Cloning of the *Schizosaccharomyces pombe* TFIID gene reveals a strong conservation of functional domains present in *Saccharomyces cerevisiae* TFIID

Alexander Hoffmann,[1] Masami Horikoshi,[1] C. Kathy Wang,[2] Stephanie Schroeder,[2] P. Anthony Weil,[2] and Robert G. Roeder[1]

[1]Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, New York 10021 USA; [2]Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, Tennessee 37232 USA

The gene encoding the *Schizosaccharomyces pombe* TATA box-binding factor (TFIID) was cloned and sequenced. The gene contains three introns and codes for a polypeptide of 231 amino acids. The cDNA-expressed protein showed both TATA box-binding and basal transcription activities. The carboxy-terminal three-quarters of *S. pombe* TFIID shares an extraordinary degree of amino acid sequence homology with a corresponding region of *Saccharomyces cerevisiae* TFIID that has been shown to be necessary and sufficient for TATA box-binding and basal transcription activities. In contrast, the amino-terminal regions of the *S. pombe* and *S. cerevisiae* TFIIDs differ markedly in amino acid sequence and composition. Structure and function relationships of TFIID are discussed in light of these data.

[*Key Words*: TATA factor; TFIID; *Schizosaccharomyces pombe*; transcription initiation; functional domains; structural motifs]

Received May 8, 1990; accepted May 24, 1990.

Among the wide variety of gene-specific promoter elements, the TATA box (consensus TATAT/AAT/A) is almost universal and the major determinant for core promoter activity and the position of transcription initiation by RNA polymerase II (for review, see Breathnach and Chambon 1981; Nakajima et al. 1988). TFIID was described originally as a general transcription initiation factor (Matsui et al. 1980) and shown to bind directly to the TATA element (Nakajima et al. 1988). Recent studies have shown that it plays a central role in both promoter activation and regulation: first, by the recruitment of other initiation factors (and RNA polymerase II) into a functional preinitiation complex and, second, by interactions with regulatory factors (for review, see Horikoshi et al. 1989, 1990). The ability of a *Saccharomyces cerevisiae* factor to substitute for the human TFIID in a functional assay (Buratowski et al. 1988; Cavallini et al. 1988) led to its purification and cloning (for review, see Lillie and Green 1989). More recently, structure–function studies have implicated an unusually large and apparently unique domain in both TATA box-binding and basal level transcription (Horikoshi et al. 1990).

The apparent evolutionary conservation of TFIID function has prompted studies of TFIID gene and protein structure in other organisms, including the fission yeast *Schizosaccharomyces pombe*. There appear to be significant differences between *S. pombe* and *S. cerevisiae*, with respect to cellular or molecular events involved in mitosis and cytokinesis (Hiraoka et al. 1984), cell cycle control and growth (Russell and Nurse 1986; Lee and Nurse 1987), and transcription initiation (Losson et al. 1985; Russell 1985) and RNA splicing (Käufer et al. 1985; Padgett et al. 1986), with *S. pombe* showing more similarities to higher eukaryotes than to *S. cerevisiae*. Consistent with this are indications that *S. pombe* and *S. cerevisiae* have diverged from each other as much as they have from humans (Matsumoto and Yanagida 1985; Russell and Nurse 1986). Here, we show that the sequence of a large functional domain of *S. cerevisiae* TFIID is highly conserved in the *S. pombe* protein, whereas other regions are apparently unrelated and possibly indicative of distinct regulatory interactions.

Results

Isolation of the S. pombe *TFIID gene and cDNA*

Southern blots of *S. pombe* DNA were probed with an *S. cerevisiae* TFIID gene fragment under low stringency

hybridization conditions. A single cross-hybridizing band was observed with each of several restriction enzymes (Fig. 1A), which suggests the presence of a single *S. pombe* gene with sequence relatedness to the *S. cerevisiae* gene (a control hybridization with the homologous genomic fragment at high stringency is shown in Fig. 1B). A 1.37-kb *Hin*dIII fragment that reproducibly cross-hybridized with the *S. cerevisiae* probe under low stringency conditions (Fig. 1A) was isolated from a subgenomic library (Materials and methods). Sequence analysis revealed stretches that had high sequence identity to *S. cerevisiae* TFIID in different reading frames but were interrupted by three putative intervening sequences that contained stop codons in all three reading frames. A probe from this fragment was used to isolate additional genomic clones, two of which were mapped and partially sequenced. Both contained the previously identified *Hin*dIII fragment with ~10 kb of additional 5′- and 3′-flanking sequences. On the basis of amino acid homology to *S. cerevisiae* TFIID, the complete gene was identified (Fig. 2), along with putative introns of 255, 52, and 70 bp that contain consensus splice donor and acceptor sites (GTANGT and TAG; Käufer et al. 1985). The two largest putative introns also contain splice recognition boxes that conform to the consensus established previously for *S. pombe* (CTA/GAT/C in the 3′ part of the intron; Langford et al. 1984; Käufer et al. 1985), whereas the third and shortest appears not to require such a recognition sequence. To verify the putative exon/intron structure and to obtain a functional cDNA, an oligo(dT)-primed cDNA library was screened and PCR products from total RNA were cloned and screened. Appropriate clones were sequenced and shown to contain the previously predicted open reading frame, which is preceded by several stop codons.

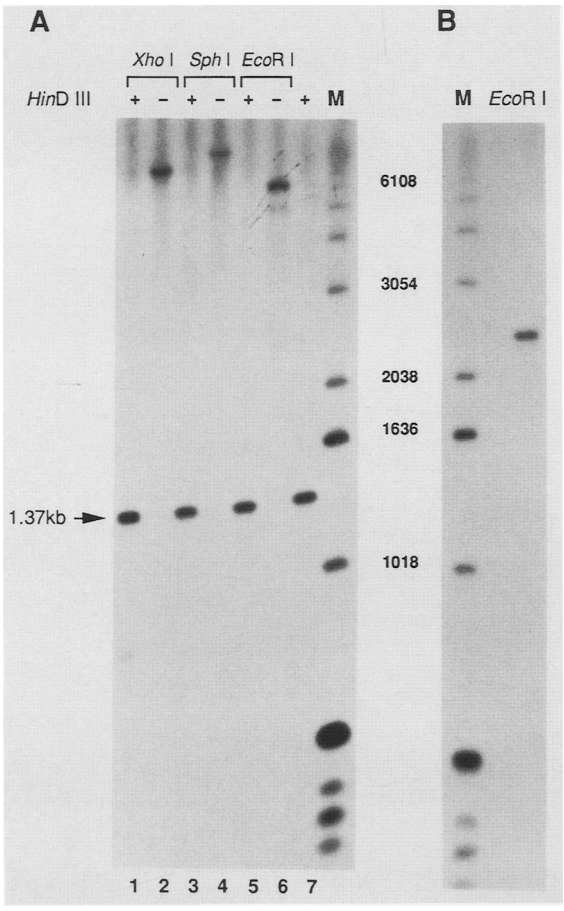

Figure 1. Detection of *S. pombe* TFIID gene on a genomic Southern blot. (*A*) *S. pombe* genomic DNA probed with *S. cerevisiae* TFIID DNA under low stringency hybridization conditions after digestion with *Xho*I, *Sph*I, *Eco*RI, or *Hin*dIII alone (lanes *2*, *4*, *6*, and *7*) or in combination with *Hin*dIII (lanes *1*, *3*, and *5*). (*B*) *S. cerevisiae* genomic DNA digested with *Eco*RI and probed with *S. cerevisiae* TFIID DNA under high stringency conditions.

Characterization of S. pombe *TFIID RNA*

Northern blot analysis (Fig. 3A) was used to detect and compare mRNAs from TFIID genes in *S. pombe* and *S. cerevisiae*. The probes did not show cross-hybridization at high stringency conditions but did reveal discrete RNAs of 1.3 kb for *S. pombe* and 1.1 kb for *S. cerevisiae* (as shown previously; Horikoshi et al. 1989). The 5′ terminus of the *S. pombe* TFIID mRNA was mapped by primer extension (Fig. 3B) to a position ~25 bp downstream from a putative TATA box on the genomic sequence (Fig. 2), indicating a 5′-untranslated leader sequence of 68 nucleotides. A putative polyadenylation signal similar to the consensus for higher eukaryotes (AATAAA; Proudfoot et al. 1976) was found 15 bp from the apparent poly(A) tail of the mRNA (Fig. 2), indicating a 3′-untranslated trailer of 529 nucleotides. Given a coding sequence of 693 nucleotides, the predicted size of the mature message is 1294 nucleotides exclusive of the poly(A) tail, in good agreement with the estimated size of the natural mRNA (Fig. 3A). In none of the relevant experiments [genomic and cDNA cloning, polymerase chain reaction (PCR) analysis on genomic DNA and total RNA, Northern mapping, primer extension, and low stringency genomic Southern analysis] has there been any evidence for additional homologous genes or alternative RNA splicing in *S. pombe* (data not shown).

Functional activity of S. pombe *TFIID*

To demonstrate that the open reading frame described above encodes a protein with TFIID activity, a corresponding cDNA was cloned into pGEM-7Zf(+) and T7 polymerase-transcribed products were translated in reticulocyte lysates. SDS–polyacrylamide gel analysis (Fig. 4A, left) revealed a single band with an apparent size of 25 kD, in good agreement with the calculated molecular weight of 25,427. The *S. pombe* RNA-encoded protein also bound specifically to the TATA box in the adenovirus major late promoter, as shown by competition with wild-type versus mutant TATA box-containing oligonucleotides (Fig. 4B). As expected, the

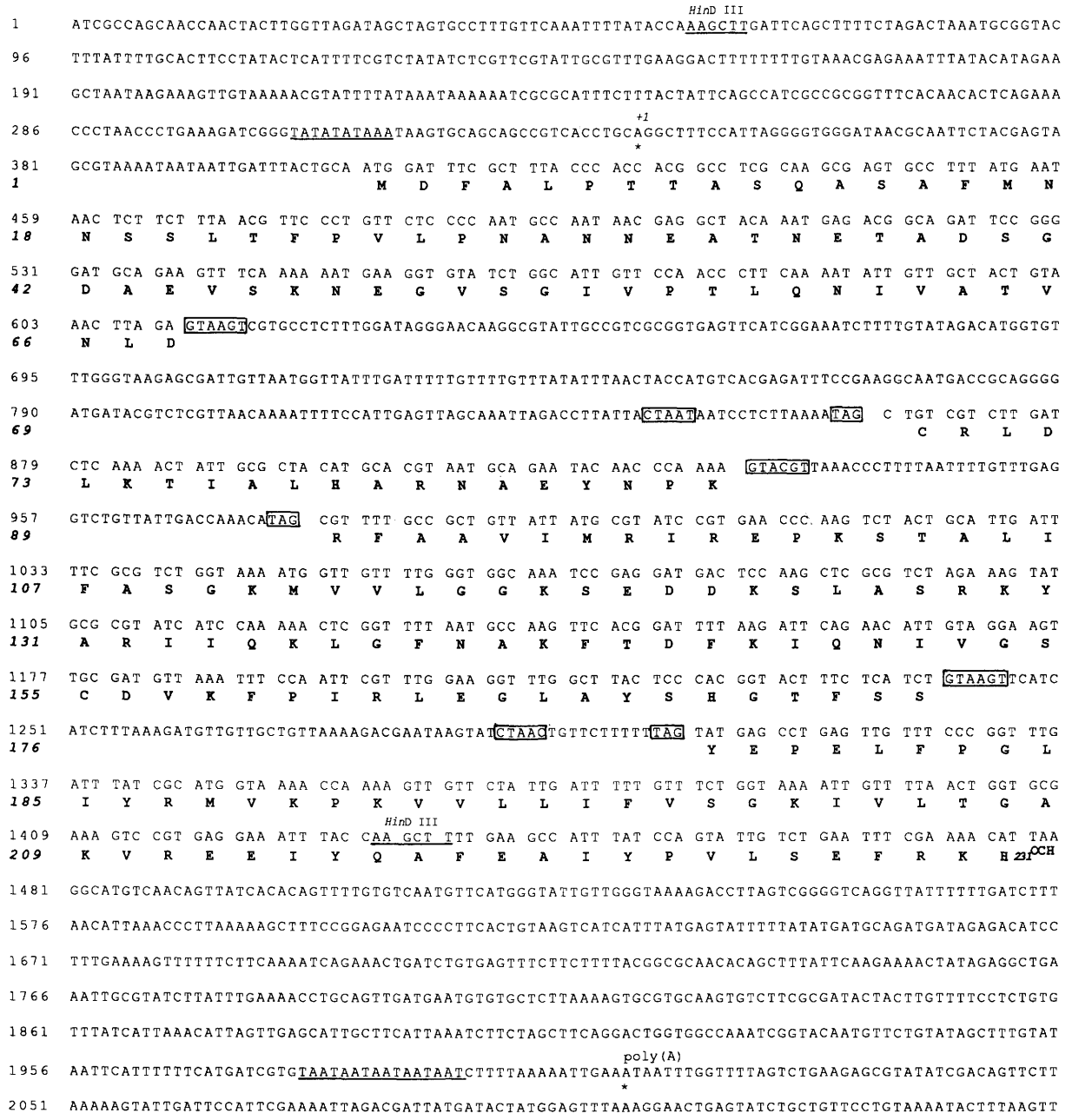

Figure 2. Nucleotide and predicted amino acid sequences of the *S. pombe* TFIID gene. The exon/intron boundaries are based on sequencing of genomic and cDNA clones and are indicated by boxes around splice consensus sequences. The *Hin*dIII sites for the genomic fragment identified in Fig. 1A are shown. Also shown are the transcription start site (+1), a putative polyadenylation signal (underlined), and the location of the poly(A) tail found in the cDNAs. The predicted translation of the TFIID open reading frame is shown.

mobility of the resulting complex was slightly faster than that observed with the 27-kD *S. cerevisiae* protein (data not shown). The *S. pombe* cDNA-encoded protein (p25) was also expressed in *Escherichia coli* (Fig. 4A, right) and showed transcriptional activity when assayed in a TFIID-deficient complementation system from human cells (Fig. 4C). This ability to substitute for the human TFIID provides the final proof that the cloned *S. pombe* gene encodes a functional protein.

Discussion

This report describes the isolation and characterization of an *S. pombe* gene encoding TFIID. The gene is present in a single copy, contains three introns, and appears to encode only a single spliced mRNA. This RNA encodes a functional TFIID, as evidenced by the ability of the in vitro-expressed protein to bind to the TATA box and to initiate transcription in conjunction with RNA poly-

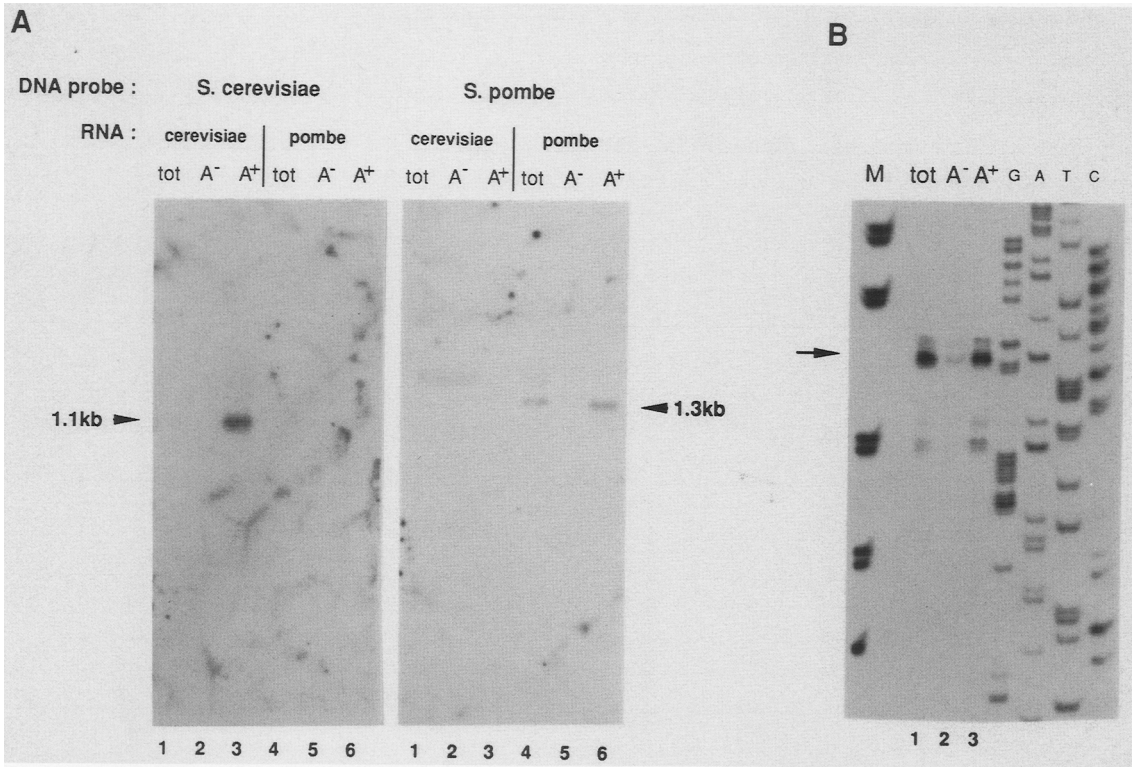

Figure 3. Mapping of *S. pombe* mRNA. (*A*) Northern blots of 20 μg of total RNA, poly(A)$^-$ RNA, and poly(A)$^+$ RNA from *S. pombe* and *S. cerevisiae* probed under high stringency conditions with both *S. pombe* and *S. cerevisiae* TFIID DNA fragments. The positions of the 1.3-kb (*S. pombe*) and 1.1-kb (*S. cerevisiae*) hybridizing species were determined from size markers (not shown). (*B*) Primer extension analysis of total, poly(A)$^-$, and poly(A)$^+$ RNA from *S. pombe*. Primer extension products were run along with size markers and a DNA sequence ladder generated from the cloned TFIID gene with the same primer used for the RNA analysis.

merase II and other general initiation factors from human cells. Sequence comparisons reveal an extraordinary degree of homology (93% sequence identity) between the 180-residue carboxy-terminal region of *S. pombe* TFIID (residues 52–231) and the corresponding region of *S. cerevisiae* TFIID (residues 61–240), whereas the remaining amino-terminal regions of the proteins show no similarity in sequence. This sequence conservation is remarkable, considering the evolutionary distance between these organisms. Moreover, at least some aspects of the mechanism of transcription initiation differ, because the distance between the TATA box and the transcription initiation site is large and variable (up to 120 bp) in *S. cerevisiae* but closer to that in higher eukaryotes (20–40 bp) in *S. pombe*. The extent of the sequence conservation between the two TFIID proteins is even higher than that reported for the eukaryote RNA polymerases (for review, see Cornelisson et al. 1988) and similar to that (91–92%) reported for histones H3 and H4 (Matsumoto and Yanagida 1985).

This sequence conservation suggests a highly conserved TFIID "core" structure that must satisfy a number of essential functions on the promoter. These include both site-specific binding to the TATA element by apparently unique types of interactions (Nakajima et al. 1988; for review, see Horikoshi et al. 1990) and the subsequent recruitment (by direct or indirect interactions) of RNA polymerase II and other general initiation factors into a functional preinitiation complex (Van Dyke et al. 1988; Buratowski et al. 1989). The extent of sequence conservation between *S. pombe* TFIID and *S. cerevisiae* TFIID is remarkably consistent with recent mutagenesis studies (Horikoshi et al. 1990), which show that amino-terminal residues 1–62 of *S. cerevisiae* TFIID are completely dispensable for basal level transcription, whereas the remainder of the molecule is absolutely essential. Even more striking than the overall conservation of amino acid sequence in the carboxy-terminal region is the near perfect conservation of the residues comprising previously described structural motifs, including the lysine repeat in the central basis core, the flanking direct repeats and the sigma homology (Fig. 5). Altogether, these studies argue strongly for a TFIID core that contains potentially overlapping domains that are essential for these general functions and are highly conserved in evolution. Below, we consider the possible roles of these domains in light of these studies. Because TFIID also can be a target for various regulatory factors (for review, see Horikoshi et al. 1989, 1990), considerations of these interactions are also relevant.

The direct repeats and the σ homology Interrupted direct repeats, indicative of an ancient duplication event, have been noted in *S. cerevesiae* TFIID (Cavallini et al.

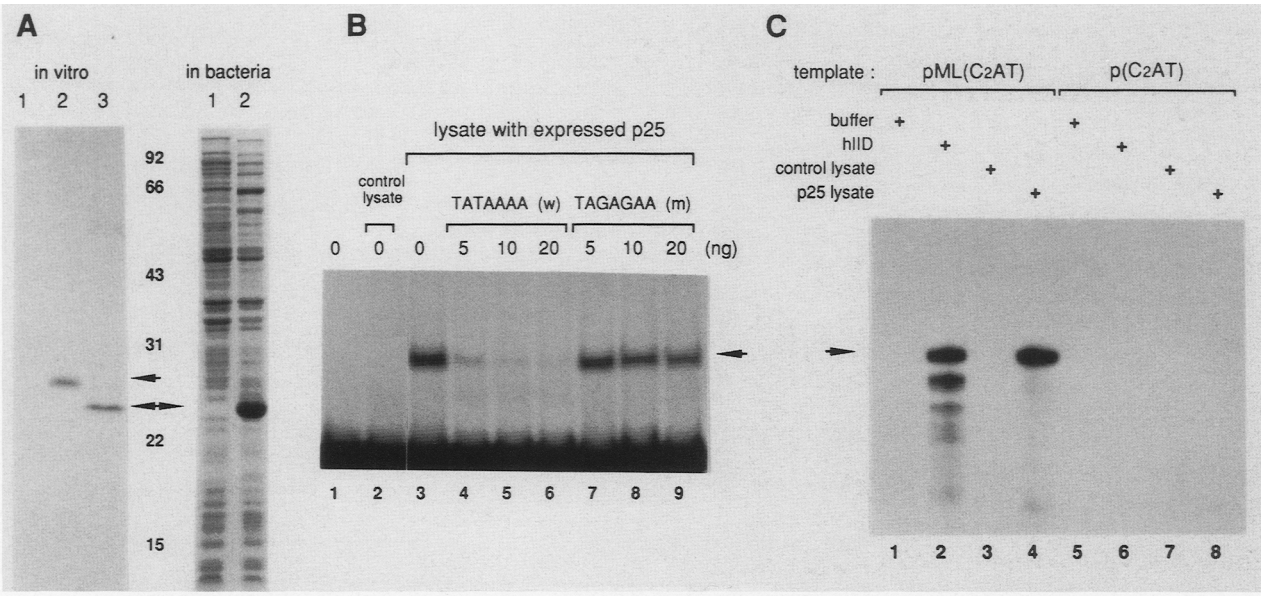

Figure 4. Expression and functional analysis of cloned *S. pombe* TFIID. (*A*) SDS-PAGE analysis of [^{35}S]methionine-labeled TFIID expressed in reticulocyte lysates and visualized by autoradiography (*left*) or Coomassie Blue-stained TFIID expressed in bacteria (*right*). (*Left*) Reticulocyte lysates were programmed with no RNA (lane *1*) or with in vitro-transcribed TFIID mRNA from *S. cerevisiae* (lane *2*) or *S. pombe* (lane *3*) cDNAs. (*Right*) *E. coli* lysates were prepared 3 hr postinduction of cells containing the PET3a plasmid alone (lane *1*) or with the *S. pombe* TFIID cDNA insert (lane *2*). (*B*) Specific binding of *S. pombe* TFIID to the TATA box. Gel mobility-shift assays with a 184-bp adenovirus major late promoter fragment used no protein (lane *1*) and reticulocyte lysates programmed with no RNA (lane *2*) or with *S. pombe* TFIID RNA (lanes *3–9*). Competing oligonucleotides (containing base pairs −45 to −15 of the adenovirus major late promoter) contained either wild-type or mutant TATA sequence (as indicated) and were added at the levels (0–20 ng) indicated. (*C*) Transcriptional activation by *S. pombe* TFIID. Transcription assays containing partially purified TFIIB, IIE/F, and RNA polymerase II were complemented with buffer (lanes *1* and *5*), with partially purified human TFIID (lanes *2* and *6*), with lysate from *E. coli* containing the PET3a plasmid with no insert (lanes *3* and *7*), and with lysate from *E. coli* containing the PET3a plasmid with the *S. pombe* TFIID cDNA insert (lanes *4* and *8*). The pML(C_2AT) template in lanes *1–4* contained the adenovirus major late promoter (sequences −400 to +10) attached to the G-less cassette construct, whereas the p(C_2AT) template in lanes *5–8* contained only the G-less cassette (Sawadago and Roeder 1985). Templates containing only promoter sequences from −50 to +10 attached to the G-less cassette behaved exactly like the pML(C_2AT) templates (data not shown). The arrow indicates the transcript resulting from specific initiation at the +1 site.

1989; Hoeijmakers 1990; Nagai 1990; Stucka and Feldmann 1990). Mutational studies (Horikoshi et al. 1990; T. Yamamoto, unpubl.) have implicated both of these repeats in DNA binding and have led to the suggestion that they could give the TFIID structure a twofold symmetry that might be important for overall binding specificity and strength. The fact that only 2 of the 25 amino acids in the more strictly defined repeat (shaded residues in Fig. 5A) differ between *S. pombe* and *S. cerevisiae* indicates a strong evolutionary pressure to preserve this structure. However, the significance of the larger, more broadly defined direct repeats in *S. cerevisiae* (Cavallini et al. 1989; Hoeijmakers 1990) is questionable, because part of the first repeat falls into the TFIID amino-terminal region that is not conserved between the two yeasts. We note also that many of the invariant residues within the smaller repeats are hydrophobic and therefore could be involved in stabilizing interactions between the two domains.

We have noted previously weak homologies between the carboxy-terminal regions of *S. cerevisiae* TFIID (residues 180–240) and a region of bacterial σ-factor involved directly in recognition of the −10 promoter element (TATAAT consensus). Consistent with the possibility that these regions are functionally related, the present analysis shows that a near absolute conservation of the relevant residues [the exception being a conservative change at position 202 in *S. pombe*; mutational analysis (T. Yamamoto, unpubl.)] has also implicated specific residues of the σ homology in TFIID binding. Because the σ homology region lies partially within the second direct repeat and is not conserved in the first, the regions of TFIID containing the repeats may not be functionally equivalent.

Basic repeat region As noted previously, *S. cerevisiae* TFIID contains an interesting repeat of basic residues, especially lysines, in the central region of TFIID (Horikoshi et al. 1989). This sequence appears capable of forming an α-helical structure, and deletion mutants in this region are incapable of binding DNA (Horikoshi et al. 1990). Strikingly, all of the charged amino acids (mainly lysines) within this 37-residue region are identical between *S. cerevisiae* and *S. pombe* TFIID, indicating that a specific structure is important for function.

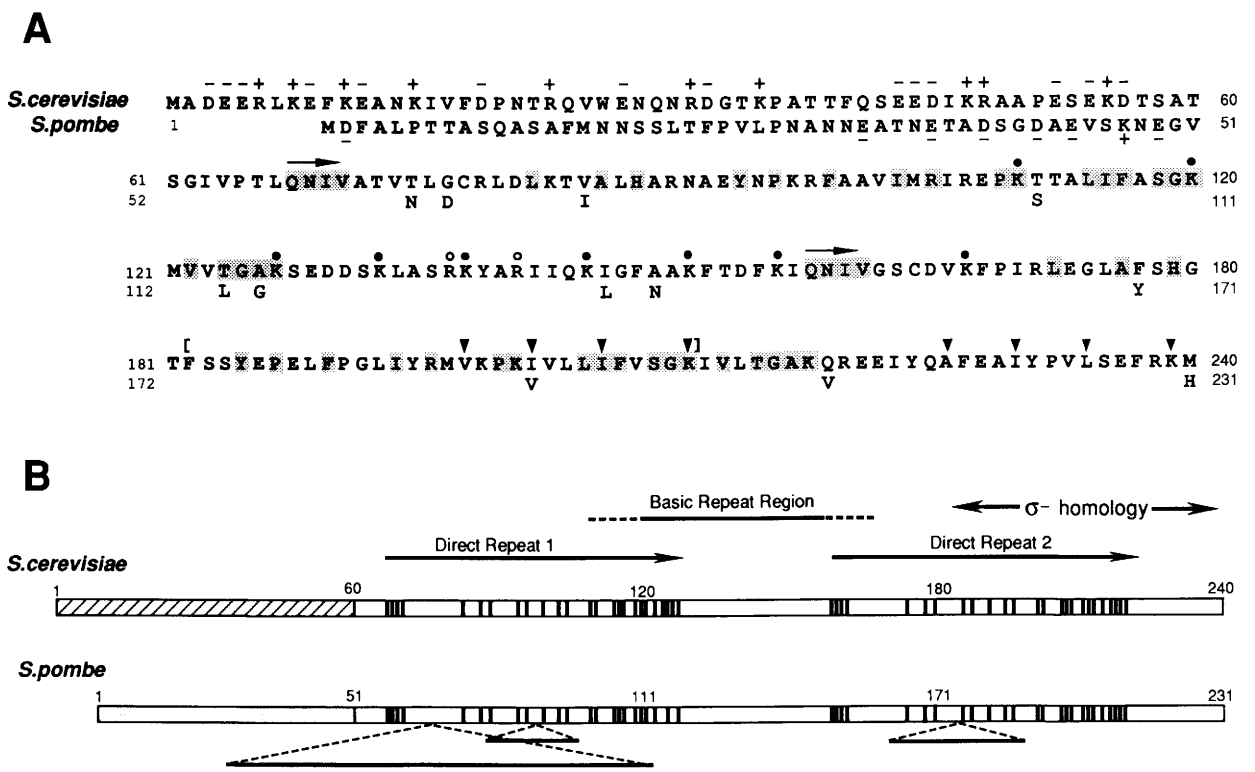

Figure 5. Sequence comparisons and a conserved core domain in yeast TFIIDs. (*A*) The sequence of *S. cerevisiae* TFIID (Horikoshi et al. 1989) is shown in the upper line, whereas only the *S. pombe* TFIID residues differing between the two yeasts are shown in the second line. The + and − symbols denote the basic and acidic residues in the highly divergent amino termini. The arrows indicate the beginning of each of the two interrupted direct repeats (Cavallini et al. 1989), with the individual residues comprising these repeats shaded. Lysines (●) and arginines (○) indicate residues in the central basic core (Horikoshi et al. 1989); this core has the potential for forming an α-helix with the positively charged residues (especially the lysine repeat) on one face. The σ-homology region (Horikoshi et al. 1989) is comprised of a region (denoted by brackets) with a weak overall sequence identity to the σ 2.3 and 2.4 domains and two small regions containing residues (▼) that are the most highly conserved in the 2.4 regions of various σ-factors. (*B*) The schematic shows both the large carboxy-terminal core that is highly conserved (93% sequence identity) between *S. pombe* and *S. cerevisiae* and the conserved motifs (above) within this core. Vertical bars in the cores indicate the residues comprising the direct repeats; the lower horizontal bars indicate positions and relative sizes of the introns in the *S. pombe* gene.

More recent studies with point mutants (T. Yamamoto, unpubl.) have shown that most of the lysine residues between the direct repeats, presumably on one face of an α-helix (Horikoshi et al. 1989, 1990), are not important either for DNA binding or for basal promoter activity. Given the strong conservation of these residues, these results are consistent with the earlier suggestion (Horikoshi 1989) that this repeat may play a role in mediating the action of certain types (e.g., acidic) of activators.

Amino-terminal regions Contrasted with these conserved structural features are the divergent amino termini. Although both hydrophilic, the 60 residue *S. cerevisiae* amino terminus has a much higher density of charged residues (14 acidic and 10 basic) than does the *S. pombe* amino terminus (7 acidic and 1 basic). On the other hand, the 51 residue *S. pombe* amino terminus contains high concentrations of serines (6) and threonines (5), as well as alanines (8) and asparagines (7), which are not found in the *S. cerevisiae* protein.

The precise separation of completely divergent amino termini from the highly conserved core domain is conspicuous and suggests a functional difference for these regions. The greater and more variable distance between the TATA box and the initation site in *S. cerevisiae* raises the possibility of a different initiation mechanism that might involve novel interactions of general factors (or other intermediates) with the amino-terminal region. Because the TATA box–initiation site distance on *S. pombe* promoters is closer to that observed for mammalian genes, it may be relevant that the amino terminus of the human TFIID, although larger, resembles that of *S. pombe* more closely (Hoffmann et al. 1990). On the other hand, and especially because the TFIID core is sufficient for basal promoter activity, it seems equally probable that the amino termini may be involved in differential interactions with regulatory factors. The availability of active regulatory factors and reconstituted cell-free systems from several organisms, including *S. cerevisiae*, should allow these possibilities to be tested.

Materials and methods

Isolation of the S. pombe *TFIID gene and cDNA*

To construct the subgenomic library (cf. Fukui and Kaziro 1985), 20 μg of *S. pombe* (strain 972h⁻) DNA was digested with *Hin*dIII and fractionated on a preparative agarose gel. Fragments in the range of 1.3–1.45 kb were electroeluted and ligated into pGEM-7Zf(+). Upon transformation of super-competent XL1-Blue cells (Stratagene), the resulting 200 colonies were divided into 24 minicultures. Plasmid DNA was prepared from each, restricted with *Hin*dIII, *Eco*RI, *Sph*I, and *Xho*I, and analyzed by Southern blot under low stringency conditions. A single positive culture had an *S. pombe* DNA fragment of the expected size (1.37 kb), and the constituent clones were similarly analyzed to yield a single positive clone. Subsequently, the 684-bp *Hin*cII–*Hin*dIII subfragment was used to screen 8×10^4 clones of a genomic library (*Sau*3A partial) in pDB248 (generous gift of Dr. Masayuki Yamamoto). Ten positive clones were isolated, and restriction mapping revealed two classes that each contained the full-length TFIID gene on different genomic fragments.

An oligo(dT)-primed cDNA library was constructed from *S. pombe* poly(A)⁺ RNA in λ II (Stratagene). Plaques (2.4×10^5) were screened by standard methods, and 18 resulting positive clones were plaque-purified and analyzed. Partial and full-length cDNAs were also obtained by PCR on total RNA with several pairs of primers.

All DNA sequence data were determined by the dideoxy chain termination method with Sequenase (U.S. Biochemicals) in each direction with both dGTP and dITP.

Southern, Northern, and primer extension analysis

Standard procedures were as described (Maniatis et al. 1982). For Southern analysis, 1-μg amounts of *S. pombe* (strain 972h⁻) or *S. cerevisiae* (strain BJ926) DNA were digested with restriction enzymes, electrophoresed on 1% agarose gels, blotted onto nitrocellulose (Schleicher & Schuell), and probed with a ^{32}P-labeled *S. cerevisiae* TFIID gene fragment (3×10^8 cpm/μg) containing base pairs 1–672 of the open reading frame. Hybridization solution (50 μl/cm²) contained 10^5 cpm/ml probe, 10× Denhardt's solution, 3× SSC, 50 mM HEPES (pH 8.4), 50 μg/ml denatured salmon sperm DNA, and either 20% (low stringency) or 50% (high stringency) deionized formamide. The blot was incubated at 60°C for 1 hr and cooled slowly to 42°C. After an additional 12–16 hr, blots were washed in 0.1× SSC, 0.1% SDS, at room temperature for 1 hr and exposed to Kodak XAR films with double screens for 3–12 hr at −70°C.

For Northern analysis, total RNA was isolated from *S. cerevisiae* and *S. pombe*, separated into poly(A)⁺ and poly(A)⁻ fractions using oligo(dT)-cellulose, fractionated on 1.5% agarose gels containing 2 M formaldehyde, transferred to nitrocellulose, UV cross-linked, and probed with random primer-labeled fragments of the *S. cerevisiae* TFIID gene (nucleotide positions 122–1616; Horikoshi et al. 1989) or the *S. pombe* TFIID gene (684-bp *Hin*cII–*Hin*dIII fragment). Blots were processed as described (Horikoshi et al. 1989).

Primer extension mapping of the 5′ end employed a ^{32}P-end-labeled oligonucleotide complementary to nucleotides 348–378 (Fig. 2) and resolution on a standard 6% polyacrylamide–urea sequencing gel prior to autoradiography.

Expression of TFIID and functional assays

An *S. pombe* cDNA (corresponding to nucleotides 51–882 of the mature mRNA) in pGEM-7Zf(+) was in vitro-transcribed with T7 RNA polymerase, and derived RNA was translated in rabbit reticulocyte lysates (Promega) according to supplier's recommendations. Mobility-shift assays were performed according to previously published conditions (Horikoshi et al. 1989), with 0.5 μl of reticulocyte lysate and a ^{32}P-end-labeled fragment containing adenovirus major late promoter sequences (−138 to +46). An *S. pombe* cDNA, corresponding to nucleotides 51–882 of the mRNA, was cloned into *Nde*I and *Bam*HI sites of the PET3a plasmid (Rosenberg et al. 1987) and expressed in *E. coli* by induction with IPTG. Lysates were prepared by sonication of cells in 20 mM Tris-HCl (pH 7.9), 500 mM NaCl, 20 mM 2-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride, 20 μg/ml leupeptin, 20 μg/ml pepstatin, and 10% (vol/vol) glycerol. Transcription assays were performed as described (Horikoshi et al. 1989), with circular plasmid templates containing the adenovirus major late promoter attached to the G-less cassette (Sawadogo and Roeder 1985).

Acknowledgments

We thank Dr. Masayuki Yamamoto for the *S. pombe* genomic library. A.H. is an Arnold and Mabel Beckman graduate fellow and thanks many, especially Polly Gregor, Hildegard Kaulen, Thomi Gerster, Shona Murphy, Michael Meisterernst, Hillel Fromm, Eric Sinn, and, above all, Alexander Gasch and Tohru Yamamoto for comments and technical advice. M.H. is an Alexandrine and Alexander L. Sinsheimer Scholar. This study was supported by National Institutes of Health grant DK-42502 to P.A.W., grants CA-42567 and UO1 AI-27397 to R.G.R., and by general funding from the Pew Trusts to The Rockefeller University.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

Note added in proof

Sequence data described in this paper have been submitted to the EMBL/GenBank Data Libraries under accession number X-53415.

References

Breathnach, R. and P. Chambon. 1981. Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50:** 349–383.

Buratowski, S., S. Hahn, P.A. Sharp, and L. Guarente. 1988. Function of a yeast TATA element-binding protein in a mammalian transcription system. *Nature* **334:** 37–42.

Buratowski, S., S. Hahn, L. Guarente, and P.A. Sharp. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56:** 549–561.

Cavallini, B., J. Huet, J.-L. Plassat, A. Sentenac, J.-M. Egly, and P. Chambon. 1988. A yeast activity can substitute for the HeLa cell TATA box factor. *Nature* **334:** 77–80.

Cavallini, B., I. Faus, H. Matthes, J.M. Chipoulet, B. Winsor, J.-M. Egly, and P. Chambon. 1989. Cloning of the gene encoding the yeast protein BTF1Y, which can substitute for the human TATA box-binding factor. *Proc. Natl. Acad. Sci.* **86:** 9803–9807.

Cornelisson, A.W.C.A., R. Evers, and J. Kock. 1988. Structure and sequence of genes encoding subunits of eukaryotic RNA polymerases. *Oxford Surv. Eukaryotic Genes* **5:** 92–131.

Fukui, Y. and Y. Kaziro. 1985. Molecular cloning and sequence analysis of a ras gene from Schizosaccharomyces pombe. *EMBO J.* **4:** 687–691.

Hiraoka, Y., T. Toda, and M. Yanagida. 1984. The NDA3 gene of fission yeast encodes beta-tubulin: A cold-sensitive nda3 mutation reversibly blocks spindle formation and chromosome movement in mitosis. *Cell* **39:** 349–358.

Hoeijmakers, J.H.J. 1990. Cryptic initiation sequence revealed. *Nature* **343:** 417–418.

Hoffmann, H., E. Sinn, T. Yamamoto, J. Wang, A. Roy, M. Horikoshi, and R. Roeder. 1990. A human TATA factor (TFIIID) contains a highly conserved core domain and a unique N-terminus with presumptive regulatory motifs. *Cell* (in press).

Horikoshi, M., C.K. Wang, H. Fujii, J.A. Cromlish, P.A. Weil, and R.G. Roeder. 1989. Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. *Nature* **341:** 299–303.

Horikoshi, M., T. Yamamoto, Y. Ohkuma, P.A. Weil, and R.G. Roeder. 1990. Analysis of structural and functional relationships of yeast TATA box-binding factor TFIID. *Cell* (in press).

Käufer, N.F., V. Simanis, and P. Nurse. 1985. Fission yeast Schizosaccharomyces pombe correctly excises a mammalian RNA transcript intervening sequence. *Nature* **318:** 78–80.

Langford, C.J., F.-J. Klinz, C. Donath, and D. Gallwitz. 1984. Point mutations identify the conserved, intron-contained TACTAAC box as an essential splicing signal in yeast. *Cell* **36:** 645–653.

Lee, M.G. and P. Nurse. 1987. Complementation used to clone a human homologue of the fission yeast cell cycle control gene cdc2. *Nature* **327:** 31–35.

Lillie, J.W. and M.R. Green. 1989. Activator's target in sight. *Nature* **341:** 279–280.

Losson, R., R.P.R. Fuchs, and F. Lacroute. 1985. Yeast promoters URA1 and URA3: Examples of positive control. *J. Mol. Biol.* **185:** 65–81.

Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor Press, New York.

Matsui, T., J. Segall, P.A. Weil, and R.G. Roeder. 1980. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J. Biol. Chem.* **255:** 11992–11996.

Matsumoto, S. and M. Yanagida. 1985. Histone gene organization of fission yeast: A common upstream sequence. *EMBO J.* **4:** 3531–3538.

Nagai, K. 1990. Cryptic initiation sequence revealed. *Nature* **343:** 418.

Nakajima, N., M. Horikoshi, and R.G. Roeder. 1988. Factors involved in specific transcription by mammalian RNA polymerase II: Purification, genetic specificity, and TATA box-promoter interactions of TFIID. *Mol. Cell. Biol.* **8:** 4028–4040.

Padgett, R.A., P.J. Grabowski, M.M. Konarska, S. Seiler, and P.A. Sharp. 1986. Splicing of messenger RNA precursors. *Annu. Rev. Biochem.* **55:** 1119–1150.

Proudfoot, N.J. and G.G. Brownlee. 1976. 3′ Non-coding region sequence in eukaryotic messenger RNA. *Nature* **263:** 211–214.

Rosenberg, A.H., B.N. Lade, D.-S. Chui, S.-W. Lin, J.J. Dunn, and F.W. Studier. 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56:** 125–135.

Russell, P. 1985. Transcription of the triose-phosphate-isomerase gene of Schizosaccharomyces pombe initiates from a start point different from that in Saccharomyces cerevisiae. *Gene* **40:** 125–130.

Russell, P. and P. Nurse. 1986. Schizosaccharomyces pombe and Saccharomyces cerevisiae: A look at yeasts divided. *Cell* **45:** 781–782.

Sawadogo, M. and R.G. Roeder. 1985. Factors involved in specific transcription by human polymerase II: Analysis by a rapid and quantitative in vitro assay. *Proc. Natl. Acad. Sci.* **82:** 4394–4398.

Stucka, R. and H. Feldmann. 1990. An element of symmetry in yeast TATA box-binding protein transcription factor IID. *FEBS Lett.* **261:** 223–225.

Van Dyke, M.W., R.G. Roeder, and M. Sawadogo. 1988. Physical analysis of transcription preinitiation complex assembly on a class II gene promoter. *Science* **241:** 1335–1338.

Cloning of the Schizosaccharomyces pombe TFIID gene reveals a strong conservation of functional domains present in Saccharomyces cerevisiae TFIID.

A Hoffmann, M Horikoshi, C K Wang, et al.

Genes Dev. 1990 4: 1141-1148
Access the most recent version at doi:10.1101/gad.4.7.1141