Nucleic Acids Research, 1993, Vol. 21, No. 17 4079–4083

The fission yeast *prp4*$^{+}$ gene involved in pre-mRNA splicing codes for a predicted serine/threonine kinase and is essential for growth

Suresh K.Alahari, Henning Schmidt[1] and Norbert F.Käufer*
Department of Bioscience and Biotechnology, Drexel University, Philadelphia, PA 19104, USA and [1]Institut für Genetik, Biozentrum Technische Universität Braunschweig, Spielmannstrasse 7, W-38106 Braunschweig, Germany

Received April 27, 1993; Revised and Accepted July 7, 1993 GenBank accession no. L10739

ABSTRACT

Only four *prp* (pre-mRNA processing) genes of the fission yeast *Schizosaccharomyces pombe* have been reported. We exploited yeast genetics and identified and isolated the *prp4* gene. Sequence analysis revealed that the splicing factor encoded by this gene contains the signature sequences that define the serine/threonine protein kinase family. This is the first kinase gene identified whose product is involved in pre-mRNA splicing. The *prp4* gene contains one intron in the kinase domain. Gene replacement studies provided evidence that this gene is essential for growth and is located on chromosome III.

INTRODUCTION

Splicing of pre-mRNA takes place in a multi-component complex which has been termed the spliceosome. Intron recognition, spliceosome assembly and the splicing reaction are dependent on many factors that interact with each other and with the pre-mRNA. Major components of this dynamic process are the small nuclear ribonucleoprotein particles (snRNP U1, U2, U4/U6, U5) which consist of RNA and associated proteins. The U4/U6 snRNA molecule may play an important catalytic role in removal of the intron sequences (1, 2, 3, 5).

The pre-mRNA splicing process requires ATP (1, 4). One set of proteins consuming ATP during pre-mRNA splicing belong to the DEAD/DEAH family, which contain signature sequences assigned to RNA-dependent helicases with ATPase activity (4, 11, 35). DEAD/DEAH signature sequences have been found in several *PRP* (pre-mRNA- processing) genes isolated from the budding yeast *Saccharomyces cerevisiae* (6, 8, 11). In mammalian splicing systems, the activity of other ATP consuming enzymes has been detected. In *in vitro* experiments it has been shown that the spliceosome and isolated parts of the spliceosome display specific kinase activity (5, 9, 10, 39). Also, phosphoproteins have been identified as spliceosomal components (7, 36). Little, however, is known about protein kinases involved in the pre-mRNA splicing process.

Four temperature sensitive (ts) *prp* (pre-mRNA processing) mutants from the fission yeast *S.pombe* have been reported (12,13). We have described the identification and characterization of the *prp4* $^{(ts)}$ mutation (13). This mutant strain does not grow at the restrictive temperature (36°C) and accumulates pre-mRNA of intron containing genes, while spliced mRNA is rapidly degraded. At the permissive temperature (25°C), however, no pre-mRNA is detected and growth behavior is normal.

Here we describe the isolation and characterization of the gene and its predicted product complementing the *prp4* mutation. The gene complementing the *prp4* mutation is an essential gene and contains the signature sequences predictive for a serine/threonine protein kinase. This is the first report of a gene whose product is involved in pre-mRNA splicing that is also predicted to be a protein kinase.

MATERIALS AND METHODS

Schizosaccharomyces pombe strains and manipulations

Strains used in this study were h^{-S} (972), h^{+N} (975) (21). The mutant strain h^{-S}, *ura4-D18*, *leu1-32*, *prp4* $^{(ts)}$ strain has been described by Rosenberg *et al.* (13). Standard classical and molecular genetics procedures and media for growth of the *S.pombe* strains used in this study have been described by Gutz *et al.* (21) and Moreno *et al.* (19). The *S.pombe* genomic library cloned into the pUR shuttle vector containing the *Amp*^{r} gene as the bacterial marker and the *ura4* gene as the yeast marker was a gift from A.Carr (25). The cDNA library was from Becker *et al.* (26). Transformation of *S.pombe* with shuttle-plasmids and the *S.pombe* genomic library was performed as previously described by Gatermann *et al.* (15).

Plasmid recovery from *S.pombe* and double stranded sequencing

Shuttle-plasmids from *S.pombe* were isolated via transformation of isolated *S.pombe* DNA into *Escherichia coli* using a protocol described by A.C.Ward (37). Genomic inserts were mapped by restriction analysis and subsequently subcloned into *E.coli* vectors

* To whom correspondence should be addressed

(23). Sequencing was performed as previously described (22). The predicted amino acid sequence of the PRP4 protein, derived from the nucleotide sequence was compared with sequences in the GenBank data base using BLAST server (40).

Nucleic acid isolation and analysis

Standard methods of cloning, subcloning and manipulating DNA *in vitro* were used (22). Procedures of DNA and RNA isolation from *S.pombe*, Southern and Northern blotting, and PCR analyses were performed exactly as described previously (13, 14, 15, 24).

RESULTS

Complementation of the *prp4* mutation

To suppress the *prp4* mutation, we transformed the *prp4* strain with a partial genomic library constructed from a wild type *S.pombe* strain. After transformation the cells were incubated for two days at the permissive temperature (25°C) and then screened for growth at the restrictive temperature (36°C). We found several growing colonies. Plasmids were isolated from these colonies and retransformed into the *prp4* strain. Those inducing growth at the restrictive temperature were further characterized.

A 2.5 kb genomic fragment suppresses the *prp4* mutation

The shuttle vector pUR which was used to construct the genomic library contains a PstI and a KpnI restriction site in the multiple cloning site (25). Restriction analysis of the isolated shuttle plasmids revealed that we had recovered two different plasmids complementing the *prp4* mutation. One plasmid contained a 4.3 kb PstI/KpnI fragment; the second plasmid released a 2.5 kb PstI/KpnI genomic fragment. A more detailed restriction map of both inserts indicated that the 4.3 kb PstI/KpnI fragment included the 2.5 kb fragment. Therefore, the smaller fragment was further characterized. The shuttle plasmid containing this 2.5 kb DNA fragment was called pSAK.

The *prp4* strain containing pSAK splices the *dis2* gene

To demonstrate that the plasmid pSAK suppresses the splicing deficiency caused by the mutation in the *prp4* locus, we performed Northern analysis. Strain *prp4*$^{(ts)}$ transformed with the plasmid pSAK was grown at 25°C to mid log phase. RNA was isolated from parts of the culture. The remaining culture was shifted to 36°C for three hours. Then RNA was isolated. In the control experiment, the *prp4*$^{(ts)}$ strain was transformed with the plasmid pUR which did not contain genomic *S.pombe* DNA (25). Total RNA was isolated from cultures grown and incubated at temperatures as described before. The RNA was electrophoresed and transferred to Nitrocellulose. The RNA-blot was probed with a labeled DNA fragment comprising the *dis2* gene. The *dis2* gene contains a 395 bp intron (33, 34). Fig.1

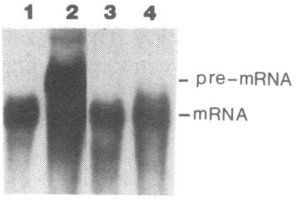
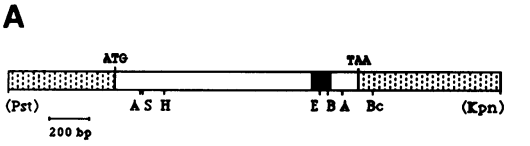


Figure 1. Northern analysis to demonstrate that the splicing defect of the *prp4* mutation is suppressed by a 2.5 kb genomic fragment. Total RNA isolated from the *prp4* strain grown at 25°C (lane 1) and then shifted and incubated for 3 h at 36°C (lane 2) and total RNA isolated from cells containing pSAK which rescues the *prp4* mutation (lane 3, RNA from cells grown at 25°C and lane 4, RNA from cells grown at 36°C) was separated on an 1.2% agarose gel and transferred on to nitrocellulose. The blot was hybridized to a 600 bp radio labeled HindIII/XbaI fragment which comprises exon1, the 396 bp intron and parts of exon2 of the *dis2* gene from *S.pombe* (33, 34). In this figure mRNA and pre-mRNA refer to spliced and unspliced transcripts of the *dis2* gene.



B

```
-556 ATTGCTAGTGATCCAAGAGCGTATGCAGTTCGATCTGGGAGGTTGCTAATGTATCGGTAATTGATTCTCTTAAAAATCA
-467 TTTAGAAAGAGCAACGTTACGCAGAGAAACGTTCTGAAGATGAACGTTGGTCTAGTATTTTTCAATGATTTCTCGTATG
-388 CGTTTATATCATGTCTTACTCCCTAATATACTTCTAATAAAAATTTTTGTATTACTAATTTATTCATTTTTCTCTTGACA
-309 GCTATATTTCAGGTTCTTAGATTTTTTTCTAATAGAAACTGCTTACCAAACTTTTTTTTCAACTCGTACTTGAAGCTAC
-230 GATCATAGCCACTTCAGTCCGTATTCCTAATATATTGAATTTTTATTCGAAAATTTGCATTTGTATGCCTCATTAATTAA
-151 AAAAAGAAAGAAAAGAGGCTTAATTGGTAAAAAACAATAGTGTTATTGTCTTGAAAAAGTACTTTTCATCAGGCATTGC
 -72 AAGCACTTCGTCAAACACCTACTTCTATGGACGTATTACATATTTATTGTAAAGAACTACAATAGACTTAAA
   1 M  S  D  D  R  F  A  E  D  E  I  I  Q  Q  R  R  K  R  R  L
   1 ATG AGT GAC GAT AGA TTT GCA GAA GAT GAA ATT ATA CAG CAA AGA CGG AAA CGG AGG TTG
  21 E  I  L  K  K  Y  Q  Q  T  G  N  G  H  S  D  L  S  I  P  E
  61 GAA ATT TTA AAG AAG TAC CAA CAG ACT GGC AAT GGC CAT TCG GAT TTG TCA ATT CCA GAG
  41 K  K  L  K  E  D  V  D  Q  V  S  T  T  K  P  I  E  A  V  P
 121 AAA AAG CTA AAA GAA GAT GTC GAC CAA GTC TCT ACA ACG AAA CCT ATC GAA GCT GTA CCA
  61 K  M  K  T  N  A  S  K  I  E  I  N  K  E  G  S  N  S  N  T
 181 AAA ATG AAA ACC AAT GCA TCT AAA ATT GAG ATT AAC AAA GAA GGT TCT AAT TCT AAT ACT
  81 K  L  D  V  T  N  S  T  T  S  D  S  P  S  I  K  S  S  V  Q
 241 AAG CTT GAT GTA ACT AAT AGT ACT ACT AGT GAT TCC CCC TCT ATC AAA TCT TCT GTA CAA
 101 I  E  D  T  E  D  D  M  F  A  D  S  P  S  P  S  V  K  R  Q
 301 ATT GAA GAT ACT GAA GAT GAC ATG TTT GCA GAT TCT CCT TCG CCT TCT GTT AAG CGG CAA
 121 N  T  G  K  G  I  S  T  L  T  R  S  F  A  D  M  Q  D  N  W
 361 AAC ACT GGG AAA GGT ATT TCC ACT CTA ACT AGG AGT TTT GCA GAT ATG CAG GAT AAT TGG
 141 D  D  I  E  G  Y  Y  K  V  V  L  M  E  E  L  D  S  R  Y  I
 421 GAC GAT ATT GAA GGT TAT TAT AAG GTA GTT CTT ATG GAG GAA TTG GAT TCT CGT TAC ATC
 161 V  Q  S  N  L  G  K  G  M  F  S  T  V  V  S  A  L  D  R  N
 481 GTT CAG TCC AAC CTT GGT AAA GGC ATG TTC TCT ACG GTG GTC AGT GCT TTG GAT CGG AAT
 181 R  N  Q  T  F  A  I  K  I  I  R  N  N  E  V  M  Y  K  E  G
 541 AGA AAT CAA ACT TTC GCT ATA AAA ATA ATT AGA AAC AAT GAG GTA ATG TAT AAG GAG GGA
 201 L  K  E  V  S  I  L  E  R  L  Q  A  A  D  R  E  G  K  Q  H
 601 CTC AAA GAA GTA TCA ATT CTT GAG CGA CTG CAA GCA GCG GAT CGT GAA GGT AAG CAG CAC
 221 I  I  H  Y  E  R  H  F  M  H  K  N  H  L  C  M  V  F  E  M
 661 ATA ATC CAC TAT GAA AGG CAT TTC ATG CAC AAA AAT CAT CTT TGC ATG GTG TTT GAG ATG
 241 L  S  L  N  L  R  D  I  L  K  K  F  G  R  N  V  G  L  S  I
 721 TTA AGT CTT AAT CTT CGG GAC ATT TTG AAA AAA TTT GGG CGT AAC GTT GGG TTG AGT ATA
 261 K  A  V  R  L  Y  A  Y  Q  M  F  M  A  L  D  L  L  K  Q  C
 781 AAG GCT GTT CGT CTG TAT GCC TAT CAA ATG TTC ATG GCA TTG GAT TTG TTG AAG CAA TGC
 281 N  V  I  E  S  D  I  K  P  D  N  M  L  V  N  E  K  R  N  I
 841 AAC GTA ATT CAT TCA GAT ATC AAA CCG GAC AAT ATG CTG GTG AAC GAA AAA AGA AAT ATT
 301 L  K  I  C  D  L  G  S  A  S  D  A  S  E  N  E  I  T  P  Y
 901 TTA AAA ATA TGT GAC CTT GGT TCT GCT TCA GAT GCC TCT GAA AAT GAA ATA ACA CCG TAC
 321 L  V  S  R  F  Y  R  A  P  E  I
 961 CTC GTG AGT CGA TTT TAT CGT GCT CCA GAA ATC A GTAAGTCTAAATAAACGCCAAATTGCTTATAA

1027 TTCCAAAAATATGGAAATTGAATTCTTCTTCTCATTACTTTTGTACTCTTTTTTCAAGATCTAACCTTTAAAATAG

 332 I   L  G  F  P  Y  S  C  P  I  D  T  W  S  V  G  C  S  L  Y
1103 TT  TTA GGC TTT CCT TAT TCG TGT CCA ATA GAT ACA TGG AGT GTT GGA TGT TCT TTA TAT
 352 E   L  Y  T  G  Q  I  L  F  P  G  R  T  N  N  Q  I  A  S  L
1179 GAG TTG TAT ACT GGT CAA ATA CTA TTT CCT GGC CGA ACA AAT AAC CAG ATT GCT TCG CTA
 372 Y   D  G  M  *
1239 TAT GAT GGA ATG TAA

1254 GGGGAAGTTCAGTCATAAAATGTTAAAAAGAAGTCAATTTTTAAATGACCATTTCGACGCGGATTTCAATTTCATACAAAT
1335 TGATCATGACCCCATTACAAATCAAGAAACGAGGAAACCTGTAAAATTCTCTAAGCCCGACCAAAGACATTCGATCTCGTT
1416 TAAAAGAAGTCCAACTTCAACTGATGAGGAATTTATAATTCGACAAGAGCTTATGGATCTTTGGAAAAATGTTTGGAGCTT
1497 AATCCAGAGAAACTGTACCGCCTGAGATGCTTGAAGCATCCTTCTTTATAAAAAATAAAACTGCGGGTTCATGGGTAACTA
1578 TCAATTTAGAGTGAAGCAGACACTGTTTCTTAGTTACGTGATCTCTCTGGTCAATCAATACCAATCAATGTATATATTTTA
1659 TCTAACCAAGTGTAACAATTTTGAATAGTAGAAACGCCAACTGGTTTGTTTGACTTACAAATCTATGTGGTTTTCAATTTT
1740 CTTCTCCAGTGTAAATAGTGTTTGTGTCGAAGAGTCTCATACTCGTTCAACAAATTAGATACTAATTGCAAAACTAATTGC
1821 TTCATTTTATTTTATCATTTGCGAAATAAAACCAATGGTAACGACATTTCGATAACTGCGTAGATGTCACATATGCTGTCA
1902 ATGGCCGTGGACATTTTAGCAATTGCCAACATTCCTAACAAAATCTCTGTGACCCAAAATTCATAAAAATATGAGAGATC
```

Figure 2. Cloning and sequence determination. **A.** Restriction map of the 2.5 kb clone complementing the *prp4* mutation. Restriction sites in parenthesis (PstI) and (KpnI) are sites in the multiple cloning site of the vector (25). A, AccI, B, BglII, Bc, BclI, E, EcoRI, S, SalI. ATG and TAA indicate start and stop codon of the ORF (open box); closed box shows the intron; dotted boxes represent 5′ and 3′ flanking areas of the *prp4* gene. **B.** Sequence of the putative kinase gene. The derived ORF is depicted in the one letter code. The most likely translational start codon M has been numbered as 1; the nucleotide sequence has also been numbered: −1 is the first nucleotide upstream of the start codon ATG, while the A of the ATG is the nucleotide numbered as 1. The signature sequences predicting a serine/threonine kinase are printed in bold letters throughout the amino acid sequence (29). The nucleotide sequences which represent the splice site consensus sequences found in *S.pombe* introns are also printed in bold letters (16).

shows that the *prp4(ts)* strain which has been transformed with the plasmid pSAK does not accumulate pre-mRNA of the *dis2* gene at the restrictive temperature, only mature mRNA was detected (Fig.1, lane 4). As expected, the mutant strain *prp4(ts)* accumulates at the restrictive temperature pre-mRNA of the *dis2* gene (Fig.1, lane 2).

DNA sequence analysis predicts that the gene suppressing *prp4* encodes a serine/threonine kinase

The complete nucleotide sequence of the 2.5 kb genomic fragment (Fig.2A) contained in pSAK was determined. The sequence is shown in Fig.2B. We find an open reading frame (ORF) comprising 375 amino acids which appears to be interrupted by a 109 bp intron (Fig.2B). The putative intron displays an architecture described for small 5'S/3'S *S.pombe* introns and has a 5' splice site and a branch-sequence conforming with the consensus sequences (16). The deduced amino acid sequence predicts a protein with a molecular mass of 42 742 Dalton and a pI of 7.1.

The predicted polypeptide derived from this nucleotide sequence was compared to GenBank data base. The search revealed that the deduced amino acid sequence contains conserved features which are predictive of the catalytic domains of a serine/threonine protein kinase. The amino acids in Fig.2B depicted in bold letters indicate protein kinase signature sequences (27, 28, 29). The first conserved sequence L G X G X F S X V beginning at position 165 (Fig.2B in bold letters) conforms to the consensus sequence for nucleotide binding proteins (27, 28, 29); at position 188 we find a lysine (K) an amino acid residue which appears to be invariant in all serine/threonine kinases; the following sequences contain all the other conserved amino acid residues identified in proteins belonging to the protein kinase family (Fig.2B, conserved residues in bold letters, 27, 28, 29).

Within the kinase domain starting at position 161 (Fig.2B) in the region where the nucleotide binding site is located, we find stretches of high sequence similarities to two *S.cerevisiae* protein kinases and a *Dictyostelium discoideum* kinase. The SPK1 kinase from *S.cerevisiae* (31) shows 67% similarity over 32 amino acids (Fig.3). While the YAK1 kinase (38) shows 57% similarity over 47 amino acids (Fig.3). DDK2, kinase-2 from *D.discoideum* shows 65% similarity to the *S.pombe* kinase in this region (Fig.3, signature sequences are underlined). In addition, starting at position 228 in Fig.2B we find a stretch of 18 amino acids with 78% identity to a stretch of amino acids in the YAK1 protein kinase from *S.cerevisiae* (Fig.3). Outside the kinase domain, in the N-terminus of the predicted protein a stretch of 64 amino acids display 25% identical and 51% similar amino acid residues with residues found in the N-terminus and also outside the kinase domain of protein kinase-2 from the slime mold *D.discoideum* (Fig.3).

To prove that the gene encoding the putative kinase contains an intron, we used oligonucleotides complementary to the 5' end and to the 3' end of the open reading frame (Fig.2B) and performed PCR (polymerase chain reactions) using a cDNA library made from a wild type *S.pombe* strain (26). With this approach we revealed a product of approximately 1 kb. This fragment was subcloned and sequenced (results not shown). The sequence was identical with the sequence shown in Fig.2B except that it was missing the predicted intron sequence.

```
Prot.   Sequences                                          %Id  %Si
        within the Kinase-Domain

PRP4    IVQSNLGKGMFSTVVSALDRNRNQTFAIKII
        I++  +G G F+TV  A++R    TFA+KII                    45   67
SPK1    IIDEVVGQGAFATVKKAIERTTGKTFAVKII


PRP4    YYKVVLMEELDSRYIVQSNLGKGMFSTVVSALDRNRNQTFAIKIIRN
        Y + VL  E + +Y+V   LG G F+ VV   +   ++  A+K+++  29   57
YAK1    YVNDVLGVEQNRKYLVLDILGQGTFGQVVKCQNLLTKEILAVKVVKS


PRP4    FMHKNHLCMVFEMLSLNL
        K+HKNHLC+VFE+LS NL                                 78   99
YAK1    FVHKNHLCLVFELLSNNL


PRP4    LGKGMFSTVVSALDRNRNQTFAIKIIRNN
        +GKG F+ V+    +  +  FA+K++R++                      31   65
DDK2    IGKGSFGKVMQVKKKGEDKIFAMKVLRKD

        Sequences
        outside the Kinase-Domain

PRP4    KEDVDQVSTTKPIEAVPKMKTNASKIEINKEGSNSNTKLDVTNSTTS
        KE+ +Q +   P    P   T ++ IE +   +++N     +N ++S
DDK2    KEQQQQQNIPAPATQTPITQTGTPTIEESQKNTDNNNINGASNEASS

PRP4    DSPSIKSSVQIEDTEDD
         + S ++S   +D ED+                                  25   51
DDK2    SPDSPNGSGNGNDDEDE
```

Figure 3. Alignment of deduced amino acid stretches of the putative protein kinase (PRP4) from *S.pombe* with deduced amino acids from kinases SPK1 (38) and YAK1 (31) from *S.cerevisiae* and with the protein kinase-2 (DDK2) from *D.discoideum* (32). Prot. protein, %Id, % Identity, %Si, % Similarity. When the amino acid residues between two sequences are identical it is indicated by the letter code, when the amino acid residues have similar biochemical features it is indicated by a + sign. Amino acid residues underlined are part of the protein kinase signature sequences indicating a nucleotide binding site (27,28,29).

Steady state transcripts of the *prp4* suppressing gene

We performed a Northern analysis using RNA isolated from wild type cells, from the *prp4(ts)* strain and from the *prp4(ts)* strain transformed with the plasmid pSAK complementing the *prp4* mutation. As probe we used the radiolabeled 1.1 kb SalI/BglII fragment which contains most part of the open reading frame (Fig.2A). In the autoradiography we detected in all lanes a somewhat broad band of about 1200 nt in size (Fig.4). It is possible that the bands include pre-mRNA and spliced mRNA of the gene. It is also possible that the gene has more than one transcriptional start and transcriptional stop site. In the *prp4(ts)* mutant strain the somewhat longer transcripts accumulate at the

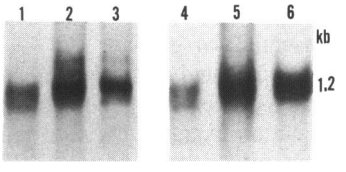


Figure 4. Northern analysis to determine transcript of the putative protein kinase gene. Total RNA isolated from a wild type strain and from the *prp4* strain transformed with the plasmid containing the 2.5 kb fragment complementing the temperature sensitive phenotype of the *prp4* strain was hybridized to the radio labeled 1.1 kb SalI/BglII fragment comprising parts of the open reading frame of the putative protein kinase and parts of the intron (Fig.2A). Lane 1, 15 μg RNA isolated from the *prp4* strain grown at 25°C; lane 2 and 3, 15 μg RNA isolated from the *prp4* strain grown at 25°C to mid log phase and then shifted for three and six hours to 36°C. Lane 4, 10 μg RNA of *S.pombe* strain 972 from cells grown at 30°C; lane 5, 10 μg RNA of the *prp4* strain transformed with pSAK and grown at 25°C; lane 6, 10 μg RNA of the pSAK transformed *prp4* strain grown at 36°C.

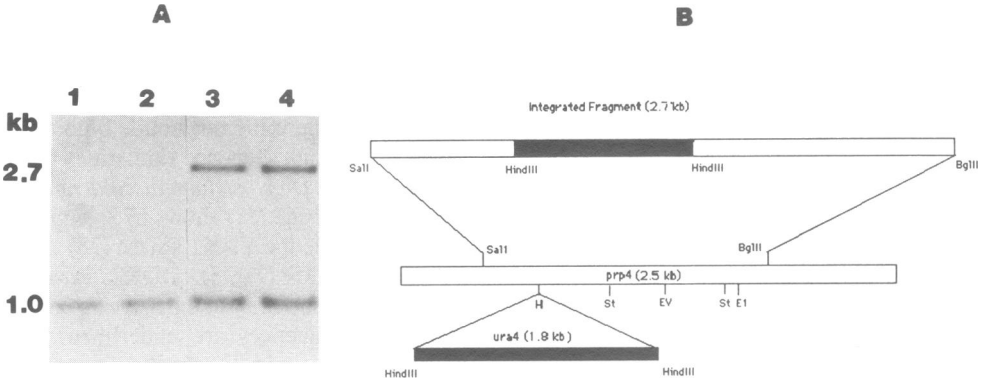


Figure 5. Southern analysis to confirm gene replacement. **A.** DNA was double digested with the restriction enzymes SalI and BglII and separated on an 0.8% agarose gel. Lane 1 and lane 2 contains DNA isolated from growing colonies after tetrad analysis, lane 3 and lane 4 contain DNA from uracil prototrophic diploid cells after transformation with the *ura4* interrupted 2.7 kb fragment shown in **B.** The DNA was transferred to nitrocellulose and hybridized to the radiolabeled 1.1 kb SalI/BglII fragment shown also under **B. B.** Construction of an interrupted *prp4: ura4* allele. The 1.8 kb HindIII *ura4* fragment was cloned into the HindIII site of the 2.5 kb *prp4* fragment. The HindIII site is in the ORF (Fig. 2). The 2.7 kb SalI/BglII fragment was isolated and transformed as described in the text. H, HindIII, St, StyI, EV, EcoRV, E1, EcoRI.

restrictive temperature (Fig.4 lane 2 and 3). This is consistent with the presence of the 109 bp intron in this gene.

The putative protein kinase gene is essential for growth and is located at the *prp4* locus

To test whether the gene encoding this putative protein kinase is essential for growth, we constructed a diploid strain h^{+N}/h^{-S} and homozygous for the *leu1-32* and *ura4-D18* alleles. We used the SalI/BglII fragment which had the functional *ura4* gene inserted into the HindIII site residing in the open reading frame of the gene (Fig.5B). This construction was transformed into the cells in its linearized form to target the genomic sequences and to replace the wild type gene with the *ura4* disrupted gene by homologous recombination (19, 20). From six independently isolated transformants tetrads were analyzed. In all tetrads no more than two spores grew up at 30°C. The viable spores were auxotrophic for uracil. When DNA isolated from these haploid growing colonies was digested with SalI and BglII and then hybridized with the 1.1 kb SalI/BglII fragment (Fig.5B) we detected a fragment of about 1 kb representing the uninterrupted genomic copy of the gene (lanes 1 and 2). DNA isolated from diploid cells contain one 2.7 kb *ura4* interrupted allele and one wild type allele (Fig.5A, lanes 3 and 4). This clearly demonstrates that this gene is essential for growth.

Early observations indicated that the *prp4* locus is linked to the *ura4* locus which has been mapped to chromosome III (20). For the complementation experiments reported above we also used a genomic library cloned in the vector pDB248′. This vector contains the *S.cerevisiae LEU2* gene as a marker which complements the *leu1* gene of *S.pombe* (30) The *leu1* gene is located on chromosome II. Transformation of this clone bank into the $prp4^{(ts)}$ strain also revealed suppressors of the *prp4* mutation and plasmids isolated contained the identical sequence reported above. We found that in some cases the suppressing plasmid was integrated in chromosome III. When these strains were crossed to $leu1^{-}, prp4^{-}$ strains tetrad analysis indicated a close linkage between $prp4^{+}$ and *LEU2* from *S.cerevisiae*. Therefore, we used one of these strains and crossed it to a wild type strain. The dissection of 42 asci yielded for each tetrad at 36°C four growing colonies. These data clearly demonstrate that the plasmid suppressing the *prp4* mutation was integrated at the *prp4* locus.

DISCUSSION

We are investigating pre-mRNA splicing in the fission yeast *S.pombe*. The fission yeast is only distantly related to the budding yeast *S.cerevisiae* but is also genetically tractable. The number and size of introns found in these two yeasts differ significantly. The majority of the *S.pombe* introns are smaller than 150 bp and about 40−50 percent of the genes contain introns. In the yeast *S.cerevisiae* only about 5% of the genes contain introns. These introns are on average significantly longer than *S.pombe* introns (16, 18). Recently it has been demonstrated with a small typical *S.pombe* intron that during the assembly of spliceosomes the 5′ end of U1 SnRNA may not only interact by complementary base-pairing with the 5′ splice site but also may interact by complementary base pairing with the 3′ AG of the intron (17). The fission yeast *S.pombe* splices efficiently the small t intron (66 bp) in the early region of the SV40 virus (14). Interestingly, the small t SV40 intron displays the features of a typical *S.pombe* intron. Therefore, we suggested that the small *S.pombe* introns might reflect the primitive, the archetype of pre-mRNA introns (15).

To the best of our knowledge this is the first report of the isolation of a gene involved in pre-mRNA splicing encoding a putative serine/threonine protein kinase. In this paper we clearly demonstrate that the product of the *prp4* gene is involved in pre-mRNA splicing and that it is essential for growth. Neither from the budding yeast *S.cerevisiae* nor from any other organism a protein kinase gene potentially involved in pre-mRNA splicing has been isolated. We do not have an indication yet in which step of the splicing process this kinase is involved. We also do not know anything about the substrate(s) of this protein kinase encoded by the *prp4* gene.

We are currently trying to isolate the protein to further investigate the biochemical features of this putative kinase. In a mammalian system a snRNP U1 associated kinase activity

phosphorylates *in vitro* the serine/arginine rich domains of the U1-70K protein (9, 39). It has been argued that kinases may be involved in the proper assembly or disassembly of the spliceosome. The ordered association or dissociation of splicing components during the splicing process might be brought about by the phosphorylation and de-phosphorylation of specific substrates at each step of the splicing process.

In any case, the *prp4* gene of *S.pombe* encodes a protein which is predicted to be a serine/threonine kinase (27, 28, 29). We have taken advantage of the genetic amenability of *S.pombe* and isolated several suppressors of the temperature sensitive *prp4* mutation. This approach will allow us to genetically determine the protein(s) which may interact with this kinase and will hopefully help to shed more light on the specific functional involvement of kinase(s) in the pre-mRNA splicing process. Experiments to complement the *prp4* mutation with a mouse cDNA library are currently in progress. It will be interesting to find out whether *S.pombe* has features in the pre-mRNA splicing process which can be studied across organismal borders like cell cycle regulation.

ACKNOWLEDGEMENTS

We are indebted to Dr Tony Carr for his genomic *S.pombe* library. We also thank Dr Rebecca Dickstein (Drexel University) for editing the manuscript. This work was supported by a grant (MV-389C) from the American Cancer Society awarded to N.F.K.

REFERENCES

1. Guthrie, C. (1991) Science, **253**, 157–163.
2. Padgett, R. A., Grabowski. P. J., Konarska, M. M., Seiler, S. and Sharp, P. A. (1986) Annu. Rev. Biochem., **55**, 1119–1150.
3. Brody, E. and Abelson, J (1985) Science, **228**, 983–967.
4. Wassarman, D. A. and Steitz. J. A. (1991) Nature, **349**, 463–464.
5. Lührmann, R., Kastner, B. and Bach, M. (1990) Biochim. Biophys. Acta, **1087**, 265–292.
6. Ruby, S. W. and Abelson, J. (1991) Trends Genet., **7**, 79–85.
7. Zahler, A. M., Lane W. S., Stolk J. A. and Roth, M. B. (1992) Genes Dev., **6**, 837–847.
8. Company, M., Arenas, J. and Abelson J. (1991) Nature, **349**, 487–493.
9. Tazi, J., Daugeron, M.-C., Cathala, G., Brunel, C. and Jeanteur, P. (1992) J. Biol. Chem., **267**, 4322–4326.
10. Behrens, S. V. and Lührmann, R. (1991) Genes Dev., **5**,1439–1452.
11. Schwer, B. and Guthrie, C. (1992) EMBO J., **11**,5033–5039.
12. Potashkin, J., Li., R. and Frendewey, D. (1989) EMBO J., **8**, 551–559.
13. Rosenberg, G. H., Alahari, S. K. and Käufer, N. F. (1991) Molec. Gen. Genet., **226**, 305–309.
14. Käufer, N. F., Simanis, V. and Nurse, P. (1985) Nature, **318**, 78–80.
15. Gatermann, K., Hoffmann, A., Rosenberg, G. H. and Käufer, N. F. (1989) Mol. Cell. Biol., **9**, 1526–1535.
16. Prabhala, G., Rosenberg, G. H. and Käufer, N. F. (1992) Yeast, **8**, 171–182.
17. Reich, C. I., VanHoy. R. W., Porter, G. L. and Wise J. A. (1992) Cell, **69**, 1159–1169.
18. Woolford, J. L., (1989) Yeast, **5**, 439–457.
19. Moreno, S., Klar, A. and Nurse, P. (1991) Methods Enzymol., **194**, 795–823.
20. Grimm, C. Kohli, J., Murray, J. and Maundrell, K. (1988) Mol. Gen. Genet., **215**, 81–86.
21. Gutz, H., Heslot, H., Leupold, U. and Loprieno, N. (1974) In King, R.C. (ed.) Handbook of Genetics. Plenum Press, New York, London, Vol.1, pp. 395–446.
22. Gatermann, K., Rosenberg, G. H. and Käufer, N. F. (1988), **10**, BioTechniques 52–53.
23. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular cloning. A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
24. Rödel, C., Kirchhoff, S. and Schmidt, H. (1992) Nucl. Acids Res., **20**, 6347–6353.
25. Barbet, N., Muriel, W. J. and Carr, A. M. (1992) Gene, **114**, 59–66.
26. Becker, D. M., Fikes J. D. and Guarente, L. (1991) Proc. Natl. Acad. Sci. USA, **88**, 1968–1972.
27. Hanks, S. K., Quinn, A. M. and Hunter, T. (1988) Science, **241**, 41–52.
28. Hunter, T. (1991) Methods Enzymol., **200**, 3–37.
29. Hanks, S. K. and Quinn, A. M. (1991) Methods Enzymol., **200**, 38–62.
30. Wright, A., Maundrell, K., Heyder, W-D., Beach, D. and Nurse, P. (1986) Plasmid, **15**, 156–158.
31. Garrett, S. and Broach, J. (1989) Genes Dev., **3**, 1336–1348.
32. Haribabu, B. and Dottin, R. (1991) Proc. Natl. Acad. Sci. USA, **88**, 1115–1119.
33. Ohkura, H., Kinoshita, N., Miyatani, S., Toda, T. and Yanagida, M. (1989) Cell, **57**, 997–1008.
34. Booher, R. and Beach, D. (1989) Cell, **57**, 1009–1016.
35. Kim, S., Smith, J., Claude, A. and Lin, R. J. (1992) EMBO J., **11**, 2319–2326.
36. Bennett, M., Michaud, S., Kingston, J. and Reed. R. (1992) Genes Dev., **6**, 1986–2000.
37. Ward, A. C. (1990) Nucl. Acids Res., **18**, 5319.
38. Stern, D. F., Zheng, P., Beidler D. V. and Zerillo C. (1991) Mol. Cell. Biol., **11**, 987–1001.
39. Tazi, J., Kornstädt, U., Rossi, F., Jeanteur, P., Cathala, G., Brunel, C. and Lührmann, R. (1993) Nature, **363**, 283–286.
40. Altschul, S. F., Gish, W., Miller, W., Meyers, F. W. and Lipman, D. J. (1990) J. Mol. Biol., **215**, 403–410.