Subunit architecture of general transcription factor TFIIH

Citation: Gibbons, B. J. et al. "Subunit Architecture of General Transcription Factor TFIIH." Proceedings of the National Academy of Sciences 109.6 (2012): 1949–1954. Copyright ©2012 by the National Academy of Sciences

As Published: http://dx.doi.org/10.1073/pnas.1105266109

Publisher: National Academy of Sciences

Persistent URL: http://hdl.handle.net/1721.1/72183

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Subunit architecture of general transcription factor TFIIH

Brian J. Gibbons[a], Edward J. Brignole[b], Maia Azubel[a], Kenji Murakami[a], Neil R. Voss[b,c], David A. Bushnell[a], Francisco J. Asturias[b], and Roger D. Kornberg[a,1]

[a]Department of Structural Biology, Stanford School of Medicine, Stanford CA 94305; [b]Department of Cell Biology, The Scripps Research Institute, La Jolla CA 92037; and [c]Department of Biological, Chemical and Physical Sciences, Roosevelt University, Chicago IL 60605

Edited by E. Peter Geiduschek, University of California, San Diego, La Jolla, CA, and approved November 30, 2011 (received for review April 4, 2011)

Structures of complete 10-subunit yeast TFIIH and of a nested set of subcomplexes, containing 5, 6, and 7 subunits, have been determined by electron microscopy (EM) and 3D reconstruction. Consistency among all the structures establishes the location of the "minimal core" subunits (Ssl1, Tfb1, Tfb2, Tfb4, and Tfb5), and additional densities can be specifically attributed to Rad3, Ssl2, and the TFIIK trimer. These results can be further interpreted by placement of previous X-ray structures into the additional densities to give a preliminary picture of the RNA polymerase II preinitiation complex. In this picture, the key catalytic components of TFIIH, the Ssl2 ATPase/helicase and the Kin28 protein kinase are in proximity to their targets, downstream promoter DNA and the RNA polymerase C-terminal domain.

cell cycle | DNA repair | xeroderma pigmentosa | tricothiodystropy | Cockayne syndrome

Five general transcription factors (GTFs), termed TFIIB, -D, -E, -F, and -H, assemble with RNA polymerase II (pol II) at a promoter prior to the initiation of transcription (1–3). TFIIH, comprising 10 subunits, with a total mass in excess of 500 kDa, is comparable in size and complexity to the polymerase itself (4). TFIIH is not only required for transcription but also plays central roles in the fundamental processes of DNA repair and cell cycle control (5–7).

Three TFIIH subunits possess catalytic activities, the DNA-dependent ATPase/helicases Ssl2 and Rad3, and the cyclin-dependent protein kinase Kin28, which is associated with the cyclin subunit Ccl1 (8). Ssl2 is responsible for unwinding promoter DNA to form the so-called "transcription bubble" at the active center of pol II engaged in transcription (9–14). Ssl2 and Rad3 are also responsible for DNA unwinding during nucleotide excision repair (NER) (12, 15). Kin28 phosphorylates the C-terminal repeat domain (CTD) of pol II at dozens of sites, in the first of many CTD phosphorylation/dephosphorylation events controlling the interaction of pol II with regulatory, splicing, and cleavage/polyadenylation factors (16–20).

Isolation of TFIIH from yeast yields the complete 10-subunit protein complex (holoTFIIH) and various subcomplexes, including a six-subunit complex of Rad3, Tfb1, Tfb2, Tfb4, Tfb5, and Ssl1 (referred to as core); a seven-subunit complex (designated as core + Ssl2); and a three-subunit kinase module comprising Kin28, Ccl1, and Tfb3 (termed TFIIK). HoloTFIIH is required for transcription, while the core + Ssl2 complex is the form of the protein involved in DNA repair (5, 21). The mammalian counterpart of TFIIK activates the cyclin-dependent protein kinases responsible for cell cycle control (22). Mutations of TFIIH subunits, and consequent deficiencies in transcription and DNA damage repair, are associated with many forms of cancer (23, 24), as well as with disease syndromes, such as xeroderma pigmentosum (XP), tricothiodystrophy (TTD) and, in some patients, the combined symptoms of XP and Cockayne Syndrome (CS) (25, 26).

Structural information on TFIIH has come from EM of single particles of human TFIIH at 38-Å resolution (27), from EM of 2D crystals of the yeast core complex at 18-Å resolution (28), from X-ray crystallography of archaeal counterparts of Ssl2 (11) and Rad3 (29–31) and of human counterparts of CDK7 (32) and cyclinH (33), and from both X-ray and NMR studies of subunit fragments and interaction domains (Table S1). Little is known of the arrangement of subunits or of the relationship of structure to function in transcription or DNA repair. Here we have exploited improved methods for the isolation of TFIIH complexes from yeast and performed EM of single particles to obtain a set of structures from which the subunit organization of holoTFIIH could be derived.

Results

Structural Analysis of HoloTFIIH and Subcomplexes Using Single-Particle EM. Improved procedures for isolation from yeast (*SI Methods*) yielded complete 10-subunit, 532 kDa holoTFIIH and a series of subcomplexes: seven-subunit, 415-kDa core + Ssl2 complex; six-subunit, 320-kDa core complex; and five-subunit, 231-kDa complex lacking both helicases, Rad3 as well as Ssl2 (designated as minimal core) (Fig. 1). All four complexes were large enough for structural study by EM of single particles and 3D reconstruction. All complexes were uniform in composition, with equimolar amounts of subunits, on the basis of SDS/PAGE and staining with Coomassie blue (Fig. 1). EM, however, disclosed heterogeneity, which could arise from both instability and conformational flexibility. Heterogeneity has been dealt with as in the past through sample preparation techniques that maximize stability (34) and through data processing methods for sorting into homogeneous subsets (35). Heterogeneous particles are preserved in stain to give images with a high enough signal-to-noise ratio to differentiate molecular conformations (36) and compositional states.

The 10-subunit holoTFIIH and seven-subunit core + Ssl2 complexes were sedimented through glycerol gradients containing glutaraldehyde cross-linker to enhance stability and minimize compositional heterogeneity (34). For comparison, the core+ Ssl2 complex was prepared in both presence and absence of cross-linker. The core and minimal core complexes were prepared in the absence of cross-linker. Upon adsorption to amorphous carbon support films, all complexes showed strongly preferred orientations, so image alignment and classification could be used to distinguish differences in subunit composition and molecular conformation, without having to simultaneously address differences in orientations. Pairs of images of negatively stained particles were recorded at tilt angles of 0 and −55 degrees, so 3D structures could be calculated by the random conical tilt recon-

Author contributions: B.J.G., E.J.B., M.A., K.M., N.R.V., D.A.B., F.J.A., and R.D.K. designed research; B.J.G., E.J.B., M.A., K.M., N.R.V., and D.A.B. performed research; B.J.G., E.J.B., and M.A. contributed new reagents/analytic tools; B.J.G., E.J.B., M.A., K.M., N.R.V., D.A.B., F.J.A., and R.D.K. analyzed data; and B.J.G., E.J.B., F.J.A., and R.D.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. E-mail: kornberg@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1105266109/-/DCSupplemental.

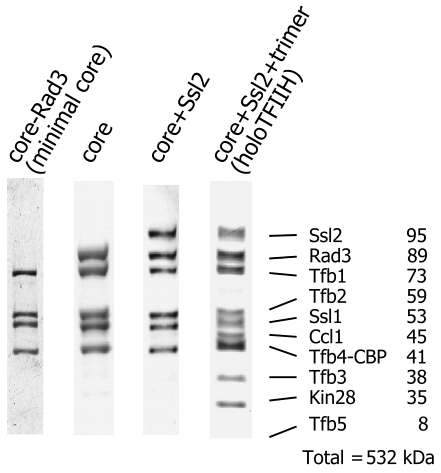

Fig. 1. SDS/PAGE of purified TFIIH complexes. Gels were stained with Coomassie blue and bands identified as TFIIH subunits (subunit names and predicted molecular weights indicated; see Table S3 for mass spectrometry analysis).

struction (RCT) method (37). The resolution of the reconstructions was estimated to be 25–35 Å by the Fourier shell correlation method (38) (Fig. S2).

For each preparation of holoTFIIH and subcomplexes, reference-free alignment, classification, and averaging produced predominant class averages as well as additional minor class averages that differed substantially from one another, likely due to compositional and conformational variability, deviations from preferred orientation, and contamination (Fig. S1). When the predominant class averages and corresponding 3D reconstructions of all four complexes were compared, a striking consistency and clear ordering was observed (Fig. 2). Each complex exhibited the features of that preceding it in the series, with extra density in a well defined region, attributable to the addition of a subunit or subunits to the structure. The locations of Rad3, Ssl2, and TFIIK and their molecular outlines were revealed in this manner.

The unfixed core TFIIH preparation produced a predominant class average that exhibited features of the minimal core preparation with extra density attributable to Rad3 (Fig. 2*B*). The unfixed core + Ssl2 preparation produced four predominant class averages that exhibited features of the core preparation, three with extra density attributable to the Ssl2 subunit in different conformations, suggesting flexibility (Fig. 2*C*). One of the class averages was identical to the class average obtained from the core preparation, due to the loss of Ssl2, suggesting its tendency to dissociate in the absence of cross-linker (Fig. S1*C*). Similarly, two predominant class averages were identified for the fixed core + Ssl2 complexes that varied only in the Ssl2 region, suggesting two distinct conformations of Ssl2 (Fig. 3 *B* and *C* "core + Ssl2 and core + Ssl2*"; class averages in Fig. S1*D*). Just as gradient fixation prevented dissociation of Ssl2 from the core, mild fixation in the holoTFIIH preparation also stabilized the association of TFIIK sufficiently to produce class averages that exhibited the features of the core + Ssl2 preparation with extra density attributable to the TFIIK kinase trimer, enabling 3D reconstruction of the complete 10-subunit TFIIH complex (Fig. 2*D*). Despite fixation, the holoTFIIH preparation also produced some class averages that were identical to the class averages obtained from the core + Ssl2 preparation, due to the loss of TFIIK from many of the particles (Fig. S1*E*). When the datasets were split into larger numbers of classes, small differences among class averages were observed due to flexibility of the connections between Rad3 and the minimal core, between Ssl2 and the core (Fig. 2*C* and Figs. S1 and S3), and between Rad3 and TFIIK (Fig. S3).

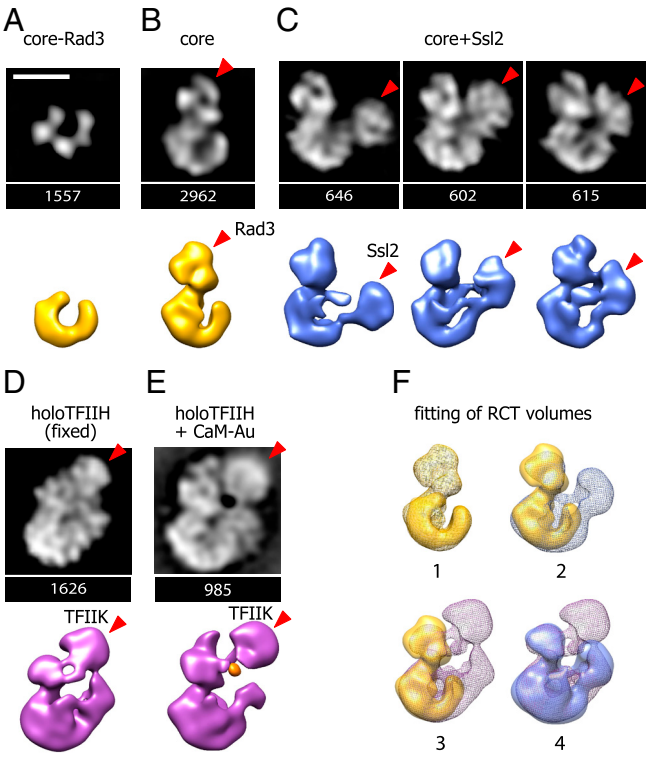

Fig. 2. TFIIH complexes. Class averages (*Upper*) and RCT reconstructions (*Lower*) of TFIIH complexes. (*A*) Minimal core, lacking Rad3. (*B*) Core, with additional density revealing the location of Rad3. (*C*) Unfixed core + Ssl2 data with additional density revealing the location of Ssl2 in several conformations indicating the flexibility of the Ssl2 subunit. (*D*) Fixed holoTFIIH with additional density revealing the location of the TFIIK kinase trimer. (*E*) Fixed holoTFIIH with Tfb3-calmodulin-Au (holoTFIIH+gold), indicating location of the TFIIK kinase trimer. In gold is shown the inverse volume corresponding to the negative density attributed to the gold particle attached to calmodulin bound to the CBP at the C terminus of Tfb3. Density that appears to be a connection between the trimer and the top of Rad3 is presumably due to Tfb3. RCT reconstructions were low-pass filtered to 35 Å and thresholded to match predicted volumes based on molecular weights (Table S4). The number of particles used to produce each class average and 3D reconstruction is indicated. The class average for fixed holoTFIIH has a less distinct appearance possibly due to inferior staining caused by residual glycerol or variability in cross-linking, because this dataset was obtained from a preparation isolated on a glycerol gradient containing glutaraldehyde. The class average for fixed holoTFIIH+gold has an improved appearance due to extensive washing of the grid to remove glycerol prior to staining. Scale bar, 10 nm (applies to all class averages and reconstructions in Figs. 2 and 3). (*F*) Fitting of RCT reconstructions. Panel 1 shows the minimal core volume fitted into the core volume, revealing the extra density for Rad3. Panel 2 shows the core volume fitted into one of the fixed core + Ssl2 volumes, revealing the extra density for Ssl2. Panel 3 shows the core volume fitted into the holoTFIIH volume. Panel 4 shows one of the fixed core + Ssl2 volumes fitted into the holoTFIIH volume, showing the extra density for the TFIIK kinase trimer. Fixed core + Ssl2 volumes are shown in *F* rather than unfixed volumes as in *C* for consistency with the fixed holoTFIIH in *F*.

Calmodulin-Au Labeling and Flexibility of Subunits. To confirm the location of TFIIK, holoTFIIH bearing a calmodulin-binding peptide at the C terminus of Tfb3 was labeled with a calmodulin-gold cluster conjugate. The preparation was sedimented through a glycerol gradient containing cross-linker, and class averages of EM images obtained as described above displayed a dense spot between the regions attributed to Rad3 and TFIIK (Fig. 2*E*). Specificity was established by absence of the dense spot in class averages obtained from control preparations that included EGTA to chelate the calcium and prevent binding of the calmodulin-gold clusters to the calmodulin-binding peptide at the C terminus of Tfb3. These results confirm the presence of Tfb3 and also the

assignment of the adjacent density to TFIIK. There appears to be a connection between TFIIK and the top of Rad3 that is presumably due to Tfb3 (Fig. 2 *D* and *E*). Besides placing the C terminus of the Tfb3 subunit of TFIIK at the interface with Rad3, the class averages revealed alternative conformations, in which TFIIK is rotated as much as 90° around its connection with Rad3, bringing it into proximity with Ssl2 (Fig. S3). Variability in the conformation of Ssl2 is apparent as well (Figs. 2 and 3 and Fig. S3). This is consistent with a proposed mechanism in which the C terminus of XPB rotates to interact with dsDNA (11) (Fig. S3).

Comparison and Docking with Previously Determined Structures. In addition to the structural analysis of holoTFIIH and subcomplexes, and calmodulin-Au labeling, a third line of support for the locations of subunits comes from comparison and docking of previous structures determined by EM and X-ray crystallography to the volumes from EM and 3D reconstruction shown here. The core TFIIH structure determined from 2D crystals (28) is similar in overall size and shape but appears to have captured a different conformation than the core volume reported here (Fig. 3*A*), and proposed locations of subunits by crude volume segmentation of the 2D crystal structure are also consistent with our subunit localizations. Observed differences in shape are likely to reflect either dramatic conformational changes or artifacts of the 2D crystallographic and/or single-particle specimen preparation methods.

The class averages and reconstructed volumes for the core + Ssl2 complex reported here are remarkably similar to those of the previously published human TFIIH structure (27) (Fig. 3*C*). This resemblance between yeast and human structures demonstrates evolutionary conservation and strongly suggests the absence of the kinase module in the human structure. Although it was reported to represent complete TFIIH, the human reconstruction has a volume that is only sufficient to accommodate the yeast core + Ssl2 without the kinase trimer (Fig. 3*C*). It should also be noted that the localization of the subunits within our structures is partly inconsistent with the human TFIIH mapping by antibody labeling (27). While the human Ssl2 homolog, XPB, was localized with antibodies in the same general region identified as Ssl2 in this study, the human Kin28 homolog, cdk7 (component of the kinase trimer) and human Ssl1 homolog, p44 were both localized with antibodies in the region that has been clearly identified as Rad3 in this study, and the Rad3 human homolog, XPD, was localized with antibodies in the region that has been clearly identified as part of the minimal core in this study. The localizations that we report here are likely to be accurate because of the large and obvious density differences between the complexes, while the antibody labeling was limited by the small size of the antibodies and correspondingly small and ambiguous density differences. It is also possible that the epitopes for antibody binding, located at the subunit termini, are distant from the bulk of the subunit mass, leading to the misassignment of density to particular subunits.

X-ray crystal structures for Rad3 homologs from *Sulfolobus tokodaii* (29), *Sulfolobus acidocaldarius* (30), and *Thermoplasma acidophilum* (31) resemble closely the three-lobed density that we have mapped here by comparison of the EM structures of the core and minimal core complexes (Fig. 2, Fig. 3, and Fig. 4). The crystal structure of an Ssl2 homolog in *Archaeoglobus fulgidus* (11) fits well into the extra density that we have mapped here by comparison of the EM structures of the core + Ssl2 and core complexes (Fig. 2, Fig. 3, and Fig. 4). Crystal structures are also available for the Kin28 and the Ccl1 human homologues Cdk7 (32) and cyclinH (33), respectively, which can be accommodated in the extra density that we have mapped here by comparison of the EM structures of the holoTFIIH and core + Ssl2 complexes and that we have assigned to the TFIIK kinase trimer, with space remaining for Tfb3 (Fig. 2, Fig. 3, and Fig. 4).

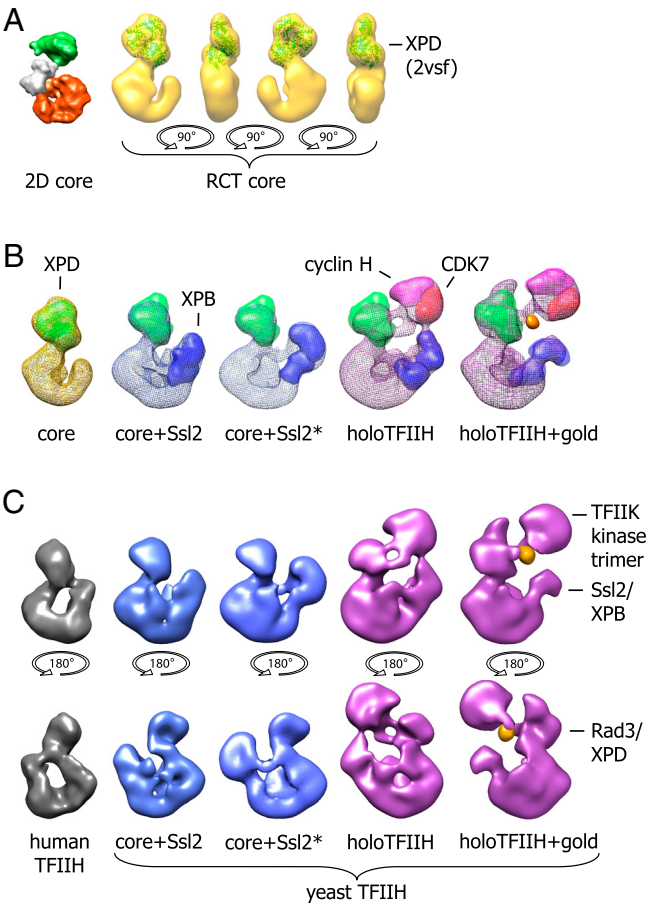

Fig. 3. Comparison of TFIIH reconstructions with previously determined EM and X-ray structures. (*A*) On the left (labeled "2D core") is shown the reconstructed volume of core TFIIH from EM of 2D crystals in stain with speculative designations based on segmentation of volumes for: Rad3 colored green; Ssl1 colored light gray; and Tfb1-4 colored orange (28). On the right (labeled "RCT core") is shown the reconstructed volume of core TFIIH from this work with four different viewing angles (see also Fig. 2). The X-ray crystal structure of XPD (PDB ID code 2VSF) is fitted [with Chimera (65), correlation coefficient = 0.98] to the Rad3 density revealed by the comparison between the core and core-Rad3 (minimal core) reconstructions (Fig. 2). XPD only fits well in one orientation, with the C-terminal region of XPD contacting the predicted location of Ssl1 (Fig. 4*B*). (*B*) Fitting of 20-Å resolution models computed [with Chimera (65)] from X-ray crystal structures. XPD (2VSF), the homolog of Rad3, is shown in green fitted to each of the TFIIH reconstructions (from left to right): core, core + Ssl2, core + Ssl2*, holoTFIIH, and holoTFIIH+gold. The two fixed core + Ssl2 classes suggest the presence of two distinct conformations of Ssl2 (referred to as core + Ssl2 and core + Ssl2*). XPB, the homolog of Ssl2, containing N- and C-terminal domains (2FWR) is shown in blue fitted to the Ssl2 density in the core + Ssl2 reconstruction. The individual N-terminal domain (2FZ4) and C-terminal domain (2FZI) are shown in blue fitted together into the Ssl2 density in the core + Ssl2* reconstruction (see also Fig. S3). Cyclin H (1JKW), the homolog of Ccl1, is shown in magenta; and CDK7 (1UA2), the homolog of Kin28, is shown in red. Both are shown fitted together into the kinase trimer density in the holoTFIIH and holoTFIIH+gold reconstructions. (*C*) Comparison of reconstructed volumes of human TFIIH from EM (27) (gray) to yeast TFIIH from this work. Yeast TFIIH structures from left to right: core + Ssl2 (blue), core + Ssl2* (blue), holoTFIIH (violet), and holoTFIIH+gold (violet). The core + Ssl2 structures show similarity to the human reconstruction but provide additional density, particularly for Ssl2. The holoTFIIH structures reveal the kinase trimer that is absent in the human TFIIH reconstruction.

In the case of Rad3, the X-ray crystal structure may be oriented in the 3D volume on the basis of additional information. The N-terminal region of the human homolog of Rad3, XPD, was shown to interact with the C-terminal region of the human homolog of Tfb3, MAT1, by four-hybrid analysis (18). As already

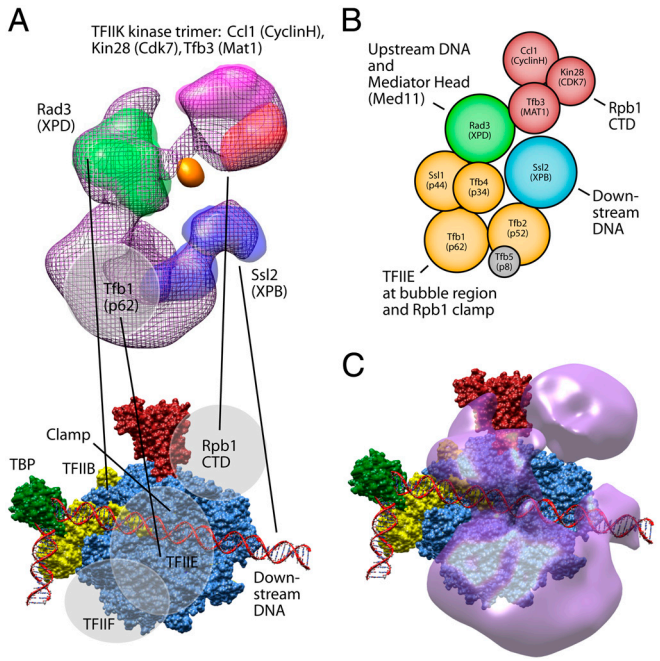

Fig. 4. Models of TFIIH and of the transcription preinitiation complex. (*A*) Interactions of TFIIH with the transcription preinitiation complex, based on structural and biochemical studies (10, 46, 63). The minimal preinitiation complex model (44), comprising RNA polymerase II (blue) with Rbp4/7 (red), TBP (green), TFIIB (yellow), and promoter DNA (red), is a superimposition of 3K7A (Pol II with TFIIB) and 1VOL (TBP and TFIIB with DNA) with the addition of 3K1F (Pol II with TFIIB including Rpb4/7) and further extension of the downstream DNA. Approximate locations of TFIIE and TFIIF are shown with transparent circles highlighting interaction regions identified in cross-linking studies (46, 51, 52, 64). (*B*) Schematic of subunit architecture of TFIIH. Each subunit is represented by a circle with the radius of a sphere corresponding to its molecular weight (Table S4). Arrangement of circles is based on evidence of interaction. Overlaps indicate evidence of interaction from two-hybrid and functional studies. Human homologs are listed in parentheses. (*C*) Preinitiation complex model based on *A*, with TFIIH volume in translucent violet. TFIIH is aligned with the Ssl2 density near the downstream DNA, the TFIIK density near the approximate region of the Rpb1 CTD, and the minimal core density containing Tfb1 near the approximate region of TFIIE. This alignment places the Rad3 density near the transcription bubble formation region and also near the proposed location of the Mediator head module. TFIIH is angled slightly outward from pol II to accommodate both TFIIE and TFIIF (see also Fig. S4).

mentioned, labeling with a calmodulin-gold conjugate placed the C terminus of Tfb3 at the interface with Rad3 (Fig. 2*E*). Other studies have shown that the C-terminal region of XPD interacts with the N-terminal region of the human homolog of Ssl1, p44 (40). Disease-causing mutations in the N- and C-terminal regions of XPD have been shown to destabilize its interactions with MAT1 and p44 and interfere with activity in transcription and DNA repair (18, 39–43). The Rad3 homolog crystal structure could be fitted in the reconstructed volume from EM with its N-terminal region near the C terminus of Tfb3 and its C-terminal region adjacent to the likely location of Ssl1, constraining it to a single orientation in the volume. When fitted in this orientation, both the dimensions and the overall trilobate shape of the crystal structure are remarkably complementary with the appearance of this density in the EM structure (Fig. 3*A*). This result illustrates the bridging of Tfb3 to Ssl1 by Rad3, and highlights the role of Rad3 in anchoring the kinase trimer to the core TFIIH (Fig. 3*B* and Fig. 4*B*).

Discussion

TFIIH is a multisubunit enzyme that plays multiple roles in transcription, DNA repair, and the cell cycle. Its subunit organization underlies these diverse roles. The main findings of this work concern the arrangement of the TFIIH subunits (Fig. 2, Fig. 3, and Fig. 4). From a nested set of subcomplex structures, we have discerned the densities due to individual subunits. This extends our understanding of the organization of TFIIH beyond previous approximate locations of subunits determined by antibody labeling with the use of affinity tags or by volume segmentation. The delineation of subunit densities, in turn, provides a basis for docking previously described X-ray structures of individual subunits. The result is a first approximation to the atomic structure of complete TFIIH.

Fine structural features cannot be determined from the 3D reconstructions alone, due to various limitations: Embedding in stain results in loss of detail and flattening, and the random conical tilt reconstruction method suffers from a missing cone of information. Nevertheless, the fitting of X-ray structures, as well as consistency with previous electron microscopy (27, 28), testify to the validity of the 3D reconstructions (Fig. 3).

Our studies of TFIIH form part of a larger effort to determine the structure of the entire pol II transcription initiation complex (PIC). Toward this end, we combined our 3D reconstruction with structures of pol II and other components previously obtained to give a preliminary picture of the PIC (Fig. 4). In previous work, the X-ray crystal structure of a pol II—TFIIB complex and that of a TBP—TFIIB—TATA box DNA complex were combined by superposition of the TFIIB components (44). The TATA box DNA was extended with straight B-form DNA without steric clash with pol II. The trajectory of the DNA downstream of the TATA box was consistent with that previously reported for PICs (45, 46). We introduced our TFIIH reconstruction in the picture on the basis of known interactions of TFIIH subunits with other components of the PIC (Fig. 4*A*). Ssl2 has been shown by protein–DNA cross-linking to interact with DNA approximately 30 base pairs downstream from the transcription start site (9, 10, 13, 46). We found two orientations of TFIIH that placed Ssl2 at this location and brought the Kin28 protein kinase in proximity to its substrate, the C-terminal domain of the Rpb1 subunit of pol II. Of these two orientations, only one was consistent with the known interaction of Tfb1 (47, 48) and Ssl2 (49, 50) with TFIIE, and the interactions of TFIIE and TFIIF with pol II and DNA in PICs from chemical cross-linking (46, 51, 52, 64). In the resulting picture (Fig. 4*C*), the two ATPase/helicases, Ssl2 and Rad3, bracket the promoter and interact on opposite sides of pol II. Rad3 is in proximity to DNA, near the location of the transcription bubble in a transcribing complex, as has been suggested on the basis of biochemical work (9). Both Rad3 and TFIIK are placed near the proposed location of the Mediator head module (53) (see Fig. S4*B*) consistent with suggested interactions that may be important for regulation of PIC assembly and transcription initiation (54). In this manner, the findings from previous studies were combined with the present results to model the relationship between TFIIH and the other components of the transcription machinery. This model of the PIC (Fig. 4*C*) represents the association of TFIIH with the other components of the PIC prior to strand separation and formation of the "transcription bubble." It is assumed that TFIIE, TFIIF, and, in some cases, TFIIA and TFIIS will also be present. The precise arrangement of components doubtless changes over the course of strand separation and the initiation of transcription.

Materials and Methods

Yeast Strains. *Saccharomyces cerevisiae* protease-deficient strain CB010 was modified by the introduction of tandem affinity purification (TAP) tags at the C termini of Tfb3, Tfb4, and Ssl2 as described (55, 56). Briefly, homologous recombination cassettes were produced by amplifying the coding sequence from pBS1479 for a linker, a calmodulin-binding peptide, a tobacco etch virus (TEV) cleavage site, dual Protein A elements, a stop codon and a Trp selection marker, with primers (from IDT) containing sequences complementary to the C-terminal regions of Tfb3, Tfb4, and Ssl2, as well as sequences for amplifica-

tion from the plasmid (Table S2). An approximately 10-fold increase in the yield of stable holoTFIIH complex was achieved through the deletion of the YOR352W gene product, found to copurify with TFIIH complexes containing Ssl2 (Table S3 and *SI Methods*). YOR352W knockout strains were generated by a variation of the homologous recombination method using disruption cassettes produced by amplifying the coding sequence from pUG6 (57) for a Kan/Geneticin selection marker cassette flanked by loxP sites (for recycling) with primers (from IDT) containing sequences complementary to the upstream and downstream regions of YOR352W as well as sequences for amplification from the plasmid (Table S2). Transformations were performed by the LiOAc method and colonies were tested by PCR with internal and external primers to confirm the insertion locus and orientation, followed by Western blotting. The subunit composition of isolated TFIIH complexes was verified via mass spectrometry (see Table S3).

TFIIH Purification. TAP-tagged HoloTFIIH and TFIIH complexes were isolated from *Saccharomyces cerevisiae* lysates with IgG affinity chromatography followed by additional steps to separate subcomplexes and improve homogeneity. HoloTFIIH and core + Ssl2 complexes were purified via TAP-tag on the Tfb3 subunit using IgG chromatography followed by ion exchange chromatography, size exclusion chromatography, and glycerol gradient centrifugation. Core and minimal core complexes were purified via TAP-tag on the Tfb4 subunit using IgG chromatography followed by hydrophobic interaction chromatography, ion exchange chromatography, and size exclusion chromatography (see *SI Methods* for details).

Calmodulin-Au Labeling. Wheat germ (*Triticum aestivum*) calmodulin (A.G. Scientific, Inc.) was resuspended in labeling buffer (25 mM Hepes pH 6.5, 0.1 mM calcium chloride) and reduced with 1 mM TCEP for 1 h at 37 °C. Reduced calmodulin (5 mg/mL) was reacted 4 h at room temperature (20 to 25 °C) with an equal mass of approximately 1.5 nm diameter monolayer-protected gold particles (approximately 1.5:1.0 molar ratio), prepared as described (58) except with 3-mercaptobenzoic acid instead of 4-mercaptobenzoic acid. The gold particles were rendered inert to further reaction by treatment with 10 mM glutathione for 1 h at 37 °C. The product was purified by electrophoresis in a 12% polyacrylamide gel and was extracted by crushing the gel band and soaking overnight at 4 °C in buffer (20 mM Hepes pH 7.6, 5% glycerol v/v, 1 mM TCEP, and 0.5 mM $CaCl_2$). Aliquots were flash frozen in liquid nitrogen and stored at −80 °C. Calmodulin-Au was added with a three fold molar excess to holoTFIIH complex and incubated for 1 h at 4 °C before gradient centrifugation or size exclusion chromatography.

Specimen Preparation and Electron Microscopy. TFIIH complexes (1–15 mg/mL) were stored until use at −80 °C in 20 mM Hepes, pH 7.6, 250 mM potassium acetate, 1 mM DTT or TCEP, 1 mM EDTA or 0.5 mM $CaCl_2$, and 5% (v/v) glycerol, or 20–30% glycerol if from glycerol gradient. Where indicated, complexes were sedimented through 10–40% glycerol gradients containing 0.1% glutaraldehyde (34) (Fig. S1 and *SI Methods*). Concentrations were adjusted to 10–50 μg/mL in degassed, filtered solution containing 20 mM Hepes, pH 7.6, 250 mM potassium acetate, 1 mM DTT or TCEP, and 1 mM EDTA or 0.5 mM $CaCl_2$, and 2–4 μL were applied to continuous carbon-coated specimen grids [prepared by floating a fresh, thin carbon film onto 300 mesh Cu/Rh grids (Ted Pella or EMS), drying, and glow discharge in the presence of amyl amine immediately before use]. Grids were washed with water, immersed in 1% uranyl acetate solution, lifted through a piece of thin carbon floating on the solution to apply a second carbon layer, blotted, and dried. Untilted and −55° tilt-pair images of the TFIIH complexes were collected manually on SO163 film (Kodak) or automatically on CCD (4 K × 4 K Gatan Ultrascan™ 4000) under low-dose conditions (each exposure approximately 15 $e^- Å^{-2}$) on a Tecnai F20 microscope (FEI) operating at either 120 kV (Scripps) or 200 kV (Stanford). Negatives were digitized with a Leafscan 45 scanner (Leaf Systems) and binned to a final pixel size of 3.0 Å on the object scale; CCD images were recorded with a pixel size of 2.29 Å.

Image Processing. The SPIDER and Web software packages (Version 17) (35) were employed. The defocus of each micrograph was calculated and power spectra were evaluated visually for evidence of drift or astigmatism. Particles from untilted micrographs and CCD frames with estimated defocus between approximately 200 nm and 600 nm were selected for further processing. Tilted micrographs were assessed on an optical diffractometer to ensure that the image was entirely underfocused. For the holoTFIIH and core+Ssl2 datasets, particles in tilt-pair micrographs were selected manually and interactively using TiltPicker (59). For the core dataset, particles in tilt-pair micrographs were picked automatically using a combination of ApDogPicker.py (59) and ApTiltAutoPicker.py (59) (details in *SI Methods*). For the minimal core dataset, tilt-pairs were picked automatically using the "DoG Picking" and "Auto Align Tilt Pairs" routines within the Appion processing pipeline (59). All individual particle images were windowed, ramped, and normalized. Images were then band-pass filtered, retaining information between 21 Å and 330 Å, and a soft-edged circular mask was applied to remove information at the corners of the images. Particles were reference-free aligned and classified using modified routines originally developed for analysis of conformational flexibility of fatty acid synthase (60). Improvements were made to this procedure to allow for mirroring of particles. An initial round of this alignment and classification routine was used to remove particles that were consistently assigned to poorly aligned classes of particles. After removal of these particles, the minimal core, core, core + Ssl2, holoTFIIH, and holoTFIIH-CaM-Au datasets were composed of 7,104, 5,783, 9,112, 6,144, and 9,400 single-particle images, respectively. These images were subjected to a second round of alignment and hierarchical ascendant classification (61), producing 2D class averages. The final in-plane rotational alignment parameters were used to calculate projection angles for the tilted particles to produce random conical tilt (RCT) reconstructions (62). These reconstructions were improved through six iterations of shift refinement to center the tilted particle images, and then band-pass filtered, retaining information between 35 Å and 300 Å.

ACKNOWLEDGMENTS. TEM data were collected at Stanford and also at the National Resource for Automated Molecular Microscopy (NRAMM). NRAMM is supported by the National Institutes of Health (NIH) National Center for Research Resources under grants RR017573 and RR023093. This work was supported in part by National Cancer Institute Grant T32 CA09151 to B.J.G. and by NIH grants AI21144 to R.D.K., R01 GM0670167 to F.J.A., and F32-DK080622 to E.J.B. Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco http://www.cgl.ucsf.edu/chimera (supported by NIH P41 RR-01081).

1. Thomas MC, Chiang CM (2006) The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* 41:105–178.
2. Koleske AJ, Young RA (1994) An RNA polymerase II holoenzyme responsive to activators. *Nature* 368:466–469.
3. Maldonado E, Drapkin R, Reinberg D (1996) Purification of human RNA polymerase II and general transcription factors. *Methods Enzymol* 274:72–100.
4. Ranish JA, et al. (2004) Identification of TFB5, a new component of general transcription and DNA repair factor IIH. *Nat Genet* 36:707–713.
5. Svejstrup JQ, et al. (1995) Different forms of TFIIH for transcription and DNA repair: holo-TFIIH and a nucleotide excision repairosome. *Cell* 80:21–28.
6. Laine JP, Mocquet V, Egly JM (2006) TFIIH enzymatic activities in transcription and nucleotide excision repair. *Methods Enzymol* 408:246–263.
7. Drapkin R, et al. (1994) Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II. *Nature* 368:769–772.
8. Tirode F, Busso D, Coin F, Egly JM (1999) Reconstitution of the transcription factor TFIIH: Assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Mol Cell* 3:87–95.
9. Douziech M, et al. (2000) Mechanism of promoter melting by the xeroderma pigmentosum complementation group B helicase of transcription factor IIH revealed by protein-DNA photo-cross-linking. *Mol Cell Biol* 20:8168–8177.
10. Kim TK, Ebright RH, Reinberg D (2000) Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* 288:1418–1422.
11. Fan L, et al. (2006) Conserved XPB core structure and motifs for DNA unwinding: Implications for pathway selection of transcription or excision repair. *Mol Cell* 22:27–37.
12. Beck BD, Hah DS, Lee SH (2008) XPB and XPD between transcription and DNA repair. *Adv Exp Med Biol* 637:39–46.
13. Spangler L, Wang X, Conaway JW, Conaway RC, Dvir A (2001) TFIIH action in transcription initiation and promoter escape requires distinct regions of downstream promoter DNA. *Proc Natl Acad Sci USA* 98:5544–5549.
14. Bradsher J, Coin F, Egly JM (2000) Distinct roles for the helicases of TFIIH in transcript initiation and promoter escape. *J Biol Chem* 275:2532–2538.
15. Rudolf J, Rouillon C, Schwarz-Linek U, White MF (2010) The helicase XPD unwinds bubble structures and is not stalled by DNA lesions removed by the nucleotide excision repair pathway. *Nucleic Acids Res* 38:931–941.
16. Valay JG, et al. (1995) The KIN28 gene is required both for RNA polymerase II mediated transcription and phosphorylation of the Rpb1p CTD. *J Mol Biol* 249:535–544.
17. Liu Y, et al. (2004) Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. *Mol Cell Biol* 24:1721–1735.
18. Sandrock B, Egly JM (2001) A yeast four-hybrid system identifies Cdk-activating kinase as a regulator of the XPD helicase, a subunit of transcription factor IIH. *J Biol Chem* 276:35328–35333.
19. Rodriguez CR, et al. (2000) Kin28, the TFIIH-associated carboxyl-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. *Mol Cell Biol* 20:104–112.

20. Qiu H, Hu C, Hinnebusch AG (2009) Phosphorylation ofthe Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters. *Mol Cell* 33:752–762.
21. Coin F, et al. (2008) Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Mol Cell* 31:9–20.
22. Kaldis P (1999) The cdk-activating kinase (CAK): from yeast to mammals. *Cell Mol Life Sci* 55:284–296.
23. Wang F, et al. (2008) DNA repair gene XPD polymorphisms and cancer risk: a meta-analysis based on 56 case-control studies. *Cancer Epidemiol Biomarkers Prev* 17:507–517.
24. Manuguerra M, et al. (2006) XRCC3 and XPD/ERCC2 single nucleotide polymorphisms and the risk of cancer: A HuGE review. *Am J Epidemiol* 164:297–302.
25. Lehmann AR (2003) DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *Biochimie* 85:1101–1111.
26. Zurita M, Merino C (2003) The transcriptional complexity of the TFIIH complex. *Trends Genet* 19:578–584.
27. Schultz P, et al. (2000) Molecular structure of human TFIIH. *Cell* 102:599–607.
28. Chang WH, Kornberg RD (2000) Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH. *Cell* 102:609–613.
29. Liu H, et al. (2008) Structure of the DNA repair helicase XPD. *Cell* 133:801–812.
30. Fan L, et al. (2008) XPD helicase structures and activities: Insights into the cancer and aging phenotypes from XPD mutations. *Cell* 133:789–800.
31. Wolski SC, et al. (2008) Crystal structure of the FeS cluster-containing nucleotide excision repair helicase XPD. *PLoS Biol* 6:e149.
32. Lolli G, Lowe ED, Brown NR, Johnson LN (2004) The crystal structure of human CDK7 and its protein recognition properties. *Structure* 12:2067–2079.
33. Kim KK, Chamberlin HM, Morgan DO, Kim SH (1996) Three-dimensional structure of human cyclin H, a positive regulator of the CDK-activating kinase. *Nat Struct Biol* 3:849–855.
34. Kastner B, et al. (2008) GraFix: Sample preparation for single-particle electron cryomicroscopy. *Nat Methods* 5:53–55.
35. Shaikh TR, et al. (2008) SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nat Protoc* 3:1941–1974.
36. Burgess SA, Walker ML, Thirumurugan K, Trinick J, Knight PJ (2004) Use of negative stain and single-particle image processing to explore dynamic properties of flexible macromolecules. *J Struct Biol* 147:247–258.
37. Radermacher M (1988) Three-dimensional reconstruction of single particles from random and nonrandom tilt series. *J Electron Microsc Tech* 9:359–394.
38. van Heel M, Schatz M (2005) Fourier shell correlation threshold criteria. *J Struct Biol* 151:250–262.
39. Lehmann AR (2008) XPD structure reveals its secrets. *DNA Repair (Amst)* 7:1912–1915.
40. Seroz T, Perez C, Bergmann E, Bradsher J, Egly JM (2000) p44/SSL1, the regulatory subunit of the XPD/RAD3 helicase, plays a crucial role in the transcriptional activity of TFIIH. *J Biol Chem* 275:33260–33266.
41. Coin F, Bergmann E, Tremeau-Bravard A, Egly JM (1999) Mutations in XPB and XPD helicases found in xeroderma pigmentosum patients impair the transcription function of TFIIH. *EMBO J* 18:1357–1366.
42. Tremeau-Bravard A, Perez C, Egly JM (2001) A role of the C-terminal part of p44 in the promoter escape activity of transcription factor IIH. *J Biol Chem* 276:27693–27697.
43. Coin F, et al. (1998) Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. *Nat Genet* 20:184–188.
44. Liu X, Bushnell DA, Wang D, Calero G, Kornberg RD (2009) Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. *Science* 327:206–209.
45. Kim TK, et al. (1997) Trajectory of DNA in the RNA polymerase II transcription preinitiation complex. *Proc Natl Acad Sci USA* 94:12268–12273.
46. Miller G, Hahn S (2006) A DNA-tethered cleavage probe reveals the path for promoter DNA in the yeast preinitiation complex. *Nat Struct Mol Biol* 13:603–610.
47. Di Lello P, et al. (2008) p53 and TFIIEalpha share a common binding site on the Tfb1/p62 subunit of TFIIH. *Proc Natl Acad Sci USA* 105:106–111.
48. Okuda M, et al. (2008) Structural insight into the TFIIE-TFIIH interaction: TFIIE and p53 share the binding region on TFIIH. *EMBO J* 27:1161–1171.
49. Okuda M, et al. (2004) A novel zinc finger structure in the large subunit of human general transcription factor TFIIE. *J Biol Chem* 279:51395–51403.
50. Lin YC, Gralla JD (2005) Stimulation of the XPB ATP-dependent helicase by the beta subunit of TFIIE. *Nucleic Acids Res* 33:3072–3081.
51. Chen HT, Warfield L, Hahn S (2007) The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. *Nat Struct Mol Biol* 14:696–703.
52. Forget D, Langelier MF, Therien C, Trinh V, Coulombe B (2004) Photo-cross-linking of a purified preinitiation complex reveals central roles for the RNA polymerase II mobile clamp and TFIIE in initiation mechanisms. *Mol Cell Biol* 24:1122–1131.
53. Cai G, et al. (2010) Mediator head module structure and functional interactions. *Nat Struct Mol Biol* 17:273–279.
54. Esnault C, et al. (2008) Mediator-dependent recruitment of TFIIH modules in preinitiation complex. *Moll Cell* 31:337–346.
55. Puig O, et al. (2001) The tandem affinity purification (TAP) method: A general procedure of protein complex purification. *Methods* 24:218–229.
56. Borggrefe T, Davis R, Bareket-Samish A, Kornberg RD (2001) Quantitation of the RNA polymerase II transcription machinery in yeast. *J Biol Chem* 276:47150–47153.
57. Guldener U, Heck S, Fielder T, Beinhauer J, Hegemann JH (1996) A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res* 24:2519–2524.
58. Ackerson CJ, Jadzinsky PD, Sexton JZ, Bushnell DA, Kornberg RD (2010) Synthesis and bioconjugation of 2 and 3 nm-diameter gold nanoparticles. *Bioconjug Chem* 21:214–218.
59. Voss NR, Yoshioda CK, Radermacher M, Potter CS, Carragher B (2009) DoG Picker and TiltPicker: software tolls to facilitate particle selection in single particle electron microscopy. *J Struct Biol* 166:205–213.
60. Brignole EJ, Smith S, Asturias FJ (2009) Conformational flexibility of metazoan fatty acid synthase enables catalysis. *Nat Struct Mol Biol* 16:190–197.
61. Frank J (1996) Three-Dimensional Electron Microscopy of Macromolecular Assemblies. *Three-Dimensional Electron Microscopy of Macromolecular Assemblie*s, ed J Frank (Oxford Univ Press, New York), 1st Ed, Vol. 1.
62. Radermacher M, Wagenknecht T, Verschoor A, Frank J (1987) Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *J Microsc* 146:113–136.
63. Venters BJ, Pugh BF (2009) A canonical promoter organization of the transcription machinery and its regulators in the Saccharomyces genome. *Genome Res* 19:360–371.
64. Eichner J, Chen H-T, Warfield L, Hahn S (2010) Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. *EMBO J* 29:706–716.
65. Pettersen EF, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 13:1605–1612.