein components (15,16).

All of the snRNPs have been implicated in the catalysis of RNA processing reactions. The U1, U2, U4/6 and U5 snRNPs are constituents of the spliceosome, the large ribonucleoprotein complex that catalyzes the splicing of messenger RNA precursors (pre-mRNAs), and each is required for pre-mRNA splicing (reviewed in refs. 17–23). The U1, U2, U4 and U5 RNAs possess a number of common structural features (7). Each of these RNAs has an unusual 2,2,7-trimethylguanosine (m_3G) 5′ 'cap' and an internal conserved sequence that serves as the binding site for the proteins that form the Sm antigen. In vertebrates the genes for these RNAs have virtually identical promoter and termination signals (24) and they are transcribed by RNA polymerase II (Pol II). The U6 RNA does not share these characteristics. It has neither an an Sm site nor an m_3G cap. The cap structure determined for the U6 RNA from HeLa cells is a unique monomethyl phosphate ester on the γ-phosphate of the encoded guanosine triphosphate that initiates transcription (25). In addition, although U6 genes share some of the promoter elements found in the genes for U1-U5 RNAs (24), U6 RNA is synthesized by Pol III (26–29).

In the fission yeast *Schizosaccharomyces pombe* the U6 RNA presents an additional peculiarity; its gene is interrupted by an intron whose structure is similar to those found in *S. pombe* pre-mRNAs (1). The *S. pombe* U6 RNA is the only known example of an snRNA that is produced from an intron-containing precursor. Unspliced U6 precursor (pre-U6 RNA) accumulates in heat shocked *S. pombe* cells (1) and it can be detected under other growth conditions as well (30). The fact that U6 RNA levels

* To whom correspondence should be addressed

+ Present address: Department of Pharmacology and Molecular Biology, University of Health Sciences, The Chicago Medical School, 333 Green Bay Road, North Chicago, IL 60064, USA

are reduced (31) and splicing of the pre-U6 RNA is impaired (30) in *S. pombe* mutants that are defective in pre-mRNA splicing suggests a common mechanism for the splicing of the pre-U6 RNA and pre-mRNAs.

The presence of a pre-mRNA-type intron in the gene for U6 RNA, an RNA that participates in the removal of introns, raises several intriguing questions concerning the origin of this intron and the reason for its maintenance in *S. pombe* (32). To investigate how prevalent intron-containing U6 genes are in the fungi, we have investigated the structure of the U6 genes in four other strains of *Schizosaccharomyces* and in three more distantly related fungal species. We have identified U6 introns in all of the *Schizosaccharomyces* but not in the other fungi. Sequence analysis of the *Schizosaccharomyces* introns revealed a conserved element that is very similar to the B box of tRNA-type Pol III internal promoters (33,34).

MATERIALS and METHODS

Fungal strains and growth conditions

The wild type *Schizosaccharomyces pombe* used for these studies was strain 972 introduced by U. Leupold. The *Schizosaccharomyces* strains *S. malidevorans*, *S. octosporus*, *S. japonicus* var. *japonicus*, *S. japonicus* var. *versatilis* and the more distantly related fungi *Endomyces geotrichum*, *Endomyces tetrasperma* and *Geotrichum* sp. were generously provided by S. Zimmerly and D. Söll, Yale University. All strains were grown in YE (0.5% yeast extract, 3% glucose), which is the standard *S. pombe* culture medium. Normal growth was at 30°C with aeration.

For the heat shock experiments, cultures were grown to mid-logarithmic phase. This was taken to be $1-2\times10^7$ cells/ml (as determined by visual cell counts under the microscope with the aid of a hemacytometer) for the fission yeast strains. Since some of the fungal strains are filamentous and others tend to aggregate, accurate cell numbers were difficult to determine in all cases. For these strains the growth stage was estimated by comparing optical density with that of the fission yeast cultures. One half of each culture was shifted to 43°C for 30 minutes (min), while the other half remained at 30°C. Cells were harvested by centrifugation, washed with sterile water and stored at −20°C.

Preparation of nucleic acids

Genomic DNA was isolated from 10 ml stationary cultures according to Moreno et al. (35) except that the an additional proteinase K (Sigma) digestion was included and the final DNA was spin-dialyzed through Sepharose CL6B (Pharmacia). RNA was isolated from frozen cells by glass bead breakage followed by phenol/chloroform extraction and ethanol precipitation (36).

Polymerase chain reactions

Fragments of the U6 genes were amplified from total, undigested genomic DNA by the polymerase chain reaction (PCR, 37). The oligodeoxynucleotide primers (Operon Technologies) used were U6HIN, 5'-CTCAAGCTTAAATTGAAACGATACAG-3', and U6BAM, 5'-CTTGGATCCGTGTCATCCTTGTGCAG-3'. The underlined bases are complementary to the *S. pombe* U6 gene (1) at positions 22−38 for U6HIN and 107−123 for U6BAM. U6HIN primes synthesis of the sense strand (with respect to the U6 RNA sequence) and U6BAM the antisense strand. The 9 noncomplementary bases at the 5' ends contain a *Hin*d III site for U6HIN and a *Bam*H I site for U6BAM.

The conditions for the PCRs were 10 mM Tris-HCl (pH 8.3 at room temperature), 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% (w/v) gelatin (Sigma, G2500), 0.2 mM of each deoxynucleotide triphosphate, 0.6 μg of each primer (final concentration = 0.7 μM), 100 ng of yeast DNA and 5 or 2.5 units of *Taq* or *Amplitaq*™ DNA polymerase (Perkin Elmer Cetus), respectively, in a final volume of 0.1 ml. The reactions were performed in a Perkin Elmer Cetus thermal cycler with the following protocol: 40 cycles of denaturation at 94°C for 1.5 min, annealing at 55°C for 3 min and polymerization at 72°C for 2 min followed by 10 min at 72°C after the final cycle. PCR products were analyzed on a $0.1\times10\times20$ cm nondenaturing 8% polyacrylamide gel (acrylamide:bisacrylamide, 30:1) in 0.5× TBE buffer (38).

Cloning of the PCR products

Standard molecular cloning procedures were used (38). One quarter of each PCR was fractionated on a preparative nondenaturing polyacrylamide gel (see above) and the DNA was visualized by ethidium bromide staining. The 120 bp fragments from the *Schizosaccharomyces* reactions were excised from the gel and eluted by soaking in 0.4 ml of 0.5 M ammonium acetate at 37°C for 16 hours. The DNA was recovered by ethanol precipitation, digested with *Hin*d III and *Bam*H I (New England Biolabs) and ligated to dephosphorylated *Hin*d III and *Bam*H I digested pBSM13 (Stratagene). The ligations were transformed into *E. coli* DH5 and positive clones were identified by restriction analysis.

DNA sequencing

Plasmid DNA from individual clones was purified by ion exchange chromatography (QIAGEN) from 3 ml cultures. The sequences of both insert strands were determined by the chain termination method (39) with the universal M13 forward and reverse primers and T7 DNA polymerase (Sequenase™ 2.0, United States Biochemicals) according to the manufacturer's instructions. Sites of strand-specific premature termination were resolved at higher temperature with *Taq* DNA polymerase (Perkin Elmer Cetus).

Primer extension reactions

Primer extension analysis of U6 RNA was performed with the U6 E2−2 primer as previously described (30).

Computer searches and sequence alignments

Similarity searches, sequence alignments and base composition analyses were performed with the IntelliGenetics Suite of programs.

RESULTS and DISCUSSION

No introns have been found in the U6 genes from a variety of animals and plants (7) and the budding yeast *Saccharomyces cerevisiae* (6). To investigate whether *S. pombe* is the only species to possess an intron-containing U6 gene, we decided to start our search with its close relatives. The strains investigated are listed in table I and include five members of the *Schizosaccharomyces* genus and three other fungal species. Descriptions of these yeasts can be found in Kreger-van Rij (40). For brevity in the text we will refer to the two varieties of *S. japonicus* by their varietal names. *S. malidevorans* has also been classified as a variety of *S. pombe* (41).

Table I. Yeast strains

Schizosaccharomyces	other species
S. pombe	*Endomyces geotrichum*
S. malidevorans	*Endomyces tetrasperma*
S. octosporus	*Geotrichum* sp.
S. japonicus var. *japonicus*	
S. japonicus var. *versatilis*	

Amplification of U6 gene sequences

Our strategy was to amplify a fragment of the U6 genes from genomic DNA prepared from each strain, clone the amplified fragments and determine their sequences. Figure 1A displays a portion of the *S. pombe* U6 gene sequence (1). The oligodeoxynucleotide primers used for the PCR (37) are shown adjacent to their target sequences. The primers were designed to be complementary to regions of the U6 gene that are very highly conserved from yeast to mammals (1,6,7). To facilitate subsequent cloning of the amplified products, the primers included restriction sites at their 5′ ends (see Materials and Methods for the exact sequences of the primers). We expected that the PCR of the *S. pombe* DNA would produce a 120 base pair (bp) product (68 bp between the primers + 2×26 bp for each primer). Since the region between the primer target sites is 100% identical between *S. pombe* and rat (1), we expected a product from an intronless U6 gene to be smaller by 50 bp (the size of the *S. pombe* intron). The production of an amplified DNA fragment larger than 70 bp would be indicative of an intron somewhere within the region between the primers.

The results of the PCR analysis of the *Schizosaccharomyces* and other fungal strains are shown in figure 1B. A product of the expected size of 120 bp was amplified from the *S. pombe* DNA (lanes 2 and 14). PCR products of approximately 120 bp were also produced from the other *Schizosaccharomyces* strains (lanes 3−6). These results indicated that a U6 gene containing an intron of approximately 50 bp is present in all of the fission yeast strains. Given the phylogenetic conservation of the targeted sequences and the fact that only a single abundant PCR product was obtained from each DNA sample, the results also imply that an intronless copy of the U6 gene is not present in any of the *Schizosaccharomyces* species tested. The PCRs from each of the other fungal species (lanes 8−10) resulted in a single, abundant 70 bp fragment that co-migrated with the PCR product of the cloned human U6 gene (lane 15). Thus, these strains appear to have U6 genes that are not interrupted in the region between the primers. Introns outside of the targeted sequences would not be detected by this assay. Neither the 120 bp nor the 70 bp bands were produced in the control reactions (lanes 11−13).

RNA analysis

To confirm the PCR results we attempted to detect pre-U6 RNA in the *Schizosaccharomyces* strains. Tani and Ohshima (1) reported that a short heat shock at 43°C causes an accumulation of pre-U6 RNA in *S. pombe*. We have also observed that a milder temperature shift produces a transient increase in pre-U6 RNA levels (30). We therefore grew each of the strains to approximately mid-logarithmic phase at 30°C and then shifted a portion of each culture to 43°C for 30 minutes. RNA was isolated from the normal and heat shocked cells and analyzed by primer extension with an oligodeoxynucleotide that is complementary to the second exon (nucleotides 121−140, ref. 1) of the *S. pombe* U6 RNA (30).

The results for the fission yeast strains are shown in figure 2. A 140 nt reverse transcript indicative of the pre-U6 RNA (30) was produced from both the normal (barely visible in lane 1) and heat shocked (lane 2) *S. pombe* RNA samples; however, the precursor signal was significantly stronger after incubation at 43°C. A band of approximately the same size was obtained from the *S. malidevorans* (lanes 5 and 6), *S. versatilis* (lanes 7 and 8) and *S. octosporus* RNAs (lanes 9 and 10). In these strains heat shock also elicits an increased accumulation of the pre-U6 RNA. The *S. malidevorans* precursor is apparently slightly smaller than that for *S. versatilis*, which in turn is slightly smaller than the pre-U6 RNA detected in *S. octosporus* (lanes 6 and 8−10). These relative sizes are consistent with the DNA sequencing results described below. No precursor is evident in *S. japonicus* (lanes 3 and 4), although mature U6 RNA (90 nt band) was detected. The shift to 43°C might not have provided a sufficient heat shock to *S. japonicus*, or this strain does not respond to heat shock in a manner similar to the other *Schizosaccharomyces* strains. The closely related strain *S. versatilis* also produces a weak pre-U6 RNA signal (lane 8). The same heat shock experiment was performed on the related fungal species. In this case temperature dependent reverse transcripts larger than that expected for mature U6 RNA were not observed (data not shown). Northern blot analysis of the RNAs from the temperature shift experiment confirmed that pre-U6 RNA is present after heat shock in *S. pombe*, *S. malidevorans*, and *S. octosporus* (data not shown).

Sequences of the fission yeast U6 gene fragments

The PCR products from the *Schizosaccharomyces* strains were cloned in the pBSM13 plasmid and sequenced by the chain termination method (39) with primers complementary to the vector. The sequences of the sense strands, with respect to the RNA, are shown in figure 3A. The introns and two bases in the exons adjacent to the splice junctions are displayed. The sequences between the ends of the primers and the intron (not shown) were identical for all the clones and matched the published *S. pombe* U6 sequence (1), as did the intron sequence from our *S. pombe* PCR clone.

The first aspect of the sequences to note is that the position of the intron is identical in all the strains. Brow and Guthrie (32) proposed that the position of the intron might indicate a site in the mature U6 RNA that is important for the function of U6 RNA in the pre-mRNA splicing reaction. The conservation of U6 intron position supports this idea. However, the *Schizosaccharomyces* strains are too closely related to justify a firm conclusion at this time. Another example of an intron-containing U6 gene outside of the *Schizosaccharomyces* genus is needed to address this hypothesis.

In addition to its position, the size of the intron is nearly conserved. Both *S. pombe* and *S. malidevorans* have 50 bp introns. In *S. octosporus* the intron is 52 bp as the result of an apparent insertion of 2 bp immediately downstream from the 5′ splice site. Two extra bases are also found in this same position in *S. japonicus* and *S. versatilis* as well as an apparent 1 bp deletion between the branch point and the 3′ splice site. These differences result in an intron length of 51 bp in these two closely related strains. We refer to the sequence differences as insertions and deletions because we have used *S. pombe* as the reference sequence. The positions of the insertions and deletions depend

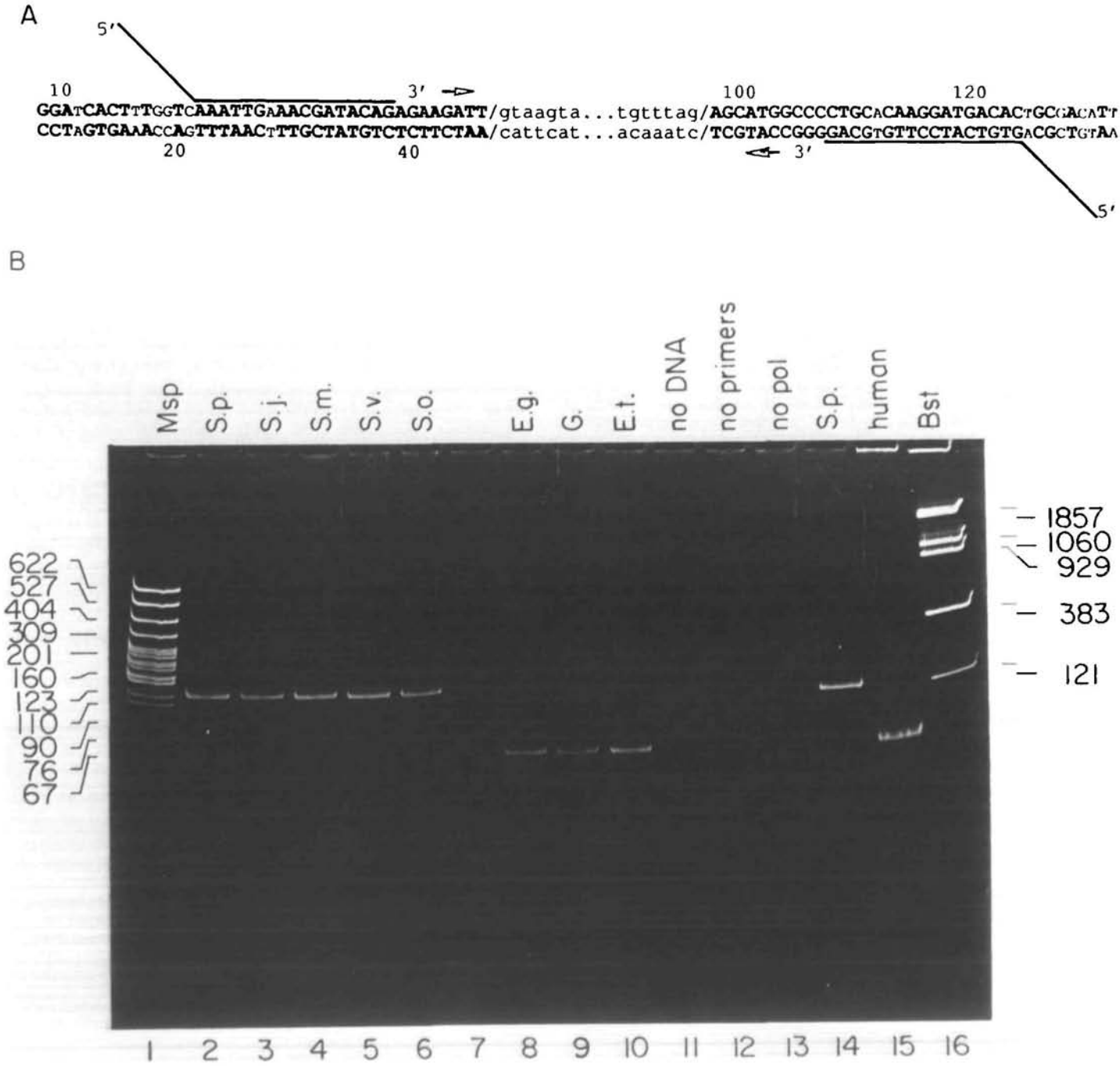

Figure 1. PCR amplification of U6 gene fragments. A. Primers and target sites. The PCR primers are represented by heavy lines aligned with the *S. pombe* U6 gene (1). The primers are identical to the U6 sequence to which they are aligned and include a restriction site at their 5′ ends, depicted as a displaced segment. Exon and intron bases are indicated by upper and lower case letters, respectively. Large upper case letters denote bases conserved between the *S. pombe* and rat U6 genes. The upper strand is the same sense as the U6 RNA with the 5′ end to the left. The numbering is with reference to the 5′ end of the RNA. Arrows show the directions of extension from the primers in the PCR. B. Analysis of the PCR products. One tenth of each PCR was fractionated by electrophoresis on an 8% nondenaturing polyacrylamide gel. Amplified DNA products were visualized by ethidium bromide. The source of the template genomic DNA is indicated above each lane by the first letters of the genus and species names (Table I). Four controls are shown: lane 11, no template; lane 12, no primers; lane 13, no DNA polymerase; and lane 15, cloned human U6 gene as template. DNA size markers were loaded in lane 1, *Msp* I digested pBR322, and lane 16, *Bst*N I digested pBR322. Sizes in base pairs are given on the left and right.

on how the sequences are aligned. The positions shown are based on the best alignment of all the sequences.

The similarity matrix shown in figure 3B indicates the relatedness of the *Schizosaccharomyces* strains with respect to their U6 sequences. *S. pombe* and *S. malidevorans* are very closely related; only 6 base differences were found, all of which lie outside of the splicing consensus regions. *S. japonicus* and *S. versatilis* are very similar to one another and diverge the most

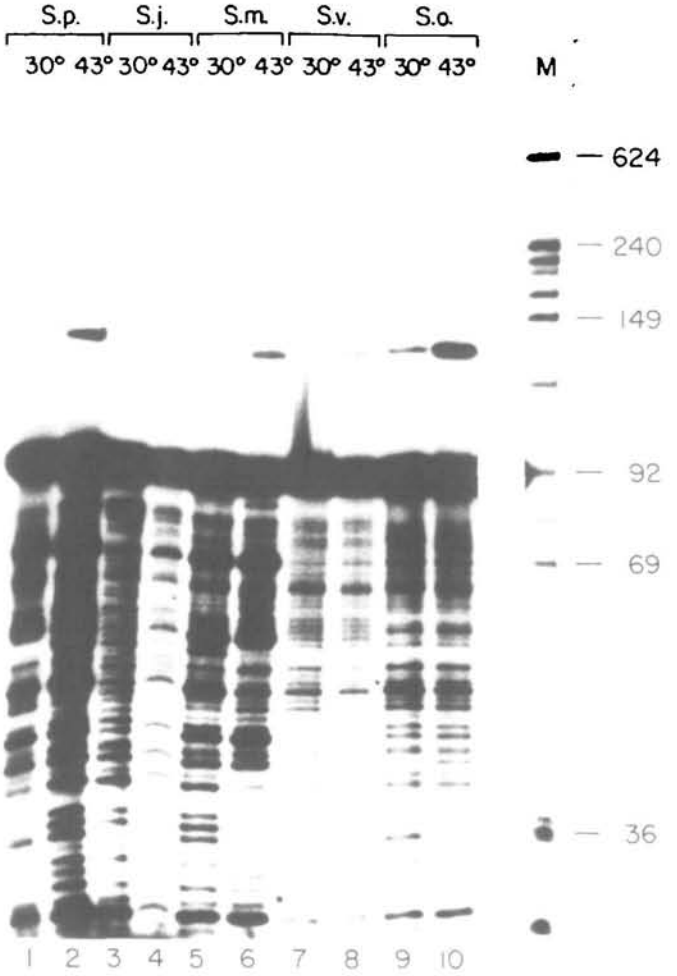

Figure 2. Effect of heat shock on pre-U6 RNA accumulation in the fission yeast strains. Whole cell RNA was prepared from normal (30°) and heat-shocked (43°) cells of each of the *Schizosaccharomyces* strains (indicated above the lanes) and analyzed by primer extension with a 5′-end-labeled primer complementary to the second exon of U6 RNA (30). Extension products were separated by electrophoresis on an 8% denaturing polyacrylamide gel and visualized by autoradiography. ^{32}P-labeled pBR322 *Msp* 1 digestion products were run as markers (M). The sizes in nucleotides of selected fragments are given on the right.

A

Alignment of the *Schizosaccharomyces* U6 Introns

```
SpU6    TT/GTAAGTA  ACAATATTTACCAAGGTTCGAGTCATACTAACTCGTTGTTTAG/AG
SmU6    TT/GTAAGTA  ACAATATTTACCAAGGCTTGAGCTATACTAACCTGTTGTTTAG/AG
SoU6    TT/GTAAGTACACCCCAGACGACCAAGGTTCGAGTCATACTAACTCGTTGTTTAG/AG
SjU6    TT/GTAAGTACCCCCCCTAAGGTCGTGAGTTCGAGTTCCACTAAC ACCCGTCTAG/AG
SvU6    TT/GTAAGTACTGCCTGGAGTACGTGAGTTCGAGTTCCACTAAC ACCCGTTTAG/AG

con     TT/GTAAGTAc--Cc-taa-taccaagGtTcGAGttatACTAACt-gttGTtTAG/AG
```

B

Similarity matrix:

	SpU6	SmU6	SoU6	SvU6	SjU6
SpU6		48	43	28	20
SmU6			37	20	12
SoU6				36	30
SvU6					46
SjU6					

Figure 3. Sequences of the fission yeast U6 introns. A. Alignments. Shown are the sense strands, with respect to the U6 RNA, of the introns and two bases in the flanking exons for the U6 genes of the five *Schizosaccharomyces* strains indicated by abbreviations to the left. The splice junctions are identified by slashes. Gaps were introduced to accommodate the best alignment of all the sequences. Identities between any two adjacent sequences are denoted by vertical bars. A consensus sequence (con) is given, with bold upper case letters designating positions of absolute conservation. The 5′ and 3′ splice sites and the branch point consensus sequence are underlined. B. Similarity matrix. Each sequence shown in A was compared with all others and the number of identical bases, according to the alignments in A, are given for each pair.

Table II. Base composition of *S. pombe* introns[a]

base	$\bar{x}$[b]	s[c]
A	29.5	6.0
C	13.6	3.5
G	16.6	2.6
T	39.7	6.1
A+T	69.2	
G+C	30.2	

[a]Base compositions are derived from 24 pre-mRNA introns in 11 published *S. pombe* gene sequences. [b]$\bar{x}$ = mean. [c]s = standard deviation.

from the *S. pombe* sequence.. The *S. octosporus* intron shares features with those of *S. japonicus* and *S. versatilis*, such as the 2 bp insertion near the 5′ end of the intron, but, overall, *S. octosporus* is more similar to *S. pombe* than to any of the other strains. The sequence similarities of the U6 introns are consistent with the classification of the *Schizosaccharomyces* proposed by Sipiczki (41), that is, three species—*pombe, octosporus* and *japonicus*—with *malidevorans* a variety of *pombe* and *versatilis* a variety of *japonicus*.

The peculiarity of an intron in a U6 gene raised the possibility that the splicing of the pre-U6 RNA might require special factors or be part of a regulatory pathway. One of the reasons for sequencing the U6 genes from the other *Schizosaccharomyces* species was that it might identify conserved intronic elements that are specific for pre-U6 RNA splicing. A consensus U6 intron sequence is shown in figure 3A. The bases that were identical in all strains are given in bold upper case letters. As expected, the highest degree of similarity among the sequences is at the 5′ and 3′ splice sites and the putative branch point region. Seven intron bases, GTAAGTA, at the 5′ splice site are absolutely conserved in all the strains. The 3′ A in this stretch is not part of the 5′ splice site consensus for *S. pombe* introns (42,43). All the strains have the sequence TAG at their 3′ splice sites. This is the consensus 3′ splice site triplet found in 75% of all known *S. pombe* pre-mRNA introns (D. Frendewey, unpublished). The putative branch point region, TACTAAC, is also 100% conserved in the U6 introns except for the 5′ T, which is a C in *S. japonicus* and *S. versatilis*. Apart from the three splicing consensus regions, three other sites exhibit absolute conservation. The conserved C near the 5′ splice site must be considered tentative since the alignment in this region is not certain. A GT is found in all the U6 genes just upstream of the 3′ splice site. The T (position −5 from the 3′ splice site) is expected; over half of all *S. pombe* pre-mRNA introns have a T at this position (D. Frendewey, unpublished). The neighboring G is rare at this position, occurring in less than 20% of *S. pombe* introns.

Figure 4. A and B box similarities in the *S. pombe* U6 gene. A. Matches between a U6 intron sequence and other *S. pombe* genes. An 11 base segment of the *S. pombe* U6 gene (shown in bold above the sequences in the left and right columns) corresponding to the region of internal sequence conservation in the fission yeast U6 introns (Fig. 3A) was compared against all *S. pombe* entries in the EMBL and GENBANK data bases. Sequences exhibiting the best matches (9 or 10 out of 11) to the U6 fragment are shown The genes to which these sequences belong are indicated beneath each match. Bases identical to the U6 sequence are denoted by a vertical bar. Numbering is with reference to the 5' end of the RNA for the left column and to the start of the protein coding sequence for the right column. B. B box similarity. The eukaryotic tRNA B box consensus (33,34) (n means any base) is shown aligned with an 11 base segment of the U6 intron consensus (Fig. 3A). Bars indicate matches and bold letters denote absolutely conserved bases. C. A box similarity. The eukaryotic tRNA A box consensus (33,34) is shown aligned with the 7SL RNA A box (56,57) and a sequence in the first exon of the U6 RNA. Numbering is from the 5' ends of the RNAs. Symbols are as in C. R = purine; Y = pyrimidine. A dot denotes length heterogeneity.

Homology to the Pol III B box

The most interesting conservation that was revealed by the U6 sequence comparison is immediately 5' to the TACTAAC sequence. The consensus in this region of the *Schizosaccharomyces* U6 genes is agGtTcGAGtt with 5 bases (upper case) out of the 9 identical in all strains (figure 3A). The 3 absolutely conserved G's are unusual. The mean G composition for *S. pombe* introns is 16.6% (table II), so a G at any position except the splice sites is uncommon. However, the overall base composition of the U6 intron does not deviate significantly from the average composition for *S. pombe* pre-mRNA introns. The base compositions shown in Table II indicate that *S. pombe* introns are similar to those in plants in having a high A+T content (44). In plants the A+T bias is essential for optimal splicing efficiency (44).

A visual inspection of the region just upstream of the branch point consensus in 24 published *S. pombe* pre-mRNA introns indicated that this sequence is not a common feature of *S. pombe* introns. We also performed a computer search of the *S. pombe* files in the GENBANK and EMBL sequence collections asking for matches to the *S. pombe* U6 intron sequence AGGTTCGAGTC. Some of the results are shown in figure 4A. No significant similarities were found to any of the *S. pombe* pre-mRNA introns in the data bases. Surprisingly, the best matches (10 out of 11 identity) were to the B box promoter elements in the 7SL RNA and $tRNA^{His}$, $tRNA^{Arg}_1$ and $tRNA^{Arg}_2$. Other tRNAs were identical to the U6 intron sequence at 8 or 9 positions in their B box regions. The B box consensus is GGTTCGANNCC (Fig. 4B), where the underlined bases are invariant in eukaryotic tRNA genes and N means any base (33,34). The 4 invariant bases are found in the *S. pombe* U6 intron sequence. Three of these, G, T and A, are conserved in all the *Schizosaccharomyces* strains; whereas the fourth, C, is present in all the strains except *S. malidevorans*. The U6 intron consensus in the region of the B box similarity matches the B box consensus at 7 out 9 unambiguous positions (Fig. 4B).

Besides the genes for 7SL RNA and tRNAs, which are Pol III transcripts, good matches (9 out of 11) were found in the protein coding sequences of two Pol II genes, *wee1*$^+$ and *mei2*$^+$ (figure 4A). The *wee1*$^+$ protein is involved in the regulation of mitosis (45) and the *mei2*$^+$ gene product is required for initiation of meiosis (46). There is no clear connection known between these genes and U6 RNA or any other Pol III transcript. Therefore, the similarities to these sequences might be chance matches.

Implications for U6 RNA synthesis in fission yeast

U6 RNA is transcribed by RNA polymerase III in vertebrates (26,27), insects (28) and budding yeast (29). The *S. pombe* U6 RNA is also likely to be a Pol III transcript (1), and recent experiments demonstrate that the *S. pombe* U6 gene can be transcribed by Pol III in a HeLa cell extract (47). The U6 genes of vertebrates and *Drosophila* have a region in the middle of their coding sequences that is similar to the A box of Pol III internal promoters (24,26,28). However, unlike other Pol III genes, internal sequences are not required for efficient transcription of vertebrate U6 genes (48–50). This suggests that the putative A box in U6 genes may not be required for transcription. Furthermore, the A box similarity is not conserved in the *S. cerevisiae* U6 gene (6) and it is split by the intron in *S. pombe* (1). All of the essential promoter elements in vertebrate U6 genes appear to be in the 5' flanking sequence (48,51–53). These motifs include an enhancer-like element (the distal element) greater than 200 bp upstream of the transcription initiation site and another conserved sequence (the proximal element) at approximately 60 bp from the coding region. Both of these elements are found in the promoters of the U1-U5 snRNAs, which are transcribed by Pol II (reviewed in 24). U6 genes have an additional promoter element, a T/A rich sequence, 30 bp upstream of the transcription start that is not a part of the other snRNA promoters. The T/A motif is required for the transcription of U6 RNA by Pol III. Its presence in a Pol II snRNA promoter switches transcription to Pol III (49,54). The budding and fission yeast U6 genes have an A/T rich sequence at the same position as in the vertebrate homologs. In addition, a sequence similar to the 3' half of the proximal element has been identified in the *S. cerevisiae* U6 gene (6). The requirement of these sequences for U6 transcription in yeast has not yet been demonstrated.

Transcription of U6 RNA requires at least some factors that are different than those involved in 5S and tRNA synthesis (29,55). The presence of a B box element in the *S. pombe* U6

intron could imply that transcription of U6 RNA in fission yeast shares features with tRNA transcription. A and B box homologies have been identified in the gene for the *S. pombe* 7SL RNA (56,57), a Pol III transcript. It is interesting that the U6 B box is identical to the 7SL B box (figure 4A) for the bases that are part of the B box consensus. We might then expect to find an A box in the *S. pombe* U6 gene. As mentioned above, the A box homology previously noticed in higher eukaryotic U6 genes is interrupted by the intron in *S. pombe*. However, a sequence in the first exon (bases 17–27 of the RNA) also resembles an A box (figure 4C); in fact, it is a better fit to the consensus (33,34). This sequence matches the 7SL A box homology region at 10 out of 12 positions and it is in approximately the same location with respect to the 5′ end of the RNA (figure 4C). This part of the *S. pombe* U6 gene shows some divergence compared to the rat U6 (1,58), but a nearly identical stretch of bases (positions 29–39 in the RNA coding sequence) is found in the *S. cerevisiae* U6 gene (6). A corresponding B box has not been identified within the *S. cerevisiae* U6 coding sequence. The separation between the putative A and B boxes in the *S. pombe* U6 gene (approximately 40 bp) is the same as in the 7SL gene (56,57) and is within the range seen for tRNA genes (33,34).

Similarity to known promoter elements is, of course, not proof of function, as the case of the vertebrate U6 A box homology demonstrates. However, the regions we have identified in the *S. pombe* U6 gene might be good targets for site directed mutagenesis experiments to define elements required for transcription. The U6 gene structure in *Schizosaccharomyces* suggests that transcription could be controlled by a mechanism that is different than the vertebrate model. It is perhaps significant in this regard that the human U6 promoter does not appear to be active in *S. pombe* (S. Lobo, N. Hernandez and D. Frendewey, unpublished results). The transcription of the *S. pombe* U6 gene by Pol III in a HeLa cell extract and in 293 cells (47) might have resulted from recognition of the A and B box similarities as promoter elements by tRNA-type transcription factors. We would predict that removal of the intron from the *S. pombe* U6 gene will severely reduce transcription efficiency both in the *S. pombe* and human systems.

Our discussion of the B box similarity in the *Schizosaccharomyces* U6 introns has focussed on its possible role in transcription. Alternatively or in addition, this conserved sequence could function at the RNA level. Although a similar sequence was not found in any of the known *S. pombe* pre-mRNA introns, the B box similarity in the U6 intron could be a signal required for the splicing of the pre-U6 RNA. For example, this region might bind a U6-specific splicing factor or a protein that helps to target this Pol III transcript to the spliceosome.

Evolution and maintenance of a U6 intron in fission yeast

The absence of introns in the known snRNA genes implies that introns in these genes could impart a selective disadvantage. Why then is a U6 intron maintained in the *Schizosaccharomyces*? The presence of the B box in the intron offers a possible explanation. If the ancestral *Schizosaccharomyces* U6 gene was transcribed by Pol III, perhaps under the control of external promoters, then the insertion of an internal Pol III promoter element might have improved Pol III transcription of this gene. The constraint against such an insertion is that it would disrupt the highly conserved U6 RNA sequence. Bringing the promoter element in as part of an intron, which will be spliced out of the pre-U6 RNA, obviates this problem. The putative A box in the first exon may have already existed in the original *Schizosaccharomyces* U6 gene since it is present in the *S. cerevisiae* U6 gene. Alternatively, the A box sequence might have been selected from mutations that occurred after the insertion of the intron. The 3′ half of the A box similarity is highly conserved from yeast to mammals (1). The 5′ half lies just outside of the conserved region, suggesting that sequence variations are tolerated in this part of the U6 RNA.

Brow and Guthrie (32) hypothesized that the *S. pombe* U6 intron may have arisen by a mishap in pre-mRNA splicing. It would then follow that an intron similar to that in the U6 gene resides in a protein coding gene. We did not find a pre-mRNA intron that shows a significant resemblance to the U6 intron by a computer search of the available *S. pombe* sequences. Southern analysis also did not reveal any related DNA (I. Barta and D. Frendewey, unpublished results). However, except for the bases required for splicing, the intron sequences would be expected to have drifted considerably since the insertion event. Thus, the origin of the U6 intron may remain a mystery and the mechanism by which it arose is still an open question. But if the apparent B box in the U6 intron proves to be an essential internal promoter element, then a plausible explanation for the maintenance of this intron in *S. pombe* will have been found.

ACKNOWLEDGEMENTS

We would like to thank G. Das and N. Hernandez for critical comments on the manuscript, the members of the Cold Spring Harbor Splicing Group for helpful suggestions and G. Kunkel for the human U6 clone. J. Duffy and P. Renna are thanked for the preparation of the figures.

This work was supported by an NIH grant (GM38242) to D. Frendewey.

REFERENCES

1. Tani, T. and Ohshima, Y. (1989) Nature **337**, 87–90.
2. Birnstiel, M.L. (1988) (ed) Structure and function of major and minor small nuclear ribonucleoprotein particles, Springer-Verlag, Berlin.
3. Bringmann , P., Appel, B., Rinke, J., Reuter, R., Theissen, H. and Lührmann, R. (1984) EMBO J. **3**, 1357–1363.
4. Hashimoto, C and Steitz, J.A. (1984) Nucl. Acids Res. **12**, 3283–3293.
5. Rinke, J., Appel, B., Digweed, M. and Lührmann, R. (1985) J. Mol. Biol. **185**, 721–731.
6. Brow, D.A. and Guthrie, C. (1988) Nature **334**, 213–218.
7. Reddy, R. and Bush, H. (1988) In M. L. Birnstiel (ed), Structure and function of major and minor small nuclear ribonucleoprotein particles, Springer-Verlag, Berlin, pp.1–37.
8. Krämer, A. (1987) Proc. Natl. Acad. Sci. USA **84**, 8408–8412.
9. Motzka, K.A. and Steitz, J.A. (1988) Proc. Natl. Acad. Sci. USA **85**, 8885–8889.
10. Tyc, K. and Steitz, J.A. (1989) EMBO J. **8**, 3113–3119.
11. Riedel, N., Wise, J.A., Swerdlow, H., Mak, A. and Guthrie, C. (1986) Proc. Natl. Acad. Sci. USA **83**, 8097–8101.
12. Tollervey, D. (1987) J. Mol. Biol. **196**, 355–361.
13. Guthrie, C. (1988) In M. L. Birnstiel (ed), Structure and function of major and minor small nuclear ribonucleoprotein particles, Springer-Verlag, Berlin, pp.196–211.
14. Lührmann, R. (1988) In M. L. Birnstiel (ed), Structure and function of major and minor small nuclear ribonucleoprotein particles, Springer-Verlag, Berlin, pp.71–99.
15. Gold, H.A., Craft, J., Hardin, J.A., Bartkiewicz, M. and Altman, S. (1988) Proc. Natl. Acad. Sci. USA **85**, 5483–5487.
16. Gold, H.A., Topper, J.N., Clayton, D.A., Craft, J. (1989) Science **245**, 1377–1380.
17. Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S.R. and Sharp, P.A. (1986) Ann. Rev. Biochem. **55**, 1019–1050.
18. Green, M.R. (1986) Ann. Rev. Genet. **20**, 671–708.
19. Sharp, P.A. (1987) Science **235**, 766–771.

20. Krainer, A. and Maniatis, T. (1988) In B. D. Hames, D. M. Glover (eds), Frontiers in Molecular Biology: transcription and splicing, IRL Press, Oxford, pp.131–206.
21. Maniatis, T. and Reed, R. (1987) Nature **325**, 673–678.
22. Steitz, J.A., Black, D.L., Gerke, V., Parker, K.A., Krämer, A., Frendewey, D. and Keller, W. (1988) In M. L. Birnstiel (eds), Structure and function of major and minor small nuclear ribonucleoprotein particles, Springer-Verlag, Berlin, pp.115–154.
23. Guthrie, C. and Patterson, B. (1988) Annu. Rev. Genet. **22** 387–419.
24. Dahlberg, J.E. and Lund, E. (1988) In M. L. Birnstiel (eds), Structure and function of major and minor small nuclear ribonucleoprotein particles, Springer-Verlag, Berlin, pp.38–70.
25. Singh, R. and Reddy, R. (1989) Proc. Natl. Acad. Sci. USA **86**, 8280–8283.
26. Kunkel, G.R., Maser, R.L., Calvet, J.P. and Pederson, T. (1986) Proc. Natl. Acad. Sci. USA **83**, 8575–8579.
27. Reddy, R., Henning, D., Das, G., Harless, M. and Wright, D. (1987) J. Biol. Chem. **262**, 75–81.
28. Das, G., Henning, D. and Reddy, R. (1987) J. Biol. Chem. **262**, 1187–1193.
29. Moenne, A., Camier, S., Anderson, G., Margottin, F., Beggs, J. and Sentenac, A. (1990) EMBO J. **9**, 271–277.
30. Potashkin, J. and Frendewey, D. (1989) Nucleic Acids Res. **17**, 7821–7831.
31. Potashkin, J., Li, R. and Frendewey, D. (1989) EMBO J. **8**, 551–559.
32. Brow, D.A. and Guthrie, C. (1989) Nature **337**, 14–15.
33. Sharp, S., DeFranco, D., Dingermann, D., Farrell, P. and Söll, D. (1981) Proc. Natl. Acad. Sci. USA **78**, 6657–6661.
34. Hofstetter, H., Kressman, A. and Birnstiel, M.L. (1981) Cell **24**, 573–585.
35. Moreno, S., Klar, A. and Nurse, P. Methods Enzymol. (in press).
36. Nischt, R., Thuroff, E. and Käufer, N.F. (1986) Curr. Genet. **10**, 365–370.
37. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R.G., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Science **239**, 487–491.
38. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
39. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463–5467.
40. Kreger-van Rij, N.J.W. (1984) (ed) The Yeasts, a taxonomic study, 3rd ed., Elsevier, Amsterdam.
41. Sipiczki, M. (1989) In A. Nassim, P. Young and B.F. Johnson (eds) Molecular Biology of the Fission Yeast, Academic Press, Florida, pp.431–452.
42. Käufer, N.F., Simanis, V. and Nurse, P. (1985) Nature **318**, 78–80.
43. Mertins, P. and Gallwitz, D. (1987) EMBO J. **6**, 1757–1763.
44. Goodall, G.J. and Filipowicz, W. (1989) Cell **58**, 473–483.
45. Russell, P. and Nurse, P. (1987) Cell **49**, 559–567.
46. Watanabe, Y., Iino, Y., Furuhata, K., Shimoda, C. and Yamamoto, M. (1988) EMBO J. **7**, 761–767.
47. Kleinschmidt, A.M., Pederson, T., Tani, T. and Ohshima, Y. (1990) J. Mol. Biol. **211**, 7–9.
48. Das, G., Henning, D., Wright, D. and Reddy, R. (1988) EMBO J. **7**, 503–512.
49. Lobo, S.M. and Hernandez, N. (1989) Cell **58**, 55–67.
50. Kunkel, G.R. and Pederson, T. (1989) Nucleic Acids Res. **17**, 7371–7379.
51. Krol, A., Carbon, P., Ebel, J. and Appel, B. (1987) Nucleic Acids Res. **15**, 2463–2478.
52. Carbon, P., Murgo, S., Ebel, J.-P., Krol, A., Tebb, G. and Mattaj, I.W. (1987) Cell **51**, 71–79.
53. Kunkel, G.R. and Pederson, T. (1988) Genes Develop. **2**, 196–204.
54. Mattaj, I.W., Dathan, N.A., Parry, H.D., Carbon, P. and Krol, A. (1988) Cell **55**, 435–442.
55. Reddy, R. (1988) J. Biol. Chem. **263**,15980–15984.
56. Ribes, V., Dehoux, P. and Tollervey, D. (1988) EMBO J. **7**, 231–237.
57. Brennwald, P., Liao, X., Holm, K., Porter, G. and Wise, J.A. (1988) Mol. Cell. Biol. **8**, 1580–1590.
58. Epstein, P., Reddy, R., Henning, D. and Busch, H. (1980) J. Biol. Chem. **255**, 8901–8906.