RNA polymerase II (RNAPII) is the multisubunit enzyme that synthesizes all mRNA precursors in eukaryotes. RNAPII is highly conserved among species, and in humans, RNAPII consists of 12 subunits, named Rpb1 to Rpb12 (16, 88). The two largest subunits, Rpb1 (220 kDa) and Rpb2 (140 kDa), form the enzyme’s catalytic center and are homologous to the β’ and β subunits of bacterial RNAP, respectively. Five subunits, Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12, are also found in RNAPI and RNAPIII. Rpb3 and Rpb11 are homologous to the α2 homodimer involved in bacterial RNAP assembly. Rpb9 was attributed a role in elongation through its action at DNA arrest sites (3). In Saccharomyces cerevisiae, Rpb4 and Rpb7 form a subcomplex that can dissociate from the enzyme upon changes in environmental conditions. Under active growth conditions, most yeast RNAPII molecules do not contain the Rpb4-Rpb7 dimer, which primarily associates during the stationary phase or following stress (11). Functional studies of RNAPII have been limited due to the lack of appropriate systems for purifying variant forms of the human enzyme.

The availability of crystal structures of both yeast (17, 18, 26, 30) and bacterial (8, 57, 84, 89) RNAPs has been invaluable for understanding many of the molecular features of the transcription reaction. For example, the structure of elongating RNAPII has revealed the positioning of the RNA-DNA duplex during the transcription reaction (30). The structures available support a model in which the DNA enters the enzyme through a channel formed by a pair of “jaws” before accessing a deep cleft, at the bottom of which is buried the Mg2+ ion-containing active site; the DNA then turns by about 90° along a wall where the upstream end exits the enzyme (30). Many loops and helices either directly contact or closely approach the RNA-DNA duplex, thus suggesting putative functions for these structural elements in the transcription reaction. Again, the lack of an efficient method for purifying variant forms of RNAPII with mutations within specific structural elements has precluded any detailed structure-function analysis of human RNAPII.

The activity of RNAPII at the various stages of the transcription reaction is regulated by several proteins, including the general transcription factors (14, 32, 65). TATA-binding protein (TBP), the TATA box-binding subunit of TFIIID, recognizes the TATA box of promoters, and its binding induces a bend of approximately 80° in the DNA (45, 46). TFIIID associates with the TBP-promoter complex and participates in RNAPII recruitment (54). TFIIF, composed of subunits RAP74 and RAP30, binds directly to RNAPII and also helps to recruit the enzyme to the promoter (15, 80). TFIIF can induce further bending and wrapping of the promoter DNA around RNAPII during preinitiation complex formation (72).
Both TFIIIE and TFIIH were shown to participate in promoter melting at the transcription initiation site (38, 66). TFIIIE, which is also composed of two subunits, TFIIE56/α and TFIIE34/β, stimulates the helicase and kinase activities of TFIIH (53, 62, 63). The helicase activity of TFIIH is required to fully open the promoter DNA between nucleotides (nt) −9 and +2 (37, 39, 44, 83), and its kinase activity phosphorylates the carboxy-terminal domain (CTD) of Rpb1 (21, 53, 75–77). The phosphorylation of the CTD on Ser5 by TFIIH has been shown to be required for the transition from initiation to early elongation and allows for the recruitment of the capping enzyme (35, 73). In yeast, the Ctk1 kinase phosphorylates the CTD on Ser2 at a later stage of the elongation phase (10), leading to the recruitment of mRNA processing factors (78).

Unlike the case for all of the general transcription factors and the prokaryotic NF-κB, it has been impossible to date to purify functionally active forms of mammalian RNAPII reconstituted from its recombinant subunits. We now report the purification of a functionally active human RNAPII by the use of doubly tagged subunits expressed in mammalian cells. Using the TAP procedure (71), we isolated a human RNAPII-containing complex which comprises the 12 subunits of core RNAPII, TFIIA, TFIIH, Fcp1, and RPAP1, a novel polypeptide of unknown function. The tagged RNAPII complex is functionally active both in vitro and in vivo. We capitalized on this purification procedure to generate two variant forms of human RNAPII by using TAP-tagged Rpb2. The first mutant has a deletion in fork loop 1, a structure that is expected to participate in maintaining the open state of the transcription bubble in elongating RNAPII (18, 30) or during RNA-DNA strand separation (85). The second mutant carries a triple alanine substitution in switch 3, a loop that directly contacts the DNA template within the DNA-RNA hybrid. Our results allowed us to detail the roles of fork loop 1 and switch 3 in transcriptional mechanisms both in vitro and in vivo.

**MATERIALS AND METHODS**

**Protein factors and antibodies.** Recombinant yeast TBP (42), human TFIIA (31), RAP74 and RAP30 (22), TFIIE56/α and TFIIE34/β (64, 67, 81), highly purified TFIIH (20, 27), and calf thymus RNAPII (36) were purified as previously described. Antibodies raised against Rpb1 (N-20; Santa Cruz) were pelleted by a 2-min centrifugation in a room-temperature table-top centrifuge and then frozen immediately in liquid nitrogen. Total RNAs were prepared by hot acidic phenol-chloroform extraction followed by ethanol precipitation. Poly(A)+ RNA purification and cDNA labeling were performed as described previously (55). Microarrays were composed of 70-mer oligonucleotides, each specific to a different yeast and human TFIIH (Invitrogen; accession number BE883306), Rpb7 (Invitrogen; accession number BE883306), Rpb2 (kindly provided by Marc Vigneron, IGBMC, Strasbourg, France) (1), and RAP30 (22) were cloned into the mammalian expression vector pMZI Rpb4 (Invitrogen; accession number BE883306), Rpb7 (Invitrogen; accession number BE883306), and TBP (30 ng), TFIIB (30 ng), RAP30 (30 ng), RAP74 (640 ng), TFIIE34 (160 ng), and TFIIE56 (320 ng), and calf thymus RNAPII or the TAP-tagged human RNAPII complex (the amounts of RNAPII used in the various experiments are detailed in the figure legends).

**DNA microarray analysis of Tet-promoter mutants.** Tet-promoter alleles were constructed as previously described (86). Mutant and isogenic wild-type cultures were grown in parallel in SC medium with 10 μg of doxycycline/ml for a total of 22 h. Dilutions were made at 16 h to ensure that mutant and wild-type optical density was matched as closely as possible to 10° cells at harvest. The cells were pelleted by a 2-min centrifugation in a room-temperature table-top centrifuge and then frozen immediately in liquid nitrogen. Total RNAs were prepared by hot acidic phenol-chloroform extraction followed by ethanol precipitation. Poly(A)+ RNA purification and cDNA labeling were performed as described previously (55). Microarrays were composed of 70-mer oligonucleotides, each specific to a different yeast and human TFIIH (Invitrogen; accession number BE883306), Rpb7 (Invitrogen; accession number BE883306), and RAP30 (22) were cloned into the mammalian expression vector pMZI Rpb4 (Invitrogen; accession number BE883306), Rpb7 (Invitrogen; accession number BE883306), Rpb2 (kindly provided by Marc Vigneron, IGBMC, Strasbourg, France) (1), and RAP30 (22) were cloned into the mammalian expression vector pMZI Rpb4 (Invitrogen; accession number BE883306), Rpb7 (Invitrogen; accession number BE883306), and TBP (30 ng), TFIIB (30 ng), RAP30 (30 ng), RAP74 (640 ng), TFIIE34 (160 ng), and TFIIE56 (320 ng), and calf thymus RNAPII or the TAP-tagged human RNAPII complex (the amounts of RNAPII used in the various experiments are detailed in the figure legends).
Abortive initiation assay. Abortive initiation assays were performed as described previously (52), with the following modifications. The DNA template (12 ng), either premelted in the −9 to +2 region or fully double-stranded, was incubated with TBP (60 ng), TFIIB (30 ng), RAP30 (30 ng), RAP74 (65 ng), TFIIE4 (40 ng), TFIIE5 (60 ng), and the TAP-tagged human RNAPII complex (the amounts of RNAPII used for the various experiments are detailed in the figure legends) in a 20 µl reaction mixture containing 750 µM ATP, 750 µM CTP, 10 µM UTP, 2.5 µCi of [α-32P]UTP, 12.5 mM MgCl2, 3 mM EGTA, and 0.64 U of RNase inhibitor/ml. The transcripts were analyzed in a 23% polyacrylamide denaturing gel containing 7 M urea.

Elongation assay. The template was prepared by annealing two complementary oligonucleotides of random sequences, one of which carried an extension of 15 CMP as described previously (66). Typically, the template (9.3 ng) was incubated with 88 ng of TAP-tagged human RNAPII Rpb2 wt or 440 ng of TAP-tagged human RNAPII Rpb2 forkl (Δ458-459)-TAP, and the reactions were performed as described previously (66). The transcripts were analyzed in 18% polyacrylamide–9% urea denaturing gels.

Nucleotide sequence accession number. The RPAP1 name and sequence were registered in the HUGO database under accession number NP_056355.

RESULTS

TAP of human RNAPII-containing complex. We have adapted the TAP procedure (69, 71), originally developed to isolate protein complexes from yeast, for the purification of human RNAPII and general transcription factors. Native conditions were maintained during the TAP procedure through the use of buffers with a low salt concentration (0.1 M Na+ and a neutral pH. As shown in Fig. 1A, which summarizes the TAP procedure, full-length cDNAs encoding polypeptides of interest were cloned in an ecdysone-inducible mammalian expression vector (61). The vector was engineered to encode polypeptides carrying the following at the C terminus: a TAP expression vector (61). The vector was engineered to encode polypeptides carrying the following at the C terminus: a TAP expression vector.

The template was prepared by annealing two complementary oligonucleotides of random sequences, one of which carried an extension of 15 CMP as described previously (66). Typically, the template (9.3 ng) was incubated with 88 ng of TAP-tagged human RNAPII Rpb2 wt or 440 ng of TAP-tagged human RNAPII Rpb2 forkl (Δ458-459)-TAP, and the reactions were performed as described previously (66). The transcripts were analyzed in 18% polyacrylamide–9% urea denaturing gels.

Nucleotide sequence accession number. The RPAP1 name and sequence were registered in the HUGO database under accession number NP_056355.

FIG. 1. TAP of human transcription factors. (A) Overview of the TAP procedure. (B) Data from a pilot experiment comparing the expression levels of a TAP-tagged polypeptide and its endogenous counterpart upon induction with ponasterone A. TFIIB-TAP expressed in EcR-293 cells after induction with different concentrations of ponasterone A (0, 1, and 3 µM) for 24 h was compared to that of endogenous TFIIB by Western blotting with an antibody raised against TFIIB.

including Rpb2, Rpb4, Rpb7, TFIIB, and RAP30, expressed them in stably transfected EcR-293 cells, and used them as baits in the TAP procedure. The eluates were analyzed by silver staining of SDS-PAGE gels, Western blotting, and/or mass spectrometry (Fig. 2B). The results indicated that all 17 polypeptides are bona fide components of the human RNAPII complex. Notably, the eluate prepared with TAP-tagged RAP30 contained both tagged RAP30 with its CBP residual domain and nontagged RAP30 (Fig. 2C). The presence of RAP30 lacking the CBP domain was confirmed by Western blotting (data not shown). These results indicate that two molecules of RAP30 are present within the complex, suggesting that TFIIF is associated with RNAPII as an αββ heterotrimer (see Discussion). Finally, our Western blot analysis with an antibody specifically directed against the N-terminal part of Rpb1 (N-20) revealed that the tagged RNAPII complex contains the hypophosphorylated form of RNAPII (IIA form) (Fig. 2D).

RPAP1 is structurally and functionally related to human RNAPII. RPAP1 is a previously identified human gene product (DKFZP727M111 protein) of unknown function. BLAST se-
sequence searches against publicly available databases revealed highly significant alignments over the full length of the RPAP1 sequence with proteins of unknown function from *Mus musculus* (accession no. BA065787), *Rattus norvegicus* (accession no. XP_230480), and to a lesser extent, *Drosophila melanogaster* (accession no. NP_648573 and CG32104-PB). In addition, these proteins show significant similarities to the *Saccharomyces cerevisiae* Ydr527wp protein (accession no. NP_010816), which is required for cell viability and was shown to bind the Rpb10 subunit of RNAPII in a two-hybrid screen (28, 43). The positions of the different alignments are shown in Fig. 3A. Notably, no particular motif or domain could be predicted with high confidence on the RPAP1 sequence. However, in view of the participation of RPAP1 in the RNAPII complex, an IMPALA analysis against the BLOCK database (74) revealed a very weak similarity (E = 0.23) between the bromodomaining-containing protein 4 (Brd4) and the RPAP1 region spanning amino acids 196 to 312 (see below). We also performed a sequence-structure comparison analysis by using the structure-guided alignment program FUGUE against the HOMSTRAD database (56). The result suggests that RPAP1 belongs to the ARM repeat protein superfamily (12). Indeed, the RPAP1 region spanning amino acids 508 to 960 shows an alignment with ARM with a Z score of 4.08 (corresponding to 95% confidence). Interestingly, both Ydr527wp and the RPAP1 *D. melanogaster* homologue were also predicted, as reported in the *Saccharomyces* Genome and FlyBase databases, respectively, to belong to the ARM repeat superfamily (http://genome-www4.stanford.edu/cgi-bin/SGD/locus.pl?locus = YDR527W and http://flybase.bio.indiana.edu/bin/fbidq.html?FBgn=0052104).

We created a Tet-promoter allele of the yeast YDR527w gene, turned off gene expression by adding doxycycline to the growth medium for 24 h, and compared gene expression in this mutant to expression patterns from 121 other Tet-promoter mutants affecting genes implicated in diverse biological processes (S. Mnaimneh, A. P. Davierwala, J. Haynes, J. Mofatt, W. T. Peng, W. Zhang, X. Yang, J. Pootoolal, G. Chua, A. Lopez, M. Trochesset, D. Morse, N. J. Krogan, S. L. Hiley, Z. Li, Q. Morris, J. Grigull, N. Mitsakakis, C. J. Roberts, J. F. Greenblatt, C. Boone, C. A. Kaiser, B. J. Andrews, and T. R. Hughes, submitted for publication). Among all of these, the effect of the tet-YDR527w allele correlated most strongly with that of tet-RPB11 (Fig. 3B). A clustering analysis did not reveal any specific group of functionally related transcripts that were up- or down-regulated uniquely in these two mutants, supporting the idea that the expression pattern observed is a consequence of an overall perturbation in transcription rather than a disruption of any individual pathway.

To further define the interaction between RPAP1 and RNAPII, we cloned the human RPAP1 cDNA and derived a human cell line that can be induced to express TAP-tagged RPAP1. An analysis of the TAP-tagged RPAP1 eluate revealed the presence of two RNAPII subunits, Rpb2 and Rpb3 (Fig. 3A, right panel). These results further demonstrate the physical interaction between RPAP1 and RNAPII. Recent TAP-tagging experiments in yeast indicated that Ydr527wp associates with Rpb2, Rpb3, and Rpb11 (33). Because Ydr527wp and RPAP1 are both physically and functionally related, the 50-kDa yeast protein encoded by the YDR527W gene was named RBA50 (the name was registered in the *Saccharomyces* Genome Database).

The TAP-tagged RNAPII complex can be recruited to promoters both in vitro and in vivo and can bind to acetylated histone H3. In order to determine whether the TAP-tagged RNAPII complex is functional, we first verified its ability to enter a preinitiation complex in vitro. The electrophoretic mobility shift assay (EMSA) results in Fig. 4A show that the human RNAPII complex can form a preinitiation complex in the presence of TBP, TFIIIB, TFIIIF, and TFIIIE on a radiolabeled probe carrying the AdML promoter. Compared to an equivalent preinitiation complex assembled with highly purified calf thymus RNAPII, the TAP-tagged RNAPII complex had a slightly lower mobility. Because our calf thymus RNAPII contains neither Fcp1 nor RPAP1 (data not shown), our results indicate that the preinitiation complex assembled with the TAP-tagged human RNAPII complex includes at least one of these two polypeptides (Fcp1 and/or RPAP1).

ChIP experiments were used to determine whether the tagged RNAPII complex can also be recruited to promoters in vivo. As shown in Fig. 4B, DNA fragments containing the CCNA1 (cyclin A) promoter (positions –50 to +50) were enriched more efficiently than DNA fragments containing an upstream region (positions –1000 to –900) of the same gene when IgG-Sepharose beads were used to bind the protein A component of the TAP tag. Our results demonstrate that TAP-tagged RNAPII is specifically recruited to the promoter of actively transcribed genes in vivo.

As mentioned above, BLAST searches against publicly available databases revealed weak similarities between the amino acid sequence of RPAP1 and a variety of proteins and domains, including Brd4. Because Brd4 was recently shown to specifically interact with histones H3 and H4 through acety-
FIG. 3. (A) Regions of homology between *Homo sapiens* RPAP1, *D. melanogaster* CG32104-PB, and *S. cerevisiae* Ydr527wp/RBA50, as determined by BLAST analyses. Boxes represent regions with significant alignments. Black boxes, E values ranging from $8 \times 10^{-37}$ to $1 \times 10^{-10}$; open boxes, E values ranging from $4 \times 10^{-5}$ to 0.1. The *M. musculus* and *R. norvegicus* orthologues, with 80% identity with the *H. sapiens* RPAP1 over the full length, were not presented. Reported protein-protein interactions of RPAP1 (this article) and Ydr527wp/RBA50 (33) with transcription factors are indicated on the right. (B) Correlations among mRNA abundance profiles for Tet-promoter mutants of *RPB11*, *YDR527w*, and *RRN3*, measured by using oligonucleotide microarrays. Each mutant was compared to an isogenic control without the tetO7 promoter; log(ratios) are plotted. Tet-RPB11 and Tet-YDR527w were assayed twice, with the replicate cultures being grown, extracted, and assayed on different dates. The correlation of Tet-RPB11 and Tet-YDR527 to a Tet-promoter mutant of *RRN3*, a gene involved in polymerase I transcription, is shown for contrast.
FIG. 4. TAP-tagged human RNAPII assembles on promoter DNA both in vitro and in vivo and binds to acetylated histones. (A) EMSA performed with 150 ng of purified calf thymus (ct) RNAPII or TAP-tagged human RNAPII in the presence of TBP, TFIIB, TFIIF, and TFIIE. A control reaction assembled in the absence of TBP was included. A radiolabeled DNA fragment containing the AdML promoter was used as a probe. (B) ChIP experiments showing that TAP-tagged RNAPII is specifically recruited to the cyclin A promoter in vivo. PCR amplification using sets of primers specific to chromosomal regions either encompassing (−50 to +50) or located upstream of (−1000 to −900) the transcription start site was used on DNA fragments enriched with IgG beads (anti-TAP). The input lanes correspond to DNA that was not subjected to immunoprecipitation. A PCR control that lacks DNA (no DNA) was included in each case. (C) Peptide binding experiment showing that TAP-tagged RNAPII is retained on acetylated histones. The results of three independent experiments were analyzed with ImageQuant software, corrected for the background, and shown relative to unmodified H3. Conventionally purified RNAPII was used as a negative control. The TAP-tagged RNAPII-bound histone tails acetylated at lysine 14 more efficiently than at lysine 9.
lated lysine residues (19), we tested the ability of the complex to bind to acetylated histones. As shown in Fig. 4C, the TAP-tagged RNAPII complex specifically bound to acetylated histone H3 in vitro, while conventionally purified RNAPII was not retained on acetylated histone H3. TAP-tagged RNAPII bound histone tails acetylated at lysine 14 more efficiently than those acetylated at lysine 9. This result suggests that RPAP1 may recruit RNAPII to chromatin through its interaction with acetylated histones.

The TAP-tagged RNAPII complex supports accurately initiated transcription in vitro. TAP-tagged RNAPII was tested for the ability to initiate transcription in vitro by use of a transcription system reconstituted with purified general initiation factors. Purified calf thymus RNAPII or purified TAP-tagged human RNAPII was incubated with a DNA template containing the AdML promoter in the presence of purified TBP, TFIIA, TFIIF, TFIIE, and TFIH. The formation of a 391-nt transcript indicated that the tagged RNAPII complex is competent for promoter-selective transcription in vitro (Fig. 5A).

When the general initiation factors were individually omitted from the transcription reaction (Fig. 5B), transcription was abolished except in the case of recombinant TFIIA or TFIIF. These data fully support our conclusion that TFIIA and TFIIF are components of the tagged RNAPII complex and establish that they are functionally active in this complex.

Purification and functional analysis of a human RNAPII variant with a site-directed mutation in fork loop 1. As part of an effort to understand the fine molecular details of the transcription reaction, we used the TAP procedure to generate and isolate RNAPII variants with site-directed mutations in structural elements predicted to have roles in the transcription reaction according to the crystal structure of the yeast enzyme. Because the Rpb2 element known as fork loop 1 (Fig. 6A) is highly conserved from yeast to humans, absent from bacterial RNAP, and proposed to be involved in maintaining the transcription bubble (17, 30) or in RNA-DNA strand separation (85), we first chose to modify this element. In yeast RNAPII, fork loop 1 amino acids Lys471 and Arg476 appear to contact the RNA phosphates at positions −5, −6, and −7 in the hybrid region. Deletion of the whole 13-amino-acid loop totally abolished the ability of Rpb2 to assemble with the other subunits of RNAPII (data not shown), probably because the deletion destabilized Rpb2, which then became misfolded. A smaller deletion of two amino acids located near the center of the loop (Δ458-459), which corresponded to amino acids Lys471 and Ala472 in the yeast enzyme, was successfully used to purify the 17-subunit RNAPII complex (Fig. 6B). This mutant was named Rpb2 forkl Δ458-459. The silver-stained SDS gel in Fig. 6B shows that the relative abundance of some components of the tagged RNAPII complex varied between the wild type and the mutant, indicating that the mutation may have slightly affected complex assembly. By quantifying the relative intensity of each subunit, we established that the most important difference concerned Rpb1, which was at the very most fivefold less abundant in the mutant than in the wild-type complex.

In order to elucidate the function of fork loop 1 in transcription, we compared Rpb2 forkl Δ458-459 with wild-type RNAPII in several different assays. The assays shown in Fig. 6C to E used fivefold larger amounts of the mutant complex than of the wild type to ensure that the abundance of the mutant complex was not underestimated (see above). As shown in Fig. 6C, Rpb2 forkl Δ458-459 was defective in runoff transcription assays when we used the AdML promoter in the presence of all of the general initiation factors.

In order to delineate with more precision the molecular defect that impairs the transcriptional activity of Rpb2 forkl Δ458-459 in vitro, we tested its activity in a promoter-independent assay. As described previously (29, 79), RNAPII can transcribe from a template carrying a 3′ C extension (C-tailed template) in the absence of the general initiation factors. The template used for Fig. 6D carried a 15-C extension on the 3′ strand. The transcribed strand was devoid of dCMP except at positions +25 to +27 (+1 corresponds to the single-stranded-double-stranded junction). In the absence of CTP, wild-type RNAPII produced a transcript of 35 nt. In the presence of CTP, the major transcript was 60 nt in length, indicating that the polymerase starts transcription 10 nt upstream of the double-stranded junction. In contrast to the RNAPII complex carrying wild-type Rpb2, Rpb2 forkl Δ458-459 was inactive in the C-tailed template elongation assay.

Figure 6E shows that Rpb2 forkl Δ458-459 was also defective in an abortive initiation assay carried out with a short double-stranded oligonucleotide carrying the AdML promoter
in the presence of GTP. A G nucleotide strategically located at +11 in the template blocked transcription at position +10 (Fig. 6E). When we used a fully double-stranded template (0/0), Rpb2 forkl Δ458-459 did not produce abortive transcripts. We rationalized that if Rpb2 forkl Δ458-459 was defective in promoter melting, a premelted template would counteract the defect in the mutant. The premelted template (−9/+2), which carried an unpaired DNA region from −9 to +2 that mimicked the fully melted promoter, supported the synthesis of abortive transcripts, although with a poor efficiency compared to that of the wild type. This finding indicated that a mutation in forkl loop 1 does not simply affect promoter melting and its maintenance, but rather affects the interaction of RNAPII with the melted promoter region.

We finally assessed the ability of Rpb2 forkl Δ458-459 to assemble into a preinitiation complex on promoter-containing DNA. EMSA experiments showed that Rpb2 forkl Δ458-459 can enter a preinitiation complex with the general initiation factors as efficiently as wild-type RNAPII (Fig. 6F).

**Human RNAPII variant with a site-directed mutation in switch 3.** The structure named switch 3 is the only loop of Rpb2 located near the DNA-RNA hybrid that is ordered in the crystal structure of elongating yeast RNAPII (30). Switch 3 makes direct contact with the template DNA near positions −5 (arginine 1122) and −2 (arginine 1129) (Fig. 6A). Since R1129 is juxtaposed with one of the Rpb2 domains that mediate the interaction with Rpb1, we reasoned that disruption of this amino acid would likely affect the assembly of RNAPII and decided to turn our attention to R1122. We produced a mutant in which amino acids R1122, S1123, and R1124 were replaced by alanine residues. In human Rpb2, these amino acids correspond to R1078, S1079, and R1080. This mutant was named Rpb2 sw3-1078.

Figure 7A shows that TAP-tagged Rpb2 with the sw3-1078 mutation is as efficient as the wild type at forming the 17-subunit RNAPII complex. However, EMSA experiments (Fig. 7B) indicated that purified Rpb2 sw3-1078 is not efficient at entering a preinitiation complex on promoter DNA in the presence of the general initiation factors. Increasing the concentration of Rpb2 sw3-1078 by a factor of 6 compared to the wild type was necessary in order to achieve efficient formation of a preinitiation complex. As expected on the basis of the EMSA results, Rpb2 sw3-1078 did not support transcription in vitro (Fig. 7C). Increasing the concentration of Rpb2 sw3-1078 by a factor of 6 compared to the wild type led to the production of runoff transcripts, but the transcription efficiency remained slightly lower than that of the wild type, indicating that the mutation in switch 3 not only affects the ability of Rpb2 sw3-1078 to assemble into a preinitiation complex, but also, to a lesser extent, affects its capacity to support transcription.

**Mutants Rpb2 forkl Δ458-459 and Rpb2 sw3-1078 are defective in vivo.** To assess the defects of Rpb2 forkl Δ458-459 and Rpb2 sw3-1078 in vivo, we used ChIP experiments to evaluate their association with both the promoter and the transcription unit of two genes that are actively transcribed in human cells, the guanine nucleotide binding protein beta polypeptide 2-like 1 (GNB2L1) gene and the ferritin light polypeptide (FTL) gene. As shown in Fig. 8, Rpb2 forkl Δ458-459 is recruited to both promoters with an efficiency similar to that of a wild-type TAP-tagged RNAPII complex, but it is not found in association with the transcription units. In contrast, Rpb2 sw3-1078 is not enriched in the promoter regions or the transcribed regions. The mutant Rpb2 forkl Δ453-465, which carries a complete deletion of the 13-amino-acid loop and does not assemble into a bona fide 17-subunit RNAPII complex, was used as a negative control that was enriched at neither the promoter nor the transcription unit of either the GNB2L1 or FTL gene. These in vivo results fully support our in vitro studies indicating that forkl loop 1 is essential for transcription but is not involved in preinitiation complex assembly, whereas switch 3 is required for efficient preinitiation complex formation.

**DISCUSSION**

**Simple, efficient purification of human RNAPII under native conditions.** Classical approaches for the purification of mammalian RNAPII and its general transcription factors have typically used many different chromatography steps, including ion-exchange chromatography. Eluting these columns with buffers containing high salt concentrations most certainly affected the polypeptide composition of the purified complexes. For example, regulatory subunits, which are often less tightly associated with the enzyme they regulate, may have been lost during the purification procedure. Seraphin and colleagues (71) developed the TAP procedure to allow for the purification of protein complexes under native conditions. Two affinity chromatography steps are used. Elution from an IgG column is performed by a proteolytic treatment with the tobacco etch virus protease while elution from a calmodulin column uses EGTA as a chelating agent. For this paper, we adapted the TAP procedure for the purification of protein complexes from human cells, using an ecdysone-inducible expression system that allows for the expression of the TAP-tagged polypeptides at near-physiological levels. The expression levels were modulated by adjusting the concentration of the inducer, ponasterone A (Fig. 1B). By preventing overexpression of the tagged polypeptides, we limited the formation of nonspecific interactions.

We report here the use of the TAP procedure with human cells for the easy and efficient purification of a complex containing the 12 subunits of RNAPII, both subunits of TFIIF, TFIIB, Fcp1, and RPAP1, a novel polypeptide of unknown function. This 17-subunit complex was purified through the alternate TAP tagging of six of its subunits, demonstrating that all of the components are part of a unique complex. Because (i) the TAP-tagged polypeptides are expressed at near-physiological levels in stably transfected human cells, (ii) protein extraction is performed under gentle conditions, and (iii) the affinity purification steps are achieved under native conditions, our results indicate that this TAP-tagged RNAPII complex is the most abundant soluble, stable form of RNAPII in human cells. If a mediator complex (70) is associated with RNAPII in human cell extracts, either its concentration is much lower than that of RNAPII or its association with RNAPII is too weak to survive double affinity purification.

The presence of the TAP tag at the C termini of subunits of the RNAPII-containing complex does not alter its activity. The TAP-tagged RNAPII complex is fully active in vitro, as it can
form a preinitiation complex with the general initiation factors on the AdML promoter, accurately initiate transcription, and elongate the transcripts (Fig. 4 and 5). In addition, both the TFIIB and TFIIF components of the complex are functionally active in transcription reactions. Notably, TAP-tagged RNApol II can be recruited to transcriptionally active promoters in vivo, an indication that the purified human RNApol II complex is biologically active.
Five different regulatory subunits are stably associated with soluble RNAPII in human cells. The soluble form of RNAPII is expected to be the form of the enzyme that is recruited to promoters upon activation signals. In support of this hypothesis, the CTD of Rpb1 in the TAP-tagged RNAPII complex is in the hypophosphorylated state, the form of the enzyme that is recruited to promoter DNA. A possible function of Fcp1 in this complex could be to maintain the CTD in this unphosphorylated state, as previously proposed (47). Our EMSAs suggested that Fcp1 can enter the preinitiation complex with the RNAPII complex. The entry of Fcp1 with RNAPII at the promoter is supported by the finding by Buratowski and col-
FIG. 7. Purification and functional analysis of TAP-tagged human RNAPII with a mutation in switch 3. (A) Silver-stained SDS gel of wild-type TAP-tagged RNAPII (Rpb2 wt) and a mutant of human RNAPII carrying a triple alanine substitution in the Rpb2 switch 3 (Rpb2 sw3-1078) domain. (B) EMSAs were performed with the AdML promoter with TFIIB, TFIIF, and TFIIE in either the presence or the absence of TBP. The amounts used for the wild type and the mutant are indicated. (C) In vitro transcription reactions (runoff) contained different amounts of Rpb2 wt and Rpb2 sw3-1078 in the presence of TBP, TFIIB, TFIIF, TFIIE, and TFIIH.
port the notion that TFIIF works as a α5β2 heterotetramer in vivo (13, 23).

Notably, both TFIIF and TFIIB can also bind to Fcp1 in a competitive manner and with opposite effects, with TFIIF stimulating and TFIIB inhibiting the phosphatase activity (9). Fcp1 may not efficiently dephosphorylate RNAPII at Ser2 of the CTD in the region downstream of the promoter (10) because TFIIF appears not to be present in this complex (50, 68). Clearly, TFIIB, TFIIF, and Fcp1 participate in an important network of regulatory interactions during RNAPII transcription.

The TAP-tagged RNAPII complex also contains at least one polypeptide that can specifically bind to acetylated histones in vitro. One putative candidate is RPAP1. Whether or not RPAP1 contains a bromodomain remains unclear because the definitive identification of a bromodomain requires structural data. An essential yeast polypeptide with regions of homology with RPAP1 was previously shown to interact with the Rpb10 subunit of RNAPII (28, 43) and also affected global gene expression in the same way as Rpb11 in our Tet-promoter mutant experiments (Fig. 3). These findings support the notion that the association of RPAP1 with RNAPII is physiologically relevant, although its exact role remains to be elucidated.

Structure-guided alignments indicated that RPAP1 could belong to the ARM repeat protein superfamily (12). ARM repeat proteins contain tandem copies of a degenerate 40-amino-acid sequence motif that forms a conserved three-dimensional structure. A single ARM repeat consists of three α helices. Such repeats are found in a variety of proteins, including the mammalian armadillo homologs β-catenin and importin α (40). β-catenin is mainly detected in adherens junctions where it links the membrane-anchored cadherin to actin filament. Cytoplasmic β-catenin is degraded by the ubiquitin-proteasome system. However, when Wnt signaling is activated, cytoplasmic β-catenin is stabilized. The increased level of β-catenin leads to its nuclear translocation, where it affects transcriptional activation. Although it was reported to be a cytoplasmic protein (41), Ydr527wp/RBA50 is found in complex with three RNAPII subunits and appears to affect transcription. Similar to β-catenin, Ydr527wp/RBA50 may have a dual cellular location and may be translocated to the nucleus.

Proteins that were previously shown to interact with RNAPII, such as the mediator and various elongation factors, were not found to be major components of our TAP-tagged RNAPII eluate. This is likely because these RNAPII-interacting proteins are associated with RNAPII in the insoluble fraction and, consequently, are mainly lost during the purification procedure. However, these complexes may represent minor forms of RNAPII complexes in our TAP-tagged eluates.

**Structure-function analysis of the human RNAPII active center.** The crystal structure of eukaryotic RNAPII was used to predict the function of many structural elements found in the active center. For example, some loops and helices that either directly contact or closely approach the DNA-RNA duplex in the active site of the elongating enzyme were suggested to regulate the formation and maintenance of both the transcription bubble and the RNA-DNA hybrid. In order to begin assessing the functions of these structural elements, we expressed TAP-tagged Rpb2 mutants carrying site-directed mutations and purified the mutant forms of RNAPII. Our data
indicate that the assembly and/or stability of the RNAPII complex is affected by mutations in the active site. For example, the full deletion of fork loop 1 led to defects in the assembly of the 17-subunit RNAPIII complex. However, a shorter mutant with a deletion of only two amino acids from Rpb2 was found to assemble correctly in solution in vivo (Fig. 6) and allowed us to analyze the function of this element during the transcription reaction. The Rpb2 fork1 Δ458-459 mutant was found to enter a preinitiation complex with the general initiation factors with an efficiency similar to that of the wild-type complex, but it was incapable of transcription initiation and elongation in vitro. When we used a premelted template in which the transcription bubble was fully open between positions −9 and +2 in our initiation assay, Rpb2 fork1 Δ458-459 was able to initiate transcription formation, but not very efficiently. These findings indicate that there may be a defect at the level of the interaction between fork loop 1 and the melted DNA at very early stages of the transcription reaction. For example, the two-amino-acid-shorter fork loop 1 in the RNAPII mutant may be unable to contact one of the DNA strands in the transcription bubble, leading to incorrect positioning of the DNA template in the active site and impairing the ability of the enzyme to initiate transcription efficiently. Since this mutant supported transcription initiation much less efficiently even when a fully open template was used, our results do not support a direct role for fork loop 1 in promoter melting.

Given the recent crystal structure of elongating yeast RNAPII showing an interaction between Lys 471 (Lys 458 in human RNAPII) and the RNA around positions −5 to −7 (85), an alternative possibility to explain the defect of Rpb2 fork1 Δ458-459 is that during transcription initiation, the mobile fork loop 1 makes crucial contacts with the first arriving ribonucleotides, making it essential for phosphodiester bond formation. According to this speculative view, fork loop 1 would keep contact with the RNA as transcription progresses and be displaced until it reaches the position seen in the crystal structure, where it would participate in RNA-DNA strand separation.

The five switches located at the bottom of the cleft of RNAPII were first proposed to be involved in coupling the closure of the mobile clamp to the presence of a DNA-RNA hybrid (18, 30). This hypothesis was based on the fact that most of the switches are disordered in the crystal structure of the 10-subunit RNAPII complex in solution (e.g., lacking the Rpb4-Rpb7 dimer), whereas they are ordered in the elongating RNAPII structure. The ordering of the switches coincided with the movement of the clamp, which was in an open state in the structure of the free enzyme to permit the entry of DNA into the cleft and in a closed state in the structure of the RNAPII-DNA-RNA ternary complex to ensure the stability of the elongating complex. Since some switches, including switch 3, directly contact the DNA in the DNA-RNA hybrid region, they were proposed to be involved in sensing the presence of nascent RNA and coupling it to the closure of the clamp on the DNA. However, the recent crystal structures of the 12-subunit RNAPII complex (2, 7) revealed that the clamp is in a closed state in the presence of the Rpb4-Rpb7 dimer. These results led the authors to propose that the DNA is first loaded onto the clamp far from the active site in a way that is similar to that of the bacterial RNA polymerase holoenzyme-promoter DNA complex (57) and that it reaches the cleft only after DNA melting has occurred. In this situation, switch 3 could only contact DNA after the opening of the promoter and entry of the DNA into the cleft.

Our results indicate that switch 3 is important for the assembly of RNAPII into a preinitiation complex on promoter DNA. Mutating amino acids 1078 to 1080 clearly reduced the ability of human RNAPII to form a preinitiation complex in the presence of the general transcription factors both in vitro and in vivo. In vitro, sixfold larger amounts of Rpb2 sw3-1078 were necessary to form a complex in gel shift experiments. Our results suggest that switch 3 is necessary for the accurate interaction of the clamp with the promoter DNA before opening of the transcription bubble and that this is the primary defect in our switch 3 mutant. Because the structure of the complete polymerase shows that a duplex DNA cannot bind switch 3 before melting, it is likely that the effect of the switch 3 mutation is either allosteric or mediated by interactions with a general initiation factor such as TFII B or TFII F. The importance of the interaction of switch 3 with the template DNA at positions −2 and −5 after promoter melting might be reflected by the fact that the transcription reactions carried out in the presence of sixfold larger amounts of the sw3-1078 mutant did not yield as many transcripts as did the wild type. This may also suggest that the preinitiation complexes formed with the sw3-1078 mutant are not fully functional, presumably because the interactions of switch 3 with the DNA template in the DNA-RNA hybrid region are impaired.

ACKNOWLEDGMENTS

We are grateful to the members of our laboratory for helpful discussions. We thank Diane Bourque for artwork and Julie Edwards for critical reading of the manuscript. We also thank Takahiro Nagase from the Kazusa DNA Research Institute for kindly providing the cDNA encoding RPAPI.

This work was supported by grants from the Canadian Institutes for Health Research (to B.C.), Genome Canada (to B.C., J.G., and T.R.H.), Genome Québec (to B.C.), the Ontario Genomics Institute (to J.G. and T.R.H.), and the Natural Sciences and Engineering Research Council of Canada (to T.R.H.). C.J. holds a studentship from the Canadian Institutes for Health Research and M.F.L. is supported by the Natural Sciences and Engineering Research Council of Canada and the Fonds Québécois de la Recherche sur la Nature et les Technologies. A.P.D. is supported by a C. H. Best Postdoctoral Fellowship. B.C. is a senior scholar from the Fonds de la Recherche en Santé du Québec.

REFERENCES

1. Acker, J., M. Wintzerith, M. Vigneron, and C. Kedinger. 1992. Primary structure of the second largest subunit of human RNA polymerase II (or B), J. Mol. Biol. 226:1295–1299.
2. Armache, K. J., H. Kettenberger, and P. Cramer. 2003. Architecture of initiation-competent 12-subunit RNA polymerase II. Proc. Natl. Acad. Sci. USA 100:6964–6968.
3. Awrey, D. E., R. G. Weilbaecher, S. A. Hemming, S. M. Ortlicky, C. M. Kane, and A. M. Edwards. 1997. Transcription elongation through DNA arrest sites. A multistep process involving both RNA polymerase II subunit RPB9 and TFII S. J. Biol. Chem. 272:14747–14754.
4. Bagby, S., S. Kim, E. Maldonado, K. I. Tong, D. Reinberg, and M. Ikura. 1995. Solution structure of the C-terminal core domain of human TFII B. similarity to cyclin A and interaction with TATA-binding protein. Cell 82: 857–867.
5. Battist, G. O., and M. Plashe. 2003. Independent recruitment in vivo by Gal4 of two complexes required for transcription. Mol. Cell Biol. 11:1301–1309.
6. Burton, Z. F., L. G. Oertel, and J. Greenblatt. 1986. Proteins that bind to RNA polymerase II are required for accurate initiation of transcription at the adenovirus 2 major late promoter. EMBO J. 5:2923–2930.
7. Bushnell, D. A., and R. D. Kornberg. 2003. Complete, 12-subunit RNA polymerase III.
polymerase II at 4.1A resolution: implications for the initiation of transcription. Proc. Natl. Acad. Sci. USA 100:6969–6973.

8. "Contact C-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

9. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

10. Cho, E. J., M. S. Kohor, M. Kim, J. Greenblatt, and S. Buratowski. 2001. Opposing effects of Ok6 kinase and Fcp1 at yeast promoter. J. Biol. Chem. 276:14962–14969.

11. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

12. Coates, J. C. 2003. Armadillo repeat proteins: beyond the animal kingdom. Trends Cell Biol. 13:463–471.

13. Conaway, J. W., and R. C. Conaway. 1989. A multisubunit transcription factor essential for accurate initiation by RNA polymerase II. J. Biol. Chem. 264:2357–2362.

14. Conaway, J. W., A. N. Kapanidis, H. Tang, D. Reinberg, and R. H. Ebright. 2000. Control of elongation by RNA polymerase II. Trends Biochem. Sci. 25:375–380.

15. Conaway, R. C., F. P. Garvey, J. P. Hanley, and J. W. Conaway. 1991. Mechanism of promoter selection by RNA polymerase II: mammalian transcription factors alpha and beta gamma promote entry of polymerase into the preinitiation complex. Proc. Natl. Acad. Sci. USA 88:6268–6272.

16. Conaway, J. W., and R. C. Conaway. 1991. Architecture of RNA polymerase II and implications for the transcription mechanism. Science 288:640–649.

17. Cramer, P., D. A. Bushnell, J. Fu, A. L. Gnat, B. Maier-Davis, N. E. Thompson, R. J. M. Edwards, P. R. David, and R. D. Kornberg. 1991. Co-translational synthesis of a putative RNA polymerase. EMBO J. 10:357–363.

18. Cramer, P., D. A. Bushnell, and R. D. Kornberg. 2001. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science 292:240–245.

19. Cramer, P., D. A. Bushnell, and R. D. Kornberg. 2001. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science 292:240–245.

20. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

21. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

22. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

23. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

24. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

25. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

26. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

27. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

28. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

29. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

30. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

31. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

32. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.

33. Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus. 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. J. Biol. Chem. 270:14962–14969.
