# Structural Studies of *Schizosaccharomyces pombe* RNA polymerase II

by

YOSHINAO AZUMA

*This Thesis Fulfilled for The Degree of Doctor of Philosophy*

in

The School of Life Science

The Department of Molecular Genetics

Graduate University for Advanced Studies

National Institute of Genetics

JAPAN

*Spring 1994*

# TABLE OF CONTENTS

# ABSTRACT

To study the molecular composition of RNA polymerase II, the enzyme was purified from the fission yeast *Schizosaccharomyces pombe*. The highly purified RNA polymerase II fraction contained at least 11 polypeptides as identified on SDS-polyacrylamide gel electrophoresis. To establish an *in vitro* reconstitution system aiming the identification of the essential subunits and core-like structure of the enzyme for transcription activity, preliminary attempts were made to set up conditions for reversible denaturation of the RNA polymerase II treated with urea. In spite of these attempts, the system was not established so far with the purified enzyme. Then the subunits, which may construct the core-like structure of RNA polymerase II, were cloned and analyzed with these genes.

The gene, *rpb1*, encoding the subunit 1 with the apparent molecular mass of 210 kDa was isolated from *S. pombe* by cross-hybridization using the corresponding gene of *Saccharomyces cerevisiae* as a probe. By analyzing both the genomic sequence and the transcript sequence, this gene was found to contain six introns and encodes a polypeptide of 1,752 amino acid residues with a molecular mass of 194 kDa. The RNA polymerase II subunit 1 of *S. pombe* contains 29 repetitions of CTD, the carboxy terminal domain consisting of heptapeptide repetitions.

For cloning the gene coding for the 40 kDa subunit 3, a part of the cDNA was isolated with RT-PCR using probes designed based on N-terminal amino acid sequences of the fragmented subunit 3 polypeptides, and the genomic DNA containing the full length of *rpb3* gene was cloned with the cDNA fragment. Nucleotide sequence of the gene, *rpb3*, indicated that it contains two introns at the 5'-proximal region as in the case of *rpb1* and *rpb2*, and encodes a protein of 297 amino acid residues in length with a calculated molecular mass of 34 kDa. The subunit 3 carries four stretches of amino acid sequence conserved among the homologues of *Escherichia coli* RNA polymerase α subunit. A metal binding motif and a leucine zipper-like motif were found in its N-terminal and C-terminal region, respectively. The *E. coli* α subunit plays a major role in not only the subunit assembly but also the

interaction with many transcription factors. To identify the function(s) of RNA polymerase II subunit 3, a number of temperature-sensitive mutants carrying mutations in the *rpb3* gene were isolated. Some of them stop growing soon after shifting temperature from the permissive 25°C to the non-permissive 37°C, implying that these mutant RNA polymerases are thermolabile. However, most of them stop growing slowly, supposedly because of assembly defect at high temperature. Some mutants showed altered cell shapes at the non-permissive temperature.

In these works, the important difference between the purified RNA polymerase II and *E. coli* RNA polymerase was found out by the reversible denaturation. And it is confirmed that the subunits 1, 2 and 3 should construct the core-like structure of RNA polymerase II, because each corresponding subunit contains rather conserved regions. But neither the essential subunits nor the core-like structure of RNA polymerase II were not decided so far. They will be next subjects for future.

# LIST OF TABLES



# LIST OF FIGURES

# LIST OF ABBREVIATION

| | |
|---|---|
| BSA | bovine serum albumin |
| CBB | Coomassie Brilliant Blue |
| CTD | C-terminal domain |
| DAPI | 4',6-diamidine-2'-phenylindole dihydrochloride |
| DMSO | dimethyl sulfoxide |
| DOC | deoxycholic acid |
| DTT | dithiothreitol |
| FPLC | fast protein liquid chromatography |
| HPLC | high performance liquid chromatography |
| kbp | kilo base pair(s) |
| kDa | kilo dalton(s) |
| LB medium | Luria Bertani medium |
| LETS buffer | LiCl-EDTA-Tris-SDS buffer |
| MOPS | 3-[N-morpholino]propanesulfonic acid |
| MPa | mega pascal |
| N.D. | Non-determined |
| PCR | polymerase chain reaction |
| PEG4000 | polyethylene glycol 4000 |
| PIM | protease inhibitor mixture |
| PMSF | phenylmethane-sulfonyl fluoride |
| ppt | precipitate |
| RT-PCR | reverse transcription-polymerase chain reaction |
| S.D. | sequence determined |
| SDS | sodium dodecyl sulfate |
| PAGE | polyacrylamide gel electrophoresis |
| SET | saline-EDTA-Tris |
| SSC | standard saline citrate |
| sup | supernatant |
| TAE buffer | Tris-acetate-EDTA buffer |
| TCA | trichloroacetic acid |
| TE buffer | Tris-EDTA buffer |
| TEDA buffer | Tris-EDTA-DTT-ammonium sulfate buffer |
| TGED buffer | Tris-glycerol-EDTA-DTT buffer |
| ts | temperature sensitive |
| V8 | *Staphylococcal aureus* V8 protease |
| YPD medium | yeast extract-peptone-glucose medium |

# INTRODUCTION

In eukaryotes, three classes of nuclear RNA polymerases, I(A), II(B) and III(C), exist, each responding for transcription of large rRNA, mRNA, and small rRNA and tRNA, respectively (for reviews see refs.1,2). Each class of purified RNA polymerase is composed of two large and about 10 small polypeptides (1,2). The genes coding for the largest and the second-largest subunits of the three classes of RNA polymerases have been cloned from a number of organisms (3,4; and other works cited there). Sequence analysis indicates that these subunits share notable homologies with the β′ and β subunits, respectively, of prokaryotic RNA polymerase. The gene encoding the third-largest subunit of the three classes of RNA polymerases has so far been cloned from four organisms (5-8). The subunit appears to have partial similarity in primary structure with the α subunit of prokaryotic RNA polymerase.

The roles of these three major subunits constructing the putative core-like structure of eukaryotic RNA polymerases can therefore be predicted from the known functions of *E. coli* RNA polymerase core subunits. The catalytic site is located on β, while β′ has non-specific binding activity to DNA; α subunit links these two large subunits into core enzyme structure (reviewed in refs. 9,10). In addition, α subunit carries the contact site with various transcription factors (reviewed in refs. 11,12). In addition to these core enzyme subunits, *E. coli* holoenzyme contains one of the multiple species of σ subunit, which plays a major role in promoter recognition (13,14). Low levels of homology to the σ subunit were found in, for example, eukaryotic TATA element-binding factor, TBP, and RPB4 subunit of *Saccharomyces cerevisiae* RNA polymerase II (15,16). , but recent progress in cloning of the genes encoding RNA polymerase subunits has provided an experimental basis for detailed characterization of each subunit polypeptide.

For the molecular anatomy of eukaryotic RNA polymerase, the RNA polymerase II of the fission yeast *Schizosaccharomyces pombe* was chosen, because *S. pombe* is such a good material for genetic analysis as *S. cerevisiae*, but

nevertheless its strategies for gene expression such as transcription initiation mechanism (17) and splicing pattern (18,19) are similar to those in higher eukaryotes. And among the three classes of RNA polymerases, RNA polymerase II is the most interesting at the point that RNA polymerase II is involved in transcriptional regulation of various gene. In this study, the RNA polymerase II was purified from *S. pombe* to define the subunit composition, but it was very difficult to determine of the role(s) for each subunit due to the complexity of its multi-subunit structure and to the low yield of purified enzymes. Then the genes coding subunits constructing the core-like structure of this enzyme was characterized at first by gene cloning and mutant analysis of the gene.

# EXPERIMENTAL PROCEDURES

## *Materials*

### Strains, plasmid and media for yeast

A wild-type haploid strain 972h$^-$, a leu$^-$ haploid strain JY265 and an ura$^-$ and leu$^-$ diploid strain JY765 of *S. pombe* were provided by Dr. M. Yamamoto (Univ. Tokyo).

YPD medium [2% Bacto-yeast extract, 1% Bacto-tryptone (Difco Laboratories), 2% glucose] and SD medium [6.7% yeast nitrogen base without amino acids (DIFCO) 2% glucose] were used in the studies as the rich and synthetic medium, respectively. 50 μg per ml of L-leucine was added to SD medium for growth of the leu$^-$ strains.

Plasmids, pRP19 and pY2413, containing the entire *S. cerevisiae RPB1* (20) and *RPB3* genes (5), respectively, were gifted from Dr. R. A. Young (MIT). YEp13 containing the *S. cerevisiae LEU* 2 gene was provided by Dr. M. Yamamoto (Univ. Tokyo). pBrpb3 plasmid was constructed in this study by cloning of the 2.3 kbp *Pst*I-*Eco*RV DNA fragment containing the *S. pombe rpb3* gene into *Pst*I-*Eco*RV position of pBluescript II KS$^+$.

### Strains, plasmids and media for bacteria

LB medium [1% Bacto-tryptone, 0.5% Bacto-yeast extract (Difco Lab.), 1% NaCl, (pH 7.0)] and M9 medium [0.6% $Na_2HPO_4$, 0.3% $KH_2PO_4$, 0.05% NaCl, 0.1% $NH_4Cl$, 0.2% glucose, 1 μg/ml thiamin, 0.1 mM $CaCl_2$, 2 mM $MgSO_4$] were used as a rich medium and a minimal medium for F' maintenance of *E. coli*, respectively.

*E. coli* strain, DH5, was used as competent cells for plasmid maintenance and *E. coli* strains, TG1 and JM109, were used for M13 phage maintenance.

pUC18 and pUC19 plasmids, pBluescript II KS+ phagemid, and M13mp18 and 19 phages were used for DNA cloning, sequencing and vector construction.

# EXPERIMENTAL PROCEDURES

## *Materials*

### Strains, plasmid and media for yeast

A wild-type haploid strain 972h$^-$, a leu$^-$ haploid strain JY265 and an ura$^-$ and leu$^-$ diploid strain JY765 of *S. pombe* were provided by Dr. M. Yamamoto (Univ. Tokyo).

YPD medium [2% Bacto-yeast extract, 1% Bacto-tryptone (Difco Laboratories), 2% glucose] and SD medium [6.7% yeast nitrogen base without amino acids (DIFCO) 2% glucose] were used in the studies as the rich and synthetic medium, respectively. 50 μg per ml of L-leucine was added to SD medium for growth of the leu$^-$ strains.

Plasmids, pRP19 and pY2413, containing the entire *S. cerevisiae RPB1* (20) and *RPB3* genes (5), respectively, were gifted from Dr. R. A. Young (MIT). YEp13 containing the *S. cerevisiae LEU* 2 gene was provided by Dr. M. Yamamoto (Univ. Tokyo). pBrpb3 plasmid was constructed in this study by cloning of the 2.3 kbp *Pst*I-*Eco*RV DNA fragment containing the *S. pombe rpb3* gene into *Pst*I-*Eco*RV position of pBluescript II KS$^+$.

### Strains, plasmids and media for bacteria

LB medium [1% Bacto-tryptone, 0.5% Bacto-yeast extract (Difco Lab.), 1% NaCl, (pH 7.0)] and M9 medium [0.6% $Na_2HPO_4$, 0.3% $KH_2PO_4$, 0.05% NaCl, 0.1% $NH_4Cl$, 0.2% glucose, 1 μg/ml thiamin, 0.1 mM $CaCl_2$, 2 mM $MgSO_4$] were used as a rich medium and a minimal medium for F' maintenance of *E. coli*, respectively.

*E. coli* strain, DH5, was used as competent cells for plasmid maintenance and *E. coli* strains, TG1 and JM109, were used for M13 phage maintenance.

pUC18 and pUC19 plasmids, pBluescript II KS+ phagemid, and M13mp18 and 19 phages were used for DNA cloning, sequencing and vector construction.

## Chemicals

Protease inhibitor mixture contained (per liter): 100 mg benzamidine-HCl (Wako Chemicals), 10 mg chymostatin (Sigma), 10 mg aprotinin (Funakoshi), 10 mg pepstatin A (Funakoshi), 10 mg leupeptin (Funakoshi), 10 mg antipain (Sigma), 50 mg N-alpha-p-tosyl-L-lysine chloromethyl ketone (Sigma), 10 mg N-tosyl-L-phenylalanine chloromethyl ketone (Sigma), and 174 mg phenylmethane-sulfonyl fluoride (PMSF) (CALBIOCHEM). α-amanitin was obtained from CALBIOCHEM.

Both fragmented, single strand calf thymus DNA (Sigma) and salmon sperm DNA (Sigma) were prepared with sonication and heat denaturation. dATP, dTTP, dCTP, dGTP, and ATP, GTP, CTP, UTP, and fluore-dUTP and fluore-prime were purchased from Pharmacia.

[α-$^{32}$P]dATP, [γ-$^{32}$P]dATP, [α-$^{35}$S]dCTP and [$^{3}$H]UTP were obtained from Amersham.

Oligo(dT) cellulose was obtained from Sigma. DEAE-Sephadex A-25 gel, Q Sepharose FF gel, a prepackaged Mono Q column and a prepackaged Superose 6 column were purchased from Pharmacia. Glass-beads No. 4 was purchased from Toshinriko.

## Enzymes and experiment kits

*E. coli* RNA polymerase was purified by DEAE Sephadex A-25 and phospho-cellulose column chromatography to the almost same purity as the highly purified RNA polymerase II in my laboratory.

Restriction enzymes, *Eco*RI, *Eco*RV, *Hind*III, *Pst*I, *Cla*I, *Bal*I, *Bam*HI, *Sac*I, *Sal*I, *Xba*I and *Eco*T22I, were purchased from Takara. RNase H, mung bean nuclease, exonuclease III, Klenow fragment, T4 DNA polymerase, T4 polynucleotide kinase and terminal deoxynucleotidyl transferase were obtained from Takara. And other enzymes were purchased from each company: AmpliTaq DNA polymerase, Perkin-Elmer; RNase A, Sigma; reverse transcriptase, Life Science; V8 protease, PIERCE.

## Chemicals

Protease inhibitor mixture contained (per liter): 100 mg benzamidine-HCl (Wako Chemicals), 10 mg chymostatin (Sigma), 10 mg aprotinin (Funakoshi), 10 mg pepstatin A (Funakoshi), 10 mg leupeptin (Funakoshi), 10 mg antipain (Sigma), 50 mg N-alpha-p-tosyl-L-lysine chloromethyl ketone (Sigma), 10 mg N-tosyl-L-phenylalanine chloromethyl ketone (Sigma), and 174 mg phenylmethane-sulfonyl fluoride (PMSF) (CALBIOCHEM). α-amanitin was obtained from CALBIOCHEM.

Both fragmented, single strand calf thymus DNA (Sigma) and salmon sperm DNA (Sigma) were prepared with sonication and heat denaturation. dATP, dTTP, dCTP, dGTP, and ATP, GTP, CTP, UTP, and fluore-dUTP and fluore-prime were purchased from Pharmacia.

[α-$^{32}$P]dATP, [γ-$^{32}$P]dATP, [α-$^{35}$S]dCTP and [$^{3}$H]UTP were obtained from Amersham.

Oligo(dT) cellulose was obtained from Sigma. DEAE-Sephadex A-25 gel, Q Sepharose FF gel, a prepackaged Mono Q column and a prepackaged Superose 6 column were purchased from Pharmacia. Glass-beads No. 4 was purchased from Toshinriko.

## Enzymes and experiment kits

*E. coli* RNA polymerase was purified by DEAE Sephadex A-25 and phospho-cellulose column chromatography to the almost same purity as the highly purified RNA polymerase II in my laboratory.

Restriction enzymes, *Eco*RI, *Eco*RV, *Hind*III, *Pst*I, *Cla*I, *Bal*I, *Bam*HI, *Sac*I, *Sal*I, *Xba*I and *Eco*T22I, were purchased from Takara. RNase H, mung bean nuclease, exonuclease III, Klenow fragment, T4 DNA polymerase, T4 polynucleotide kinase and terminal deoxynucleotidyl transferase were obtained from Takara. And other enzymes were purchased from each company: AmpliTaq DNA polymerase, Perkin-Elmer; RNase A, Sigma; reverse transcriptase, Life Science; V8 protease, PIERCE.

Experiment kits were purchased from each company: Random Primed DNA Labeling Kit, Boehringer Mannheim; GENE CLEAN II Kit, BIO 101; Sequenase Ver. 2.0 DNA Sequencing kit, USB; AutoRead and AutoCycle Sequencing kit, Pharmacia; ECL direct nucleic acid labeling and detection system, Amersham; The Silver Stain, Bio-Rad.

## Buffers

TGED buffer contained 50 mM Tris-HCl (pH 7.8 at 4°C), 25% glycerol, 0.1 mM EDTA, and 1 mM DTT. Buffer A contained 75 mM Tris-HCl (pH 7.8 at 4°C), 6% glycerol, 0.15 mM EDTA, 1.5 mM DTT, and 150 mM ammonium sulfate, while buffer B consists of 50 mM Tris-HCl (pH 7.8 at 4°C), 0.1 mM EDTA, 1 mM DTT and 200 mM ammonium sulfate. TEDA buffer contained 50 mM Tris-HCl (pH 7.8 at 4°C), 0.1 mM EDTA, 1 mM DTT, and 50 mM ammonium sulfate. 8 M urea buffer was made with the TEDA buffer.

50 x TAE buffer contained: 242 g of Tris base, 57.1 ml of glacial acetic acid, 100 ml of 0.5 M EDTA per liter. 20 x SSC buffer contained: 175.3 g of NaCl and 88.2 g of sodium citrate per liter (pH 7.0 at 25°C). Low stringency hybridization buffer contained: 30% formamide, 6 x SSC, 1 x Denhardt's, 20 mM sodium phosphate buffer (pH 7.8, at 25°C), 0.2% SDS and 0.2 mg/ml single strand salmon sperm DNA. Low stringency hybridization washing buffer contained: 3 x SET, 0.1% SDS and 0.1% sodium pyrophosphate. High stringency hybridization buffer contained: 50% formamide, 5 x SSC, 10 x Denhardt's, 50 mM sodium phosphate buffer (pH 7.8, at 25°C), 0.2% SDS and 0.1 mg/ml single strand salmon sperm DNA. High stringency hybridization washing buffer contained: 1 x SET, 0.1% SDS and 0.1% sodium pyrophosphate. 100 x Denhardt's contained: 2% Ficoll, 2% polyvinyl pyrrolidone and 2% BSA. 20 x SET contained: 3 M NaCl, 20 mM EDTA and 0.4 M Tris-HCl (pH 7.8, at 25°C). LETS buffer contained: 0.1 M LiCl, 10 mM EDTA, 10 mM Tris-HCl (pH 7.4 at 4°C), 0.2% SDS, and 0.1% diethyl pyrocarbonate. 5 x RNA gel running buffer contained: 0.2 M MOPS (pH 7.0 at 25°C), 50 mM sodium acetate, and 5 mM EDTA.

Electro-elution sample buffer contained: 0.125 M Tris-HCl (pH 6.8), 0.1% SDS and 5% glycerol. V8-digestion buffer contained: 50 mM Tris-HCl (pH 7.8 at 4°C) and 0.1% SDS.

# *Methods*

## Purification of RNA polymerase II

*S. pombe* cells were grown at 30°C with aeration to 5 x $10^7$ cells per ml. About 250 g of the cells was harvested by centrifugation (8,000 x g for 10 min at 4°C) from 25 l of the culture, suspended in two volumes (about 500 ml) of buffer A containing 1% volume of protease inhibitor mixture (PIM), and disrupted repeatedly with Mini Lab (RANNIE) under 90 MPa pressure for 4 min. Crude extract was obtained after removing particulate materials by centrifugation at 28,000 x g for 30 min at 4°C. RNA polymerase II was precipitated by the addition of Polymin P as described in a method developed for RNA polymerase purification from wheat germ (21). One percent volume of 10% Polymin P solution was added to the crude extract and after stirring for 1 hour at 4°C, the precipitates were recovered by centrifugation at 15,000 x g for 20 min, and then extracted with 1 l of buffer B containing 1% volume of PIM. After centrifugation at 15,000 x g for 20 min, 36.1 g of solid ammonium sulfate per 100 ml (final, 2.4 mM) was added to the supernatant. After stirring for 20 min at 4°C, the precipitates were collected by centrifugation (28,000 x g for 20 min) and resuspended in 200 ml of TGED containing PIM.

The sample (Polymin P eluate) was centrifuged at 28,000 x g for 20 min and loaded onto a DEAE-Sephadex A-25 (300 ml) column equilibrated with TGED containing 100 mM ammonium sulfate and 1 mM PMSF. After washing the column with six times volume of TGED containing 100 mM ammonium sulfate and 1 mM PMSF, proteins were eluted with a 1.8 l linear gradient of ammonium sulfate from 100 to 400 mM in TGED buffer. Each fraction was assayed for non-specific transcription activity. Pooled RNA polymerase II fractions eluted at about 150 mM

ammonium sulfate were applied onto a Q Sepharose FF column (20 ml; Pharmacia), and eluted with 100 ml linear gradient of 150-400 mM ammonium sulfate in TGED. RNA polymerase II was eluted at about 250 mM ammonium sulfate. In some experiments, RNA polymerase II was further purified by Superose 6 gel filtration column chromatography in FPLC system (Pharmacia) in TGED containing 100 mM ammonium sulfate and concentrated by Mono Q Sepharose in FPLC system with linear gradient of 200 mM-450 mM ammonium sulfate in TGED.

## Non-specific transcription assay

Reaction mixture contained in 40 μl: 50 mM Tris-HCl (pH 7.9), 2 mM $MnCl_2$, 0.5 mM DTT, 50 mM ammonium sulfate, 0.5 mM each of ATP, GTP, CTP, 7 μM of UTP, 0.2 μCi [$^3$H]UTP (Amersham), 3 μl of the enzyme in TGED, and 2 μg of heat denatured calf thymus DNA. RNA synthesis was carried out for 20 min at 30°C and terminated by addition of 0.9 ml stop solution containing 5% trichloroacetic acid and 10 mM pyrophosphate. TCA-insoluble precipitates were collected on a GF/C glass-filter (Whatman), and washed three times with stop solution and then with 100% ethanol. The filters were counted for radioactivity with a liquid scintillation spectrometer (Beckman LS 1800). For selective inhibition of the RNA polymerase II, α-amanitin was added at 50 μg/ml. One unit of enzyme activity represents the incorporation of 1 nmole labeled substrate into TCA precipitates under these standard assay conditions. Protein concentrations were determined according to Lowry et al. (22)

## Reversible and irreversible denaturation of RNA polymerase II

RNA polymerase in TGED containing 50 mM ammonium sulfate was denatured by adding 8 M urea buffer at various concentrations followed by incubation at 0°C for 3 hours. Renaturation was carried out by dilution of the urea-treated enzyme with TEDA buffer and incubation at 0°C for 12 hours. The recovery of

enzyme activity was measured for 10 min under the non-specific transcription assay conditions. The final concentration of urea was adjusted to 0.2 M.

## Hybridization

Southern and Northern hybridization were performed according to the pubrished methods (23). DNA fragments transferred onto a nylon filter were immobilized by exposure to 265 nm UV light after the filter was dried up at room temperature. For low or high stringency hybridization conditions, pre-hybridization of the filter was carried out in low stringency hybridization buffer or high stringency hybridization buffer for 1 hour at 37°C or 42°C, respectively. For hybridization under each condition, labeled probe was added to the hybridization buffer used for pre-hybridization, and the filter in the buffer was incubated for 10 hours at the same temperature. After hybridization, the filter was washed twice using low stringency hybridization washing buffer or high stringency hybridization washing buffer for 20 min both at 42°C.

## Cloning of the *rpb1* gene

For cloning the *S. pombe rpb1* gene, Southern analysis was carried out (24) using the *S. cerevisiae RPB1* as a probe. *S. pombe* genomic DNA fragments digested with several restriction enzymes were separated within a 0.7% agarose gel containing 1 x TAE buffer. After depurination of double strand DNA fragments with 0.2 M HCl and denaturation to single strand DNAs with 0.5 M NaOH, the DNA fragments were transferred onto a Hybond-N nylon filter (Amersham) from the gel using capillary transfer method in 10 x SSC. Cross-hybridization was carried out using the *S. cerevisiae RPB1* gene labeled with Random Primed DNA Labeling Kit (Boehringer Mannheim) as a probe under low stringency conditions. Positive DNA fragments were eluted from an agarose gel and cloned into M13 phage. Cross-hybridization with filters which had been transferred M13 plaques was carried under the same low stringency conditions.

A λ EMBL3 phage library carrying 8.5-20 kbp (kilo base-pairs) inserts of total *S. pombe* DNA was constructed from partial *Sau*3AI digests according to Kaiser and Murray (25). Plaque hybridization was carried out for the phage library by the method of Anderson and Young (26). In brief, 1.8 x $10^4$ of plaques were transferred onto Hybond-N nylon filters, and the phage DNAs were depurinated with 0.2 M HCl and denatured with 0.5 M NaOH on the filters. Plaque hybridization was done under the high stringency conditions described in the previous section using a part of *S. pombe rpb1* gene labeled with [α-$^{32}$P]dATP and Random Primed DNA Labeling Kit (Boehringer Mannheim) as a probe. Starting from positive λ EMBL3 phage clones, the genomic DNA fragments were subcloned into M13 phages or pUC plasmids.

## Cloning of the *rpb3* gene

To isolate the *S. pombe rpb3* gene encoding the RNA polymerase II RPB3, a part of *rpb3* cDNA was amplified with RT-PCR using the synthetic primers and AmpliTaq DNA polymerase (Perkin-Elmer) for 25 cycles under conditions: denaturation for 1 min at 94°C; annealing for 1.5 min at 42°C; and elongation for 2 min at 73°C. For cloning of the complete *rpb3* gene, Southern hybridization was performed for the whole *S. pombe* DNA under high stringency conditions using the PCR product labeled with [α-$^{32}$P]dATP and Random Primed DNA Labeling Kit (Boehringer Mannheim) as a probe.

## DNA sequencing

For systematic sequencing of the *rpb1* and *rpb3* genes, deletion series of M13 phage clones carrying these genes were constructed by single-strand DNA deletion method using single-strand specific 3' to 5' exonuclease activity of T4 DNA polymerase (27). Deletion series of pUC and pBluescript II clones were constructed by double-strand deletion method using a combination of double-strand specific 3' to 5' exonuclease activity of *E. coli* exonuclease III and single strand specific 5' to 3' exonuclease activity of mung bean nuclease (28).

Sequencing was performed by the dideoxy chain termination method (29). Sequencing of the *rpb1* gene was carried out using Sequenase Ver. 2.0 DNA sequencing kit (USB) and [a-$^{35}$S]dCTP (Amersham) as a label (30). For sequencing of the *rpb3* gene, an automatic DNA sequencer (Pharmacia) were used with AutoRead and AutoCycle sequencing kit, and fluore-labeled primers (31). PCR-amplified cDNA were sequenced after cloning or directly (32). The mutant *rpb3* genes were sequenced using fluore-primers with the DNA sequencer (Pharmacia).

## Preparation of poly (A)+ RNA

Poly(A)$^+$ RNA was prepared from *S. pombe* by the standard RNA extraction procedure (23). In brief, 2 g of wild type *S. pombe* cells were harvested from 200 ml of YPD culture grown to late log phase and disrupted with 8 g of glass beads in 6 ml on LETS buffer with a Vortex mixer chilled in ice. After phenol treatment and ethanol precipitation, total nucleic acid precipitates were resuspended in 1 ml of TE buffer. About 1 mg of poly (A)$^+$ RNA was purified by oligo(dT) cellulose column chromatography.

## Northern analysis

Northern hybridization analysis was performed after denaturation of poly(A)$^+$ RNA by glyoxal and dimetyl sulfoxide (23). Ten μg of poly (A)$^+$ RNA were denatured with glyoxal and DMSO, and separated on an agarose gel containing 1 x RNA gel running buffer and 2.2 M of formaldehyde. The RNA was transferred onto a Hybond-N nylon filter by the same method as employed for Southern transfer. Hybridization was carried out under high stringency conditions using a DNA probe labeled with Random Primed DNA Labeling kit.

## Primer extension analysis

For primer extension analysis (33), primers were designed based on the sequence of genomic DNA, and end-labeled at 5' termini with [γ-$^{32}$P]ATP by T4 polynucleotide kinase activity. Two μg of poly(A)$^+$ RNA and 2 pmol of crude $^{32}$P-

labeled primer were mixed and incubated at 70°C for 10 min, and the mixture was quickly cooled on ice bath. After addition of 100 units of reverse transcriptase (Life Science) and the 5 x primer extension buffer, which contained 250 mM Tris-HCl (pH8.3 at 25°C), 375 mM KCl, 15 mM $MgCl_2$ and 0.5 mM each of 4 dNTPs, elongation was carried out at 37°C for 30 min and then 42°C for 30 min. The reaction was terminated by incubation at 70°C for 10 min. The mixture was treated with 0.5 μg of heat-treated RNase A (Sigma) and 30 units of RNase H (Takara) at 37°C for 30 min, and RNase-resistant products were analyzed by PAGE.

## Analysis of 3' and 5' termini of mRNA

To determine 3' termini of transcripts, poly(A) junction sites of mRNA were amplified by RT-PCR and sequenced. To determine 5' termini of transcripts, poly(dA) was added at cDNA 3' termini with terminal deoxynucleotidyl transferase and cDNA and poly(dA) junction regions were amplified by PCR. The PCR products were sequenced directly (32) or after cloning.

## Isolation of RNA polymerase II subunits

About 30 μg of RNA polymerase II RPB3 polypeptides was dissociated with 0.1 % SDS and fractionated by SDS-polyacrylamide gel electrophoresis according to the method of Leammli (34) using a discontinuous separation gel consisting of 15 (bottom) and 7.5% (top) polyacrylamide.

Electro-elution of proteins from the gel was performed by the modified method of Toda and Ohashi (35). In brief, about 0.2 ml homogenized gel pieces of the RPB3 polypeptides gel band was incubated in 1 ml of the electro-elution sample buffer for 30 min at room temperature and then boiled for 5 min at 70°C. The gel suspension was poured into a disposable column φ7 x 60 mm (Muromachi Kagaku) sealed with dialysis tube at the bottom. After electrophoresis for 3 hours at 2 mA using the Leammli's electrode reservoir solution (34), the RPB3 polypeptide was eluted from the gel pieces and recovered into 50 μl buffer in the dialysis tube together with CBB dye. The recovery was more than 70%.

## Amino acid sequencing

Isolated polypeptides was dialyzed against V8-digestion buffer for 2 hours at 4°C and then treated with 1 mg of V8 protease (PIERCE) per 10 mg sample protein for 12 hours at 37°C. Digested polypeptide fragments were separated by reverse phase chromatography on RPC18 column in SMART system (Pharmacia) with 0.1% trifluoroacetic acid (Wako Chemicals) containing either 10 or 70% acetonitril. Amino acid sequence of the purified fragments was determined with an Applied Biosystem model 977 linked with on-line HPLC model 120 protein sequencer.

## Mutagenesis of the *rpb3* gene

Mutagenesis of *rpb3* was carried out under a lower fidelity condition of PCR amplification in the presence of 0.5 mM $MnCl_2$ (36, 37). The PCR was performed for 30 cycles under the conditions: denaturation for 1 min at 94°C; annealing for 2 min at 57°C; and elongation for 3 min at 73°C. Under these conditions, the efficiency of mutagenization might be one base exchange happened in about 500 bases long. The mutated DNA fragments were cloned into plasmid between *Bal*I and *Eco*RV sites of pBrLEU2b3.

## Transformation of *S. pombe*

Electro-transformation was done by the slightly modified method of M. T. Hood et al. (38). In brief, the cells were grown to middle log phase ($3 \times 10^7$ cells/ml) in YPD at 30°C and washed with TE buffer containing 10 mM Tris-HCl (pH 7.5 at 4°C) and 0.1 mM EDTA. The cells were resuspended in TE to give a final concentration of $5 \times 10^9$ cells per ml. One μg of fragment DNA was added to 50 μl of the cell suspension, mixed with 50 μl of 60% PEG4000 and transferred to a 0.2 cm gapped cuvette. Electroporation was performed under the conditions of 25 μF, 8.5 kV/cm and 800 Ω with a BIO-RAD Gene Pulsar set.

Protoplast method was carried out along the method of D. Beach et al (39). The cells harvested at log phase were washed twice using 1.2 M Sorbitol and resuspended

to 2 x $10^8$ cells/ml using 1.2 M Sorbitol and 5 mg/ml Novozyme. After incubation for 30 min at 30°C, cells were washed three times using 1.2 M Sorbitol and 10 mM Tris-HCl (pH 7.6, at 25°C) and resuspended to 5 x $10^8$ cells/ml using 1.2 M Sorbitol, 10 mM Tris-HCl (pH 7.6, at 25°C), 10 mM $CaCl_2$ and 10 μg/ml DNA for transformation. After incubation for 15 min at 25°C, the cell suspension was diluted ten folds using 10 mM Tris-HCl (pH 7.6, at 25°C), 10 mM $CaCl_2$ and 20% PEG4000, and the cells were collected by centrifugation for 20 min at 25°C at 3,000 x g. The cells were resuspended in 1.2 M Sorbitol, 10 mM Tris-HCl (pH 7.6, at 25°C), 10 mM $CaCl_2$, 0.5 mg/ml yeast extract and 5 μg/ml leucine, incubated at 32°C for 30 min, and spread onto a SD plate containing 1.2 M Sorbitol.

## Gene analysis of *rpb3* mutants

Each mutant strain was grown in 10 ml of YPD at 25°C until about 5 x $10^7$ cells per ml. Cells were harvested by centrifugation at 5,000 x g for 10 min at 25°C, washed with water and treated with 1 ml of 0.5 mg/ml Novozyme (Novo Industrials) at 30°C for 30 min. After adding 1 g of alkaline-washed glass beads, 10 μl of 10 mg/ml RNase A and 10 μg of 0.5 M EDTA, the cell suspension was mixed heavily for 10 min, and mixed again after adding phenol saturated with TE buffer. The genomic DNA was prepared from the water phase.

To clone the mutant *rpb3* genes, PCR amplification was performed using specific primers for the *rpb3* gene and each genomic DNA as a template. Each PCR product was cloned into a multi-cloning site of pUC19. DNA sequencing was performed with an automatic DNA sequencer (Pharmacia) using seven species of synthetic fluore-primers specific for *rpb3* and two universal fluore-primers for the multi-cloning site of pUC vector.

For Southern analysis, 3 mg of mutant genomic DNA were digested with *Eco*RI and *Eco*RV, and separated on an agarose gel, and transferred onto a Hybond-N nylon filter (Amersham) by capillary transfer method (24). Hybridization was carried out using *S. pombe rpb3* and to *S. cerevisiae LEU2* gene under high

stringency condition with ECL direct nucleic acid labeling and detection system (Amersham) .

## DAPI staining

DAPI staining was carried out by the modified method of Moreno et al (40). In brief, cultured cells were fixed at 0°C for 20 min with addition of 10% volume of glutaraldehyde. After washing twice with cold water, the cells were resuspended in 1 μg/ml DAPI solution and diluted to about 4 x $10^4$ cells per μl. The cell suspension was spread onto a slide glass, and observed with a fluorescence microscope (OLYMPUS, BH2).

## Growth analysis of mutants

Wild-type and mutant S. pombe were grown at 25°C in SD medium. The steady state of log-phase culture was maintained for more than 24 hours by continuous supply of fresh medium. The cell cultures were divided into two flasks and the incubation was continued at either permissive temperature, 25°C, or non-permissive temperature, 37°C.

# RESULTS AND DISCUSSION

## *CHAPTER I*
## *Purification of RNA polymerase II*

### Purification of RNA polymerase II

Cell lysates were prepared from the wild-type *S. pombe* strain 972h$^-$ by disruption with a high pressure French press (Mini Lab). RNA polymerase II was purified in five steps (summarized in Fig. 1): precipitation of nucleic acids and associated proteins from cell extracts with Polymin P; elution of proteins from the Polymin P precipitates with ammonium sulfate solution; separation of RNA polymerase II from other class RNA polymerases by DEAE-Sephadex column chromatography (Fig. 2). RNA polymerase I activity was a major part of total RNA polymerase activities and piled up with the RNA polymerase II activity (data not shown, Fig. 2). To achieve better the separation of RNA polymerase II from RNA polymerase I, the ammonium sulfate concentration was reduced to just only 100 mM by adding TGED buffer according conductivity of the elute. Under this condition, RNA polymerase I flew through the DEAE-Sephadex column; fractionation by Q Sepharose FF column chromatography (Fig. 3); and final-step purification by Superose 6 gel filtration (Fig. 4). During the enzyme purification, the RNA polymerase II activity was determined by measuring single-strand DNA-dependent and α-amanitin-sensitive incorporation of radioactive UTP into acid-insoluble-fraction (non-specific transcription assay). Starting from 25 l of cell culture, about 1.1 mg of RNA polymerase II was obtained at the step of Q Sepharose chromatography (Table 1). Inhibition of 50% activity of RNA polymerase II was achieved at 2.5 μg/ml α-amanitin. The α-amanitin sensitivity of *S. pombe* RNA polymerase II was as low as that from *S. cerevisiae* (40), but lower than that of RNA polymerase II from higher eukaryotes, suggesting that the conformation of α-amanitin-associating site on yeast RNA polymerases II is different from those of higher eukaryotes (41).

## Components of purified RNA polymerase II

The highly purified *S. pombe* RNA polymerase II contained more than eleven polypeptides as analyzed by SDS-polyacrylamide gel electrophoresis (35). The molecular masses of these components were estimated to be 210 (subunit 1), 150 (subunit 2), 40 (subunit 3), 34, 25, 22, 20, 16, 15, 13.5 and 13 kDa (Fig. 5A). The relative stoichiometry of the three large subunits, RPB1, 2 and 3, was 1:1:2, as measured with a laser densitometer ULTROSCAN XL (LKB) and both silver- and Coomassie brilliant blue-stained SDS-polyacrylamide gels (Fig. 5B). This molar ratio is identical with that of the *S. cerevisiae* enzyme (42). After prolonged storage of the purified RNA polymerase II, RPB1 was partially converted to a degradation product of 190 kDa, supposedly lacking the C-terminal repeated domain (CTD).

## Reversible and irreversible denaturation of RNA polymerase II with urea

As a preliminary attempt for establishment of the reconstitution system of RNA polymerase II, reversible denaturation was tried. The purified RNA polymerase II was treated with urea at various concentrations and then diluted to reduce urea concentration to 0.2 M, which did not interfere with the RNA polymerase II activity. Q Sepharose FF and gel filtration fractions of RNA polymerase II were used as a crude and a purified enzymes, respectively. *E. coli* RNA polymerase was used as a positive control of reversible denaturation. The concentration of both *E. coli* RNA polymerase and *S. pombe* RNA polymerase II was adjusted to 0.3 units per μl. By the simple dilution method, more than 50% of activity was recovered in the case of *E. coli* RNA polymerase treated with 6M urea (data not shown).

RNA polymerase II was not inactivated up to 0.5 M urea (Fig. 6). At the urea concentration between 1.0-2.5 M, the RNA polymerase II lost activity. The activity was, however, perfectly recovered after the denatured enzyme was diluted with TEDA buffer and stored more than 10 min (Fig. 6). It should be noted that the RNA polymerase II treated with from 0.25 to 2.5 M urea regained the activity 20% higher than that of the untreated enzyme. The result may suggest that some inhibitory

component(s) was removed from the RNA polymerase II or that some inactive enzyme forms were reactivated after one cycle of denaturation-renaturation. When the denaturation was done at urea concentrations above 3.0 M, RNA polymerase II activity was not recovered. It appears that, once RNA polymerase II was dissociated into individual subunits, or lost one or more of the essential subunits for the transcription activity, the reassemble did not take place only by simple dilution of urea. On the other hand, reversible denaturation was done by dialysis with TGED buffer containing 50 mM ammonium sulfate or TEDA buffer after treatment with several concentration of urea. RNA polymerase activity was not recovered in this case, either.

For renaturation of the RNA polymerase II treated with 4 M urea, several reagents (listed in Table 2) were added into the dilution buffer for renaturation, but none of them was effective to recover the enzyme activity.

To investigate the core-like structure of RNA polymerase II, the RNA polymerase II denatured in 2 M and 4 M urea was applied to gel filtration in the buffer containing 2 M urea. Multi subunit complexes were observed in both cases, but several small components were dissociated from the complexes in the 4 M urea treatment (Fig. 7). One of the prospects come from these facts is that the core-like structure of RNA polymerase II is saved after the treatment with 4 M urea, but the enzyme has no transcription activity even if the concentration of urea is diluted because the structure is irreversibly denatured. Otherwise, the core-like structure loose some essential subunits with the treatment of 4 M urea, and reassemble is not happened.

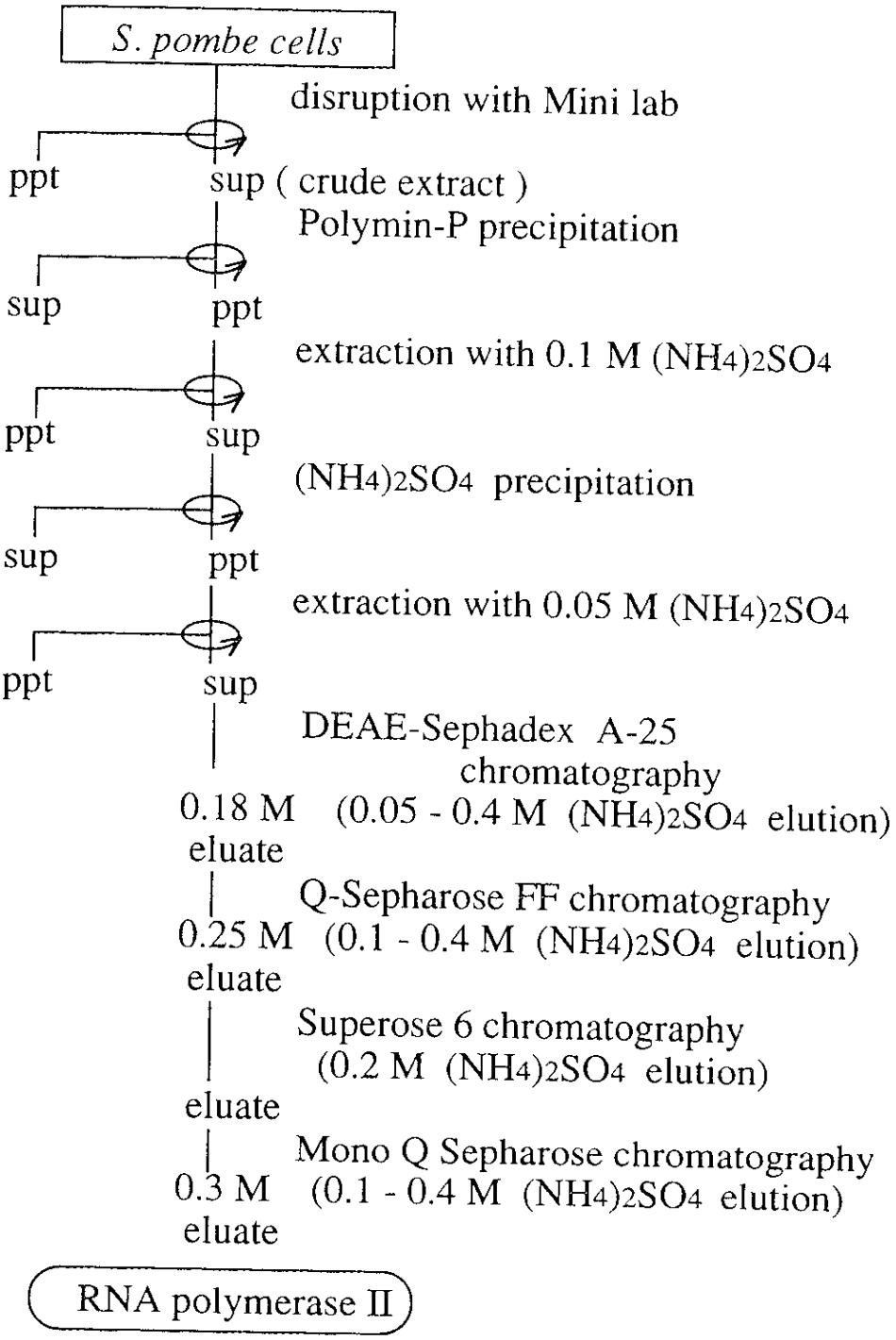



Fig. 1. Flowchart of RNA polymerase II purification.
"ppt" and "sup" represent precipitate and supernatant, respectively.

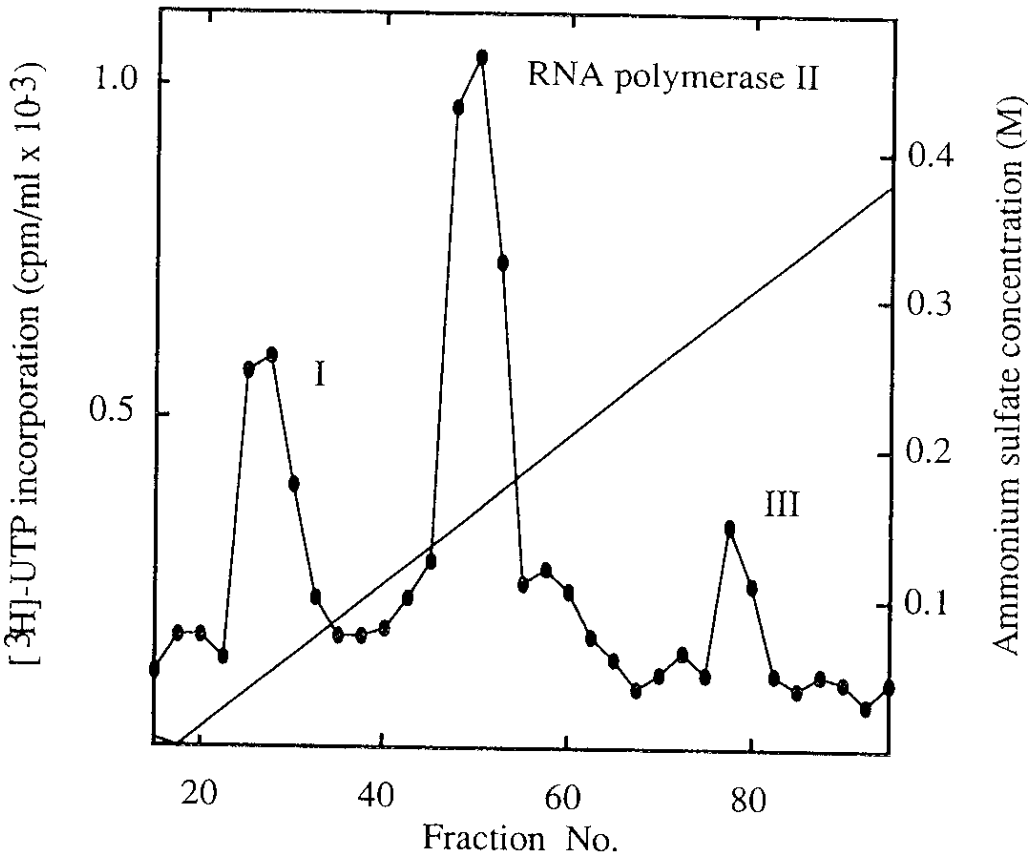



Fig. 2. Separation of RNA polymerase II from RNA polymerases I and III.
The Polymin P eluate was fractionated on a DEAE-Sephadex column (φ2.6 x 57 cm) at a flow rate of 200 ml/hour. Non-specific transcription assay was performed for 3 μl aliquots of 20 ml fractions under the standard reaction conditions described in Materials and Methods. ●—●, RNA polymerase activity; ·····, ammonium sulfate concentration. RNA polymerase activity in the flow-through fractions was resistant to 50 μg/ml α-amanitin.

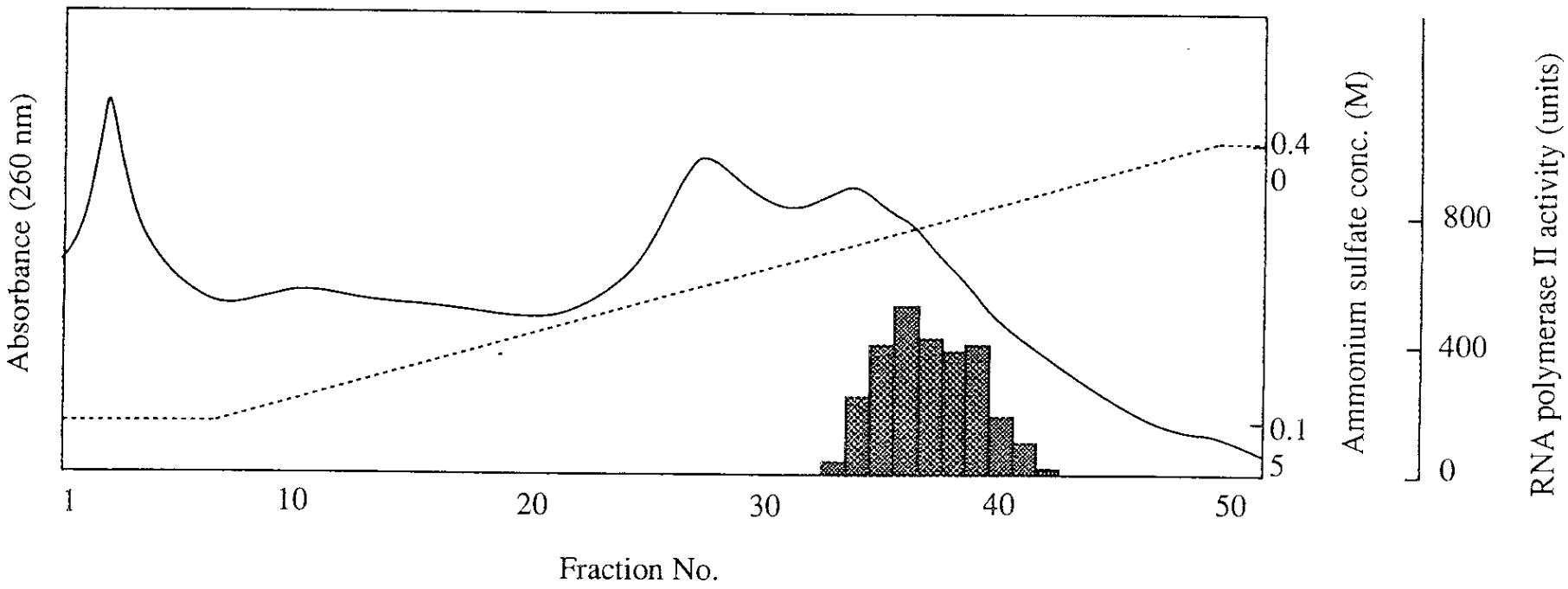



Fig. 3. Q-Sepharose FF column chromatography of RNA polymerase II.
DEAE Sephadex A-25 fraction was applied onto Q-Sepharose FF column (φ1.6 x 15 cm). Elution was carried out using 100 ml of a 0.15-0.40 M linear gradient of ammonium sulfate at a flow rate of 150 ml per hour. An aliquot (3 μl) of each fraction (2 ml) was assayed for non-specific transcription activity as described in Experimental Procedures.
——, absorbance at 260 nm; ·····, ammonium sulfate concentration; bar graph, total activity of each fraction.

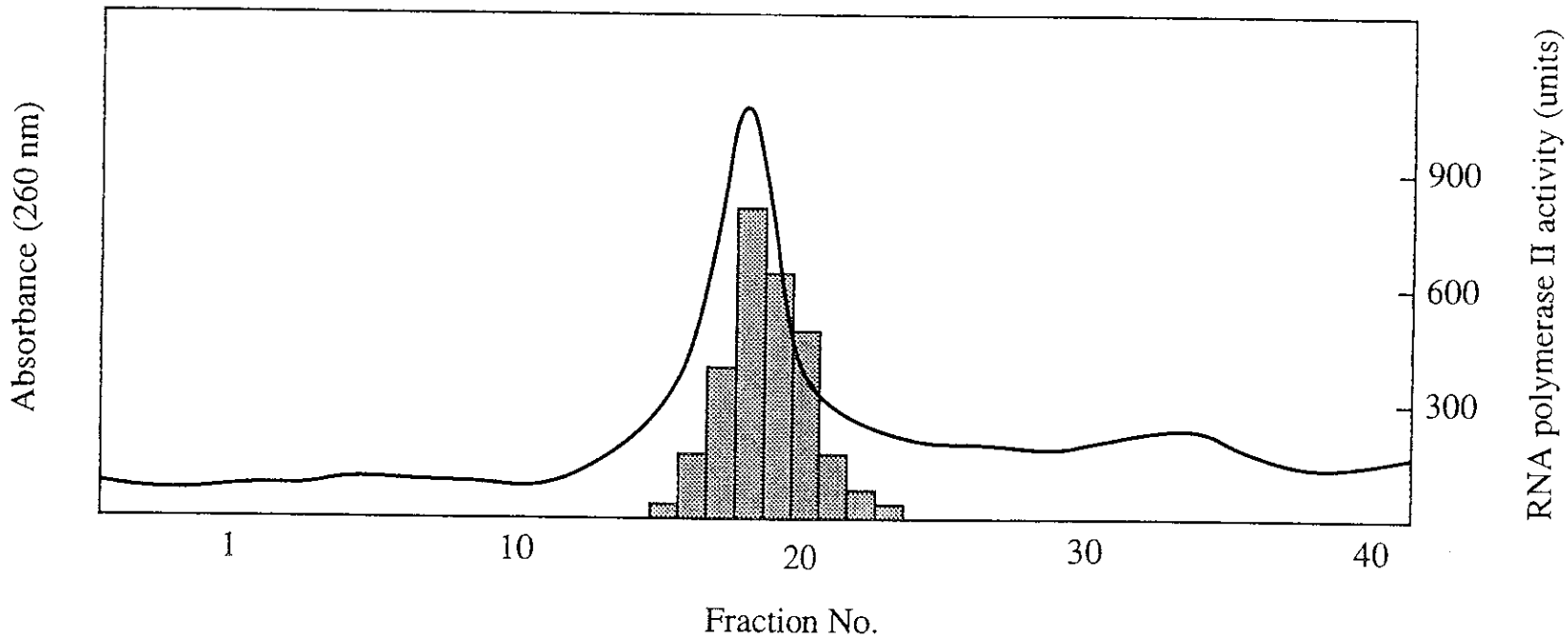



Fig. 4. Superose 6 column chromatography of RNA polymerase II.
Q Sepharose fraction was applied onto Surerose 6 column (φ1.6 x 30 cm) and eluted with TGED buffer containing 100 mM ammonium sulfate at a flow rate of 30 ml per hour. An aliquot (3 μl) of each fraction (0.8 ml) was assayed for non-specific transcription acitivity as described in Experimental Procedure.——, absorbance at 260 nm; bar graph, total activity of each fraction.

Table 1. Purification of RNA polymerase II from *S. pombe*.

| Purification step | Total protein (mg) | Total activity (units) | Recovery (%) | Specific activity (units/mg) | Purity (fold) |
|---|---|---|---|---|---|
| Crude extract | 10600 | 18000 | 100 | 0.17 | 1.0 |
| Polymin P eluate | 2200 | 13000 | 73 | 0.59 | 3.7 |
| $(NH_4)_2SO_4$ | 1500 | 12000 | 68 | 0.80 | 4.9 |
| DEAE-Sephadex | 19 | 4900 | 27 | 25.8 | 154 |
| Q-Sepharose FF | 1.1 | 4100 | 23 | 373 | 2200 |
| Superose 6 | 0.5 | 3300 | 18 | 660 | 3900 |

RNA polymerase II was purified from 250 g (wet weight) of *S. pombe*.

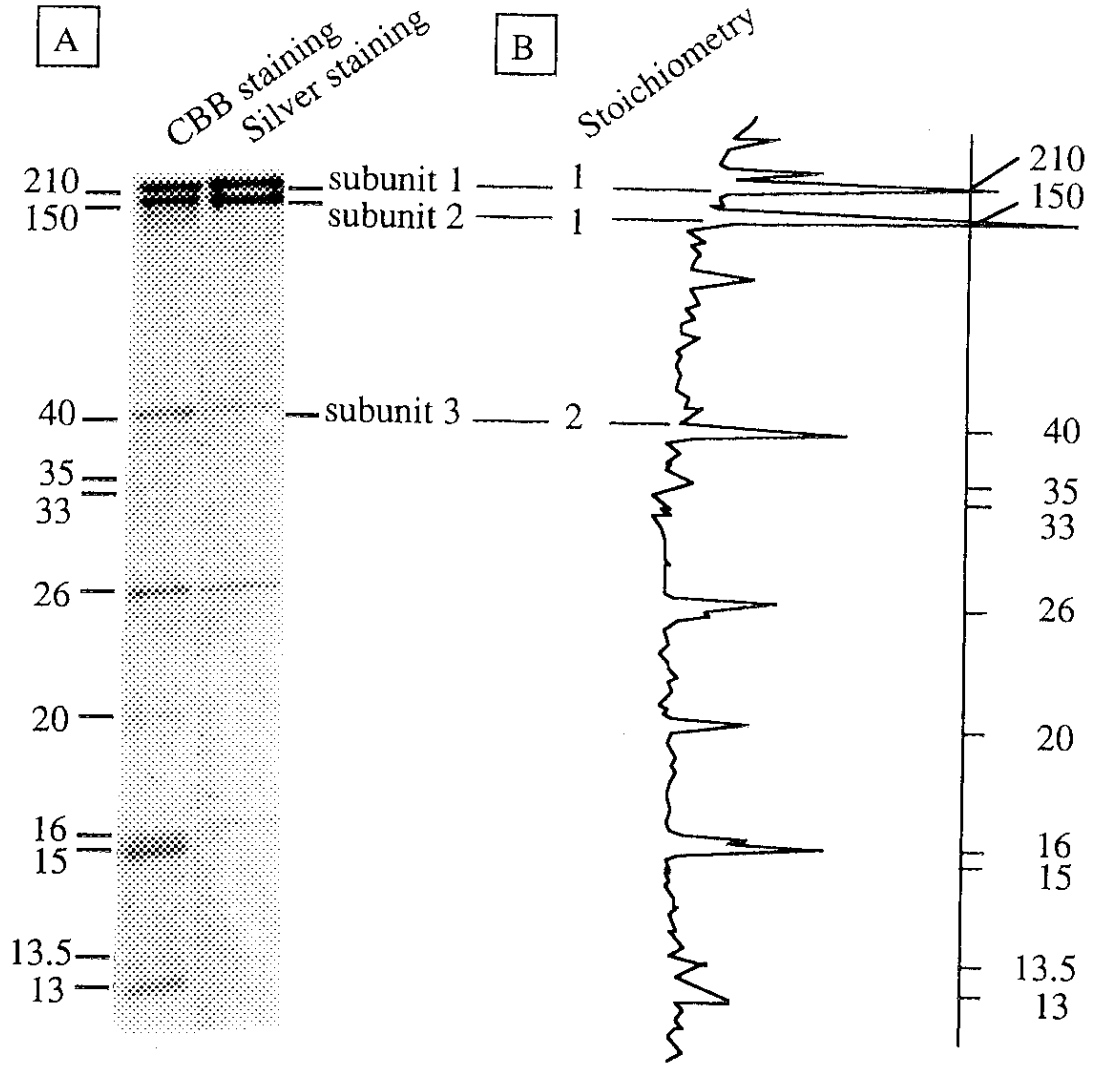



Fig. 5. Polypeptides composition of purified RNA polymerase II.
(A) SDS-PAGE analysis of the RNA polymerase II fraction. The Superose 6 fraction was electrophoresed on a polyacrylamide gradient gel (5-15%) under denaturation conditions (22) and stained with Coomassie brilliant blue (CBB) or silver stain. (B) Scanning of the stained RNA polymerase II subunits bands. The CBB stained gel was scanned at 633 nm with a LKB ULTROSCAN XL laser densitometer. Stoichiometry of the three large subunits was calculated from peak areas of the subunit bands.

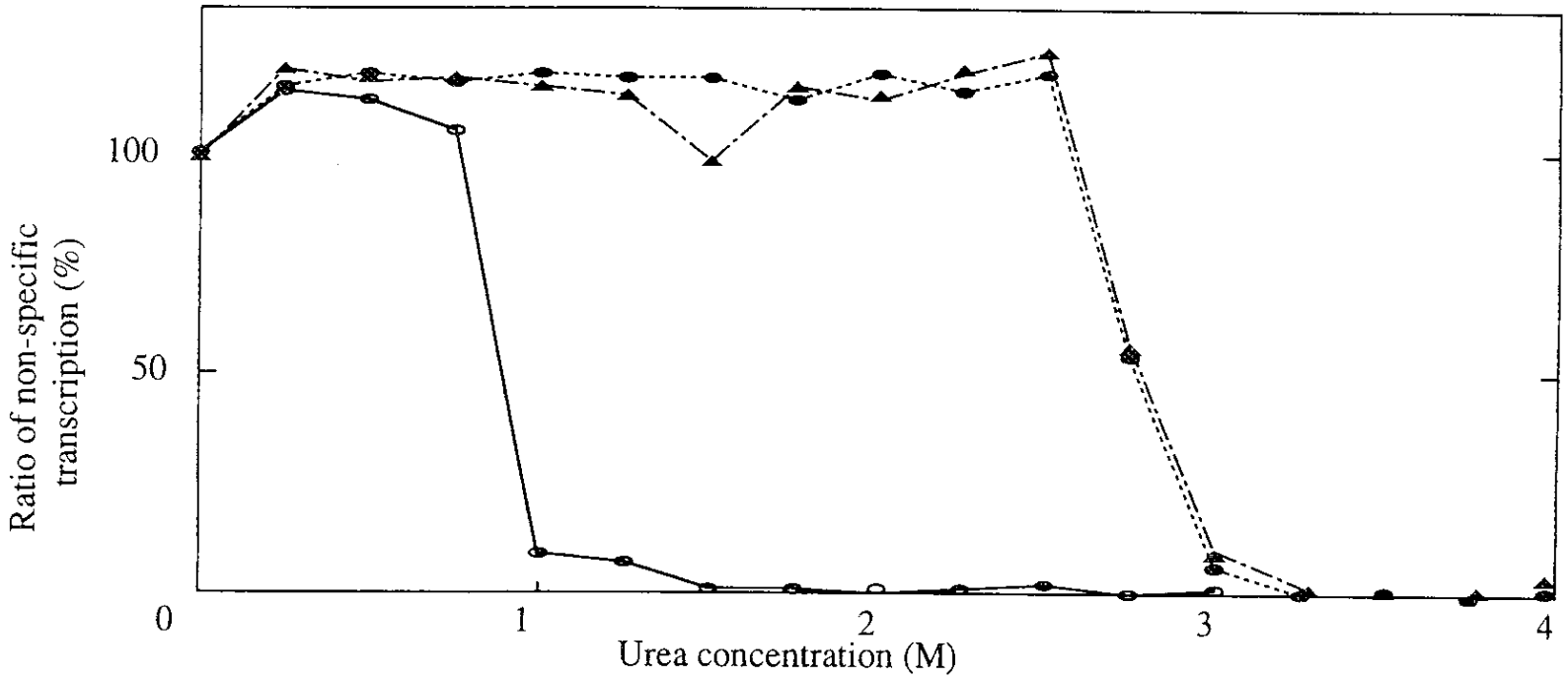



Fig. 6. RNA polymerase II activity after urea treatment.
The purified RNA polymerse II was treated with urea at various concentration as indicated in the figure at 0°C for 3 hours and diluted with TEDA buffer at 4°C for 18 hours ( o—o, without dilution; ▲···▲, 4-fold dilution; ●···● ,12-fold dilution). Urea treatment was carried out and dilution was kept.

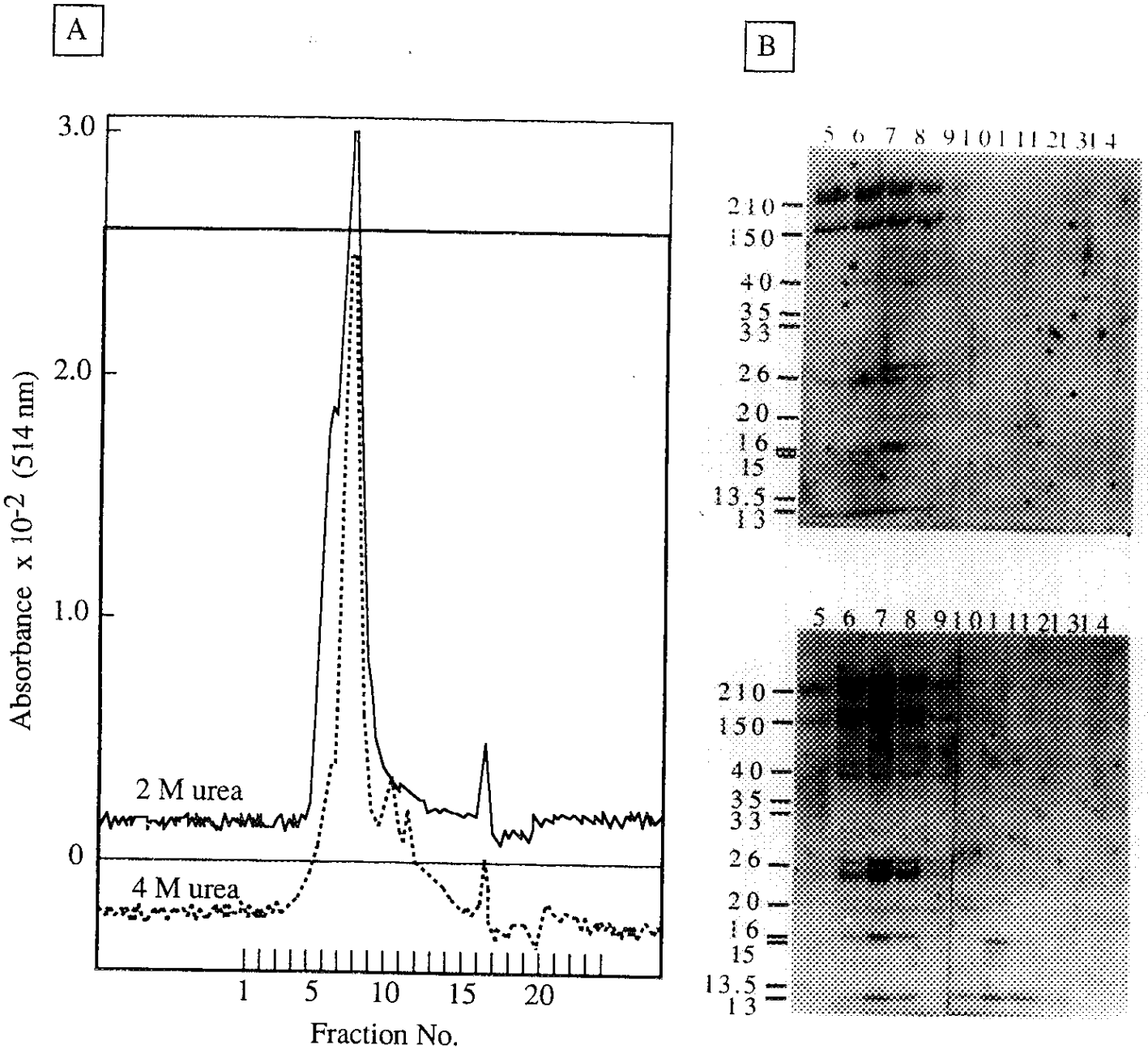



Fig. 7. Gel filtration of RNA polymrase II treated with urea.
RNA polymerase II (Superose 6 column fraction, 25 μg) was denatured in 2 M or 4M urea for 3 hours at 0°C and applied onto a Superose 6 column (φ0.4 x 30 cm). Elution was carried out with TEDA buffer containing 2 M urea and 10% glycerol at a flow rate of 5 ml per hour). (A) Elution patterns. Absorbance of elute at 514 nm was detected with micro flow cell in SMART system (Pharmacia). ——, 2 M urea-treated RNA polymerase II; ·····, 4 M urea-treated RNA polymerase II. (B) SDS-PAGE analysis of the fractions. Each fraction was electrophoresed on a polyacrylamide gradient gel (5-15% acrylamide) under the denaturation condition.

Table 2. List of reagents tested for RNA polymerase II renaturation.

| Reagent or protein | Concentration | Time(hour) |
|---|---|---|
| $MnCl_2$ | 8, 40 mM | 1 |
| $MgCl_2$ | 40, 200 mM | 20 |
| | 20 - 320 mM | 1 |
| $ZnCl_2$ | 40, 200 mM | 20 |
| NTPs | 0.2, 1 mM | 20 |
| DTT | 20, 100 mM | 1 |
| | 100 mM | 20 |
| NaF | 20, 100 mM | 20 |
| | 10 - 160 mM | 1 |
| Triton X-100 | 0.004, 0.02% | 20 |
| | 0.01 - 0.16% | 1 |
| | 0.02 - 0.11% | 2, 6 |
| | 0.005 - 0.15% | 6 |
| Tween 20 | 0.02 - 0.11% | 2, 6 |
| | 0.005 - 0.15% | 6 |
| DOC | 0.02 - 0.11% | 2, 6 |
| Glycerol | 20, 60% | 1 |
| PEG4000 | 1.6, 8% | 20 |
| DMSO | 4, 20% | 1 |
| | 20% | 20 |
| Ethylenglycol | 2.4, 12% | 20 |
| | 0.5 - 8% | 1 |
| *E. coli* proteins | | |
| DnaK, ATP | 40, 200 ng/ml | 1 |
| BSA | 0.4, 2 mg/ml | 20 |
| σ subunit | 20, 100 ng/ml | 1 |
| | 40, 200 ng/ml | 20 |
| α subunit | 40, 200 ng/ml | 20 |
| Phosphocellulose | | 20 |
| DEAE Sephadex | | 20 |

# *CHAPTER 2*
# *Cloning and Characterization of the Gene for the Largest Subunit of RNA Polymerase II*

## Cloning of the *rpb1* gene

For cloning the *S. pombe* gene for the largest subunit (subunit 1) of RNA polymerase II, a *S. cerevisiae* DNA fragment containing the entire *RPB1* gene encoding subunit 1 of RNA polymerase II was used as a hybridization probe. The cloning strategy is summarized in Fig. 8. Initially, total *S. pombe* DNA was digested with various restriction enzymes and analyzed by Southern hybridization using a *S. cerevisiae* probe (an *Eco*RI-*Hind*III fragment of the pRP19, ref. 20) under low stringency conditions. As shown in Fig. 9, one major and several minor bands were identified for all the restriction enzymes used. First, a 2.2 kbp *Hind*III-*Pst*I fragment (see lane 4 in Fig. 9) was size-selected by electro-elution from an agarose gel, and cloned into M13 phage mp18. The sequence determination of the cloned *S. pombe* DNA indicated that a high degree of amino acid (aa) sequence homology exists between the open reading frame in this fragment (C-terminal region downstream from aa residue 1,238; see Figs. 11 and 13) and a part of *S. cerevisiae* RNA polymerase II subunit 1 (C-terminal region downstream from aa residue 1,235; see Fig. 13). The result strongly suggested that the cloned fragment was a part of subunit 1 gene of *S. pombe* RNA polymerase II. In order to clone the entire gene, a genomic library of *S. pombe* DNA was constructed using a phage vector and screened it using the cloned *S. pombe* DNA as a hybridization probe. Screening of approximately 1.8 x $10^4$ plaques under high stringency conditions yielded 12 positive clones. The restriction map analysis showed that all these clones shared some identical DNA restriction fragments, presumably originated from the same chromosomal locus.

## Structure of the *rpb1* gene

Nucleotide (nt) sequence was determined for a continuous DNA segment of 7,079 bp including the above-mentioned *Hind*III-*Pst*I fragment. The outline of the

sequence is illustrated in Fig. 10, while details are described in Fig. 11. An uninterrupted open reading frame spanning nt position 656 to 5,554 (the nt position 1 was set to the first base of the putative initiation codon; see Figs. 10 and 11) encodes a polypeptide of 1,633 amino acids in length with a high degree of aa sequence homology to the region from aa position 118 to the C terminus of the *S. cerevisiae* RNA polymerase II subunit 1 (Fig. 13). In the upstream region from nt position 655, the reading frame with homology to the rest (N-terminal proximal region of 117 aa residues) of the *S. cerevisiae* subunit is interrupted six times. Since a set of the consensus sequence for intron-exon junctions was found at both 5' and 3' boundaries of each interrupting sequence (see Fig. 24), mRNA sequence of the corresponding regions were analyzed .

For analysis of the mRNA sequence, cDNA was synthesized using a synthetic primer a with the sequence of (5')GCGGCCGCGAATTC(T)17(3'), which is capable of hybridizing to mRNA poly(A)$^+$ tail, and amplified a portion of the cDNA (nucleotide position 14 to 779 in the corresponding genomic DNA sequence) by PCR using primers #3 and #8 (see Figs. 10 and 11 for positions and sequences of primers used in this study). A single DNA band was detected when PCR products were analyzed by gel electrophoresis. Direct sequencing of the PCR product demonstrated that none of the six interrupting sequences was present in the amplified region of cDNA. It was therefore concluded that six introns exist in the N-terminal region of this gene (for details see Figs. 10, 11 and 24), and that the subunit 1 of *S. pombe* RNA polymerase II is composed of 1,752 aa residues with the molecular mass of 194 kDa. The gene organization is in sharp contrast with the fact that the subunit 1 gene of *S. pombe* RNA polymerase I contains no intron (24).

The aa sequence between position 1,460 and 1,548 showed relatively weak homology to the *S. cerevisiae* RNA polymerase II subunit 1. Then cDNA sequence of this region was analyzed. For this purpose, the cDNA synthesized as described above was subjected to PCR using primers #9 and #12 (see the primer positions in Fig. 10), and a part of the region (nt position 4,571 to 5,032 in the corresponding genomic DNA sequence) was sequenced using primers #10 and #12. The cDNA

sequence agreed completely with the genomic sequence, indicating that the diverged aa sequence is not due to the insertion of intron in this region.

Detailed Southern analysis of genomic DNA digested with various restriction enzymes showed that this gene is present as a single copy in the *S. pombe* genome (data not shown). Furthermore, gene disruption experiment by insertion of the *S. pombe ura4* gene showed that the gene is essential for viability (data not shown). This gene was then proposed to designate as *rpb1*, according to the nomenclature proposed for the *S. cerevisiae* RNA polymerase genes by Nonet et al. (43).

## Transcription organization

Northern analysis demonstrated that the size of the *rpb1* transcript is about 5.6 kbp in length (Fig. 12A). The start site of the transcript determined by primer extension experiment using primer #1 was located at nt position -347 (Fig. 12B; see Figs. 10 and 11 for primer #1). When primer #2 (Figs. 10, 11 and 12B) was used for primer extension, a consistent result was obtained (data not shown). The result of the sequence determination of the *rpb1* cDNA 5'-flanking region cloned with primer a and #2 agreed with this conclusion.

The 5'-flanking region upstream from the protein coding region contains six ATG codons before the putative ATG start codon at nt position 1 (see Fig. 11). It was concluded that ATG at nt position 1 is the start codon from the following reasons: 1) The predicted start codon is located within the first exon, and no ATG codon exists in the further upstream in the same open reading frame following a TGA stop codon (nt position -69); 2) no consensus sequences for intron-exon junctions can be found between the transcription start site at nt -347 and the putative start codon; 3) the N-terminal proximal region of the predicted *S. pombe* RNA polymerase II subunit 1 from the initiator Met to the first conserved domain (domain A) is as large as that of *S. cerevisiae* RNA polymerase II subunit 1 (Fig. 13; a significant homology can be found in this region between *S. pombe* and *S. cerevisiae*); and, 4) the sequence near the putative initiation codon fits well to the Kozak's rules (44) for the consensus sequence for translational initiation.

To determine the nucleotide sequence near the 3' end of the transcript, cDNA was synthesized as described above, and a 3'-terminal region of the *rpb1* transcript was amplified with using primer b with the sequence (5')CGCCGGCGCTTAAGTTT(3'), which hybridizes to the end of cDNA started from primer a, and primer #11 specific to *rpb1* gene (see Figs. 10, 11). The PCR product was digested with *Bam*HI (the cleavage site is located at nt position 4,832 within the coding region) and *Eco*RI (the cleavage site is within primer b), and cloned the resulting *Bam*HI-*Eco*RI fragment into M13 phage mp19 for sequencing. The amplified fragments from two independent clones contained the junction points between *rpb1* transcript and poly(A) tail, at position 5,654 or 5,655 in one clone, and positions 5,659, 5,660 or 5,661 in another clone (the ambiguities are due to the presence of multiple A residues in the genomic sequence at the junction point). The length of mature mRNA (about 5.6 kbp) predicted from the sequence analysis is in good agreement with the result obtained by Northern analysis (see above).

## Structure of the RNA polymerase II subunit 1

The predicted aa sequence of *S. pombe* RNA polymerase II subunit 1 was compared with those of other eukaryotes (Figs. 13 and 14) and those of other β' homologues of prokaryotes and archibacteria (Fig.14). Eight structural domains (domain A to H) conserved among the largest subunits of all three species of eukaryotic RNA polymerases were also identified in the *S. pombe* RNA polymerase II subunit 1. These domains have significant homologies to the corresponding regions of *E. coli* β', strongly suggesting that these domains are involved in some common and essential functions associated with the RNA polymerases.

Domain A has a putative zinc-binding site with the consensus sequence of $CX_2CX_9HX_2H$ (aa position 69 to 85 in *S. pombe* sequence) (45,46,47). The functional importance of this motif in the subunits 1 of RNA polymerase I and II was confirmed by isolation of *S. cerevisiae* ts mutants with mutations in this region (47,48). In domains C and D, there are two different sequences (position 354 to 384, and position 499 to 507, respectively, in the *S. pombe* sequence) with homologies to

the sequences conserved within *E. coli* DNA polymerase I and T7 DNA polymerase (49). The former sequence has a single two-helix motif and might play some roles in DNA binding as discussed previously (49,50). Domain F is believed to be involved in binding of α-amanitin, a potent inhibitor of eukaryotic RNA polymerase II, since aa substitution of Asn to Asp at position 793 within this domain of mouse RNA polymerase II subunit 1 renders the RNA polymerase insensitive to this drug (51). This position and the surrounding sequences are highly conserved among the RNA polymerase II subunits 1 from all the organisms so far examined. However, the Asn residue at corresponding sites in the two yeast subunits (position 775 in *S. pombe* and position 769 in *S. cerevisiae*) are substituted for Ile or Ser, respectively. Interestingly, both *S. cerevisiae* and *S. pombe* RNA polymerase II are less sensitive to inhibition by α-amanitin than RNA polymerase II from other higher eukaryotes (52; M. Yamagishi, unpublished observations).

Besides the eight conserved domains, a unique C-terminal repetition of a heptapeptide (CTD) with the unit sequence of YSPTSPS is highly conserved (an exception is an unusual structure found in the CTD region of the largest subunits of two RNA polymerase II species from Trypanosoma; see refs. 23,35). The CTD does not exist in the β' subunit of prokaryotic RNA polymerases. Deletion experiments for the hamster, mouse, and *S. cerevisiae* RNA polymerase II subunit 1 genes revealed that loss of most of the CTD repeats causes lethal effect on cell growth. *S. cerevisiae* mutants containing deletions shorter than half the length of CTD exhibit conditional lethal phenotypes (54,55,56). These observations altogether suggest that CTD plays an indispensable function for cell viability. Several lines of experiment indicate that CTD is needed for transcriptional activation by trans-activating factors such as *S. cerevisiae* GAL4 (57,58) and for interaction between RNA polymerase II and general transcription factors *in vitro* (59). The number of the repetitions is, however, variable among different species: 17 in *Plasmodium falciparum* (60), 26 or 27 in *S. cerevisiae* (49,55), 41 in *Arabidopsis* (61), 42 in *Caenorhabditis elegans* (62), 45 in *Drosophila* (50) and 52 in mouse and hamster (63,55). The subunit 1 of *S. pombe* RNA

polymerase II was now found to have 29 repeat units (Fig. 15). The repetition ends exactly after the final repeat.

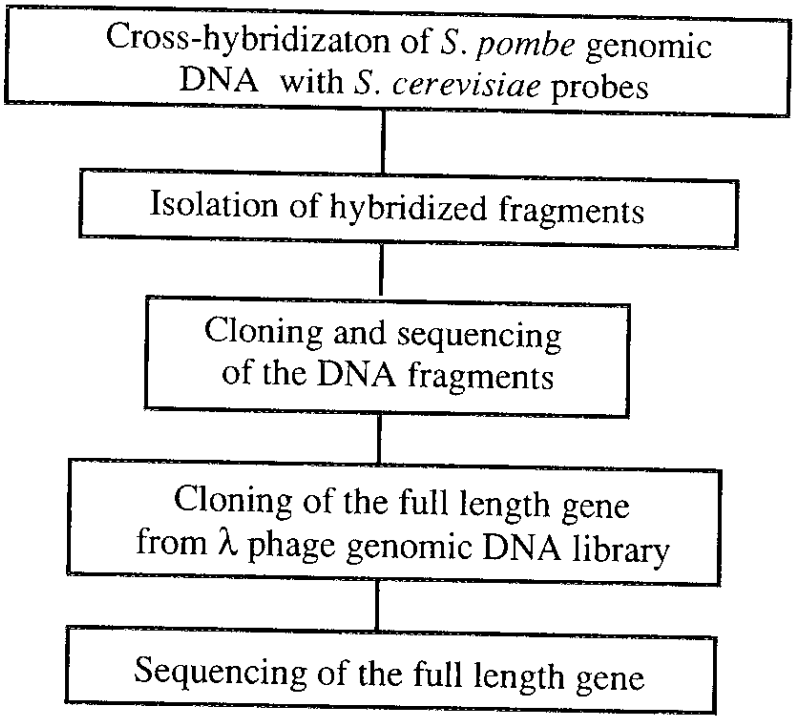



Fig. 8. Cloning strategy for the RNA polymerase II subunit 1 gene. The gene encoding RNA polymerase II subunit 1 was isolated by cross-hybridization using *S. cerevisiae* probes.

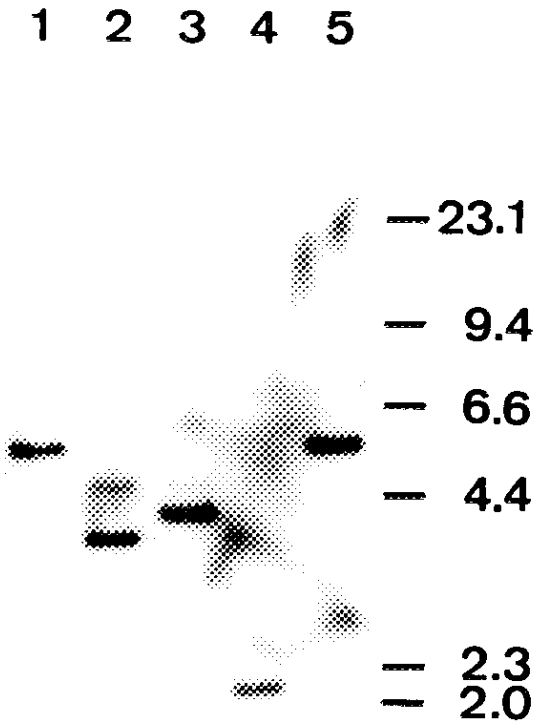



Fig. 9. Southern hybridization of *S. pombe* genomic DNA using *S. cerevisiae RPB1* gene.
Five μg of total *S. pombe* DNA were digested with various restriction enzymes, and subjected to Southern analysis under the low stringency condition. Restriction enzymes used are: *Eco*RI (lane 1); *Eco*RI/ *Hind*III (lane 2); *Hind*III3); *Hind*III/*Pst*I (lane 4); and *Pst*I (lane 5).

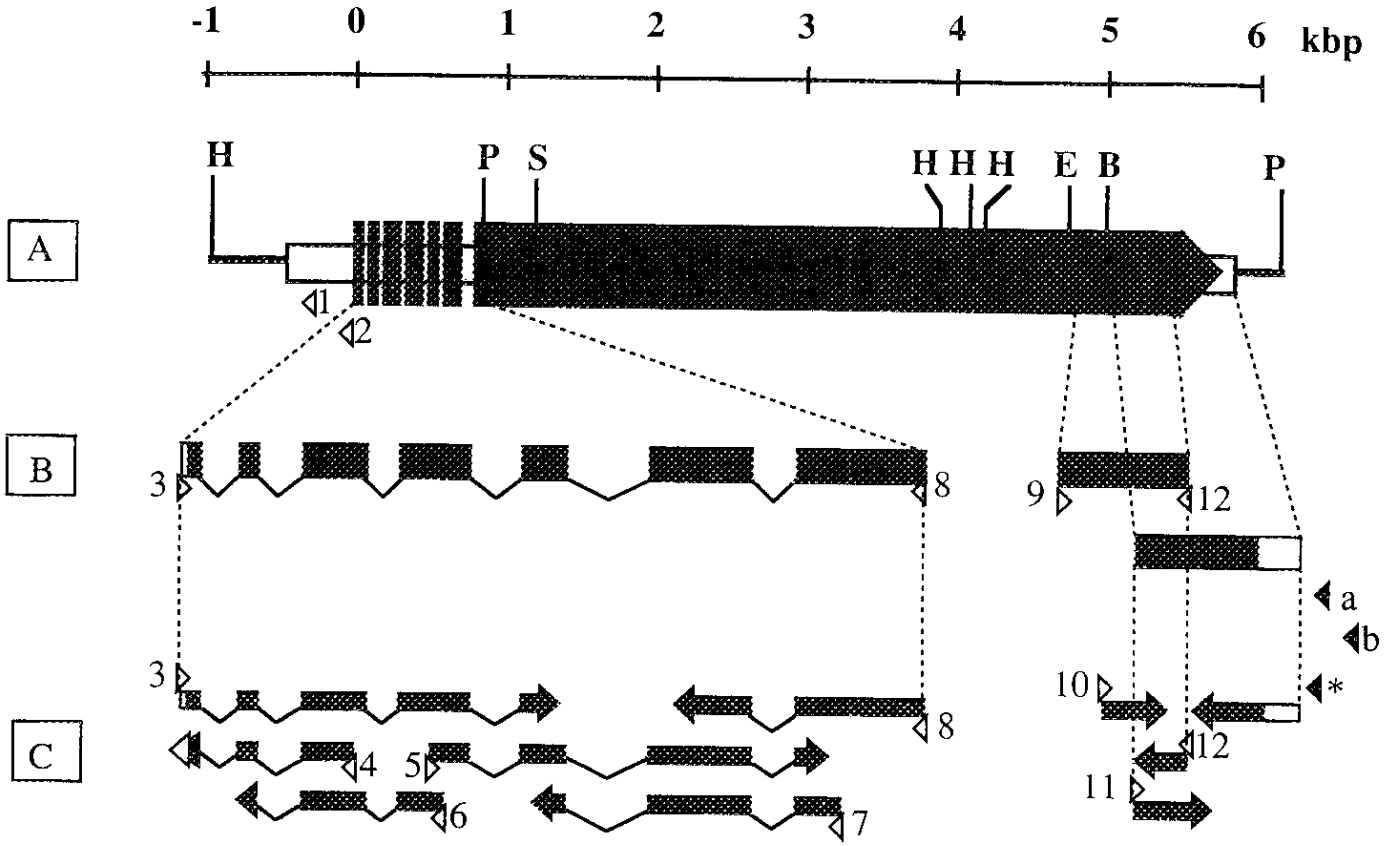
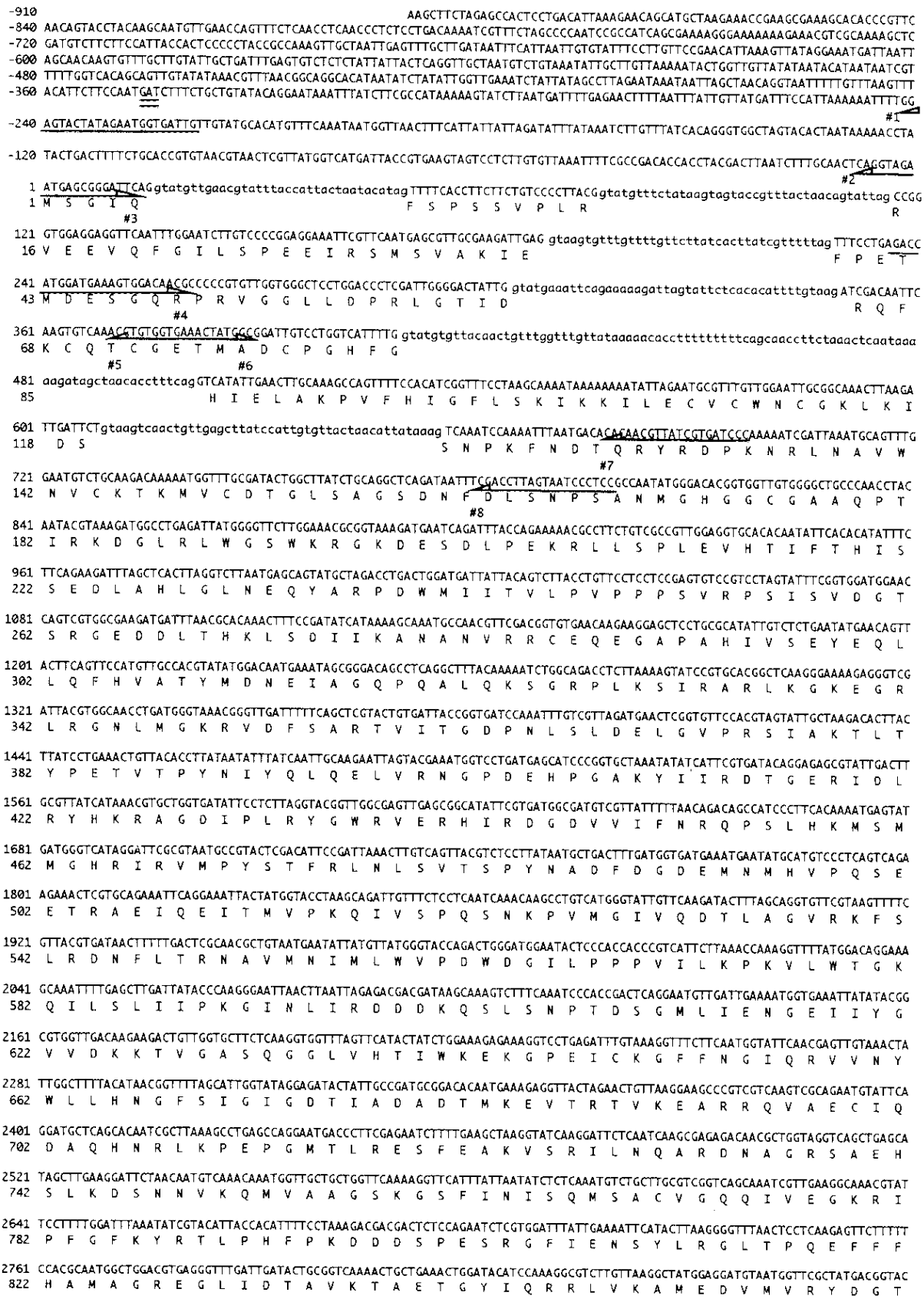



Fig. 10. Structure of the *rpb1* gene.
(A) Restriction map of the *rpb1* gene. Filled and open boxes represent exons and non-coding sequences including six introns, respectively. Restriction enzyme sites mapped are: H, *Hind*III; P, *Pst*I; S, *Sac*I; E, *Eco*RI; B, *Bam*HI. (B) PCR amplified cDNA sequence. Gaps connected with thin lines represent genomic DNA sequences which are not present in cDNA. Positions and directions of primers used for PCR and DNA sequencing are shown by triangles below the PCR products. (C) Sequencing strategy of the PCR products. Thick lines represent stretches sequenced, and arrows under the lines represent primers used for the direct sequencing. The sequence with an asterisk was determined after cloning into mp19. Thin lines represent gaps which are not present in cDNA.



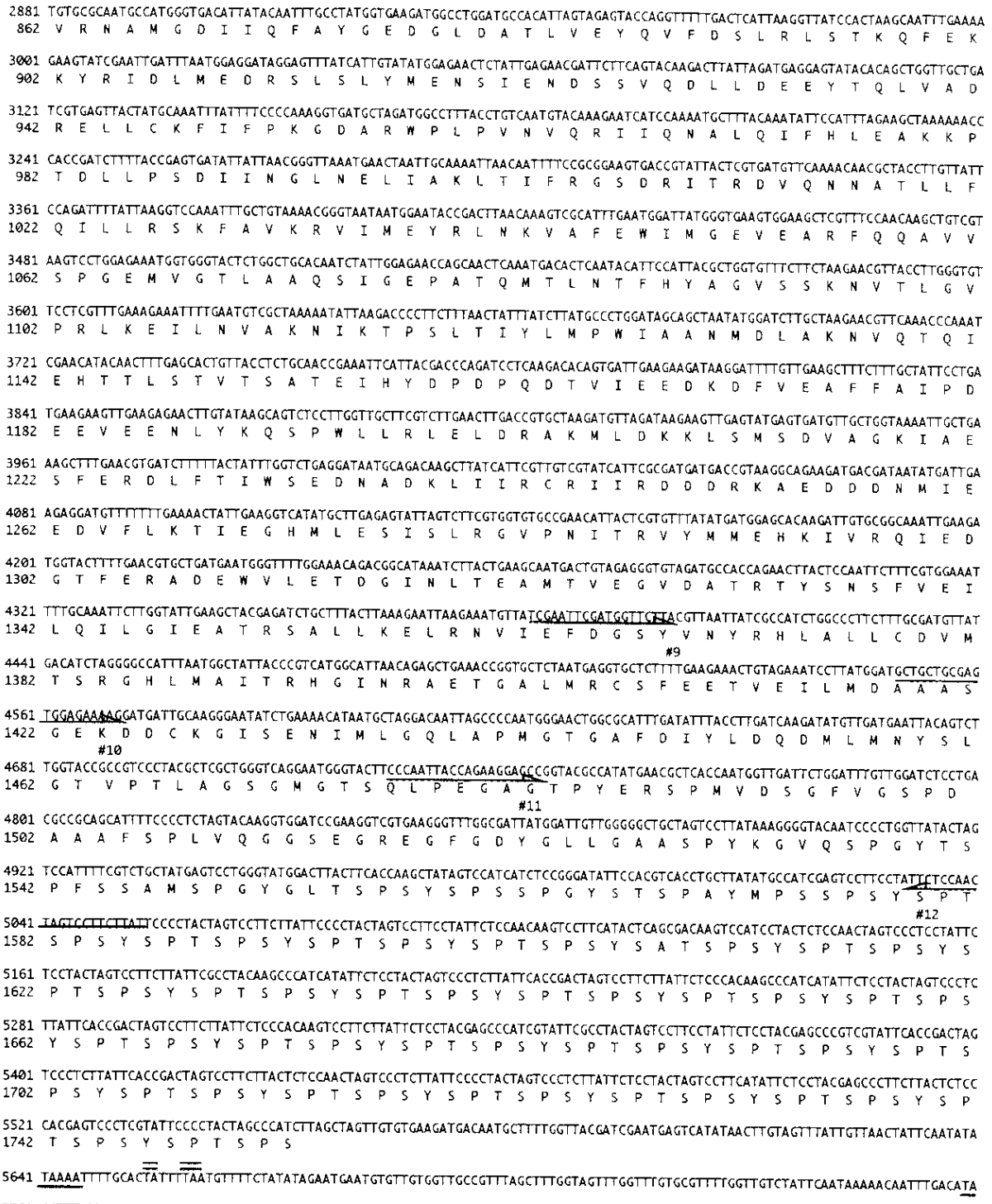


Fig. 11. Nucleotide and predicted amino acid sequence of the *rpb1* gene. The coding sequence of the RNA pol ymerase II subunit 1 starts at nt postion 1 and ends at nt postion 5,552. This sequence is interrupted near N-terminal proximal region by six introns, indicated by small letters. The 5' and 3' ends of the transcript are indicated by double underlines and double overlines, respectively. The positions, directions and names of the primers are indicated by arrowheads. Site for polyadenylation is underlined.

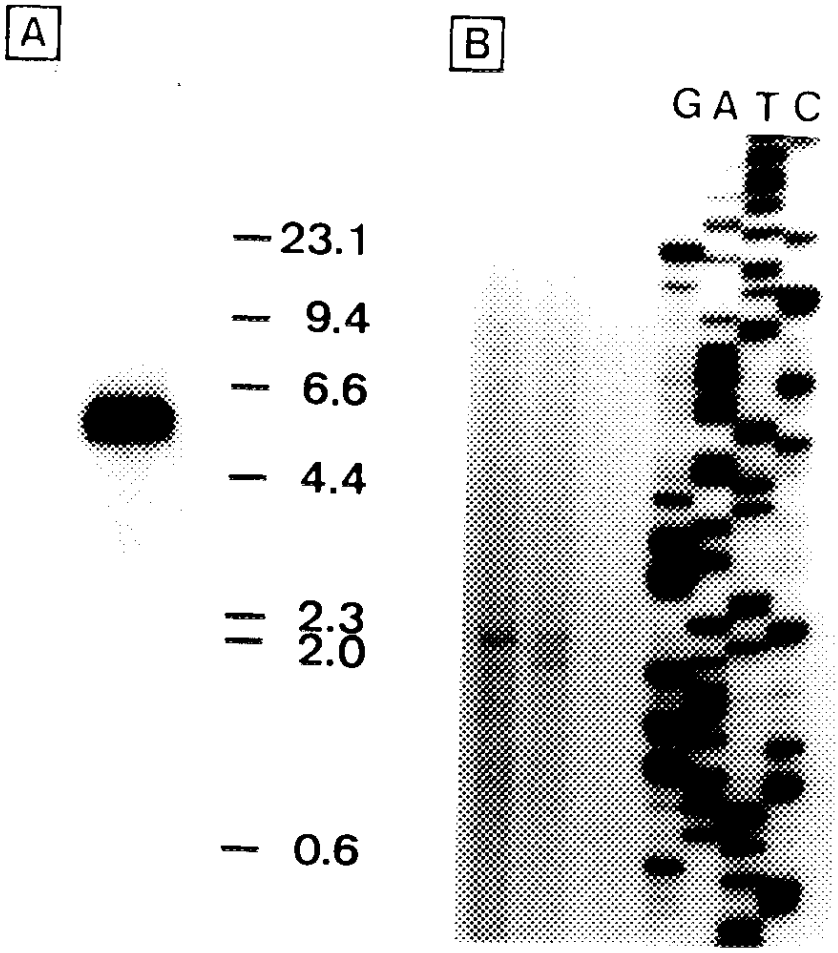



Fig. 12. RNA analysis of *rpb1* gene.
(A) Northern blot analysis of the *rpb1* transcript. Seven μg of mRNA was subjected to Northern analysis. The blot was hybridized with a 5.4 kb *Pst*I fragment containing a part of *rpb1* (see Fig. 10). (B) Primer extension analysis. $^{32}$P-labeled primer #1 (see Fig. 11) was hybridized to five μg of mRNA and elongated by AMV reverse transcriptase. Products were electrophoresed along DNA sequence ladders obtained using a subcloned *rpb1* DNA fragment and primer #1 (lanes 3-7). Two different mRNA preparations were analyzed (lanes 1 and 2).


S.p M----S-GIQFSPSSVPLRRVEEVQFGILSPEEIRSMSVAK--IEFPETMDESGQRPRVGGLLDPRLGTI 63
S.c M----V-GQQYS--SAPLRTVKEVQFGLFSPEEVRAISVAK--IRFPETMDETQTRAKIGGLNDPRLGSI 61
A.th M----DTRFPFS--PAEVSKVRVVQFGILSPDEIRQMSVIH--VEHSETTEKG--KPKVGGLSDTRLGTI 60
C.e M----ALVGVDF--QAPLRIVSRVQFGILGPEEIKRMSVAH--VEFPEVYENG--KPKLGGLMDPRQGVI 60
M. MHGGGPPSG-DS--ACPLRTIKRVQFGVLSPDELKRMSVTEGGIKYPETTEGG--RPKLGGLMDPRQGVI 65
D.m M----STPT-DS--KAPLRQVKRVQFGILSPDEIRRMSVTEGGVQFAETMEGG--RPKLGGLMDPRQGVI 61
S.p DRQFKCQTCGETMADCPGHFGHIELAKPVFHIGFLSKIKKILECVCWNCGKLKIDSSNP--KFNDTQRYR 131
S.c DRNLKCQTCQEGMNECPGHFGHIDLAKPVFHVGFIAKIKKVCECVCMHCGKLLLDEHNE--LMRQALAIK 129
A.th DRKVKCETCMANMAECPGHFGYLELAKPMYHVGFMKTVLSIMRCVCFNCSKILADE-VCRSLFRQAMKIK 129
C.e DRRGRCMTCAGNLTDCPGHFGHLELAKPVFHIGFLTKTLKILRCVCFYCGRLLIDKSAPRVLEILKKTGT 130
M. ERTGRCQTCAGNMTECPGHFGHIELAKPVFHVGFLVKTMKVLRCVCFFCSKLLVDSNNPKIKDILAKSKG 135
D.m DRTSRCQTCAGNMTECPGHFGHIDLAKPVFHIGFITKTIKILRCVCFYCSKMLVSPHNPKIKEIVMKSRG 131
S.p DPKNRLNAVWNVCKTKMVCDTGLSAGSDNFDLSNPS--------ANMGHGGCGAAQPTIRKDGLRLWGSWK 194
S.c DSKKRFAAIWTLCKTKMVCETDVPSEDD------PT--------QLVSRGGCGNTQPTIRKDGLKLVGSWK 186
A.th NPKNRLKKILDACKNKTKCDGGDDIDDVQSHSTDEP-------VKKSRGGCGAQQPKLTIEGMKMIAEYK 192
C.e NSKKRLTMIYDLCKAKSVCEGAAEKEEGMPDDPDDPMND----GKKVAGGCGRYQPSYRRVGIDINAEWK 196
M. QPKKRLTHVYDLCKGKNICEGGEEMDNKFGVEQPEG--DEDLTKEKGHGGCGRYQPRIRRSGLELYAEWK 203
D.m QPRKRLAYVYDLCKGKTICEGGEDMDLTKENQQPDP--N----KKPGHGGCGHYQPSIRRTGLDLTAEWK 195
S.p --RGKDESDL--------PEKRLLSPLEVHTIFTHISSEDLAHLGLNEQYARPDWMIITVLPVPPPSVRP 254
S.c KDRATGDADE--------PELRVLSTEEILNIFKHISVKDFTSLGFNEVFSRPEWMILTCLPVPPPPVRP 248
A.th --IQRKKNDEPDQLPEPAERKQTLGADRVLSVLKRISDADCQLLGFNPKFARPDWMILEVLPIPPPPVRP 260
C.e --KNVNEDTQ--------ERKIMLTAERDLEVFQQITDEDILVIGMDPQFARPEWMICTVLPVPPLAVRP 256
M. ---HVNEDSQ--------EKKILLSPERVHEIFKRISDEECFVLGMEPRYARPEWMIVTVLPVPPLSVRP 262
D.m ---HQNEDSQ--------EKKIVVSAERVWEILKHITDEECFILGMDPKYARPDWMIVTVLPVPPLAVRP 254
S.p SISVDGTSRGEDDLTHKLSDIIKANANVRRCEQEGAPAHIVSEYEQLLQFHVATYMDNEIAGQPQALQKS 324
S.c SISFNESQRGEDDLTFKLADILKANISLETLEHNGAPHHAIEEAESLLQFHVATYMDNDIAGQPQALQKS 318
A.th SVMMDATSRSEDDLTHQLAMIIRHNENLKRQEKNGAPAHIISEFTQLLQFHIATYFDNELPGQPRATQKS 330
C.e AVVTFGSAKNQDDLTHKLSDIIKTNQQLQRNEANGAAAHVLTDDVRLLQFHVATLVDNCIPGLPTATQKG 326
M. AVVMQGSARNQDDLTHKLADIVKINNQLRRNEQNGAAAHVIAEDVKLLQFHVATMVDNELPGLPRAMQKS 332
D.m AVVMFGAAKNQDDLTHKLSDIIKANNELRKNEASGAAAHVIQENIKMLQFHVATLVDNDMPGMPRAMQKS 324
S.p GRPLKSIRARLKGKEGRLRGNLMGKRVDFSARTVITGDPNLSLDELGVPRSIAKTLTYPETVTPYNIYQL 394
S.c GRPVKSIRARLKGKEGRIRGNLMGKRVDFSARTVISGDPNLELDQVGVPKSIAKTLTYPEVVTPYNIDRL 388
A.th GRPIKSICSRLKAKEGRIRGNLMGKRVDFSARTVITPDPTINIDELGVPWSIALNLTYPETVTPYNIERL 400
C.e GRPLKSIKQRLKGKEGRIRGNLMGKRVDFSARTVITADPNLPIDTVGVPRTIAQNLTFPEIVTPFNVDKL 396
M. GRPLKSLKQRLKGKEGRVRGNLMGKRVDFSARTVITPDPNLSIDQVGVPRSIAANMTFAEIVTPFNIDRL 402
D.m GKPLKAIKARLKGKEGRIRGNLMGKRVDFSARTVITPDPNLRIDQVGVPRSIAQNLTFPELVTPFNIDRM 394
S.p QELVRNGPDEHPGAKYIIRDTGERIDLRYHKRAGDIPLRYGWRVERHIRDGDVVIFNRQPSLHKMSMWGH 464
S.c TQLVRNGPNEHPGAKYVIRDSGDRIDLRYSKRAGDIQLQYGMKVERHIMDNDPVLFNRQPSLHKMSMMAH 458
A.th ---VRLVFISFSETKYIIRDDGQRSDLRYLKKSSDQHLELGYKVERHLQDGDFVLFNRQPSLHKMSIMGH 467
C.e QELVNRGDTQYPGAK----ENGARVDLRYHPRAADLHLQPGYRVERHMKDGDIIVFNRQPTLHKMSMWGH 462
M. QELVRRGNSQYPGAKYIIRDNGDRIDLRFHPKPSDLHLQTGYKVERHMCDGDIVIFNRQPTLHKMSMWGH 472
D.m QELVRRGNSQYPGAKYIVRDNGERIDLRFHPKSSDLHLQCGYKVERHLRDDDLVIFNRQPTLHKMSMWGH 464
S.p RIRVMPYSTFRLNLSVTSPYNADFDGDEMNMHVPQSEETRAEIQEITMVPKQIVSPQSNKPVMGIVQDTL 534
S.c RVKVIPYSTFRLNLSVTSPYNADFDGDEMNLHVPQSEETRAELSQLCAVPLQIVSPQSNKPCMGIVQDTL 528
A.th RIRIMPYSTFRLNLSVTSPYNADFDGDEMNMHVPQSFETRAEVLELMMVPKCIVSPQANRPVMGIVQDTL 537
C.e RVKILPWSTFRMNLSVTSPYNADFDGDEMNLHLPQSLETRAEIEEIAMVPRQLITPQANKPVMGIVQDTL 532
M. RVRILPWSTFRLNLSVTTPYNADFDGDEMNLHLPQSLETRAEIQELAMVPRMIVTPQSNRPVMGIVQDTL 542
D.m RVKVLPWSTFRMNLSCTSPYNADFDGDEMNLHVPQSMETRAEVENIHITPRQIITPQANKPVMGIVQDTL 534
S.p AGVRKFSLRDNFLTRNAVMNIMLWVPDWDGILPPPVILKPKVLWTGKQILSLIIPKGINLIRD-----DD 599
S.c CGIRKLTLRDTFIELDQVLNMLYWVPDWDGVIPTPAIIKPKPLWSGKQILSVAIPNGIHLQRF-----DE 593
A.th LGCRKITKRDTFIEKDVFMNTLMWWEDFDGKVPAPAILKPRPLWTGKQVFNLIIPKQINLLRYSAWHADT 607
C.e CAVRMMTKRDVFIDWPFMMDLLMYLPTWDGKVPQPAILKPKPLWTGKQVFSLIIPGNVNVLRTHSTHPDS 602
M. TAVRKFTKRDVFLERGEVMNLLMFLSTWDGKVPQPAILKPRPLWTGKQIFSLIIPGHINCIRTHSTHPDD 612
D.m TAVRKMTKRDVFITREQVMNLLMFLPTWDAKMPQPCILKPRPLWTGKQIFSLIIPGNVNMIRTHSTHPDE 604
S.p KQS----LSNPTDSGMLIENGEIIYGVVDKKTVGASQGGLVHTIWKEKGPEICKGFFNGIQRVVNYWLLH 665
S.c GTT----LLSPKDNGMLIIDGQIIFGVVEKKTVGSSNGGLIHVVTREKGPQVCAKLFGNIQKVVNFWLLH 659
A.th ETG----FITPGDTQVRIERGELLAGTLCKKTLGTSNGSLVHVIWEEVGPDAARKFLGHTQWLVNYWLLQ 673
C.e EDSGPYKWISPGDTKVIIEHGELLSGIVCSKTVGKSAGNLLHVVTLELGYEIAANFYSHIQTVINAWLIR 672
M. EDSGPYKHISPGDTKVVVENGELIMGILCKKSLGTSAGSLVHISYLEMGHDITRLFYSNIQTVINNWLLI 682
D.m EDEGPYKWISPGDTKVMVEHGELIMGILCKKSLGTSAGSLLHICFLELGHDIAGRFYGNIQTVINNWLLF 674
S.p NGFSIGIGDTIADADTMKEVTRTVKEARRQVAECIQDAQHNRLKPEPGMTLRESFEAKVSRILNQARDNA 735
S.c NGFSTGIGDTIADGPTMREITETIAEAKKKVLDVTKEAQANLLTAKHGMTLRESFEDNVVRFLNEARDKA 729
A.th NGFTIGIGDTIADSSTMEKINETISNAKTAVKDLIRQFQGKELDPEPGRTMRDTFENRVDQVLNKARDDA 743
C.e EGHTIGIGDTIADQATYLDIQNTIRKAKQDVVDVIEKAHNDDLEPTPGNTLRQTFENKVNQILNDARDRT 742
M. EGHTIGIGDSIADSKTYQDIQNTIKKAKQDVIEVIEKAHNNELEPTPGNTLRQTFENQVNRILNDARDKT 752
D.m EGHSIGIGDTIADPQTYNEIQQAIKKAKDDVINVIQKAHNMELEPTPGNTLRQTFENKVNRILNDAHDKT 744
S.p GRSAEHSLKDSNNVKQMVAAGSKGSFINISQMSACVGQQIVEGKRIPFGFKYRTLPHFPKDDDSPESRGF 805
S.c GRLAEVNLKDLNNVKQMVMAGSKGSFINIAQMSACVGQQSVEGKRIAFGFVDRTLPHFSKDDYSPESKGF 799
A.th GSSAQKSLAETNNLKAMVTAGSKGSFINISQMTACVGQQNVEGKRIPFGFDGRTLPHFTKDDYGPESRGF 813
C.e GSSAQKSLSEFNNFKSMVVSGSKGSKINISQVIACVGQQNVEGKRIPFGFRHRTLPHFIKDDYGPESKGF 812
M. GSSAQKSLSEYNNFKSMVVSGAKGSKINISQVIAVVGQQNVEGKRIPFGFKHRTLPHFIKDDYGPESRGF 822
D.m GGSAKKSLTEYNNLKAMVVSGSKGSNINISQVIACVGQQNVEGKRIPYGFRKRTLPHFIKDDYGPESRGF 814
S.p IENSYLRGLTPQEFFFHAMAGREGLIDTAVKTAETGYIQRRLVKAMEDVMVRYDGTVRNAMGDIIQFAYG 875
S.c VENSYLRGLTPQEFFFHAMGGREGLIDTAVKTAETGYIQRRLVKALEDIMVHYDNTTRNSLGNVIQFIYG 869
A.th VENSYLRGLTPQEFFFHAMGGREGLIDTAVKTSETGYIQRRLVKAMEDIMVKYDGTVRNSLGDVIQFLYG 883
C.e VENSYLAGLTPSEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKAMESVMVNYDGTVRNSLAQMVQLRYG 882
M. VENSYLAGLTPTEFFFHAMGGREGLIDTAVKTAETGYIQRRLIKSMESVMVKYDATVRNSINQVVQLRYG 892
D.m VENSYLAGLTPSEFYFHAMGGREGLIDTAVKTAETGYIQRRLIKAMESVMVNYDGTVRNSVGQLIQLRYG 884
S.p EDGLDATLVEYQVFDSLRLSTKQFEKKYRIDLME-DRSLSLYM----------ENSIENDSSVQDLL 931
S.c EDGMDAAHIEKQSLDTIGGSDAAFEKRYRVDLLNTDHTLDPSLL----------ESGSEILGDLKLQVLL 929
A.th EDGMDAVWIESQKLDSLKMKKSEFDRTFKYEIDDENWNPTYLSD----------EHLEDLKGIRELRDVF 943
C.e EDGLDGMWVENQNMPTMKPNNAVFERDFRVSVAQNAIKLMDLTDNKFLRKNYSEDVVREIQESEDGISLV 952
M. EDGLAGESVEFQNLATLKPSNKAFEKKFRFDYTNERALRRTLQE----------DLVKDVLSNAHIQNEL 952
D.m EDGLCGELVEFQNMPTVKLSNKSFEKRFKFDWSNERLMKKVFTD----------DVIKEMTDSSEAIQEL 944
S.p DEEYTQLVADR-ELLCKFIFPKGDARWPLPVNVQRIIQNALQIFHLEAKKPTDLLPSDIINGLNELIAKL 1000
S.c DEEYKQLVKDR-KFLREV-FVDGEANWPLPVNIRRIIQNAQQTFHIDHTKPSDLTIKDIVLGVKDLQENL 997
A.th DAEYSKLETDRFQLGTEI-ATNGDSTWPLPVNIKRHIWNAQKTFKIDLRKISDMHPVEIVDAVDKLQERL 1012
C.e ESEWSQLEEDR-RLLRKD-FPRGDAKIVLPCNLLRLIWNAQKIFKVDLRNAVNLSPLHVISGVRELSKKL 1020
M. EREFERMREDR-EVLRVI-FPTGDSKVVLPCNLLRMIWNAQKIFHINPRLPSDLHPIKVVEGVKELSKKL 1020
D.m EAEWDRLVSDR-DSLRQI-FPNGESKVVLPCNLQRMIWNVQKIFHINKRLPTDLSPIRVIKGVKTLLERC 1012
S.p TIFRGSDRITRDVQNNATLLFQILLRSKFAVKRVIMEYRLNKVAFEWIMGEVEARFQQAVVSPGEMVGTL 1070
S.c LVLRGKNEIIQNAQRDAVTLFCCLLRSRLATRRVLQEYRLTKQAFDWVLSNIEAQFLRSVVHPGEMVGVL 1067
A.th LVVPGDDALSVEAQKNATLFFNILLRSTLASKRVLEEYKLSRERFEWVIGEIESRFLQSLVAPGEMIGCV 1082
C.e IIVSGNDEISKQAQYNATLLMNILLRSTLCTKNMCTKSKLNSEAFDWLLGEIESRFQQAIAQPGEMVGAL 1090
M. VIVNGDDPLSRQAQENATLLFNIHLRSTLCSRRMAEEFRLSGEAFDWLLGEIESKFNQAIAHPGEMVGAL 1090
D.m VIVTGNDRISKQANENATLLFQCLIRSTLCTKYVSEEFRLSTEAFEWLVGEIETRFQQAQANPGEMVGAL 1082

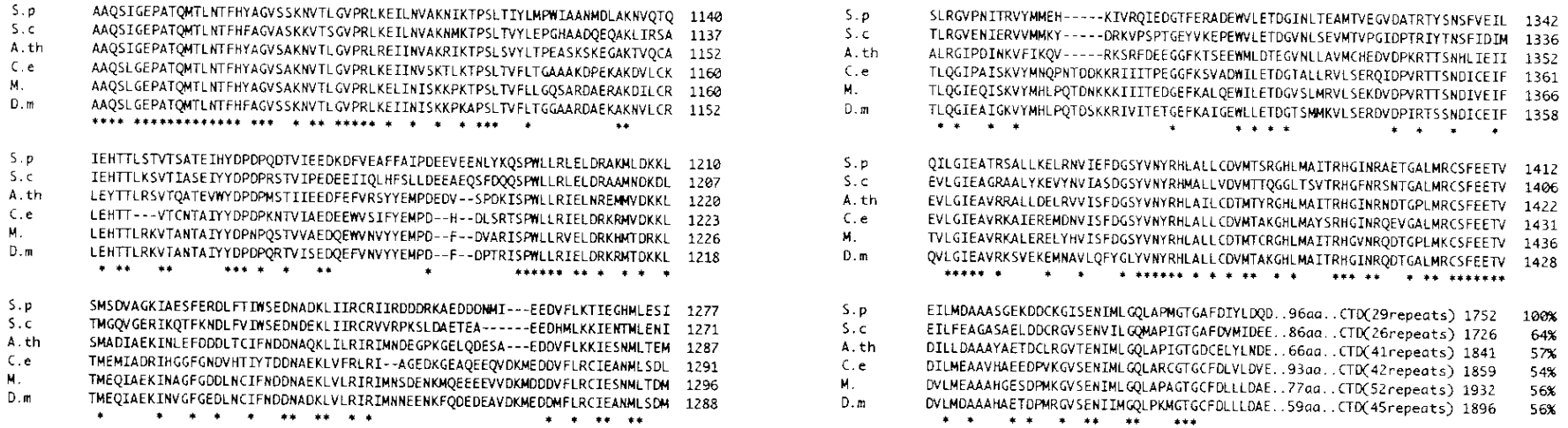


Fig. 13. Multiple alignment of the RNA polymerase II subunits 1 from vairous organisms. The alignment was performed with a computer program, TreeAlign, produced by J. Hein (53). CTD (numbers of the repetition units are shown in parentheses) and adjacent diversed regions between domain H and CTD (numbers of amino acid are indicated) are not included in this alignment. Positions with identical amino acids are indicated by asterisks under the alignment, while the eight conserved domains are indicated by lines over the alignment. Overall identities of amino acid sequences of the RNA polymerase II subunit 1 between *S. pombe* and other test species are shown at the end of the alignment. Species examined are: S.p., *Schizosaccharomyces pombe*; S.c., *Saccharomyces cerevisiae*; A.t., *Arabidopsis thaliana*; C.e., *Caenorhabditis elegance*; M.m., *Mus musculus*; D.m., *Drosophila melanogaster*.

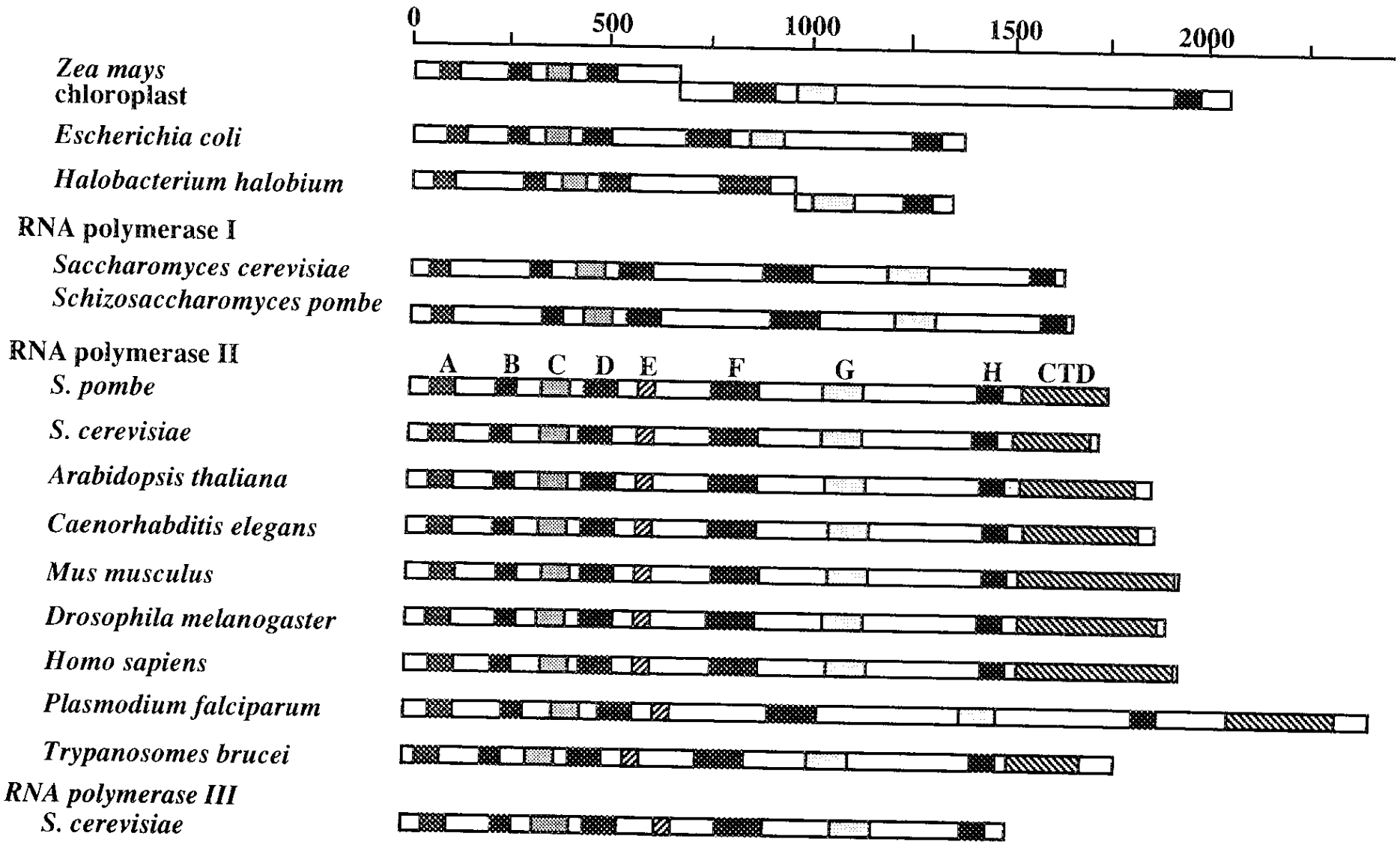



Fig. 14. Schematic alignment of β' homologues.
Nine conserved regions were found in eukaryotic RNA polymerase subunit 1 and its homologues from prolkaryotes and archaebacteria.

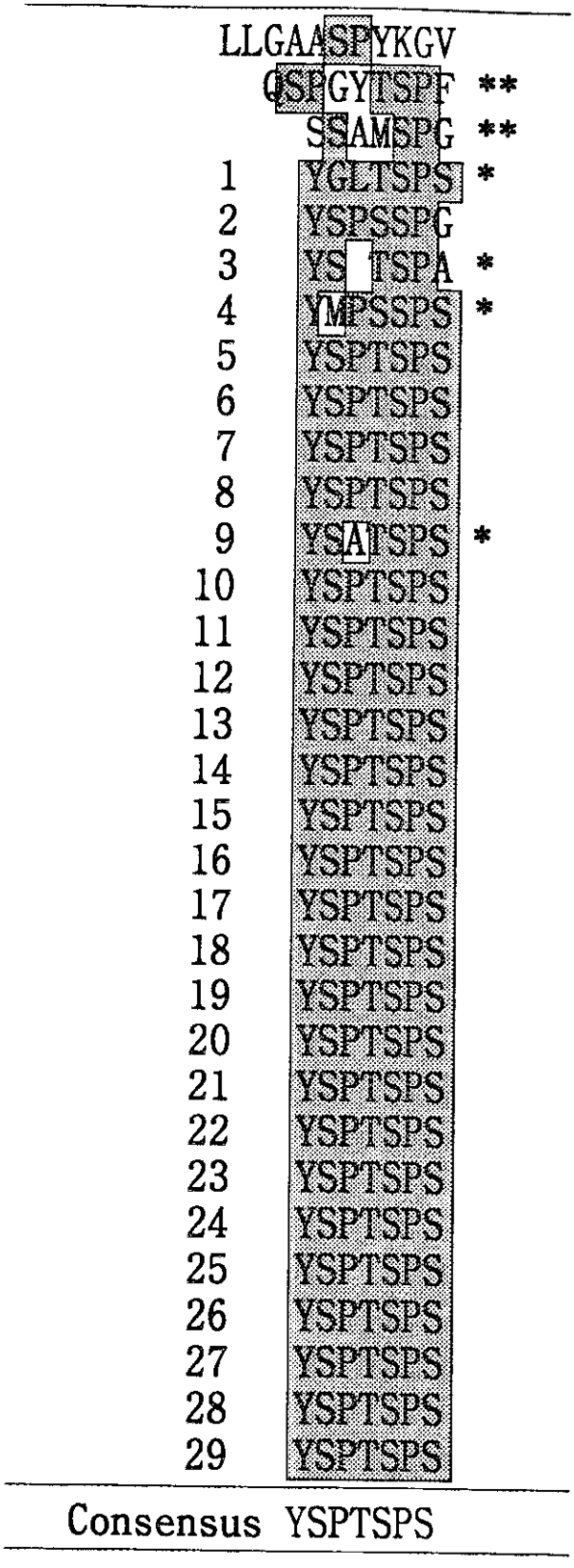



Fig. 15. CTD sequence of *S. pombe* RNA polymerase II subunit 1. Amino acid sequence from position 1,524 to the C-terminal end contains 29 repetitions of the YSPTSPS sequence. Repeat units with one or two mismatches to the consensus sequence are indicated by asterisks. Two units with sequences similar to the consensus exist upstream of CTD, as indicated by double asterisks.

# *CHAPTER 3*
# *Gene Cloning and characterization of the third-largest subunit*

## Cloning of the *rpb3* gene

To clone the *S. pombe* gene encoding the third largest subunit (subunit 3) of RNA polymerase II, cross-hybridization method was tried using the *S. cerevisiae RPB3* gene as a probe. But the *RPB3* gene did not hybridize to *S. pombe* genomic DNA under the condition used in the cloning of the largest subunit gene. Then another strategy was employed as described in Fig. 16.

At first, RNA polymerase II (Q Sepharose fraction) was applied to SDS-PAGE (Fig. 17A) and the polypeptides were eluted from gel pieces with electro-elution. Amounts of the eluted polypeptides were calculated from staining intensity on SDS-polyacrylamide gel (Fig. 17B). The isolated subunit 3 polypeptide was partially digested with *Staphylococcal aureus* V8 protease and the polypeptide fragments generated were separated by reverse phase column chromatography using RPC18 column in SMART system (Pharmacia) (Fig. 18). After three cycles of chromatography, the peak fractions of absorbance at 214 nm UV rays were applied to a protein sequencer (Applied Biosystem). Then N-terminal amino acid sequences were determined for three fragments ( see Fig. 20).

From aa sequences, IPTVAID and LGMIPLD, two 20-mer oligodeoxy-nucleotide primers were designed with the sequences of (5')ATICCIACIGTI-GCIATIGA(3') and (5')(A/G)TCIA(A/G)IGGIATCATICCI(3'), and RT-PCR was carried out with *S. pombe* cDNA as a template. A single major product of amplified cDNA was obtained . After DNA sequence analysis, this fragment of 97 bp (F1 in Fig. 19) was found to have one open reading frame, which is homologous in aa sequence to a part of *S. cerevisiae* RNA polymerase II RPB3 (60% identity from aa 41 to 73, Fig. 23B).

For isolation of a complete genomic DNA fragment carrying the entire RPB3 gene, whole *S. pombe* genomic DNA was digested with various restriction enzymes

and analyzed by Southern hybridization using the PCR-amplified cDNA fragment F1 as a probe and under high stringency conditions (for details see Experimental procedure). As shown in Fig. 21, one major band was identified for all the restriction enzymes used. A 2.0 kbp *Pst*I-*Eco*RI fragment (lane 4 in Fig. 21; indicated as F2 in Fig. 19) was size-selected by agarose gel electrophoresis, and cloned into M13mp18. The sequence determination indicated the presence of all the aa sequences determined for the RPB3 fragments within this cloned DNA fragment (see Fig. 21). A region of high degree of aa sequence homology (56% identity; the conserved domain A in Fig. 23) was found between this major open reading frame and the *S. cerevisiae* RNA polymerase II RPB3 (aa position from 46 to 248). The result strongly suggested that the cloned DNA fragment contained a part of the RNA polymerase II subunit 3 (*rpb3*) gene of *S. pombe*. In order to clone the entire gene, we next cloned a *Pst*I-*Eco*RV 2.5 kbp genomic DNA fragment (F3 in Fig. 19) into M13mp18 using the *Pst*I-*Eco*RI fragment as a hybridization probe.

## Structure of the *rpb3* gene

Nucleotide (nt) sequence was determined for the continuous genomic DNA segment of 2348 bp between *Pst*I and *Eco*RV sites. The outline of this sequence is illustrated in Fig. 19, and the details are described in Fig. 20. The coding frame spanning nt position 988 to 1966 encodes a polypeptide of 297 amino acids in length with a high degree of aa sequence homology to the *S. cerevisiae* RNA polymerase II RPB3 (see Fig. 23 for aa sequence comparison). Then this gene was designated as *rpb3*, according to the nomenclature proposed for the *S. cerevisiae* RNA polymerase genes by Nonet et al. (64). From sequence analysis of the cDNA clone F1 covering a part of *rpb3* mRNA and the genomic DNA clones, the *rpb3* gene was found to contain at least one intron (nt position from 1161 to 1206). This finding is in good agreement with the direct aa sequence analysis of one V8 digested fragment covering aa 40-48.

Detailed analysis of the exon-intron organization was then performed by sequencing several other cDNA clones covering different parts of *rpb3* mRNA using

primers designed from the genomic DNA sequence of *rpb3*. Results indicated that the coding frame of *rpb3* was interrupted by two introns (another at nt position from 1068 to 1110). In both 5' and 3' boundaries of these two introns, there are the consensus sequences of the *S. pombe* intron-exon junctions (boxed in Fig. 20). In sharp contrast to *S. cerevisiae*, the *rpb1*, *rpb2* and *rpb3* genes of *S. pombe* all contain several introns(see Fig. 24). Detailed Southern analysis of the genomic DNA digested with various restriction enzymes showed that this gene is present as a single copy in the *S. pombe* genome (Fig. 21).

## Transcription organization

Northern analysis demonstrated that the size of *rpb3* transcript is about 1.2 kbp in length (Fig. 22A). The start site of the transcript determined by primer extension analysis (Fig. 22B) using primer #1 (nt 877-894) was located at nt position 792 (as marked in Fig. 20). that ATG at nt position 988 is concluded as the start codon because: 1) the predicted start codon is located within the first exon (no intron was found between the transcription start site and this putative start codon); 2) no ATG codon exists in upstream of the same open reading frame; 3) the N-terminal proximal region of RPB3 from the initiator Met to the first conserved domain (domain A in Fig. 23) is as large as that of *S. cerevisiae* RNA polymerase II RPB3, even though homology is not so high in this region between *S. pombe* and *S. cerevisiae*; and 4) the sequence near the putative initiation codon fits well to the Kozak's rules (44) of the consensus sequence for translational initiation.

To determine the nt sequence of the 3' terminal region of *rpb3* transcript, the region was amplified f rom one of our cDNA clones using primer #2 (nt 1833-1852 in Fig. 20) and the primer b with the sequence (5')CGCCGGCGC-TTAAGTTT(3') (same as used for the *rpb1* gene cloning). From sequence analysis of 3 independent clones generated from the PCR products, the poly(A) tail was identified after nt position 2043 or 2044 as marked in Fig. 20, and thus the putative poly(A) signal might be located from 2029 to 2037 (double underline in Fig. 20). The length of mature mRNA (1172 b plus poly(A) tail) predicted from the cDNA sequence

analysis is in good agreement with the result (1.2 kbp) obtained by Northern analysis (shown in Fig. 22).

## Structure of the subunit 3

The predicted aa sequence of *S. pombe* RNA polymerase II subunit 3 was compared with those of other organisms: *S. cerevisiae* RPB3 (5) and RPC40 (6); *H. sapiens* RPB33 (7); *T. thermophila* cojC (8); and *E. coli* α subunit (65). As shown in Fig. 23, four structural domains (domain A to D) are conserved within these subunits from eukaryotic RNA polymerases except *S. cerevisiae* RPC40, which lacks the domain B. The domains A and D exist even in the α subunit of *E. coli* RNA polymerase, suggesting that these two domains are involved in some common and essential functions associated with the RNA polymerase. On the other hand, domains B and C are conserved only in eukaryotic RNA polymerases and thus considered to be involved in function(s) specific for eukaryotes.

An *E. coli* mutant defective in RNA polymerase assembly has a mutation *rpoA*112 in domain A of the α-subunit gene (66,67), and a *S. cerevisiae* mutant *rpb3-1* affecting RNA polymerase II assembly carries double mutations in the *RPB3* gene, one in domain A and another in domain C (68). Thus, the domain A may play an important role in subunit-subunit contact of RNA polymerase. As shown in Fig. 24 the domain B has a putative metal-binding sequence, $CXCX_3CX_2C$ (aa position 90 to 99 in the *S. pombe* sequence). Since this domain B sequence is not present in the corresponding subunits of RNA polymerases I and III, this motif may be related to function(s) specific for RNA polymerase II. According to the recent studies with the α-subunit C-terminal deletion mutants of *E. coli* RNA polymerase (69), the C-terminal truncated α containing the domain D can still be assembled into enzymatically active pseudo-core enzymes, suggesting that the domain D is required for the formation and/or stability of RNA polymerase. In this domain D sequence of *S. pombe* (and also *H. sapiens*), a leucine zipper-like motif was found. This finding may support a role of the domain D in subunit-subunit contact.

The C-terminal region of *E. coli* RNA polymerase α subunit downstream from aa 235 carries the protein-protein contact site I with the class I activator proteins (43,70). Since two RPB3 of eukaryotic RNA polymerases lack the corresponding region, the molecular communication between RNA polymerase and transcription factors may be different between prokaryotes and eukaryotes. However, it is not excluded yet that this putative leucine zipper motif in the domain D plays a role in contact with transcription factors.

## Isolation of ts mutants

pBrpb3 plasmid was constructed by cloning of the 2.3 kbp DNA fragment containing the full length of wild type *S. pombe rpb3* gene into pBluescript II KS+ between *Pst*I and *Eco*RV sites (Fig. 25). pBrLEU2b3 plasmid was made by replacing the *Cla*I-*Cla*I region of pBrpb3 with 2.2 kbp *Xho*I-*Sal*I fragment of YEp13 containing the entire *LEU2* gene of *S. cerevisiae*. Mutagenesis of *rpb3* was carried out by lower fidelity PCR in the presence of 0.5 mM $MnCl_2$. To construct plasmid, pBrLEU2rpb3', the mutated DNA fragments were replaced with *Bal*I-*Eco*RV region of pBrLEU2b3. The DNA fragment, rLEU2rpb3', for transformation was prepared from pBrLEU2rpb3 after digestion with *Eco*T22I.

Five μg each of the DNA fragment, rLEU2rpb3' was used to transform JY265 strain of *S. pombe* by either spheroplast method or electroporation (Fig. 25). By each method, 5000 and 6000 $leu^+$ transformants were obtained on SD medium plates, respectively. Total 9100 independent transformants were tested for growth on SD plates at both 25°C and 37°C for the permissive and non-permissive temperature, respectively. After 3-day-incubation, temperature-sensitive clones were re-streaked from the 25°C plate onto two SD plates and re-checked for growth at 25°C and 37°C. Total 178 temperature-sensitive mutants were isolated, and 68 of them were examined for viability, growth rate and reversion frequency. From the characteristics thus analyzed, these mutants were classified roughly into 16 groups. One representative of each group ( total sixteen ts mutants) was analyzed in details for its phenotype (Fig. 26, summarized in Table. 3)

## Characterization of ts mutants

To confirm the length and copy number of the *rpb3* DNA fragment in the 68 mutants, inserted DNA fragments were analyzed with PCR using specific primers for the inserted DNA fragment or for the boundary regions of homologous recombination. Most of the 68 mutants carried full length DNA fragment as a single copy in its genomic DNA but some carried either partially deleted DNA fragment or multiple copies of DNA fragments (data not shown). Some of the 68 mutants were also analyzed by Southern hybridization using specific probes for *rpb3* and *LEU2* genes (Fig. 27). The copy numbers determined for the tested mutants were summarized in Table 3. It should be noted that most of the mutants showing unstable growth on SD plate medium carried multi copies of the DNA fragment.

To determine the mutation sites of nine mutants each carrying a single copy of the fragment, the inserted DNA fragments were cloned into pUC19. The results of sequence analysis are shown in Fig. 28 and schematically illustrated in Fig. 29, and the RPB3 polypeptides encoded by the mutant *rpb3* genes are shown in Fig. 30 and summarized in Fig. 31. All these mutant *rpb3* genes carried multiple mutations, but many mutations were clustered in the N-terminal region of RPB3 polypeptides. Since the *S. cerevisiae* mutant *rpb3-1* affecting RNA polymerase II assembly carried one of the two mutations in this region, these mutants carrying the mutations in the N-terminal region may have defect in RNA polymerase II assembly.

As an initial effort to analyze the defective function of mutant RNA polymerase II, the growth rate was measured for these mutants after shift of the culture temperature from permissive temperature, 25°C, to non-permissive temperature, 37°C (Fig. 33). Some of the mutants stopped growing immediately after the temperature shift. The RNA polymerase II in these mutants may have defect(s) in function at non-permissive temperature. On the other hand, some of the mutants stopped growing slowly, implying that the assemblies of RNA polymerase II and/or of multi-transcription machinery including some transcription factors are defective in these mutants.

By microscope observation, some mutants showed abnormal cell shapes (for example, 7 times longer in cell size) at the non-permissive temperature (Fig. 32). After incubation at non-permissive temperature for one day, none of the sixteen mutants regained viability even at permissive temperature. When these mutant cells were stained with DAPI and observed with a fluorescence microscope, however, their nuclei still existed in these mutants.

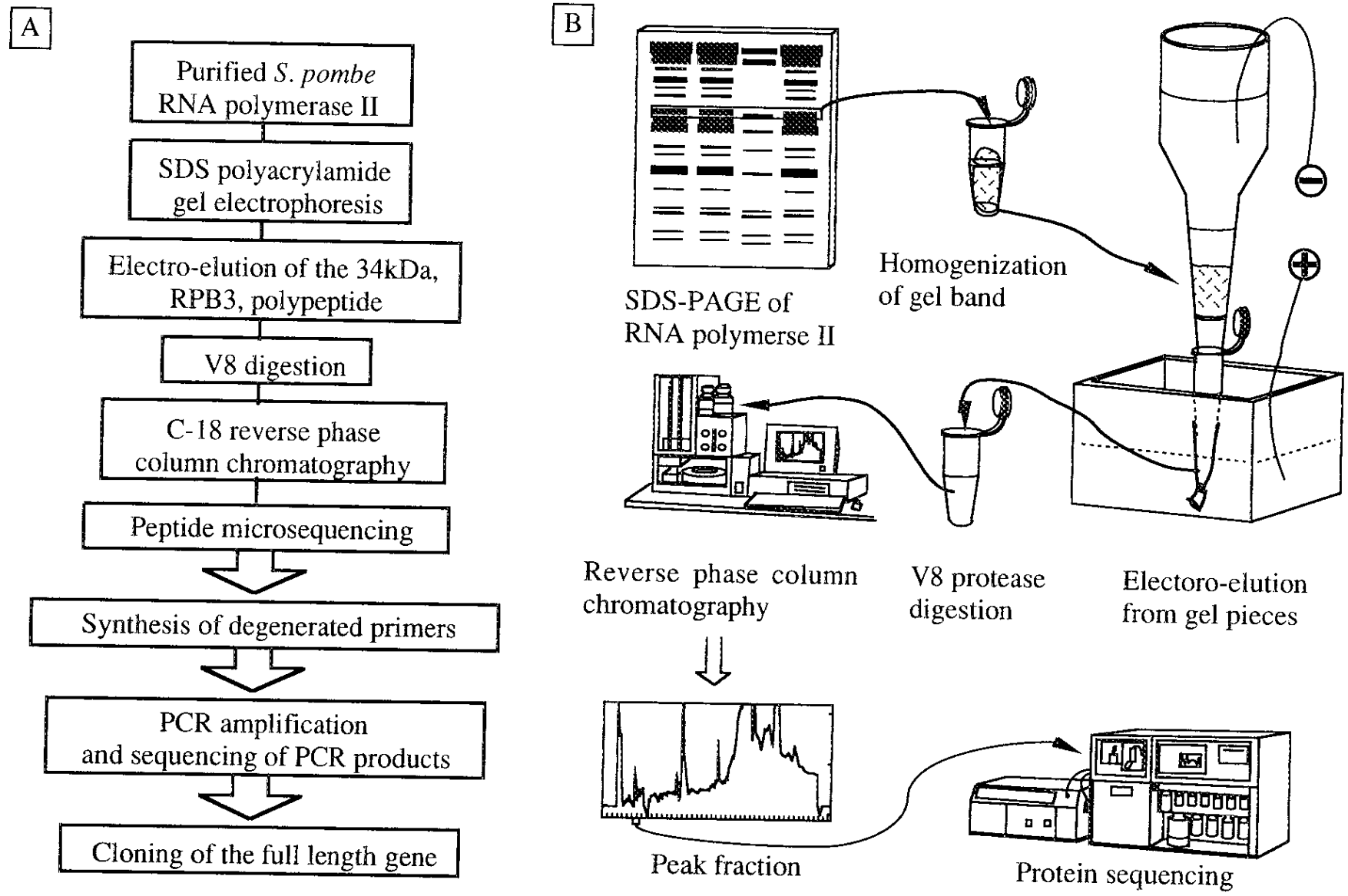



Fig. 16. Cloning strategy of the *rpb3* gene.
(A) Flowchart of the strategy. (B) Schematic flowchart of this strategy.

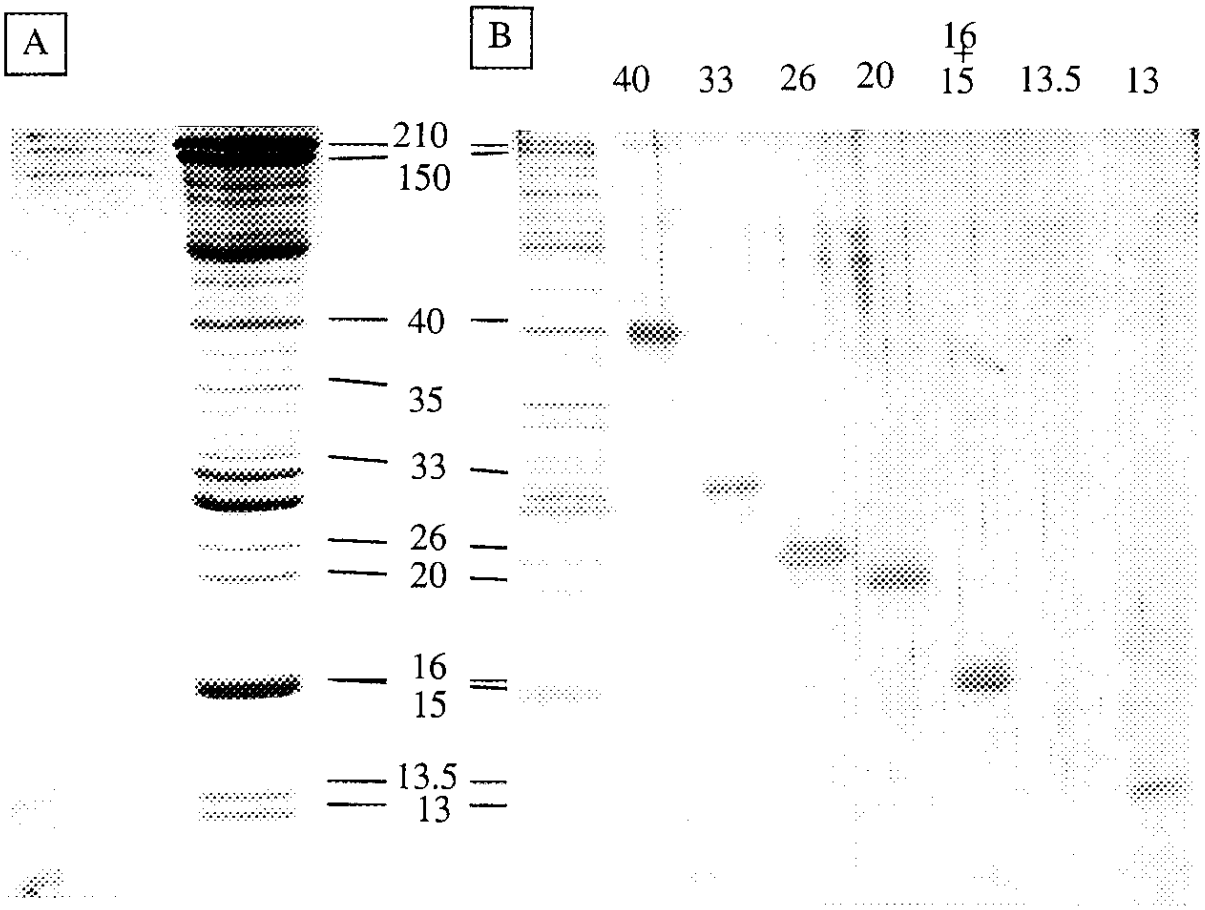
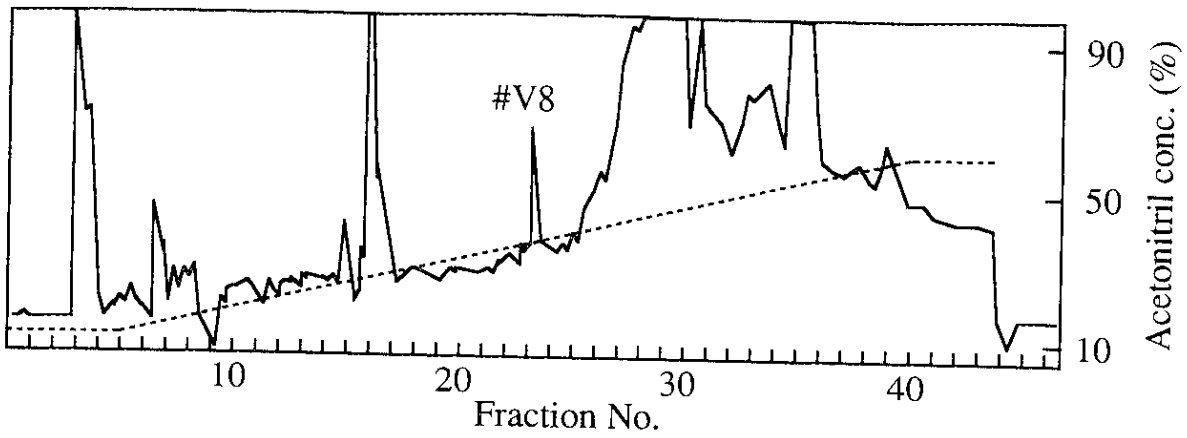


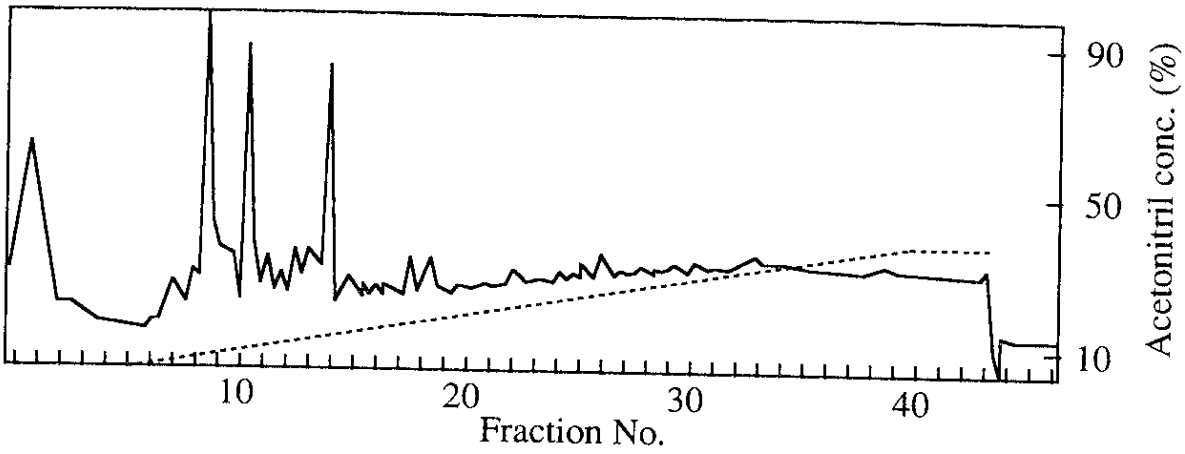

Fig. 17. Subunit isolation by electrophoresis.
(A) Subunits of the RNA polymerase II (Q Sepharose fractions, right) were separated by SDS-PAGE. The highly purified RNA polymerase II (Superose 6 fraction, left) was run in parallel as a control marker to identify the tightly associated subunits. (B) SDS-PAGE analysis of the eluted polypeptides. The recovery of the polypeptides measured after silve staining was more than 70%.

A

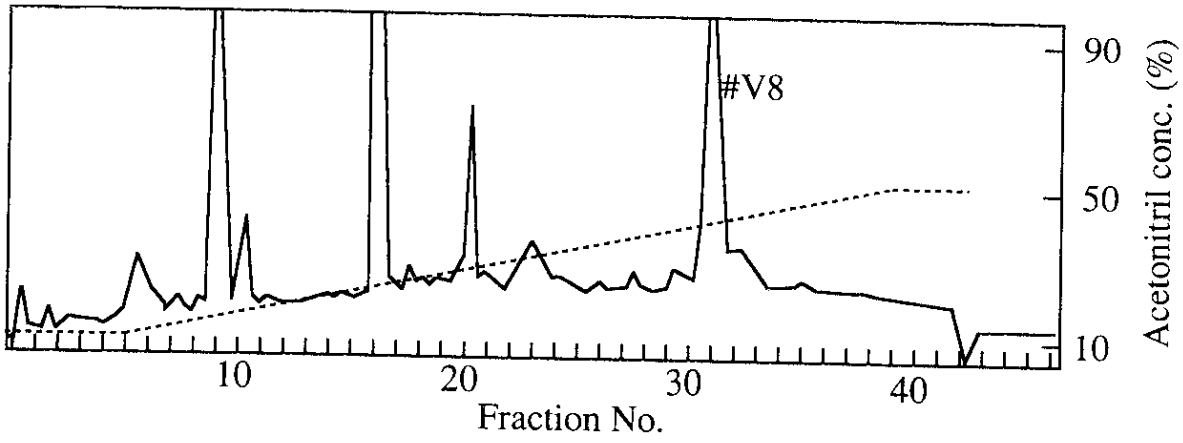

#V8
90
50
10
Acetonitril conc. (%)
10
20
30
40
Fraction No.

B

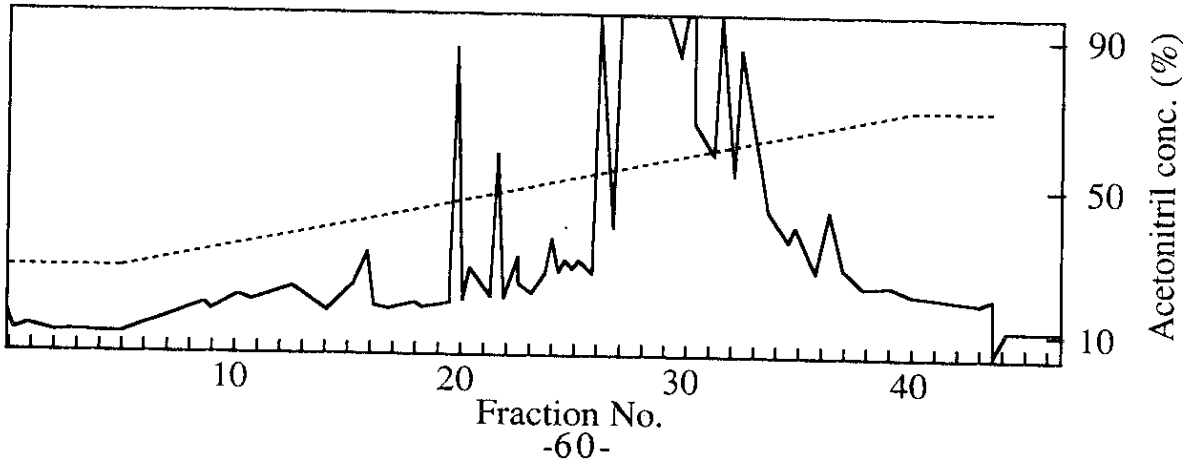
[1]


90
50
10
Acetonitril conc. (%)
10
20
30
40
Fraction No.

[2]


#V8
90
50
10
Acetonitril conc. (%)
10
20
30
40
Fraction No.

[3]


90
50
10
Acetonitril conc. (%)
10
20
30
40
Fraction No.

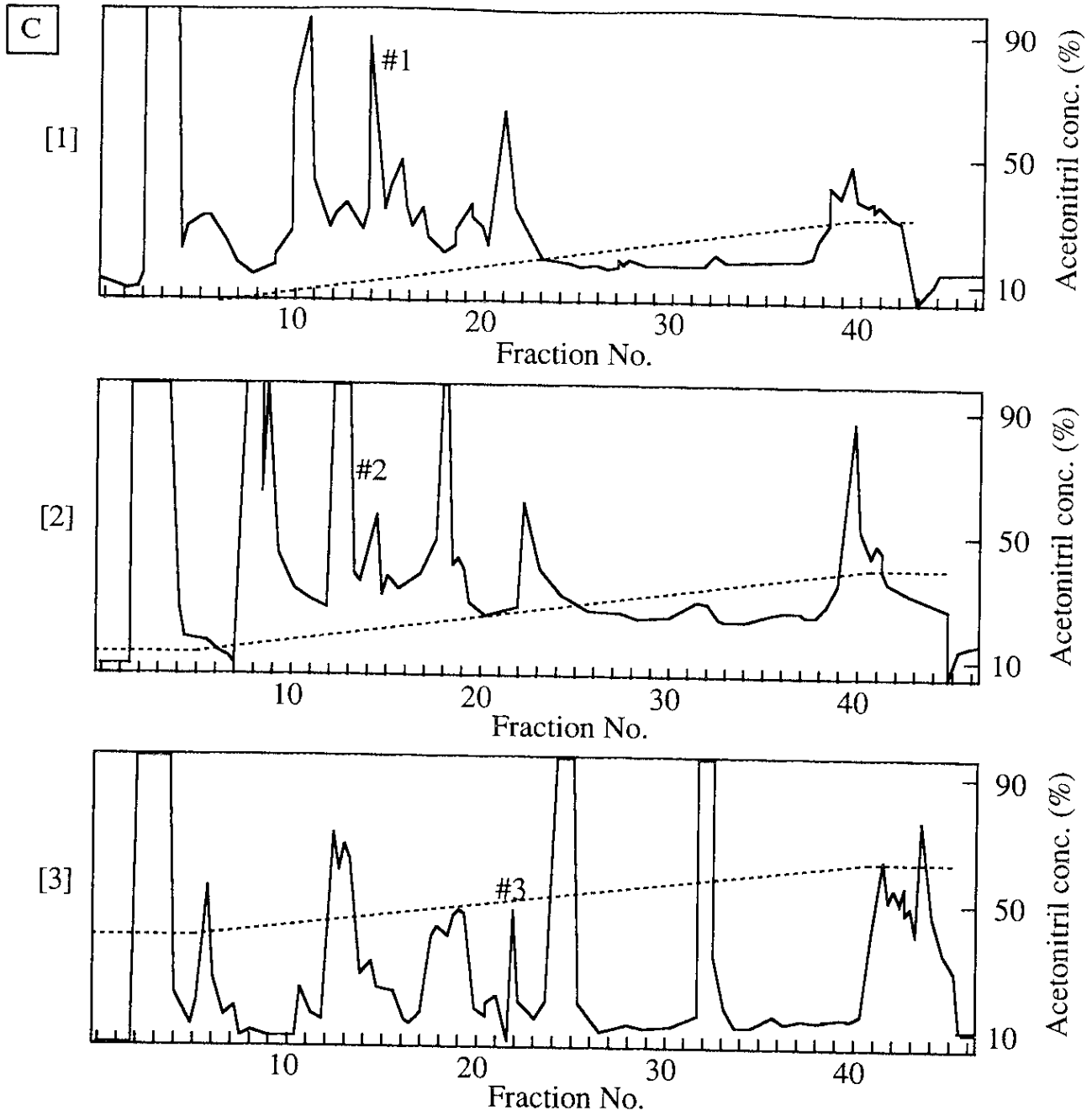



Fig. 18. Separation of V8 fragment of the 40 kDa polypeptide by C-18 reverse phase column chromatography.
(A) First separation using 10-60% acetonitril gradient. (B) Second separation. [1], fraction No. 4-15 were mixed and rechromatographed using 0-40% acetonitril gradient. [2], fraction No. 16-22 were rechromatographed using 10-50% acetonitril gradient. [3], fraction No. 23-36 were rechromatographed using 30-70% acetonitril gradient. (C) Third separation. [1], fraction No. 7-16 of B1 fractionation were rechromatographed using 0-30% actonitril gradient. [2], fraction No. 8-21 of B2 were rechromatographed using10-40% acetonitril gradient. [3], fraction No. 25-36 of B3 were rechromatographed using 40-60% acetonitril gradient. N-terminal amino acid sequences of the digested subunit 3 polypeptide were determined for #1, #2 and #3 marked in panel C1, C2 and C3, respectively. #V8 shows V8 protease fraction.

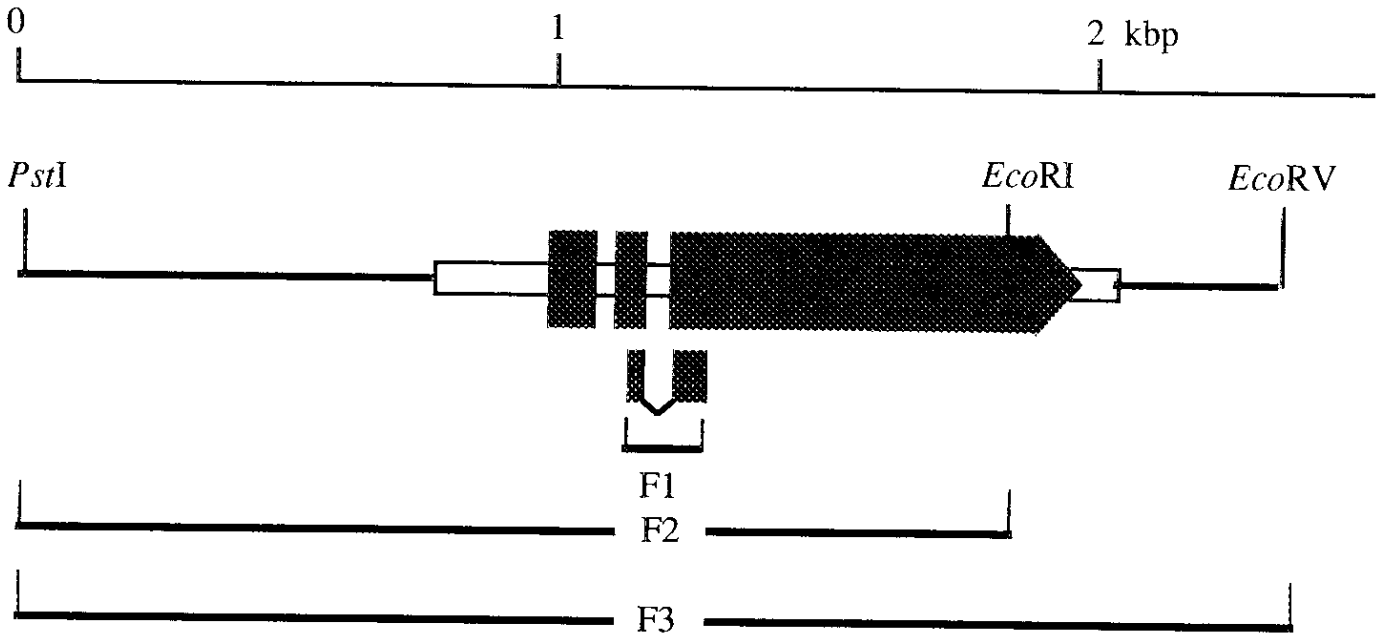



Fig. 19. Physical map of the *S. pombe rpb3* gene and clones isolated in this study.
Open box shows *rpb3* transcript including full boxes showing the coding regions. F1 is the PCR amplified cDNA fragment and F2 and F3 are genomic DNA fragments cloned using F1 and F2 probes, respectively.

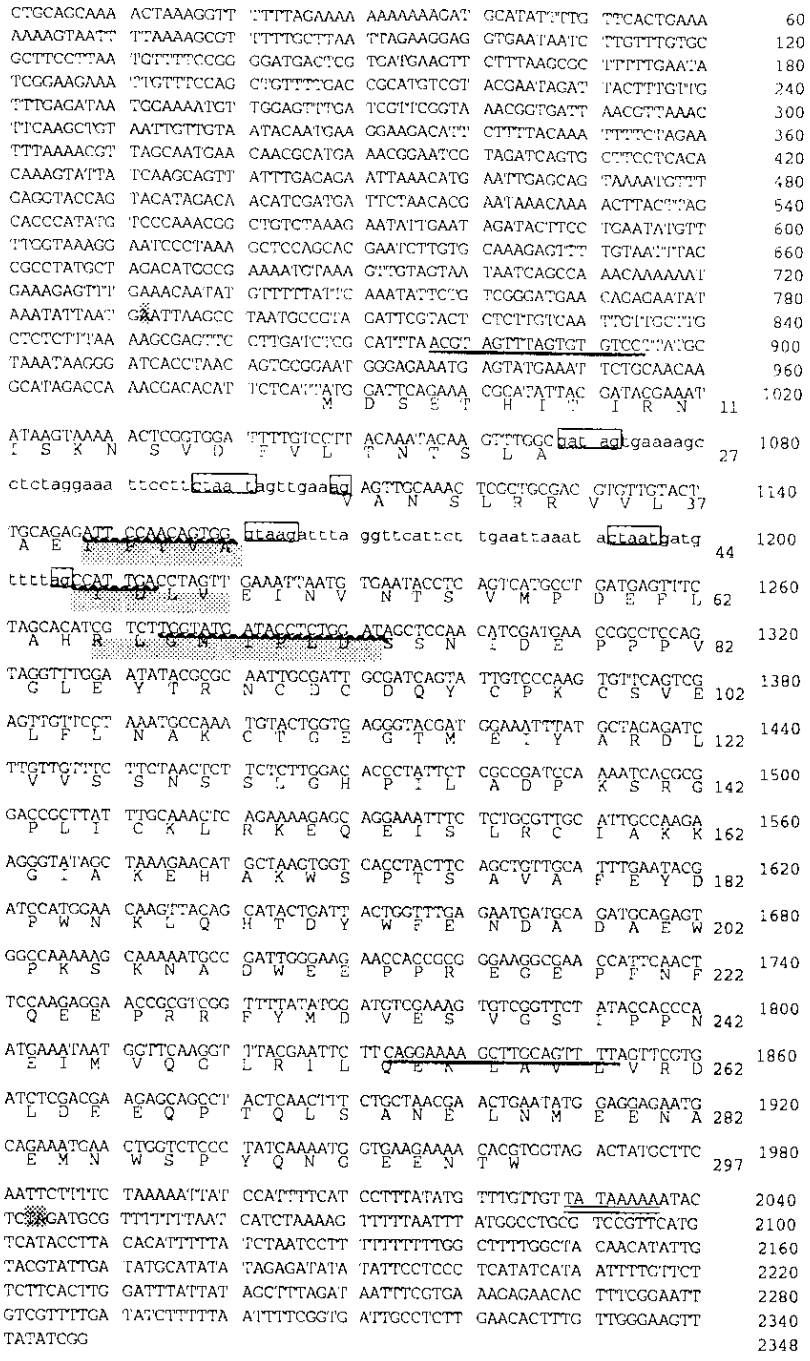


Fig. 20. Nucleotide sequence of the rpb3 gene and a predicted amino acid sequence. The coding sequence of RNA polymerase II subunit 3 starts at nt postion 988 and ends at nt postion 1966. This sequence is interrupted near the N-terminal proximal region by two introns which are indicated by small letters. Both the 5' and 3' ends of the transcript are indicated by shadow boxes on nt sequence and putative poly(A) signal is shown by a double underline. The sequenced fragments of amino acid are shown by shadow boxes on amino acid sequences, and the primers designed from those amino acid sequences were indicated by waved lines.

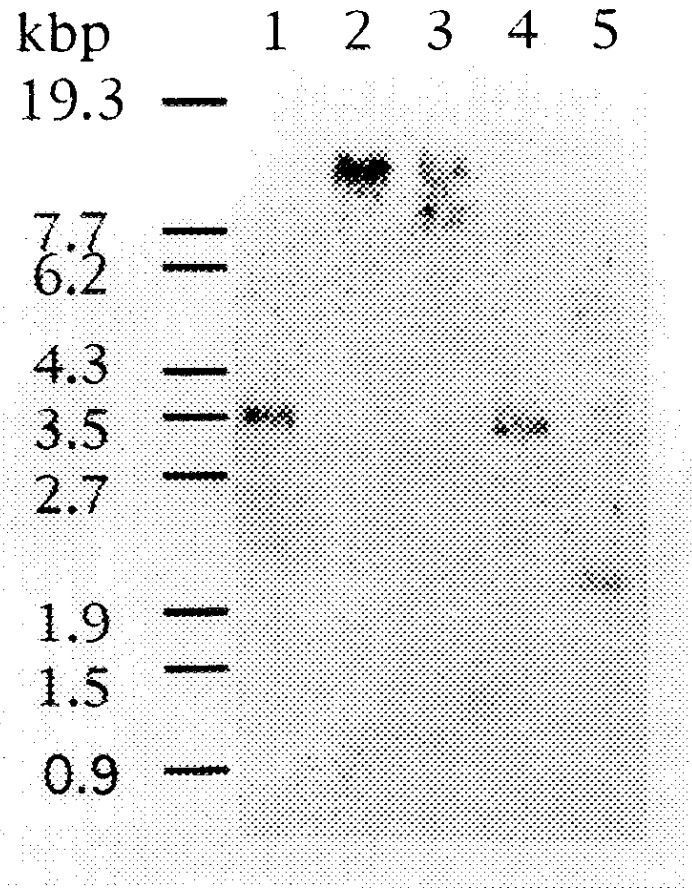



Fig. 21. Southern hybridization of *S. pombe* genomic DNA.
Three μg of total *S. pombe* DNA were digested with various restriction enzymes, and subjected to Southern analysis under high stringency hybridization conditions using the PCR-amplified cDNA fragment F1 as a probe. Restriction enzymes used are: *Eco*RI (lane 1); *Bam*HI (lane 2); *Pst*I (lane 3); *Eco*RI/*Bam*HI (lane 4); and *Eco*RI/*Pst*I (lane 5).

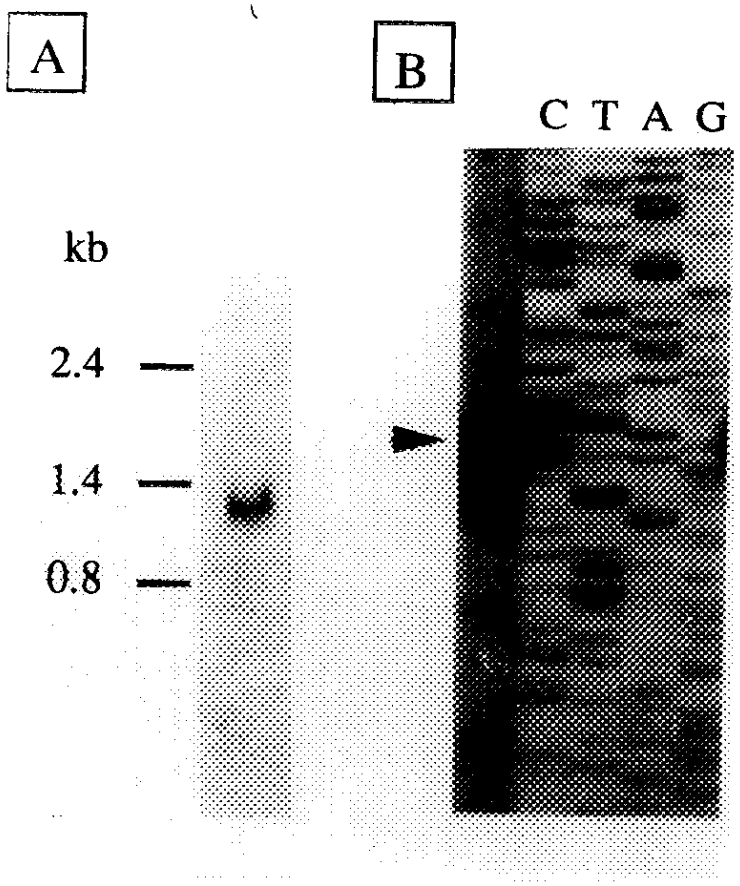



Fig. 22. RNA analysis of the *rpb3* gene. (A) Northern analysis of *rpb3* transcript. Seven μg of poly(A)$^+$ RNA was used for hybridization with the cDNA fragment F1 containing a part of *rpb3* as a probe. (B) Primer extension analysis. Five μg of poly (A) RNA was used for primer extension using the primer, #301.

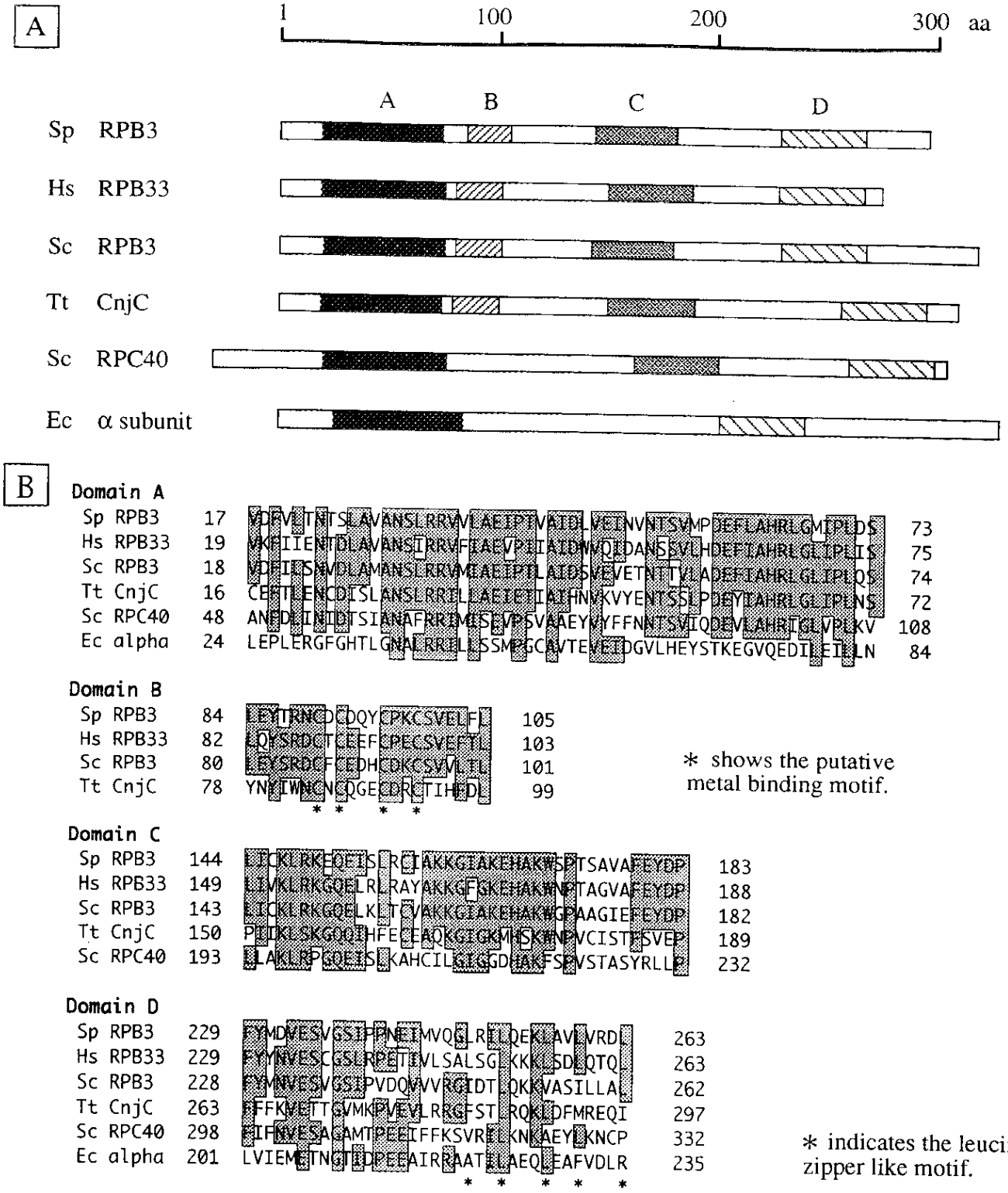
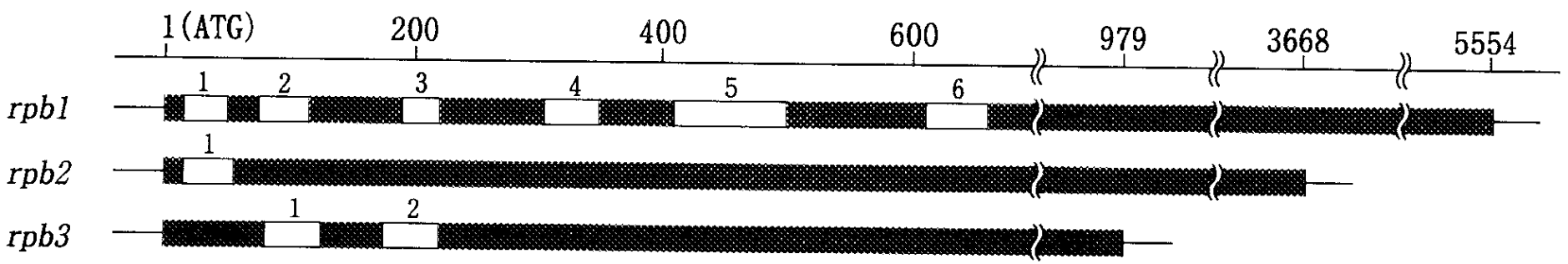



Fig. 23. Multiple alignment of the a homologue subunits. The alignment was performed with a computer program, Gene Works (IntelliGenetics Inc.). (A) Schematic alignment of conserved domains within the a-homologues. Sp RPB3, *Schizosaccharomyces pombe* RNA polymerase II subunit 3 (this study); Hs RPB33 (7), *Homo sapiens* RNA polymerase II Mr 33 subunit; Sc RPB3 (5) *Saccharomyces cerevisiae* RNA polymerase II subunit 3; Tt conjC (8), *Tetrahymena thermophila* a cojugation-specific a-homologous protein; Sc RPC40 (6), *S. cerevisiae* RNA polymerase I and III Mr 40 subunit; Ec alpha (29), *Escherichia coli* RNA polymerase a subunit. (B) Amino acids sequence comparison of the four conserved domains. Identical amino acids are boxed. The putative motifs are marked by *.

A




B

| Intron | | Nt position | 5' | | | | | | Branch | | | | | | | | | | | | 3' | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *rpb1* | 1 | 16 | G | T | A | T | G | T | - - 19 - - | A | C | T | A | A | T | - - 3 - - | T | A | G | | |
| | 2 | 77 | G | T | A | T | G | T | - - 22 - - | A | C | T | A | A | C | - - 5 - - | T | A | G | | |
| | 3 | 190 | G | T | A | A | G | T | - - 20 - - | A | C | T | T | A | T | - - 6 - - | T | A | G | | |
| | 4 | 305 | G | T | A | T | G | A | - - 23 - - | T | C | T | C | A | C | - - 9 - - | A | A | G | | |
| | 5 | 410 | G | T | A | T | G | T | - - 73 - - | G | C | T | A | A | C | - - 6 - - | C | A | G | | |
| | 6 | 609 | G | T | A | A | G | T | - - 28 - - | A | C | T | A | A | C | - - 6 - - | A | A | G | | |
| *rpb2* | 1 | 17 | G | T | A | A | G | T | - - 18 - - | T | T | T | A | A | C | - - 7 - - | T | A | G | | |
| *rpb3* | 1 | 81 | G | T | A | A | G | T | - - 24 - - | T | C | T | A | A | T | - - 7 - - | A | A | G | | |
| | 2 | 273 | G | T | A | A | G | A | - - 26 - - | A | C | T | A | A | T | - - 8 - - | T | A | G | | |
| | Consensus | | G | T | A | n | G | t | - - - -- - - | n | C | T | r | A | y | - - - - - | at | A | G | | |
| | | A | | | 94 | 50 | | 13 | | 66 | | | 76 | 100 | 3 | | 13 | 100 | | | |
| | | G | 100 | | 3 | 7 | 99 | 3 | | 20 | | | 16 | | | | 1 | | 100 | | |
| | | C | | | | 10 | | 3 | | | 90 | | 4 | | 64 | | 9 | | | | |
| | | T | | 100 | 3 | 33 | 1 | 81 | | 14 | 10 | 100 | 3 | | 33 | | 77 | | | | |

Fig. 24. Intron locations and exon-intron boundary sequences within the three large subunit genes of *S. pombe* RNA polymerase II. (A) Locations of introns. Introns are shown by open bars. (B) Exon-intron boundary sequences. Exon-intron boundary sequences in the *rpb1*, *rpb2* and *rpb3* genes are compared with those of 85 introns in *S. pombe* 45 genes transcripted by RNA polymerase II.

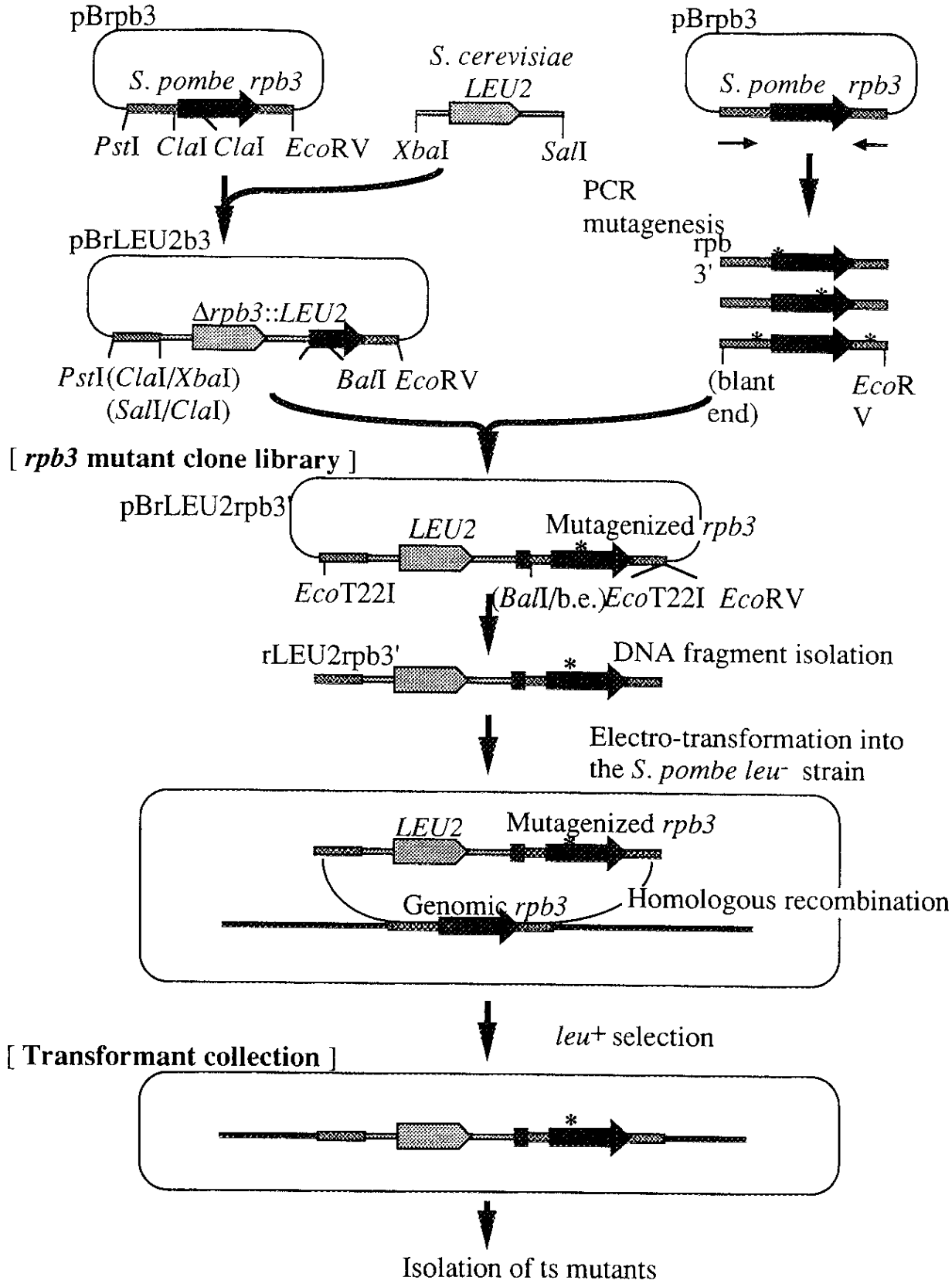



Fig. 25 Strategy for mutagenesis of the *rpb3* gene.

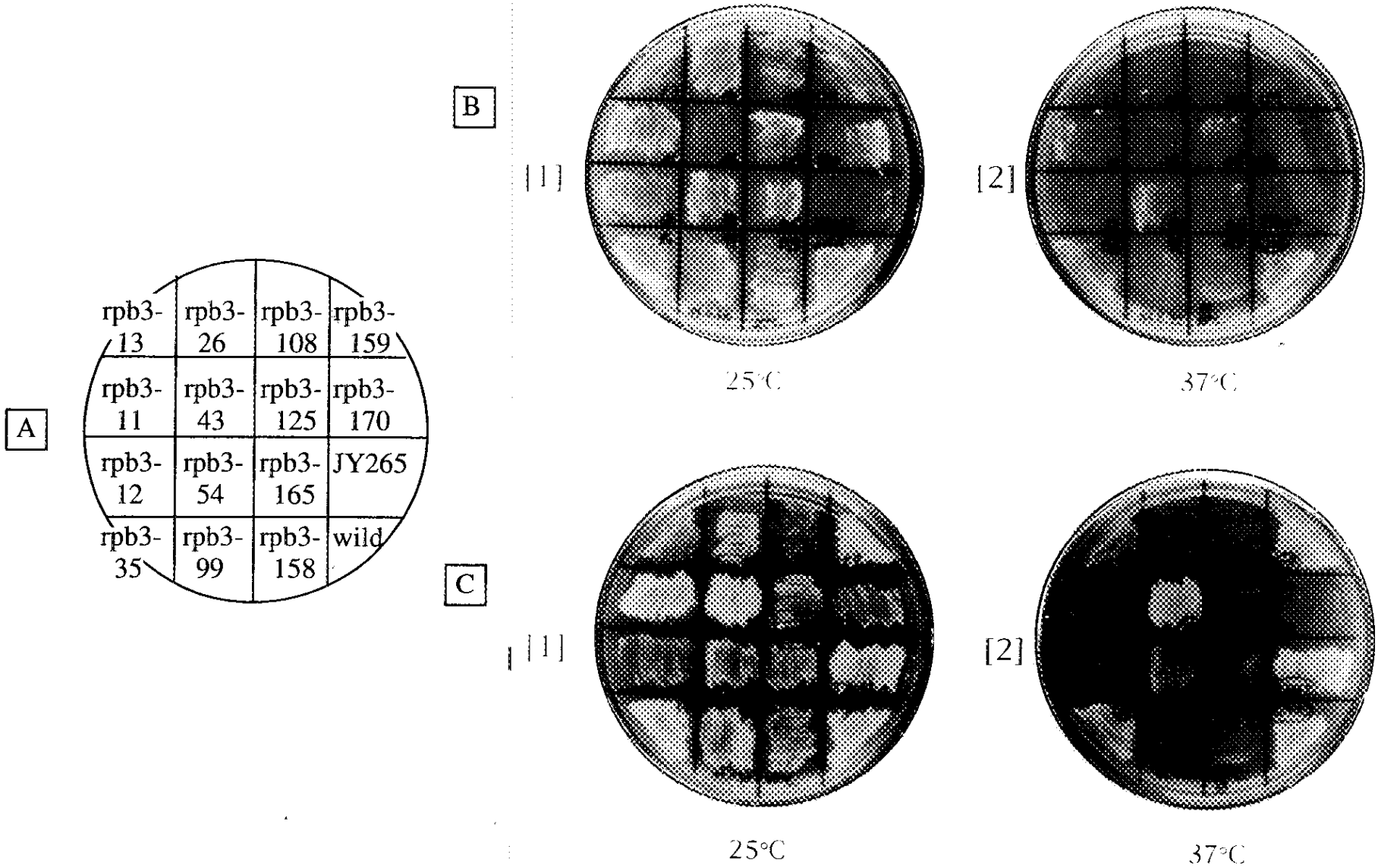



Fig. 26. Isolation of ts mutants.
(A) Analyzed strains. (B) Growth on SD plate. [1] Growth on a synthetic medium at permissive temparature 25°C, for 4 days. [2] Growth on a synthetic medium at non-permissive temparature 37°C, for 5 days. (C) Growth on YPD plate. [1] Growth on a rich medium at permissive temparature 25°C, for 4 days. [2] Growth on a rich medium at non-permissive temparature 37°C, for 5 days.

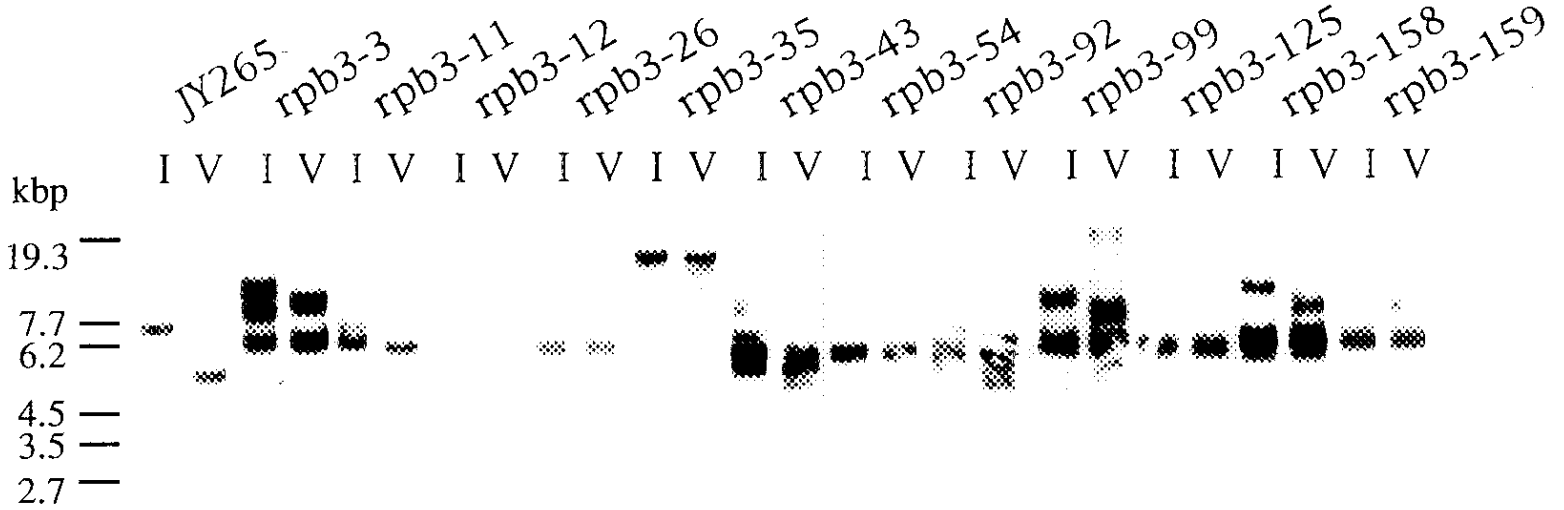
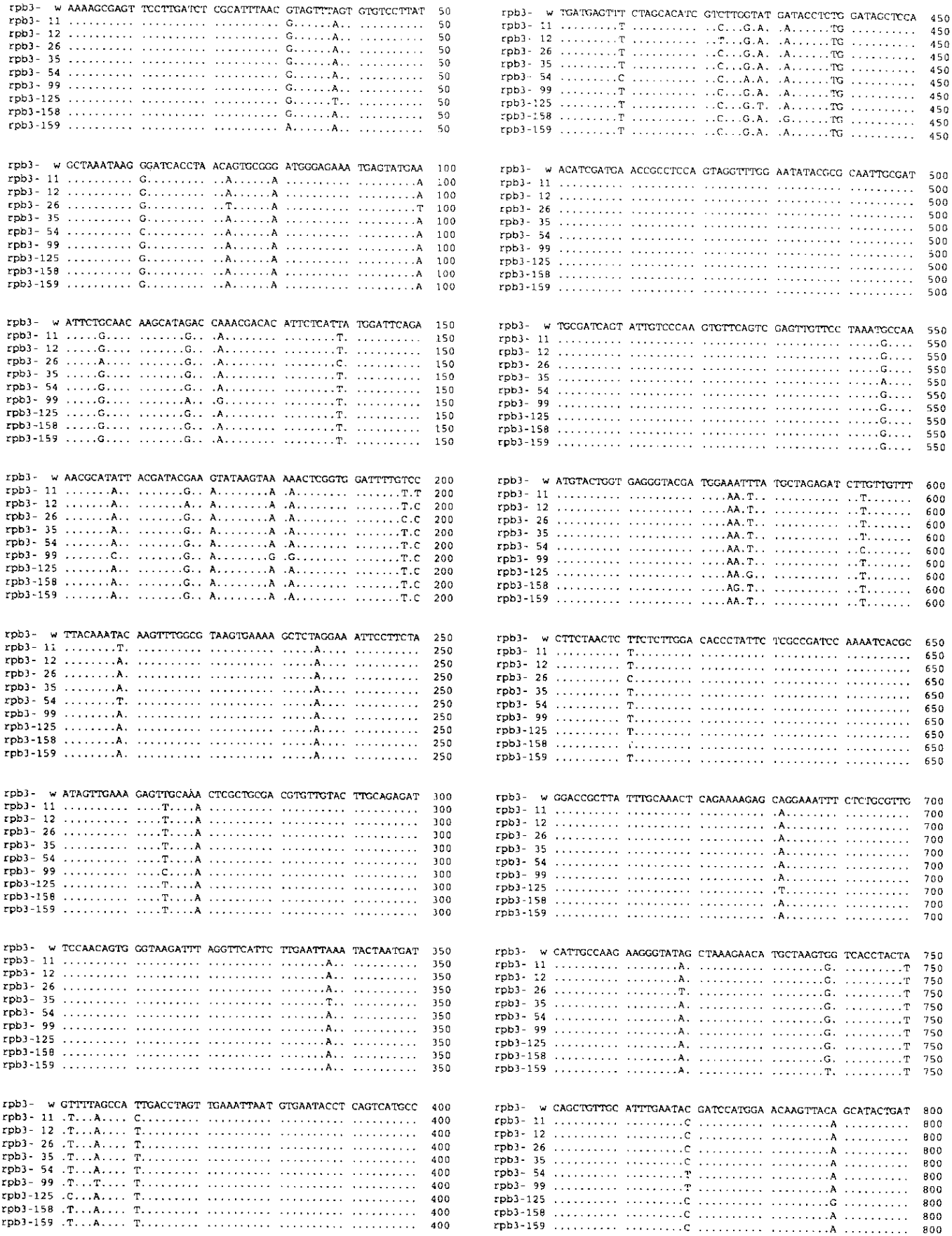



Fig. 27. Southern analysis of genomic DNA from mutants.
Restriction enzymes used are: I, *Eco*RI; V, *Eco*RV.



rpb3- w AAAAGCGAGT TCCTTGATCT CGCATTTAAC GTAGTTTAGT GTGTCCTTAT 50
rpb3- 11 ..... ..... ..... G.....A.. ..... 50
rpb3- 12 ..... ..... ..... G.....A.. ..... 50
rpb3- 26 ..... ..... ..... G.....A.. ..... 50
rpb3- 35 ..... ..... ..... G.....A.. ..... 50
rpb3- 54 ..... ..... ..... G.....A.. ..... 50
rpb3- 99 ..... ..... ..... G.....A.. ..... 50
rpb3-125 ..... ..... ..... G.....T.. ..... 50
rpb3-158 ..... ..... ..... G.....A.. ..... 50
rpb3-159 ..... ..... ..... A.....A.. ..... 50

rpb3- w TGATGAGTTT CTAGCACATC GTCTTGGTAT GATACCTCTG GATAGCTCCA 450
rpb3- 11 .....T ..... ..C...G.A. .A.....TG ..... 450
rpb3- 12 .....T ..... ..T...G.A. .A.....TG ..... 450
rpb3- 26 .....T ..... ..C...G.A. .A.....TG ..... 450
rpb3- 35 .....T ..... ..C...G.A. .A.....TG ..... 450
rpb3- 54 .....C ..... ..C...A.A. .A.....TG ..... 450
rpb3- 99 .....T ..... ..C...G.A. .A.....TG ..... 450
rpb3-125 .....T ..... ..C...G.T. .A.....TG ..... 450
rpb3-158 .....T ..... ..C...G.A. .G.....TG ..... 450
rpb3-159 .....T ..... ..C...G.A. .A.....TG ..... 450

rpb3- w GCTAAATAAG GGATCACCTA ACAGTGCGGG ATGGGAGAAA TGAGTATGAA 100
rpb3- 11 ..... G..... ..A.....A ..... .....A 100
rpb3- 12 ..... C..... ..A.....A ..... .....A 100
rpb3- 26 ..... G..... ..T.....A ..... .....T 100
rpb3- 35 ..... G..... ..A.....A ..... .....A 100
rpb3- 54 ..... C..... ..A.....A ..... .....A 100
rpb3- 99 ..... G..... ..A.....A ..... .....A 100
rpb3-125 ..... G..... ..A.....A ..... .....A 100
rpb3-158 ..... G..... ..A.....A ..... .....A 100
rpb3-159 ..... G..... ..A.....A ..... .....A 100

rpb3- w ACATCGATGA ACCGCCTCCA GTAGGTTTGG AATATACGCG CAATTGCGAT 500
rpb3- 11 ..... ..... ..... ..... ..... 500
rpb3- 12 ..... ..... ..... ..... ..... 500
rpb3- 26 ..... ..... ..... ..... ..... 500
rpb3- 35 ..... ..... ..... ..... ..... 500
rpb3- 54 ..... ..... ..... ..... ..... 500
rpb3- 99 ..... ..... ..... ..... ..... 500
rpb3-125 ..... ..... ..... ..... ..... 500
rpb3-158 ..... ..... ..... ..... ..... 500
rpb3-159 ..... ..... ..... ..... ..... 500

rpb3- w ATTCTGCAAC AAGCATAGAC CAAACGACAC ATTCTCATTA TGGATTCAGA 150
rpb3- 11 .....G.... .......G.. .A..... .....T. ..... 150
rpb3- 12 .....G.... .......G.. .A..... .....T. ..... 150
rpb3- 26 .....A.... .......G.. .A..... .....C. ..... 150
rpb3- 35 .....G.... .......G.. .A..... .....T. ..... 150
rpb3- 54 .....G.... .......G.. .A..... .....T. ..... 150
rpb3- 99 .....G.... .......A.. .G..... .....T. ..... 150
rpb3-125 .....G.... .......G.. .A..... .....T. ..... 150
rpb3-158 .....G.... .......G.. .A..... .....T. ..... 150
rpb3-159 .....G.... .......G.. .A..... .....T. ..... 150

rpb3- w TGCGATCAGT ATTGTCCCAA GTGTTCAGTC GAGTTGTTCC TAAATGCCAA 550
rpb3- 11 ..... ..... ..... ..... .....G.... 550
rpb3- 12 ..... ..... ..... ..... .....G.... 550
rpb3- 26 ..... ..... ..... ..... .....G.... 550
rpb3- 35 ..... ..... ..... ..... .....A.... 550
rpb3- 54 ..... ..... ..... ..... .....G.... 550
rpb3- 99 ..... ..... ..... ..... .....G.... 550
rpb3-125 ..... ..... ..... ..... .....G.... 550
rpb3-158 ..... ..... ..... ..... .....G.... 550
rpb3-159 ..... ..... ..... ..... .....G.... 550

rpb3- w AACGCATATT ACGATACGAA GTATAAGTAA AAACTCGGTG GATTTTGTCC 200
rpb3- 11 .....A.. .....G.. A.....A .A..... .....T.T 200
rpb3- 12 .....A.. .....A.. A.....A .A..... .....T.C 200
rpb3- 26 .....A.. .....G.. A.....A .A..... .....C.C 200
rpb3- 35 .....A.. .....G.. A.....A .A..... .....T.C 200
rpb3- 54 .....A.. .....G.. A.....A .A..... .....T.C 200
rpb3- 99 .....C.. .....G.. A.....G .G..... .....T.C 200
rpb3-125 .....A.. .....G.. A.....A .A..... .....T.C 200
rpb3-158 .....A.. .....G.. A.....A .A..... .....T.C 200
rpb3-159 .....A.. .....G.. A.....A .A..... .....T.C 200

rpb3- w ATGTACTGGT GAGGGTACGA TGGAAATTTA TGCTAGAGAT CTTGTTGTTT 600
rpb3- 11 ..... ..... ....AA.T.. ..... ..T..... 600
rpb3- 12 ..... ..... ....AA.T.. ..... ..T..... 600
rpb3- 26 ..... ..... ....AA.T.. ..... ..T..... 600
rpb3- 35 ..... ..... ....AA.T.. ..... ..T..... 600
rpb3- 54 ..... ..... ....AA.T.. ..... ..C..... 600
rpb3- 99 ..... ..... ....AA.T.. ..... ..T..... 600
rpb3-125 ..... ..... ....AA.G.. ..... ..T..... 600
rpb3-158 ..... ..... ....AG.T.. ..... ..T..... 600
rpb3-159 ..... ..... ....AA.T.. ..... ..T..... 600

rpb3- w TTACAAATAC AAGTTTGGCG TAAGTGAAAA GCTCTAGGAA ATTCCTTCTA 250
rpb3- 11 .....T. ..... ..... .....A.... ..... 250
rpb3- 12 .....A. ..... ..... .....A.... ..... 250
rpb3- 26 .....A. ..... ..... .....A.... ..... 250
rpb3- 35 .....A. ..... ..... .....A.... ..... 250
rpb3- 54 .....T. ..... ..... .....A.... ..... 250
rpb3- 99 .....A. ..... ..... .....A.... ..... 250
rpb3-125 .....A. ..... ..... .....A.... ..... 250
rpb3-158 .....A. ..... ..... .....A.... ..... 250
rpb3-159 .....A. ..... ..... .....A.... ..... 250

rpb3- w CTTCTAACTC TTCTCTTGGA CACCCTATTC TCGCCGATCC AAAATCACGC 650
rpb3- 11 ..... T..... ..... ..... ..... 650
rpb3- 12 ..... T..... ..... ..... ..... 650
rpb3- 26 ..... C..... ..... ..... ..... 650
rpb3- 35 ..... T..... ..... ..... ..... 650
rpb3- 54 ..... T..... ..... ..... ..... 650
rpb3- 99 ..... T..... ..... ..... ..... 650
rpb3-125 ..... T..... ..... ..... ..... 650
rpb3-158 ..... T..... ..... ..... ..... 650
rpb3-159 ..... T..... ..... ..... ..... 650

rpb3- w ATAGTTGAAA GAGTTGCAAA CTCGCTGCGA CGTGTTGTAC TTGCAGAGAT 300
rpb3- 11 ..... ....T....A ..... ..... ..... 300
rpb3- 12 ..... ....T....A ..... ..... ..... 300
rpb3- 26 ..... ....T....A ..... ..... ..... 300
rpb3- 35 ..... ....T....A ..... ..... ..... 300
rpb3- 54 ..... ....T....A ..... ..... ..... 300
rpb3- 99 ..... ....C....A ..... ..... ..... 300
rpb3-125 ..... ....T....A ..... ..... ..... 300
rpb3-158 ..... ....T....A ..... ..... ..... 300
rpb3-159 ..... ....T....A ..... ..... ..... 300

rpb3- w GGACCGCTTA TTTGCAAACT CAGAAAAGAG CAGGAAATTT CTCTGCGTTG 700
rpb3- 11 ..... ..... ..... .A..... ..... 700
rpb3- 12 ..... ..... ..... .A..... ..... 700
rpb3- 26 ..... ..... ..... .A..... ..... 700
rpb3- 35 ..... ..... ..... .A..... ..... 700
rpb3- 54 ..... ..... ..... .A..... ..... 700
rpb3- 99 ..... ..... ..... .A..... ..... 700
rpb3-125 ..... ..... ..... .T..... ..... 700
rpb3-158 ..... ..... ..... .A..... ..... 700
rpb3-159 ..... ..... ..... .A..... ..... 700

rpb3- w TCCAACAGTG GGTAAGATTT AGGTTCATTC TTGAATTAAA TACTAATGAT 350
rpb3- 11 ..... ..... ..... .....A.. ..... 350
rpb3- 12 ..... ..... ..... .....A.. ..... 350
rpb3- 26 ..... ..... ..... .....A.. ..... 350
rpb3- 35 ..... ..... ..... .....T.. ..... 350
rpb3- 54 ..... ..... ..... .....A.. ..... 350
rpb3- 99 ..... ..... ..... .....A.. ..... 350
rpb3-125 ..... ..... ..... .....A.. ..... 350
rpb3-158 ..... ..... ..... .....A.. ..... 350
rpb3-159 ..... ..... ..... .....A.. ..... 350

rpb3- w CATTGCCAAG AAGGGTATAG CTAAAGAACA TGCTAAGTGG TCACCTACTA 750
rpb3- 11 ..... .....A. ..... .....G. .....T 750
rpb3- 12 ..... .....A. ..... .....G. .....T 750
rpb3- 26 ..... .....T. ..... .....G. .....T 750
rpb3- 35 ..... .....A. ..... .....G. .....T 750
rpb3- 54 ..... .....A. ..... .....G. .....T 750
rpb3- 99 ..... .....A. ..... .....G. .....T 750
rpb3-125 ..... .....A. ..... .....G. .....T 750
rpb3-158 ..... .....A. ..... .....G. .....T 750
rpb3-159 ..... .....A. ..... .....T. .....T 750

rpb3- w GTTTTAGCCA TTGACCTAGT TGAAATTAAT GTGAATACCT CAGTCATGCC 400
rpb3- 11 .T...A.... C..... ..... ..... ..... 400
rpb3- 12 .T...A.... T..... ..... ..... ..... 400
rpb3- 26 .T...A.... T..... ..... ..... ..... 400
rpb3- 35 .T...A.... T..... ..... ..... ..... 400
rpb3- 54 .T...A.... T..... ..... ..... ..... 400
rpb3- 99 .T...T.... T..... ..... ..... ..... 400
rpb3-125 .C...A.... T..... ..... ..... ..... 400
rpb3-158 .T...A.... T..... ..... ..... ..... 400
rpb3-159 .T...A.... T..... ..... ..... ..... 400

rpb3- w CAGCTGTTGC ATTTGAATAC GATCCATGGA ACAAGTTACA GCATACTGAT 800
rpb3- 11 ..... .....C ..... .....A ..... 800
rpb3- 12 ..... .....C ..... .....A ..... 800
rpb3- 26 ..... .....C ..... .....A ..... 800
rpb3- 35 ..... .....C ..... .....A ..... 800
rpb3- 54 ..... .....T ..... .....A ..... 800
rpb3- 99 ..... .....T ..... .....A ..... 800
rpb3-125 ..... .....C ..... .....G ..... 800
rpb3-158 ..... .....C ..... .....A ..... 800
rpb3-159 ..... .....C ..... .....A ..... 800

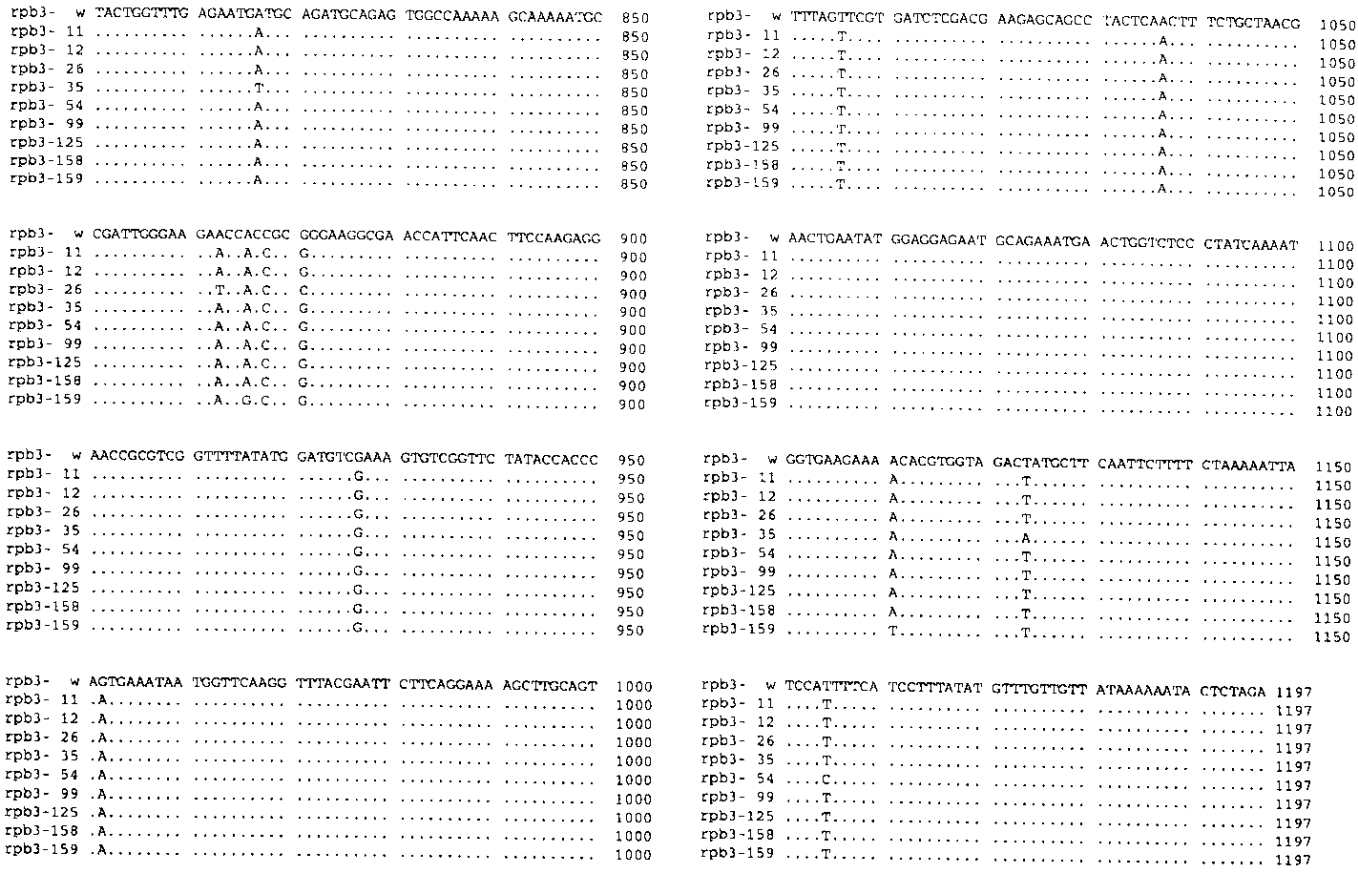



Fig. 28. DNA sequences of the *rpb3* genes from ts mutants.

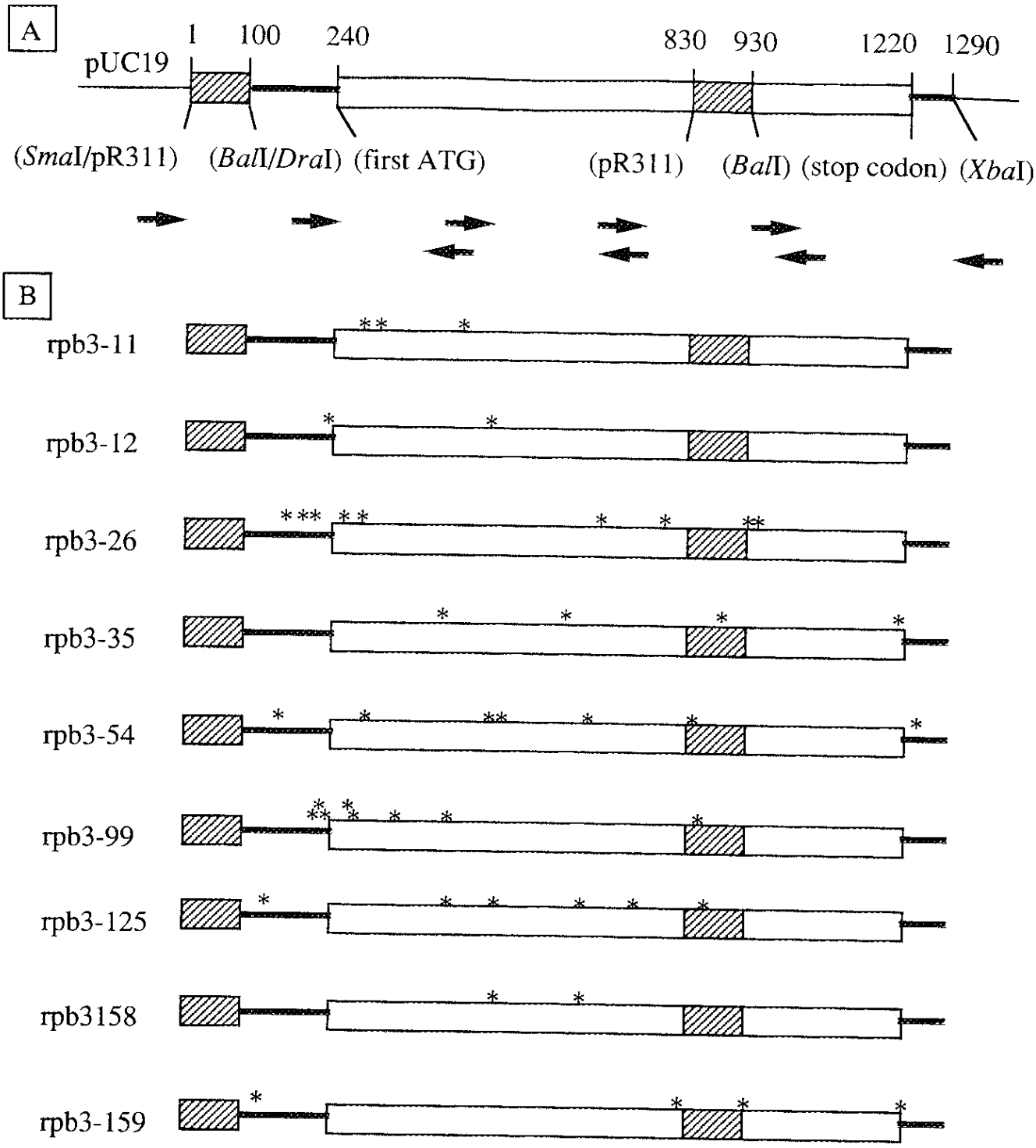



Fig. 29. Mutation sites of the *rpb3* genes.
(A) Cloned fragments of genomic DNA from mutant *S. pombe*. Position, 240, shows the first ATG. Nine arrows show fluore-labeled primers for automatic DNA sequencing. (B) Locations of mutations. Asterisks show the mutation sites.

```
RPB3-  W MDSETHITIR NISKNSVDFV LTNTSLAVAN SLRRVVLAEI PTVAIDLVEI   50
RPB3- 11 .......... .......... F..S...... .......... ....T.....   50
RPB3- 12 .........Q .......... .......... .......... ..........   50
RPB3- 26 .......... .........A .......... .......... ..........   50
RPB3- 35 .......... .......... .......... .......... ..........   50
RPB3- 43 .......... ...T...... .......... .......... ..........   50
RPB3- 54 .......... .......... ...S...... .......... ..........   50
RPB3- 99 ......L... ...RD..... .......... .......... ..........   50
RPB3-125 .......... .......... .......... .......... ..........   50
RPB3-158 .......... .......... .......... .......... ..........   50
RPB3-159 .......... .......... .......... .......... ..........   50

RPB3-  W NVNTSVMPDE FLAHRLGMIP LDSSNIDEPP PVGLEYTRNC DCDQYCPKCS  100
RPB3- 11 .......... .......... .......... .......... ..........  100
RPB3- 12 .......... .....F.... .......... .......... ..........  100
RPB3- 26 .......... .......... .......... .......... ..........  100
RPB3- 35 .......... .......... .......... .......... ..........  100
RPB3- 43 .......... .......... P......... .......... ..........  100
RPB3- 54 .......... ......D... .......... .......... ..........  100
RPB3- 99 .......... .......... .......... .......... ..........  100
RPB3-125 .......... .......L.. .......... .......... ..........  100
RPB3-158 .......... ........V. .......... .......... ..........  100
RPB3-159 .......... .......... .......... .......... ..........  100

RPB3-  W VELFLNAKCT GEGTMEIYAR DLVVSSNSSL GHPILADPKS RGPLICKLRK  150
RPB3- 11 .......... .......... .......... .......... ..........  150
RPB3- 12 .......... .......... .......... .......... ..........  150
RPB3- 26 .......... .......... .......... .......... ..........  150
RPB3- 35 ......T... .......... .......... .......... ..........  150
RPB3- 43 .......... .......... .......... .......... ..........  150
RPB3- 54 .......... .......... .......... .......... ..........  150
RPB3- 99 .......... .......... .......... .......... ..........  150
RPB3-125 .......... ......M... .......... .......... ..........  150
RPB3-158 .......... ......V... .......... .......... ..........  150
RPB3-159 .......... .......... .......... .......... ..........  150
```

```
RPB3-  W EQEISLRCIA KKGIAKEHAK WSPTSAVAFE YDPWNKLQHT DYWFENDADA  200
RPB3- 11 .......... .......... .......... .......... ..........  200
RPB3- 12 .......... .......... .......... .......... ..........  200
RPB3- 26 .......... .......... .......... .......... ..........  200
RPB3- 35 .......... .......... .......... .......... ......V...  200
RPB3- 43 .......... .......... .......... .......... ..........  200
RPB3- 54 .......... .......... .......... .......... ..........  200
RPB3- 99 .......... .......... .......... .......... ..........  200
RPB3-125 .L........ .......... .......... .......R.. ..........  200
RPB3-158 .......... .......... .......... .......... ..........  200
RPB3-159 .......... .......... L......... .......... ..........  200

RPB3-  W EWPKSKNADW EEPPREGEPF NFQEEPRRFY MDVESVGSIP PNEIMVQGLR  250
RPB3- 11 .......... .......... .......... .......... ..........  250
RPB3- 12 .......... .......... .......... .......... ..........  250
RPB3- 26 .......... .D..P..... .......... .......... ..........  250
RPB3- 35 .......... .......... .......... .......... ..........  250
RPB3- 43 .......... .......... .......... ...K...... ..........  250
RPB3- 54 .......... .......... .......... .......... ..........  250
RPB3- 99 .......... .......... .......... .......... ..........  250
RPB3-125 .......... .......... .......... .......... ..........  250
RPB3-158 .......... .......... .......... .......... ..........  250
RPB3-159 .......... .......... .......... .......... ..........  250

RPB3-  W ILQEKLAVLV RDLDEEQPTQ LSANELNMEE NAEMNWSPYQ NGEENTW   297
RPB3- 11 .......... .......... .......... .......... .......  297 (3)
RPB3- 12 .......... .......... .......... .......... .......  297 (2)
RPB3- 26 .......... .......... .......... .......... .......  297 (3)
RPB3- 35 .......... .......... .......... .......... .......  297 (2)
RPB3- 43 .......... .......... .......... .......... .......  297 (3)
RPB3- 54 .......... .......... .......... .......... .......  297 (2)
RPB3- 99 .......... .......... .......... .......... .......  297 (3)
RPB3-125 .......... .......... .......... .......... .......  297 (4)
RPB3-158 .......... .......... .......... .......... .......  297 (2)
RPB3-159 .......... .......... .......... .......... ....I..  297 (2)
```

Fig. 30. Amino acid sequences of the mutant *rpb3* gene products.
Mutated amino acids are shown, while unaltered amino acids are shown by dots.

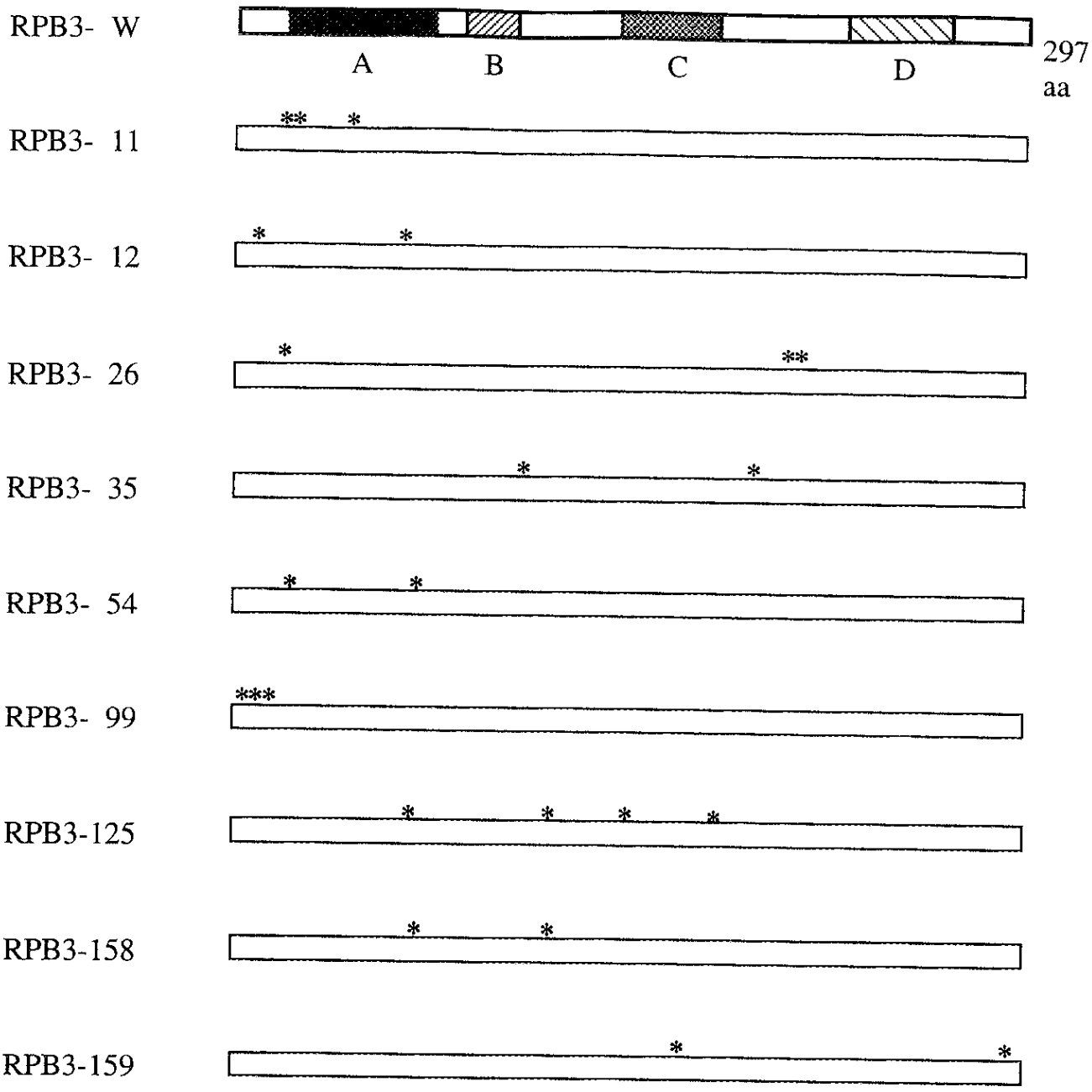



Fig. 31. Locations of mutated amino acids in RPB3 polypeptide.
Asterisks show mutation sites in each mutant. The first column shows the locations of four coserved regions.

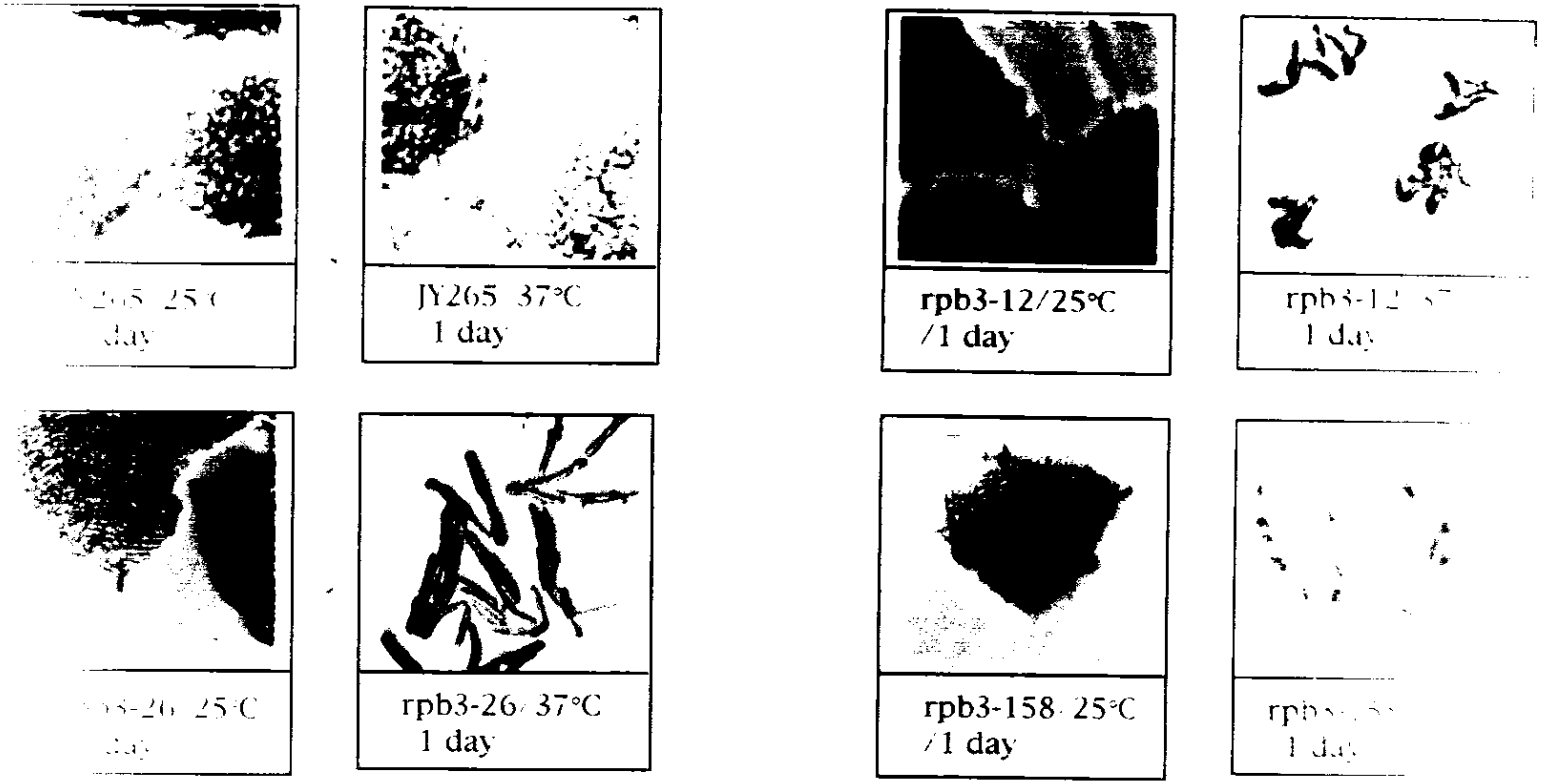



Fig. 32. Cell shapes of *rpb3* ts mutants.
Wild-type (JY265) and three ts mutants (rpb3-12, rpb3-26, rpb3-158) were grown at the indicated temperature for one day.

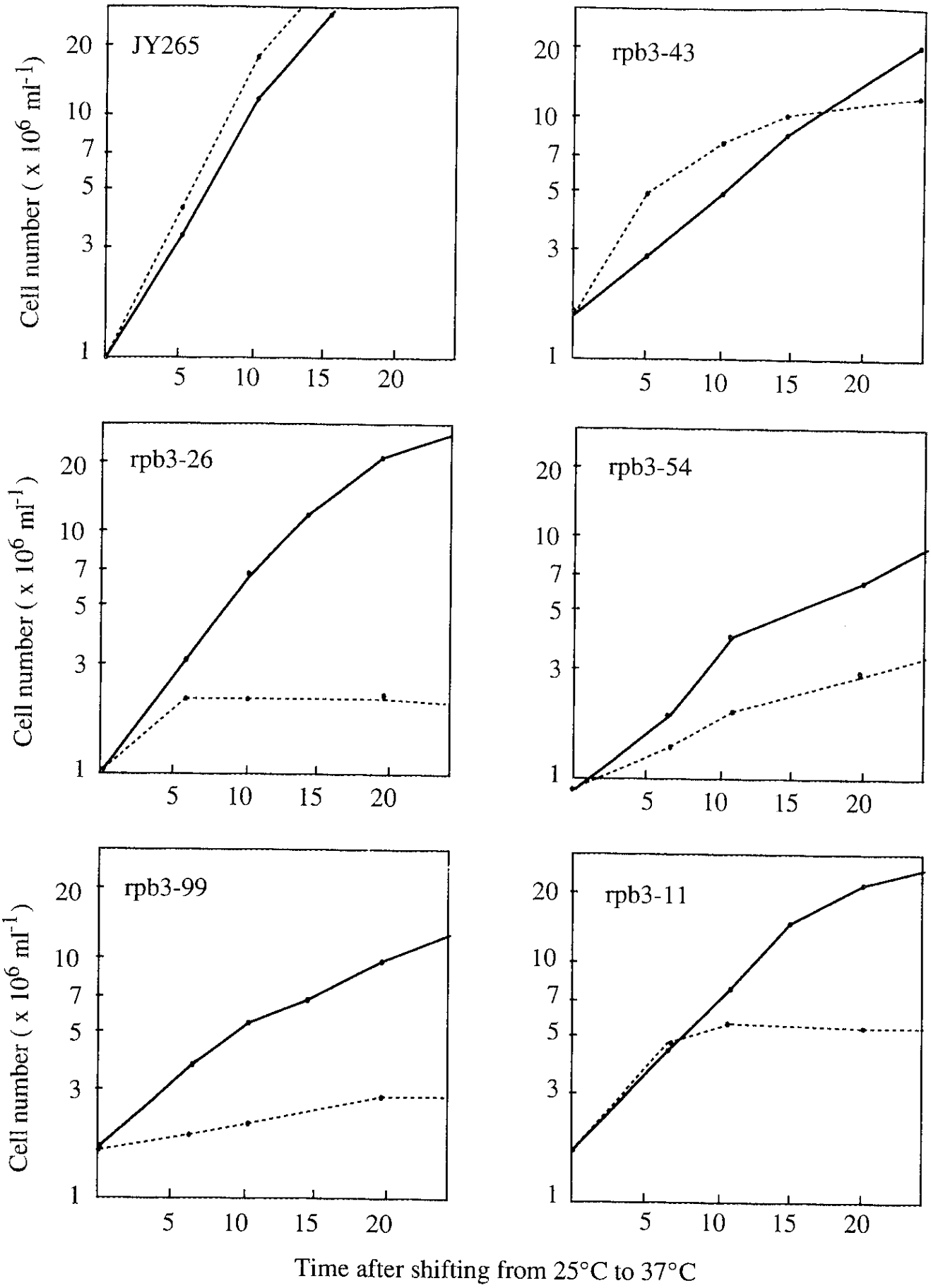



Fig. 33. Growth curve of *rpb3* mutants in SD liquid medium.
Permissive temperature, 25°C (•——•); non-permissive temperature, 37°C (•·····•).

Table 3. List of *rpb3* mutants.

| Strains | rpb3 copy number | Mutation | Temperature sensitivity | | Cell shape | Phenotypes | |
|---|---|---|---|---|---|---|---|
| | | | YPD | SD | (37。C) | Doubling time (hours) | Stop of growth (hours) |
| 3 | 3 | N.D. | TS | TS | Bent | 4.8 | Slowly |
| 11 | 1 | S.D. | TS | TS | Bent | 4.2 | Rapidly |
| 12 | 1 | S.D. | TS | TS | Bent | N.D. | |
| 13 | N.D. | N.D. | TS | TS | | 5.4 | Slowly |
| 26 | 1 | S.D. | TS | TS | Long | 4.1 | Rapidly |
| 35 | 1 | S.D. | TS | | | N.D. | |
| 43 | 2 | N.D. | TS | | | 5.4 | Slowly |
| 54 | 1 | S.D. | TS | TS | | 6.0 | Slowly |
| 92 | 2 | N.D. | TS | TS | | N.D. | |
| 99 | 1 | S.D. | TS | TS | Long | 6.0 | Rapidly |
| 108 | N.D. | N.D. | TS | TS | | N.D. | |
| 125 | 1 | S.D. | TS | TS | Bent | 7.0 | Slowly |
| 158 | 1 | S.D. | TS | TS | | N.D. | |
| 159 | 1 | S.D. | TS | | | N.D. | |
| 165 | N.D. | N.D. | TS | TS | Long | N.D. | |
| 170 | N.D. | S.D. | TS | TS | | N.D. | |

S.D. shows Sequence Determined, N.D. shows Non-Determined.

# CONCLUSION

## Purification of *S. pombe* RNA polymerase II

RNA polymerase II was purified from the wild type *S. pombe* in five steps: precipitation with Polymin P precipitation; elution from the Polymin P precipitates; separation of RNA polymerase II from RNA polymerases I and III by DEAE-Sephadex column chromatography; purification by Q Sepharose FF; and by Superose 6 gel filtration column chromatography. The highly purified *S. pombe* RNA polymerase II contained more than eleven polypeptides as analyzed by SDS-polyacrylamide gel electrophoresis. The molar ratio of the three large subunits, subunits1, 2 and 3, was 1:1:2, which is identical with that of the *E. coli* and *S. cerevisiae* enzymes. In order to identification of the strict meaning 'subunits', essential subunits of RNA polymerase II for non-specific transcription activity, the *in vitro* reconstitution system is necessary. For this purpose, renaturation of urea-treated RNA polymerase II was carried out as a preliminary attempt.

In this attempt, RNA polymerase II was not inactivated up to 0.5 M urea, RNA polymerase II lost activity but the activity was recovered perfectly after dilution of urea between 1.0 and 2.5 M. Above 3.0 M urea, RNA polymerase II activity was irreversibly lost. And the denatured RNA polymerase II with 4 M urea has been never reactivated even if adding many kinds of reagents. From the result of gel filtration, multi subunit complexes of RNA polymerase II were observed but several small components released from the complexes after 4 M urea treatment. Establishment of *in vitro* reconstitution system may be started with this multi subunit complexes polished with 4 M urea as a core-like structure of RNA polymerase II. At that time, some of several small subunits may be also key subunits. Alternatively, in spite of the sufficient subunits remaining in the complex for non-specific transcription activity, some of the subunits may be irreversibly denatured not to possess the activity. In any case, the gene coding each subunit should be necessary to go on analyzing the RNA polymerase II, the gene cloning has been done about the subunits constructing the core-like structure at first.

## Isolation and characterization of the *rpb1* gene

The *S. pombe* gene coding for the largest subunit (subunit 1) of RNA polymerase II was cloned using the *S. cerevisiae* corresponding gene, *RPB1*, as a probe in cross-hybridization under low stringency conditions. The sequence determination of the entire genomic DNA and of parts of cDNA indicated that this *rpb1* gene has six introns in the N-terminal region and encodes a product of 1,752 amino acid residues with the molecular mass of 194 kDa. The subunit 1 polypeptide has high homology with the largest subunit of *S. cerevisiae* RNA polymerase II. From Southern analysis and gene disruption experiment, it was found that this *rpb1* gene exists as a single copy in the *S. pombe* genome and is essential for cell viability. Northern analysis of *rpb1* transcript and sequence determinations of its 3' and 5' terminal regions indicated that the size of the *rpb1* transcript is about 5.6 kbp in length.

Among the subunit 1 of *S. pombe* RNA polymerase II and other β' homologues, nine structurally conserved domains (domains A to H and CTD) were identified: Domain A, a putative zinc-binding site with the consensus sequence of $CX_2CX_9HX_2H$; Domains C and D, the conserved sequences within *E. coli* DNA polymerase I and T7 DNA polymerase; Domain C, a single two-helix motif for putative DNA binding; Domain F, the putative α-amanitin binding site; CTD, highly conserved unique repetition with the unit sequence of YSPTSPS among RNA polymerase II subunits 1, while 29 repeat units exist in subunit 1 of *S. pombe* RNA polymerase II.

## Isolation and characterization of the *rpb3* gene

The cDNA fragment coding for the third largest subunit (subunit 3) of *S. pombe* RNA polymerase II was cloned by RT-PCR using primers designed from the amino acid sequences of V8 fragments of subunit 3. A genomic DNA fragment carrying the entire subunit 3 gene (*rpb3*) was isolated by hybridization using this cDNA fragment. The sequence determination indicated that the coding frame of *rpb3* is interrupted by two introns and this gene encodes subunit 3 of 297 amino acids in

length, which is highly homologous to the *S. cerevisiae* RNA polymerase II RPB3. Southern and Northern analysis indicated that the *rpb3* gene is present as a single copy in *S. pombe* haploid cell and the size of *rpb3* transcript is about 1.2 kbp in length.

Among the RNA polymerase subunit 3 from various organisms, four structural conserved domains (domains A to D) were found: domains A and D exist in the α subunit of *E. coli* RNA polymerase; and domains B and C are conserved only in eukaryotic RNA polymerases. Domain A may play a role in subunit-subunit contact of RNA polymerase. Domain B with a putative metal-binding sequence, $CXCX_3CX_2C$, exists only in RNA polymerase II, but not in RNA polymerase I nor III. Domain D with a leucine zipper-like motif may be required for the formation and/or stability of RNA polymerase. Subunit 3 of eukaryotic RNA polymerase lacks the sequence corresponding to the C-terminal region of *E. coli* RNA polymerase α subunit carrying the contact site I for some transcription activators.

To analyze the function of the *rpb3* gene, the *rpb3* gene was mutagenized by lower fidelity PCR and transferred into *S. pombe* by spheroplast method and electroporation method. Total 178 temperature-sensitive mutants were isolated from about nine thousand transformants. PCR analysis and Southern analysis were carried out for 68 stable ts mutants. Most of the tested mutants carried a single copy of the full-length DNA fragment in their genome DNA. For nine mutants, the mutation sites were determined after cloning and sequencing. All the mutant *rpb3* genes carried multiple mutations, but many mutations were clustered in the N-terminal region of RPB3 polypeptides. Since the *S. cerevisiae* mutant *rpb3-1* affecting RNA polymerase II assembly carried one of the two mutations in this region, those mutants carrying the mutations in N-terminal region may be RNA polymerase II assembly defective mutants.

Upon shift-up of temperature from permissive temperature, 25°C, to non-permissive temperature, 37°C, some of the mutants stop growing immediately. The RNA polymerase II in these mutants may have defect(s) in function at non-permissive temperature. On the other hand, some of the mutants stop growing

slowly, implying that the assembly of RNA polymerase II is defective in these mutants. Some mutants show abnormal cell shapes (for example, 7 times longer in cell size) at the non-permissive temperature. After incubation at non-permissive temperature for one day, none of the sixteen mutants regained viability at permissive temperature, but their nuclei still existed by the observation of DAPI stained cells.

From these results, the subunit 3 must be an important subunit to assemble RNA polymerase II and/or to interact with transcription factors. Then the collection of the mutants may be useful to investigate about RNA polymerase II.

# ACKNOWLEDGMENTS

This thesis fulfillment was supported by helpful and progressive advises of Prof. Akira Ishihama and Dr. Masahiro Yamagishi. I wish to thank deeply Dr. A. Ishihama and Dr. M. Yamagishi. I also thank Dr. N. Fujita, Dr. T. Toyoda and Dr. K. Nagata for the just hitting answers for my strange questions, and thank other members in the Department of Molecular Genetics and in National Institute Genetics for much conversation giving me many ideas.

# REFERENCES


1. Sentenac, A. (1985) CRC Crit. Rev. Biochem., **18**, 31-90.
2. Sentenac, A. and Hall, B.D. (1982) In Strathern, J.N., Jones, E.W. and Broach, J.R. (eds.),The Molcular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 561-606.
3. Azuma, Y., Yamagishi, M., Ueshima, R. and Ishihama, A. (1991) Nucleic Acids Res., **19**, 461-468.
4. Kawagishi, M., Yamagishi, M., Ishihama, A. (1993) Nucleic Acids Res., **21**, 469-473.
5. Kolodziej, P. and Young, R.A. (1989) Mol. Cell. Biol. **9**, 5387-5394.
6. Mann, C., Buhler, J,-M., Treich, I. and Sentenac, A. (1987) Cell, **48** 627-637.
7. Pati, U.K. and Weissman, S.M. (1990) J. Biol. Chem., **265**, 8400-8403.
8. Martindale, D.W. (1990) Nucleic Acids Res., **18**, 2953-2960.
9. Yura, T. and Ishihama, A. (1989) Annu. Rev. Genet., **13**, 59-97.
10. Ishihama, A. (1986) Adv. Biophy., **21**, 163-173.
11. Ishihama, A. (1992) Mol. Microbiol., **6**, 3283-3288.
12. Ishihama, A. (1993) J. Bacteriol., **175**, in press.
13. Helmann, J.D. and Chamterlin, M.J. (1988) Annu. Rev. Biol., **57**, 839-872.
14. Ishihama, A. (1988) Trends Genet., **4**, 282-286.
15. Horikoshi, M., Wang, C. K., Fujij, H., Cromlish, J.A., Weil ,P.A. and Roeder, R.G. (1989) Nature, **341**, 299-303.
16. Woychik, N.A. and Youmg, R.A. (1989) Mol. Cell. Biol., **9**, 2854-
17. Russell, P. (1985) Gene, **40**, 125-130.
18. Kaufer, N.F., Simanis, V. and Nurse, P. (1985) Nature, **318**, 78-80.
19. Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) Annu. Rev. Biochem., **55**, 1119-1150.
20. Nonet, M., Scafe, C., Sexton, J. and Young, R. (1987) Mol. Cell. Biol., **7**, 1602-1611.
21. Jendrisak, J. J. and Burgess, R. R.(1975) Biochemistry, **14**, 4639-4645.
22. Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Raudall, R.J. (1951) J. Biol. Chem., **193**, 265-275.
23. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Moleclar Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
24. Yamagishi, M. and Nomura, M. (1988) Gene, **74**, 503-515.

25. Kaiser, K. and Murray, N. (1986) In Glover, D.M. (ed.), DNA Cloning: A Practical Approach. IRL Press, Oxford, Vol.I, pp.1-47.
26. Anderson, M.L.M. and Young, B.D. (1985) In Hames, B.D. and Higgins, S.J. (eds.), Nucleic Acid Hybridization: A Practical Approach. IRL Press, Oxford, pp. 73-111.
27. Dale, R. M. K., McClure, B. A. and Honchins, J. P. (1985) Plasmid, **13**, 31-40.
28. Yanisch-perron, C.,Vieira, J. and Messing, J. (1985) Gene, **33**, 103-119.
29. Sanger, H., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA, **74**, 5463-5467.
30. Tabor, S. and Richardson, C. C. (1987) Proc. Natl. Acad. Sci. USA, **84**, 4767-4771.
31. Moon, I. S. and Krause, M. O. (1991) Methods in Molecular and Cellular Biology, **2**, pp123
32. Carothers, A.M., Urlaub, G., Mucha, J., Grunberer, D. and Chasin, L.A. (1989) BioTechniques, **7**, 494-499.
33. Domdey, H., Apostol, B., Lin, R.-J., Newman, A., Brody, E. and Abelson, J. (1984) Cell, **39**, 611-621.
34. Laemmli, U.k. (1970) Nature (London), **227**, 680-685.
35. Toda, T. and Ohashi, M. (1986) J. Biol. Chem., **261**, 12455-12461.
36. Leung, D. W., Chen, E., Goeddel, D. V. (1989) Technique, **1**, 11-15.
37. Igarashi, K. and Ishihama, A. (1991) Cell, **65**, 1015-1022.
38. Mark T. Hood and C. atachow. (1990) Nucleic Acids Res., **18**, 688
39. David Beach and Paul Nurse. (1981) Nature, **290**,
40. Sergio Moreno, Jacqueline Hayles and Paul Nurse (1989) Cell, **58**, 361-372.
41. Schultz, L.D. and Hall, B.D. (1976) Proc. Nat. Acad. Sci. USA, **73**, 1029-1033.
42. Bartolomei, M.S. and Corden, J.L. (1987) Mol. Cell. Biol., **7**, 586-594.
43. Kolodziej, P.A., Woychik, N., Liao, S-M., Young, R.A. (1990) Mol. Cell. Biol., **10**, 1915-1920.
44. Kozak, M. (1989) J. Cell. Biol., **108**, 229-241.
45. Sweetser, D., Nomet, M. and Young, R.A. (1987) Proc. Natl. Acad. Sci. USA, **84**, 1192-1196.
46. Memet, S., Gouy, M., Marck, C., Sentenac, A. and Buhler, J.-M. (1987) J. Biol. Chem., **263**, 2830-2839.
47. Himmelfarb, H.J., Simpson, E.M. and Friesen, J.D. (1987) Mol. Cell. Biol., **7**, 2155-2164.
48. Wittekind, M., Dodd, J., Vu, L., Kolb,J.M., Buhler, J.-M., Sentenac, A. and Nomura, M. (1988) Mol. Cell. Biol., **8**, 3997-4008.

49. Allison, L.A., Moyle, M., Shales, M. and Ingles, C.J. (1985) Cell, **42**, 599-610.
50. Jokerst, R.S., Weeks, J.R., Zehring, W.A. and Greenleaf, A.L. (1989) Mol. Gen. Genet., **215**, 266-275.
51. Greenleaf, A. L. (1983) J. Biol. Chem., **258**, 13403-13406.
52. Bartolomei, M.S. and Corden, J.L. (1987) Mol. Cell. Biol., **7**, 586-594.
53. Schultz, L. and Hall, B.D. (1976) Proc. Natl. Acad. Sci. USA, **73**, 1029-1033.
54. Corden, J.L. (1990) Trends Biochem. Sci., **15**, 383-387.
55. Allison, L.A., Wong, J.K.C., Fitzpatrick, V.D., Moule, M. and Ingles, C.J. (1988) Mol. Cell.Biol., **8**, 321-329.
56. Nonet, M., Sweetser, D. and Young, R.A. (1987) Cell, **50**, 909-915.
57. Bartolomei, M.S., Halden, N.F., Cullen, C.R. and Corden, J.L. (1988) Mol. Cell Biol., **8**, 330-339.
58. Allison, L.A. and Ingles, C.L. (1989) Proc. Natl. Acad. Sci. USA, **86**, 2794-2798.
59. Scafe, C., Chao, D., Lopes, J., Hirsch, J.P., Henry, S. and Young, R.A. (1990) Nature, **347**, 491-494.
60. Koleske, A. J., Buratowski, S., Nonet, M. and Young. R. A. (1992) Cell, **69**, 883-894.
61. Li, W.-B., Bzik, D.J., Gu, H., Tanaka, M., Fox, B.A. and Inselburg, J. (1989) Nucleic Acids Res., **17**, 9621-9636.
62. Nawrath, C., Schell, J. and Koncz, C. (1990) Mol. Gen. Genet., **223**, 65-75.
63. Bird, D.M. and Riddle, D.L. (1989) Mol. Cell. Biol., **9**, 4119-4130.
64. Ahearn, J.M. Jr., Bartolomei, M.S., West, M.L., Cisek, L.J. and Corden, J.L. (1987) J. Biol. Chem., **262**, 10695-10705.
65. Nonet, M., Scafe, C., Sexton, J. and Young, R. (1987) Mol. Cell. Biol., **7**, 602-1611.
66. Meek, D.W. and Hayward, R.S. (1984) Nucleic Acids Res., **12**, 5813-5821.
67. Ishihama, A., Shimamoto, N., Aiba, H., Kawakami, K., Nishimoto, H.,Tsugawa, A. and Uchida, H. (1980) J. Mol. Biol. **137**, 137-150.
68. Igarashi, K., Fujita, N. and Ishihama, A. (1990) Nucleic Acids Res., **18**, 5945-5948.
69. Kolodziej, P.A., Young, R.A. (1991) Mol. Cell. Biol., **11**, 4669-4678.
70. Igarashi, K., Fujita, N. and Ishihama, A. (1991) J. Mol. Biol., **218**, 1-6.
71. Igarashi, K., Hanamura, A., Makino, K., Aiba, H., Aiba, H., Mizuno, T., Nakata, A. and Ishihama, A. (1991) Proc. Natl. Acad. Sci. USA, **88**, 8958-8962.