

# How to limit the speed of a motor: the intricate regulation of the XPB ATPase and translocase in TFIIH

Jeannette Kappenberger<sup>†</sup>, Wolfgang Koelmel<sup>†</sup>, Elisabeth Schoenwetter, Tobias Scheuer, Julia Woerner, Jochen Kuper<sup>✉\*</sup> and Caroline Kisker<sup>✉\*</sup>

Rudolf Virchow Center for Integrative and Translational Bioimaging, Institute for Structural Biology, University of Würzburg, 97080 Würzburg, Germany

Received December 19, 2019; Revised September 28, 2020; Editorial Decision September 30, 2020; Accepted November 13, 2020

## ABSTRACT

The superfamily 2 helicase XPB is an integral part of the general transcription factor TFIIH and assumes essential catalytic functions in transcription initiation and nucleotide excision repair. The ATPase activity of XPB is required in both processes. We investigated the interaction network that regulates XPB via the p52 and p8 subunits with functional mutagenesis based on our crystal structure of the p52/p8 complex and current cryo-EM structures. Importantly, we show that XPB's ATPase can be activated either by DNA or by the interaction with the p52/p8 proteins. Intriguingly, we observe that the ATPase activation by p52/p8 is significantly weaker than the activation by DNA and when both p52/p8 and DNA are present, p52/p8 dominates the maximum activation. We therefore define p52/p8 as the master regulator of XPB acting as an activator and speed limiter at the same time. A correlative analysis of the ATPase and translocase activities of XPB shows that XPB only acts as a translocase within the context of complete core TFIIH and that XPA increases the processivity of the translocase complex without altering XPB's ATPase activity. Our data define an intricate network that tightly controls the activity of XPB during transcription and nucleotide excision repair.

## INTRODUCTION

The ATPase activity of the superfamily 2 (SF2) helicase XPB is indispensable for two main cellular processes, transcription and nucleotide excision repair (NER) (1–4). XPB is part of the general transcription factor IIIH (TFIIH) (5–7), which is functionally divided into core TFIIH and the CAK complex. Besides XPB, the core complex contains a second SF2 helicase, XPD, as well as the p62, p52, p44,

p34 and p8 subunits. The CAK (cyclin-dependent-kinase (CDK)-activating kinase) complex completes the TFIIH architecture and consists of MAT1, Cyclin H and the kinase CDK7. Holo-TFIIH orchestrates the three differential enzymatic activities of XPB, XPD and CDK7. The individual roles of the three enzymes present in TFIIH differ depending on the cellular processes that TFIIH is involved in. During RNA polymerase II (RNAPII)-based transcription the ATPase activity of XPB is required for promotor opening, while XPD solely assumes a scaffolding function (6,8–11). Phosphorylation of the C-terminal domain (CTD) of the Rpb1 subunit of RNAPII by CDK7 within the CAK complex promotes the initiation and elongation in transcription (12–15), whereas the same complex assumes an inhibitory effect on the NER activity and was shown to dissociate from core TFIIH prior to the incision of the damaged strand (16–18). In contrast to transcription initiation, both the ATPase activity of XPB and the ATP-dependent helicase activity of XPD are essential for the NER pathway to anchor TFIIH to the site of the damaged DNA and to open the DNA duplex in the context of the lesion (6,8–11). The significance of an intact TFIIH complex for NER is reflected in its association with three severe autosomal recessive disorders. Mutations in the *ERCC3* (XPB) or *ERCC2* (XPD) genes can cause xeroderma pigmentosum (XP), trichothiodystrophy (TTD) and sometimes combined features of XP and Cockayne syndrome (CS) (19–22). No disease mutations in other core TFIIH subunits are known, except for mutations in the *GTF2H5* (p8; also referred to as TTD-A) gene, which are known to cause TTD (23).

So far, all patient mutations found in the *ERCC3* gene are located in domains other than its two RecA like domains (HD1 and HD2), which harbor all seven conserved helicase motifs (22). A crystal structure of an archaeal homolog of XPB from *Archaeoglobus fulgidus* revealed in addition to the helicase domains the presence of a damage recognition domain (DRD) and two unique motifs named the thumb like motif (ThM) and the RED motif (4). The human XPB protein is substantially longer than its archaeal homo-

\*To whom correspondence should be addressed. Tel: +49 931 31 80381; Fax: +49 931 31 76999; Email: caroline.kisker@virchow.uni-wuerzburg.de  
Correspondence may also be addressed to Jochen Kuper. Tel: +49 931 31 80391; Fax: +49 931 31 76999; Email: jochen.kuper@virchow.uni-wuerzburg.de  
<sup>†</sup>The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

logue and contains N- and C-terminal extensions. In order to perform its ATPase activity, XPB is suggested to bind the DNA duplex by bringing its ThM and RED motifs into close proximity to each other (4) and mutations in either motif were shown to exhibit a reduced DNA-dependent ATPase activity comparable to a mutation in the Walker A motif (K346 in human XPB) (9). However, rather than being a ‘true’ helicase, XPB more likely functions as an ATP-dependent DNA-translocase and it was shown that its ATPase activity in transcription is utilized to track along the non-template strand in a 5′-3′ direction (24,25). Analogous to the stimulation of XPD’s helicase activity by p44, the enzymatic activity of XPB is also regulated by its binding partner p52 (6). The importance of the interaction between XPB and p52 is demonstrated by the severe effects of an XPB mutation (F99S) found in XP/CS patients. The F99S mutation was shown to weaken the interaction between XPB and p52, resulting in a highly reduced ATPase activity of XPB (6). However, so far, it is not known how this regulation is achieved. The lack of structural information further hindered an understanding of the functional interaction between XPB and p52. Recently, cryo-EM structures of the human and yeast TFIIH complex at resolutions between 3.5 and 4.7 Å (26–29) were solved and shed light on the intricate interactions within TFIIH. The network between XPB and p52 is expanded by interactions to p8 and in the human TFIIH model it was shown that L21 in p8, which causes TTD when mutated to a proline, is located at the HD2 interface of XPB (26). How the intricate network of protein interactions regulates XPB activity is, however, poorly understood so far.

In this study, we functionally dissected the regulatory events in XPB activation utilizing proteins from the eukaryotic fungus *Chaetomium thermophilum* as a model system. We solved the crystal structure of a p52/p8 complex and used this structure in combination with the current cryo-EM models to generate point mutations for the investigation of crucial p52/p8 interactions with XPB. We show that p52 not only acts as an activator of XPB but that also p8 in combination with p52 directly stimulates XPB’s ATPase activity further. Intriguingly, our data demonstrate that p52/p8 not only function as activators but at the same time repress the maximum activity that can be achieved when XPB is stimulated by DNA. This speed limiting effect can be observed in the absence of other core TFIIH components but also in the context of complete core TFIIH. Finally, we show that XPB’s ATPase activity does not directly lead to the translocase function of XPB. XPB requires the presence of all core TFIIH components to act as a translocase and the addition of XPA increases the processivity of the TFIIH translocase but not its ATPase activity. Our results suggest that XPB’s activity is tightly regulated at different levels to fine-tune its activity for the different processes in NER and transcription.

## MATERIALS AND METHODS

### Expression and purification

The genes encoding ctXPB, ctXPD, ctp62, ctp52, ctp44, ctp34, ctp8 and ctXPA were cloned from cDNA from *C. thermophilum*. The cDNA sequence of ctXPB was

codon-optimized for expression in *Escherichia coli* (ATG:biosynthetics). CtXPB was inserted into both, the pBADM-11 vector (EMBL) containing an N-terminal hexa-histidine tag with a TEV cleavage site, and the pFast-Bac vector (Invitrogen) containing a C-terminal Twin-Strep tag and a 10×-histidine tag (The latter ctXPB was only used for the pull-down assays). Ctp44, ctp34, ctp8 and ctXPA were inserted individually into the pBADM-11 vector containing an N-terminal hexa-histidine tag with a TEV cleavage site. CtXPD was inserted into the pBADM-11 vector with or without an N-terminal hexa-histidine tag. Ctp62 and ctp52 were each inserted into the pETM-11 vector (EMBL) without tag. The constructs ctp52.1–321, ctp52.121–514 and ctXPB.60–345 were cloned utilizing the full length cDNA sequences via sequence and ligation independent cloning (30). For ctp52.121–514 a 28 amino acid long linker region of ctp52 (amino acid 322–349), that is not present in human p52, was replaced by a short linker (sequence SNGNG) via sequence and ligation independent cloning (30). The Walker A variants ctXPD\_K48R and ctXPB\_K392R as well as all ctp52 variants were generated via site directed mutagenesis (31). CtXPB from the pFast-Bac vector was expressed utilizing the baculovirus system. Bacmids were prepared in SF21 cells in EX-CELL 420 medium (Sigma-Aldrich) at 27°C. The cell culture medium containing the viruses was harvested after approximately 72 h. Hi5 cells were grown in EX-CELL 405 medium (Sigma-Aldrich) at 27°C until a density of  $0.5 \times 10^6$  cells/ml was reached followed by transfection with 10% (v/v) virus containing medium. Protein expression was allowed for 72 h at 27°C. CtXPB from the pBADM-11 vector was expressed in Rosetta™ 2 (DE3) cells (Merck Millipore). CtXPD was expressed in ArcticExpress (DE3) RIL cells (Agilent). Ctp62/ctp44, ctp52/ctp8 as well as ctp52/ctp34 were co-expressed in BL21 CodonPlus (DE3) RIL cells (Agilent). Ctp8, ctp52.1–321, ctp52.121–514 and ctXPB.60–345 were separately expressed in BL21 CodonPlus (DE3) RIL cells. CtXPA was expressed in LOBSTR BL21 (DE3) RIL cells (Kerastast). The *E. coli* cells were grown at 37°C in either Lennox broth (XPB, p8, XPB.60–345, p52/p8, p52/p34, p52.121–514) or Terrific broth (XPB, XPD, p62/p44) medium (Carl Roth) to reach a final OD<sub>600</sub> of 0.6 or 1.2, respectively. For the seleno-methionine containing variants of ctp52.1–321 and ctp52.121–514 the cells were grown in M9 minimal medium supplemented with L-seleno-methionine at 37°C to a final OD<sub>600</sub> of 0.6 (32). When the final OD<sub>600</sub> was reached, protein expression was induced by addition of 0.5 mM IPTG for pETM-11 vectors or 3.3 mM arabinose for pBADM-11 vectors accompanied by a temperature reduction to 15°C and protein expression was allowed overnight. For expression of XPD, the temperature was reduced to 30°C once an OD<sub>600</sub> of 0.6 was reached and the cells were allowed to grow until an OD<sub>600</sub> of 1.2 was reached followed by induction and protein expression at 11°C.

For purification of the ctXPD/ctp44/ctp62 complex, ctp44 and ctp62 were co-purified via immobilized metal affinity chromatography (IMAC) using Ni IDA beads (Macherey-Nagel) followed by size exclusion chromatography (SEC) via a HiLoad 16/600 Superdex 200 prep grade column (Cytiva) with 20 mM Hepes pH 7.5, 250 mM

NaCl, and 1 mM TCEP. The elution fractions containing ctp44/ctp62 were pooled and directly added to the cleared lysate of ctXPD. This mixture was subjected to another IMAC, followed by SEC and anion exchange chromatography (AEC). For SEC, the same column and buffer as above was used. For AEC the mixture was applied to a MonoQ 5/50 GL column (Cytiva), with buffers containing 20 mM Hepes pH 7.5, 50/1000 mM NaCl, and 1 mM TCEP. To obtain the ctp44/ctp62 complex in the absence of ctXPD, ctp44 and ctp62 were purified via IMAC and SEC as described above, followed by AEC using a MonoQ 5/50 GL column with the same buffers as for the ctXPD/ctp44/ctp62 complex. To obtain ctXPD, the protein was purified via IMAC, followed by SEC and AEC. The SEC buffer contained 20 mM Hepes pH 7.5, 150 mM NaCl, 5 mM MgCl<sub>2</sub>, and 1 mM TCEP. The AEC buffers contained 20 mM HEPES pH 7.5, 80/1000 mM NaCl, 5 mM MgCl<sub>2</sub>, and 1 mM TCEP. CtXPB (Hi5 cells) was purified using a 5 ml StrepTrap HP column (Cytiva) followed by SEC using a HiLoad 16/600 Superdex 200 prep grade column. For SEC, the buffer contained 20 mM Hepes pH 8.0, 200 mM NaCl and 1 mM TCEP. CtXPB (expressed in *E. coli*), ctXPB.60–345, ctp52/ctp34, ctp52/ctp8, ctp52.1–321, ctp52.121–514 and ctp8 were purified via IMAC (Ni TED or Ni IDA, Macherey-Nagel) and SEC using a HiLoad 16/600 Superdex 200 prep grade column (Cytiva). For ctXPB and ctp52/ctp34, the SEC buffer contained 20 mM HEPES pH 7.5, 250 mM NaCl, and 1 mM TCEP. For ctXPB.60–345, ctp52.1–321 and ctp52.121–514, the SEC buffer contained 20 mM Tris–HCl pH 7.5 and 250 mM NaCl. For ctp52/ctp8 and ctp8 the SEC buffer contained 20 mM Hepes pH 8 and 375 mM NaCl. All variants were expressed and purified as described for the wild type proteins. In case of ctp8, two versions were used. For the co-crystallization with ctp52.121–514, ctp8 containing an N-terminal hexa-histidine tag and a TEV cleavage site was used. For the other studies, the tag was removed by incubation with TEV protease after IMAC and the protein solution was subjected to another IMAC and the protein solution was subjected to another IMAC to remove the TEV protease and uncut protein prior to SEC. CtXPA was purified via IMAC and AEC using a MonoQ 5/50 GL column with buffers containing 20 mM Tris–HCl pH 8, 50/1000 mM NaCl and 2 mM TCEP.

All proteins were concentrated to 50–1000  $\mu$ M, flash frozen in liquid nitrogen, and stored at  $-80^{\circ}\text{C}$ .

### Crystallization

Crystallization was performed via the vapour diffusion method at  $20^{\circ}\text{C}$ . For ctp52.1–321, a seleno-methionine derivative was crystallized at concentrations of 4–7 mg/ml. The reservoir solution consisted of 100 mM calcium acetate, 12.5% (w/v) PEG 8000, and 100 mM HEPES pH 7.5. For ctp52.121–514, a seleno-methionine derivative was crystallized at protein concentrations of 5–6 mg/ml. The reservoir solution consisted of 7.5% (w/v) PEG 4000 and 100 mM HEPES pH 7.0. For the ctp52.121–514/ctp8 complex, ctp52.121–514 at concentrations of 5–6 mg/ml was incubated for 1 h at  $4^{\circ}\text{C}$  with a twofold excess of ctp8 prior to crystallization. The reservoir solution consisted of 150 mM

ammonium chloride, 15% (w/v) PEG 4000, and 100 mM Tris–HCl pH 8.5.

### Structure solution and refinement

X-ray diffraction data were processed with XDS (33). Diffraction data for the ctp52.121–514/ctp8 complex were anisotropy-corrected using the STARANISO server (34). Initial phases were obtained using SHARP (35) utilizing the anomalous signal of the dataset from the seleno-methionine derivative of ctp52.121–514. The resulting electron density map was used to build an initial model of ctp52.121–514. This model was used to solve the phase problem for the ctp52.1–321 and ctp52.121–514/ctp8 datasets via molecular replacement with Phaser (36). The models for ctp52.1–321 and ctp52.121–514/ctp8 were manually completed and corrected with Coot (37). For the latter, the available structure of a Tfb5/Tfb2 minimal complex (PDB code: 3DOM), was used as additional template for manual model building. The ctp52.1–321 structure was refined with REFMAC5 (38) using automated twin refinement. The ctp52.121–514/ctp8 structure was refined with BUSTER (39).

### ATPase assay

CtXPB's ATPase activity was quantified utilizing an *in vitro* ATPase assay in which ATP consumption is coupled to the oxidation of NADH via pyruvate kinase and lactate dehydrogenase activities. Activities were measured at  $30^{\circ}\text{C}$  in 100  $\mu$ l solution containing 1.3 U pyruvate kinase, 1.9 U lactate dehydrogenase (PK/LDH, Sigma), 1.6 mM phosphoenolpyruvate, 0.24 mM NADH, 10 mM KCl, 2.5 mM MgCl<sub>2</sub>, 1 mM TCEP, and 20 mM Tris–HCl pH 8. The assay was carried out under saturating concentrations of ATP using ctXPB wild type and ctXPB.K392R at a concentration of 250 nM. Ctp52, ctp52 variants, ctp8 and ctp34 were added in a twofold molar excess to ctXPB. When core ctTFIIH was assembled, all subunits were present at an equimolar ratio and a final concentration of 250 nM. CtXPA was added in a 4:1 stoichiometric ratio to ctXPB or core ctTFIIH. If not stated otherwise, double stranded (dsDNA) (fw: 5'-AGCTACCATGCCTGCACGAATTAAGCAATTCGTAATCATGGTCATAGC-3'; rv: 5'-GCTATGACCATGATTACGAATTGCTTAATTCGTGCAGGCATGGTAGCT-3') was used at a final concentration of 125 nM for ctXPB or 1  $\mu$ M for core ctTFIIH. The sample was preincubated until a stable base line was achieved. Enzyme catalysis was initiated by the addition of 2.5 mM ATP. The activity profiles were measured at 340 nm using a Clariostar plate reader (BMG Labtech) and 384-well F-bottom  $\mu$ Clear™ microplates (Greiner Bio-One). Initial velocities were recorded and ATP consumption was determined using the molar extinction coefficient of NADH. Curves were fitted with GraphPad Prism. All measurements were carried out at least in triplicates, using at least two independently purified batches of the two helicases XPB and XPD as well as their Walker A variants. Example time courses are shown in Supplementary Figure S1A. SDS PAGE analyses of the complexes used in this study are shown in Supplementary Figure S2.

### Triplex displacement assay

Core ctTFIIH dsDNA-translocase activity was detected using a well established triplex disruption assay (25,28). DsDNA-translocase activity was measured by displacement of a fluorescently labeled triplex forming oligonucleotide (TFO) from a triple helix DNA substrate. The assay was carried out at 30°C in 50 µl solution containing 120 mM KCl, 8 mM MgCl<sub>2</sub>, 16 mM HEPES-KOH pH 8, 4% (v/v) glycerol, 0.8 mM TCEP, 2.2 mM phosphoenolpyruvate, 1.8 U pyruvate kinase (Sigma-Aldrich) and 150 nM triplex DNA. The triplex DNA consists of a 52 bp dsDNA strand with a 5' dabcy label (fw: 5'-GTCTTCTTTTAAACACTATCTTCCTGCTCATTTCTTTCTTCTTTCTTTTCTT-3'; rv: 5'-Dabcy-AAGAAAAGAAAGAAGAAAGAAATGAGCAGGAAGATAGTGTTTAAAGAAGAC-3') and a 5' Cy3-tagged TFO (5'-Cy3-TTCTTTCTTCTTCTTCTTTCTTT-3'). The triplex DNA was annealed as previously described (28). Correct triplex formation was confirmed via native PAGE. The baseline was recorded for 10 to 15 minutes prior to addition of 2 mM ATP. TFO displacement was measured for 60 min at an excitation wavelength of 520–540 nm and an emission wavelength of 590–620 nm with a gain of 1900 using a Clariostar plate reader (BMG LABTECH) in 384-well F-bottom FLUOTRAC™ high binding microplates (Greiner Bio-One). When core ctTFIIH was assembled, all subunits were present in equimolar amounts with a final concentration of 500 nM. CtXPA was added at a fourfold molar excess to core ctTFIIH or ctXPB. For the minimal XPB complex, ctp52, ctp8 and ctp34 were added at a twofold molar excess to ctXPB (500 nM). The release of Cy3-tagged TFO corresponds to the slope of increasing fluorescence. Curves were fitted with GraphPad Prism. All measurements were carried out in at least eleven replicates using two independently purified batches of the two helicases XPB and XPD as well as their Walker A variants. Example time courses are shown in Supplementary Figure S1B. SDS PAGE analyses of the complexes used in this study are shown in Supplementary Figure S2.

### Pull down assay

10 µl Strep-Tactin Sepharose beads (IBA) were equilibrated with 100 µl buffer containing 150 mM NaCl, 1 mM TCEP and 20 mM HEPES pH 7.5. 2 µM Twin-Strep-tagged ctXPB was added and incubated for 45 minutes with the beads. 16 µM ctp52/ctp8 or ctp52.E359K/ctp8 were added and incubated for another 45 minutes. As control, 16 µM ctp52/ctp8 was incubated with the beads in the absence of ctXPB. Unbound protein was washed off with buffer and bound protein was eluted in 20 µl buffer supplemented with SDS loading dye and boiled for 5 min at 95°C. Elution fractions were analysed by SDS PAGE. The intensities of the ctp52/p8 gel bands were quantified using ImageJ and graphs were generated via GraphPad Prism. Three technical replicates were performed.

### Fluorescence polarization measurements

DNA binding was analysed by fluorescence polarization employing a duplex DNA with a Cy3 la-

bel (fw: 5'-Cy3-AGCTACCATGCCTGCACGAATTAAGCAATTCGTAATCATGGTCATAGC-3' rv: 5'-GCTATGACCATGATTACGAATTGCTTAATTCGTGCAGGCATGGTAGCT-3'). Assays were carried out in 50 µl solution with 20 mM HEPES pH 7.5, 10 mM KCl, 2.5 mM MgCl<sub>2</sub>, 1 mM TCEP, and 6 nM DNA at room temperature. CtXPB was used at concentrations of 0.5–500 nM as indicated. Ctp52/ctp8 was added at a 2:1 stoichiometric ratio to ctXPB. Fluorescence polarization was detected at an excitation wavelength of 540 nm and an emission wavelength of 590 nm with a Clariostar plate reader (BMG LABTECH) using 384-well F-bottom Fluotrac high binding microplates (Greiner Bio-One). The gain was adjusted to a well containing buffer and DNA only. Curves were fitted with GraphPad Prism. Three technical replicates were performed.

### CD spectroscopy

Circular dichroism (CD) spectra were recorded under constant nitrogen flush using a Jasco J-810 spectropolarimeter over 190–260 nm. Measurements were performed at 20°C in eight scans with a speed of 50 nm/min and a band width of 2 nm. The proteins were diluted to 4 µM in a buffer containing 19 mM dipotassium phosphate and 1.2 mM monopotassium phosphate at pH 8. All samples were centrifuged for 20 minutes at 30 000 g prior to the measurement. The buffer spectrum was used as a baseline and subtracted from all protein spectra.

### Native PAGE

Native PAGE was performed using Tris/glycine gels consisting of 12.5 mM Tris, 96 mM glycine pH 8.9, 12% (v/v) ROTIPHORESE® Gel 30 (37.5:1) (Roth), 0.07% (w/v) APS, and 0.1% (v/v) TEMED. All samples contained 10 µM protein in 4 mM Tris, 16 mM HEPES pH 8, 350 mM NaCl. The samples were incubated on ice for 1 h, supplemented with 5× loading dye (0.08% (w/v) Ponceau S and 50% (v/v) glycerol) and loaded on the gel. Electrophoresis was performed at 4°C in a buffer containing 12.5 mM Tris and 96 mM glycine pH 8.9 for 90 min at 100 V followed by 3:45 h at 200 V. Protein bands were visualized by Coomassie G-250 staining.

To confirm correct triplex DNA annealing, native PAGE was performed using Tris-acetate gels consisting of 40 mM Tris pH 5.5, 40 mM acetate, 5 mM magnesium acetate, 1 mM MgCl<sub>2</sub>, 8% (v/v) ROTIPHORESE® Gel 40 (29:1) (Roth), 0.1% (w/v) APS and 0.1% (v/v) TEMED. All samples contained 1 µM triplex DNA supplemented with 5× loading dye. Electrophoresis was performed in a buffer containing 40 mM Tris pH 5.5, 40 mM acetate, 5 mM magnesium acetate, and 1 mM MgCl<sub>2</sub> for 55 min at 100 V. DNA was visualized using a Pharos FX™ Plus imager (Bio-Rad).

## RESULTS

### The ctp52/ctp8 structure

Several cryo-EM structures of TFIIH have recently been published shedding light on the molecular architecture

**Table 1.** Data collection and refinement statistics

	p52.1–321	p52.121–514/p8	p52.121–514 (ano)
<b>Data collection</b>			
Wavelength (Å)	0.97973	0.96770	0.97977
Spacegroup	<i>P</i> 3 <sub>1</sub>	<i>P</i> 2 <sub>1</sub>	<i>P</i> 1
Unit cell parameters			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	104.3, 104.3, 165.0	74.2, 86.0, 92.2	60.5, 83.2, 86.0
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 120	90, 94.9, 90	82.7, 80.0, 77.5
Resolution range (Å)	46.97–2.80 (2.89–2.80)	19.92–2.68 (2.95–2.68)	49.16–2.60 (2.69–2.60)
Unique reflections	49 518 (4628)	17 307 (865)	48 667 (4455)
<i>R</i> <sub>merge</sub> (%)	16.2 (282.9)	28.0 (142.9)	20.1 (134.2)
<i>R</i> <sub>pim</sub> (%)	5.3 (90.1)	14.0 (71.5)	8.1 (53.9)
<i>I</i> / $\sigma$ <i>I</i>	10.7 (1.0)	5.7 (1.2)	9.4 (1.5)
CC (1/2)	0.997 (0.325)	0.983 (0.318)	0.995 (0.494)
Multiplicity	10.4 (10.8)	4.9 (4.8)	7.0 (7.1)
Completeness (%)			
spherical	100.0 (100.0)	53.4 (10.8)	98.7 (98.0)
ellipsoidal	-	88.7 (49.5)	-
<b>Refinement</b>			
Resolution range (Å)	46.97–2.80	19.92–2.68	
Unique reflections	46 992	17 099	
Number of atoms	8572	5606	
<i>R</i> <sub>work</sub> (%)	18.6	21.6	
<i>R</i> <sub>free</sub> (%)	20.9	24.0	
Mean <i>B</i> -factor (Å <sup>2</sup> )	88.9	53.2	
RMS deviations			
Bond lengths (Å)	0.0054	0.008	
Bond angles (°)	1.4375	0.95	
Ramachandran statistics (%)			
Favored	81.17	95.20	
Allowed	15.96	3.93	
Outliers	2.87	0.87	

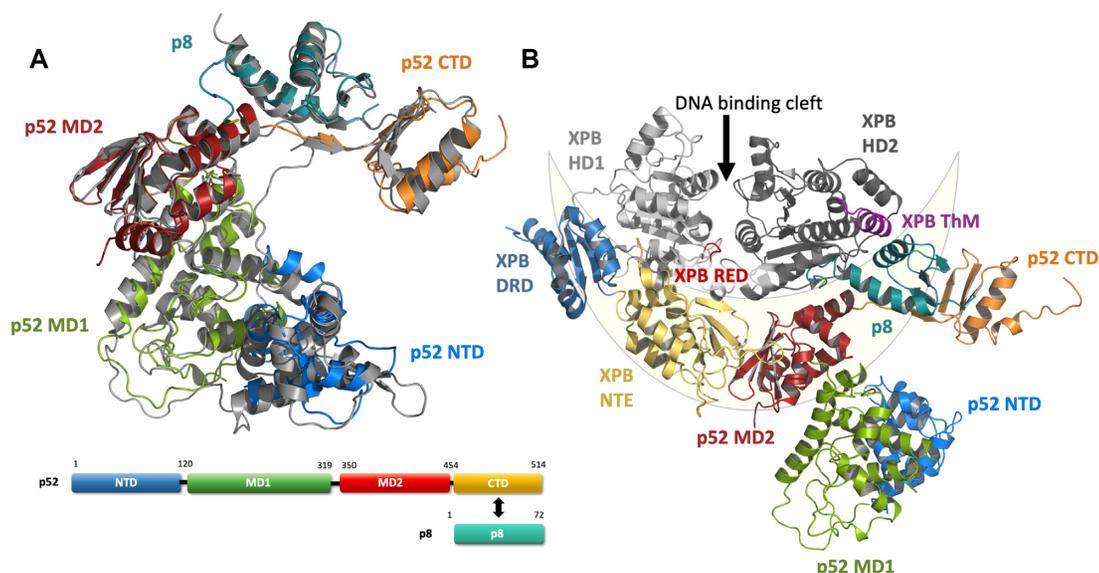
Values in parentheses correspond to highest resolution shell.

of the core TFIID components including the important p52/p8 complex (26–29). However, no full atomic model of p52, the major regulator of the XPB enzyme, alone or in complex with p8 beyond a resolution of 3.5 Å was available so far. We pursued the structural characterization of p52 from the fungal model organism *Chaetomium thermophilum* (ctp52) using x-ray crystallography. Ctp52 shares 35% and 39% sequence identity with the human p52 and yeast Tfb2 sequences, respectively (Supplementary Figure S3). We solved two crystal structures of ctp52. The first structure encompasses amino acids 1–321 (ctp52.1–321) and was solved at a resolution of 2.8 Å (Table 1, Supplementary Figure S4A). The second structure encompasses the C-terminal part starting at amino acid 121 (ctp52.121–514) and was solved in complex with ctp8 at a resolution of 2.7 Å (Table 1, Supplementary Figure S4B). In the latter structure a 28 amino acid long linker region of ctp52 (amino acids (aa) 322–349), that is not present in human p52, was replaced by a short linker (sequence SNGNG). Due to the overlap of both structures, we were able to generate a full structural model of ctp52 at atomic resolution (Figure 1A, Supplementary Figure S4C). The model shows that ctp52 is organized into four distinct domains, an N-terminal domain (NTD; aa 1–120), two middle domains (MD1; aa 121–319 and MD2; aa 350–454), and a C-terminal domain (CTD; aa 455–514), that interacts with p8 (40). The linker region that was replaced in ctp52.121–514, bridges MD1 with MD2 and could not be observed in our structure due to its flexibility. The ctp52.1–321 model comprises the NTD

and MD1 domains and the ctp52.121–514 model the MD1, MD2, and CTD domains. Both crystal structures were superimposed via the MD1 domain (rmsd of 0.87 Å) thereby generating the full-length ctp52 model.

The ctp52 structure is predominantly  $\alpha$ -helical and consists of 22  $\alpha$ -helices and 15  $\beta$ -strands (Figure 1A, Supplementary Figure S4C). None of the newly identified domains present in ctp52 (NTD, MD1 and MD2) displays significant structural homologies to any known protein domains. The interaction between ctp52 and ctp8 is mediated by  $\beta$ -strands 12–15 of p52 and  $\beta$ -strands 1 to 3 of p8 and can be readily compared to the crystal structure of the minimal yeast Tfb5/Tfb2 complex (40) (PDB code: 3DOM) with an rmsd of 1.10 Å for the CTD of ctp52 and an rmsd of 0.97 Å for ctp8 (PDBsum).

A superposition of the ctp52/ctp8 structure with the cryo-EM model of human p52/p8 within TFIID (26) shows that the MD2 domain was found to be oriented differently compared to our model (Supplementary Figure S4D). While the antiparallel  $\beta$ -sheet of MD2 ( $\beta$ -strands 8–11) is pointing towards MD1 and is involved in the formation of the dimer interface in our ctp52.121–514/ctp8 structure (Supplementary Figure S4B), it is pointing towards the N-terminal extension (NTE) of XPB in the TFIID model (Figure 1B). This is most likely due to the different conditions and interactions in the crystallization experiments compared to the cryo-EM experiment. However, after the rearrangement of the individual ctp52 domains according to the model of human p52/p8 embedded in



**Figure 1.** Structure and domain architecture of ctp52 and ctp8. (A) Structure and domain architecture of full length ctp52 and ctp8. Ctp52 is comprised of four domains: NTD (blue), MD1 (green), MD2 (red) and CTD (orange). Ctp8 is shown in mint green. Ctp8 and ctp52 CTD adopt the same fold. The domains have been rearranged to fit to the p52 conformation found in the cryo-EM structure of human TFIIH (shown in grey). (B) Ctp52 and ctp8 modelled into the TFIIH cryo-EM structure from Greber *et al.* (26) from which only XPB, p52 and p8 are shown. The lunate-like ring that encircles the XPB helicase domains is schematically depicted in light yellow.

TFIIH an overall rsmld of 1.6 Å was obtained, indicating a high degree of structural similarity. Thus, we generated a ctp52/ctp8/XPB hybrid model (Figure 1B), which contains the high resolution p52/p8 information obtained from our crystal structure embedded in the core TFIIH arrangement as observed in the cryo-EM structure of human TFIIH (6NMI, (26)).

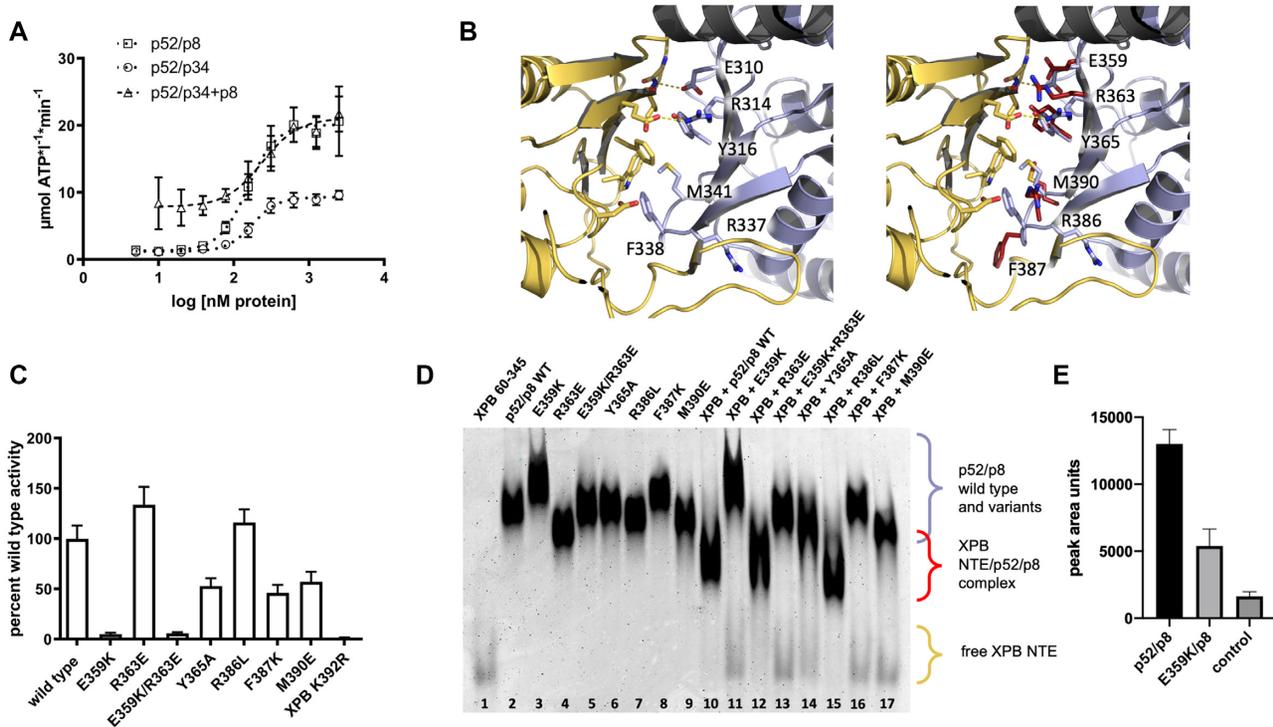
The main and most relevant interaction between the MD2 of p52 and the NTE of XPB seems to be mediated by an orthogonal packing of their respective  $\beta$ -sheets (Figure 1B). Furthermore, MD2 of p52 is also interacting with the second helicase domain, HD2, of XPB in the ctp52/ctp8/XPB hybrid model (Figure 1B). This interaction involves twelve residues of the MD2 domain located mainly in  $\alpha$ -helix 20 of p52 and the  $\beta$ -turns that connect  $\beta$ -strand 8 to 11. The other side of the interface is formed by eleven residues of XPB's HD2. Based on the cryo-EM structure of human TFIIH (PDB code: 6NMI), we observed that both interfaces combined comprise an approximate surface area of 1580 Å<sup>2</sup>. Importantly, an extension of the interface is achieved through the presence of p8 which also interacts with the HD2 of XPB thereby providing an additional surface area of approximately 516 Å<sup>2</sup>.

The hybrid structure illustrates that four entities including p8 and the MD2 of p52, together with the NTE and the DRD domains of XPB form a lunate-like ring (Figure 1B), which embraces the two RecA like domains of XPB from one side, while the DNA approaches the HD domains from the other side (27,28). DNA binding, a functional Walker A motif and an intact RED (arginine-glutamate-aspartate) motif as well as the thumb-like motif (ThM) seem to be prerequisites for the activity of XPB during NER (4,6,9). In addition, the lunate-like ring may be required to position the two RecA like domains of XPB in a conformation which stimulates its ATPase activity. Interestingly, p8 directly in-

teracts with the ThM motif of XPB. Due to the importance of XPB's ATPase activity in transcription and DNA repair, we pursued further studies to decipher the functional influence of the individual components and interactions of this intricate network.

### The ATPase activity of ctXPB is sequentially stimulated by ctp52 and ctp8

To investigate which components are necessary for the full activation of XPB, we used the proteins from *C. thermophilum* due to their superior stability. It was initially shown by Egly and colleagues, that p52 stimulates the ATPase activity of XPB and that p8 could aid in this process (6,41). Hence, we aimed to analyze the activation mechanism. To this end, we first performed titration experiments using a co-purified ctp52/ctp34 complex. Ctp34 was used to enhance the stability of ctp52 and to ensure a native ctp52 conformation in accordance with the composition in TFIIH. Ctp52/ctp34 activates ctXPB in a concentration dependent manner, reaching a  $V_{max}$  of 9.3  $\mu\text{mol ATP}^{-1}\cdot\text{min}^{-1}$  and an  $EC_{50}$  value of 178 nM (Figure 2A, open circles). When ctp8 is added to the ctXPB/ctp52/ctp34 complex, the basal activity level is 7.9  $\mu\text{mol ATP}^{-1}\cdot\text{min}^{-1}$  and reaches a maximum of 21.1  $\mu\text{mol ATP}^{-1}\cdot\text{min}^{-1}$  with an  $EC_{50}$  value of 238 nM (Figure 2A, open triangles) indicating that ctp8 is directly able to further enhance the ctp52-mediated activation of ctXPB. We then investigated whether a co-purified ctp52/ctp8 complex alone is able to activate ctXPB. Here, we observed a full activation with 20.1  $\mu\text{mol ATP}^{-1}\cdot\text{min}^{-1}$  and an  $EC_{50}$  value of 153 nM (Figure 2A, open squares), showing that ctp52/ctp8 is sufficient for ctXPB activation and that ctp34 does not alter the activation cascade. Example time courses for the ATPase measurements are shown in Supplementary

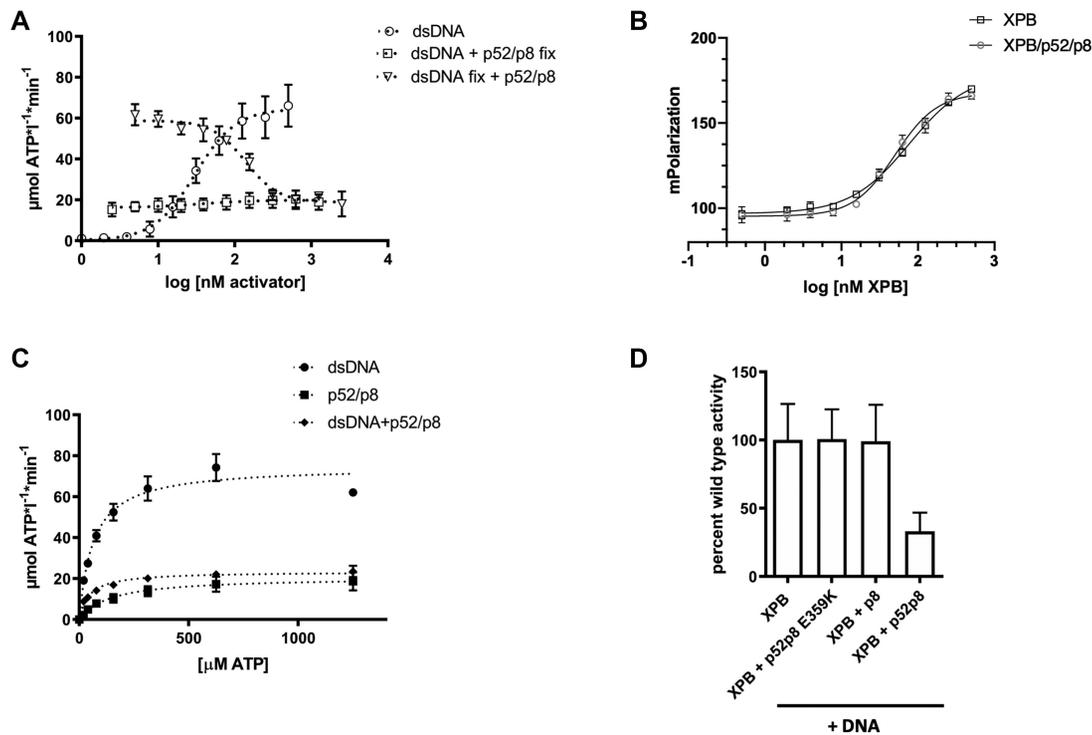


**Figure 2.** Activation of XPB's ATPase by p52/p8 and interaction studies of ctp52 MD2 and ctXPB NTE. (A) ATPase activity profile of ctXPB activated by increasing amounts of different ctp52 complexes. Ctp52 complexes were titrated to a fixed amount of ctXPB. The x-axis shows the logarithmic concentration of the respective ctp52 complex. The y-axis shows the NADH-consumption-derived ATP turnover rate in  $\mu\text{mol ATP per liter per minute}$ .  $N \geq 3$ . (B) Structural insight into the human p52 (light blue) – XPB (yellow) interface taken from 6NMI. On the left hand side, key residues in the interface are labeled. On the right hand side, the corresponding residues in ctp52 (red) are superimposed to the human counterparts and labeled. The depicted ctp52 residues were mutated and analysed with respect to their influence on ctXPB's ATPase activity (C) and in terms of their interaction with the ctXPB NTE (D). (C) Effect of ctp52/ctp8 wild type and variants on the ATPase activity of ctXPB.  $N = 8$ . (D) Native PAGE analysis of the interaction between the ctp52/ctp8 variants and ctXPB NTE (ctXPB 60–345). Individual proteins are shown in lanes 1–9. Ctp52 wild type and variants in complex with ctp8 and ctXPB NTE are shown in lanes 10–17. (E) Pull-down of ctp52/ctp8 and ctp52.E359K/ctp8 with Streptavidin-tagged ctXPB as bait. The pull down fractions were loaded on SDS PAGEs and the SDS PAGE band strength of p52/p8 wild type and E359K was analysed with ImageJ to quantify the interaction capability with full length XPB. Mean values with SD were derived with GraphPad Prism. To exclude unspecific binding to the beads, p52/p8 was loaded as a control to the beads without XPB.  $N = 3$ .

Figure S1A and the SDS PAGE analysis confirms the integrity and purity of the used complexes (Supplementary Figure S2).

Based on our ctp52 crystal structure modelled into the most recent cryo-EM map we were able to identify conserved ctp52 residues in the MD2 domain that interact with the ctXPB NTE. We chose these residues for functional mutagenesis studies since it was also shown that some of these residues (E359, R363) led to phenotypic consequences when they are mutated (41). We therefore generated the following ctp52 variants: E359K, R363E, E359K/R363E, Y365A, R386L, F387K and M390E (Figure 2B). All residues are located in a  $\beta$ -sheet that interacts with a similar motif within the ctXPB NTE. We thus targeted the entire interface from the ctp52 side ranging from a hydrophobic area comprised of Y365, F387 and M390 to the charge complementary patch constituted by E359 and R363. R386 displays different side chain conformations in the cryo-EM structure compared to the crystal structure and was thus chosen to assess which conformation is relevant (Figure 2B). Additionally, Smurnyy *et al.* observed that this residue leads to triptolide resistance in mammalian cells when mutated to a leucine, hinting towards a crucial role in the activation

of XPB's ATPase (42). All ctp52 variants were subjected to CD spectroscopy to ensure that no misfolded variants were used in the analysis (Supplementary Figure S5). Based on the observation that ctp52/ctp8 leads to a higher activation of XPB's ATPase, than ctp52 alone, we investigated whether those variants were still able to activate ctXPB in the ctp52/ctp8 context (Figure 2C). The most severe impact on ctXPB activation was observed for the E359K variant that was only able to activate ctXPB to 5% of the wild type level, whereas the adjacent R363E variant had no effect and displays 134% of wild type activity. In line with the result for the single E359K variant, the E359K/R363E variant displays a comparable decrease to 6% activation. The Y365A, F387K, and M390E variants all display a medium phenotype with an activation profile of 53%, 46% and 57%, respectively, indicating at least partial activation. Mutation of R386 to a leucine (R386L) has no effect showing an activation of 116% compared to the wild type proteins. To further investigate whether these variants still interact with the NTE of ctXPB we performed native PAGE analysis using the ctp52 variants in complex with ctp8 and the NTE of ctXPB comprising residues 60–345 (ctXPB NTE, Figure 2D). All ctp52/ctp8 variant complexes form a single band



**Figure 3.** Influence of DNA and ATP binding on ctXPB's ATPase activity. (A) Influence of dsDNA and ctp52/ctp8 on the ATPase activity of ctXPB. Increasing amounts of dsDNA were titrated onto fixed concentrations of ctXPB (open circles) and ctXPB with ctp52/ctp8 (open squares). Increasing amounts of ctp52/ctp8 were titrated onto fixed concentrations of ctXPB plus dsDNA (open triangles). The x-axis shows the logarithmic concentration of the respective activator complex that was titrated. The y-axis shows the NADH-consumption-derived ATP turnover rate in  $\mu\text{mol ATP per liter per minute}$ .  $N \geq 4$ . (B) Fluorescence polarization measurements to investigate the influence of ctp52/ctp8 on ctXPB's dsDNA binding ability. Increasing amounts of ctXPB (open squares) and ctXPB with ctp52/ctp8 (open circles) were titrated onto fixed amounts of fluorescently labeled dsDNA. The calculated  $K_D$  values are listed in Table 2.  $N = 3$ . (C) ATPase activity profile with Michaelis-Menten kinetics. Increasing amounts of substrate (ATP) were titrated onto different complexes of ctXPB with its activators dsDNA (circles), ctp52/ctp8 (squares) and dsDNA plus ctp52/ctp8 (rhombus). The x-axis shows the concentration of the respective activator complex. The y-axis shows the NADH-consumption-derived ATP turnover rate in  $\mu\text{mol ATP per liter per minute}$ .  $V_{\max}$ ,  $K_M$ ,  $k_{\text{cat}}$  and  $k_{\text{cat}}/K_M$  values for the different ctXPB activation complexes are listed in Table 3.  $N \geq 3$ . (D) ATPase activity of ctXPB in the presence of dsDNA with or without different ctp52/ctp8 complexes.  $N \geq 8$ .

in the native PAGE, however, due to the different charge properties of the variants they are located at slightly different heights (Figure 2D, lanes 3–9) compared to the wild type complex (Figure 2D, lane 2). Combining ctXPB NTE with wild type ctp52/ctp8 clearly leads to complex formation (Figure 2D, lane 10), whereas the variants that displayed reduced activation capabilities in Figure 2C (E359K, R363E/E359K, F387K, and M390E) also show no interaction in the native PAGE experiment (Figure 2D, lanes 11, 13, 16–17). Variant Y365A, which displayed a partial activation of ctXPB's ATPase (Figure 2C), shows also a partial interaction in the native PAGE (Figure 2D, lane 14) indicating that the interface is affected by the mutation but binding is only partially disrupted. The ctp52/ctp8 variants that were able to activate ctXPB's ATPase like wild type (R363E and R386L, Figure 2C) were capable of forming a complex with ctXPB NTE (Figure 2D, lane 12 and 15).

Since the cryo-EM structure of holo-TFIIH suggests the presence of a second XPB/p52 interface located in the HD2 domain of XPB and the MD2 of p52 (Figure 1B) (26) we analysed the E359K variant towards its interaction with full length ctXPB. Albeit decreased, we could indeed observe significant complex formation (Figure 2E). This result suggests that disrupting one of the interfaces is sufficient to

prevent p52/p8 mediated XPB activation, whereas complex formation is partially maintained.

### ctp52/ctp8 limits ctXPB DNA dependent ATPase activity

We next aimed to investigate the influence of DNA binding on the ATPase activity of ctXPB. In the sole presence of double stranded DNA (dsDNA) the activity of ctXPB reached a  $V_{\max}$  of  $64.6 \mu\text{mol ATP}^{-1} \cdot \text{min}^{-1}$  (Figure 3A, open circles). Interestingly, this is about three times higher than the activation of ctXPB via ctp52/ctp8 (See open squares in Figure 2A for comparison). Surprisingly, in the presence of ctp52/ctp8 and dsDNA the induced ATPase activity is limited to  $19.9 \mu\text{mol ATP}^{-1} \cdot \text{min}^{-1}$  (Figure 3A, open squares) which corresponds well to the activity achieved with ctp52/ctp8 in the absence of dsDNA (Figure 2A, open squares). This result indicates that ctp52/ctp8 negatively controls the dsDNA-dependent activation of XPB, which is illustrated further when we titrated increasing amounts of ctp52/ctp8 to an equimolar ctXPB/DNA complex (Figure 3A, open triangles). At 125 nM, dsDNA ctXPB is already completely activated with an activity of  $58.8 \mu\text{mol ATP}^{-1} \cdot \text{min}^{-1}$ . This activity decreases in a ctp52/ctp8 concentration dependent man-

**Table 2.**  $K_D$  values for XPB calculated from fluorescence polarization measurements in Figure 3B

Complex	$k_D$ (nM XPB)
XPB/DNA	77
XPB/p52/p8/DNA	51

ner until it reaches again  $18.9 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$  with an  $EC_{50}$  value of 138 nM indicating a high binding affinity between ctXPB and ctp52/ctp8 and once an equimolar complex is formed, downregulation is complete. To further decipher the mechanism of activity limitation we investigated whether the binding of ctp52/ctp8 influences the DNA binding capacity of ctXPB. We therefore pursued fluorescence polarization measurements comparing ctXPB alone to the ctXPB/ctp52/ctp8 complex in the presence of DNA (Figure 3B and Table 2). The obtained  $k_D$  values were 77 nM for ctXPB and 51 nM for the ctXPB/ctp52/ctp8 complex. These similar values indicate that there is no major effect on dsDNA binding caused by the presence of ctp52/ctp8. Since dsDNA binding did not seem to be affected, we evaluated whether ATP binding is altered by the presence of ctp52/ctp8. Michaelis-Menten experiments to assess the apparent  $K_M$  for ATP (Figure 3C and Table 3) led to  $K_M$  values of  $62.9 \mu\text{M}$  (ctXPB/dsDNA),  $140.0 \mu\text{M}$  (ctXPB/ctp52/ctp8), and  $45.0 \mu\text{M}$  (ctXPB/ctp52/ctp8 with excess dsDNA). These values indicate a positive effect on the  $K_M$  for ATP in the presence of dsDNA. However, the corresponding  $V_{\text{max}}$  value of ctXPB with excess dsDNA ( $74.9 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ ) is around three times higher than the  $V_{\text{max}}$  of ctXPB/ctp52/ctp8/dsDNA ( $23.4 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ ). The latter value fits well with the  $V_{\text{max}}$  of ctXPB/ctp52/ctp8 without dsDNA ( $20.7 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ ) indicating that ctp52/ctp8 predominantly affects the  $V_{\text{max}}$  of the ctXPB/dsDNA complex. In addition, we calculated the  $k_{\text{cat}}/K_M$  ratios for ctXPB/ctp52/ctp8 and ctXPB/ctp52/ctp8/dsDNA with  $0.7 \mu\text{M ATP}^*\text{min}^{-1}$  and  $1.8 \mu\text{M ATP}^*\text{min}^{-1}$ , respectively. Both are significantly lower than the  $k_{\text{cat}}/K_M$  ratio of the ctXPB/dsDNA complex with  $4.8 \mu\text{M ATP}^*\text{min}^{-1}$ , indicating that also the catalytic efficiency is significantly decreased by p52/p8 when dsDNA is bound to XPB.

To further substantiate these observations, we investigated whether the ctp52\_E359K/ctp8 complex can repress the dsDNA-mediated ctXPB activation (Figure 3D). Our data clearly show that the  $V_{\text{max}}$  of the ctXPB/ctp52\_E359K/ctp8 complex is not repressed and reaches XPB/dsDNA like activity. Ctp8 by itself also does not exert any effect (Figure 3D). To assess the influence ctp8 might have on the DNA activation mechanism, we titrated dsDNA onto ctXPB/ctp52/ctp34 and ctXPB/ctp52/ctp34/ctp8 complexes and compared them to the dsDNA activation of ctXPB alone (Figure 4A). The ctXPB/ctp52/ctp34 complex is activated twofold by dsDNA from 7.1 to  $17.0 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$  (Figure 4A, open triangles) and the ctXPB/ctp52/ctp34/ctp8 complex displays a similar activation from 14.3 to  $24.6 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$  (Figure 4A, open rhombs). The values for ctXPB/ctp52/ctp34/ctp8 are very similar to the maximum level of activation reached by ctp52/ctp8

in the presence or absence of dsDNA (Figure 3C, 23.4 and  $20.7 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ , respectively). It seems that ctp8 in this case partially compensates for the presence of dsDNA since the maximum activity value of the ctXPB/ctp52/ctp34 complex with excess of dsDNA ( $17.0 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ , Figure 4A open triangles) is similar to the value of the ctXPB/ctp52/ctp34/ctp8 complex ( $24.6 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ , Figure 4A open rhombs). It is important to note, however, that none of these complexes reached the rate of dsDNA induced activity in the sole presence of ctXPB ( $64.6 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ , Figure 4A, open circles).

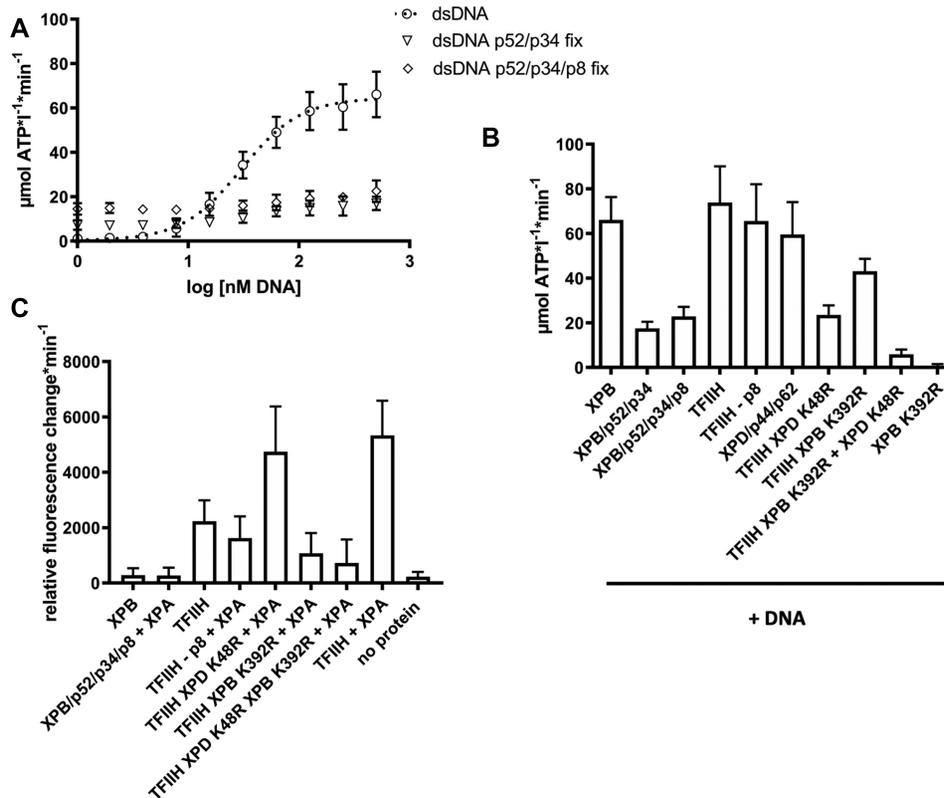
We further investigated the ATPase activity of ctXPB in the broader context of core ctTFIIH by adding a ctXPD/ctp44/ctp62 complex to ctXPB/ctp52/ctp34/ctp8. In core ctTFIIH, the overall dsDNA induced ATPase activity is higher than for ctXPB alone with  $73.8 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$  compared to  $66.1 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ , respectively (Figure 4B). To delineate the influence of the two motor proteins on core ctTFIIH activity, we investigated core ctTFIIH with either the Walker A variant of XPD or XPB (ctTFIIH XPD\_K48R or ctTFIIH XPB\_K392R, respectively). The ctTFIIH XPD\_K48R complex displayed an activity of  $23.6 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ , which is almost identical to the activity of ctXPB/ctp52/ctp34 and ctXPB/ctp52/ctp34/ctp8 in the presence of dsDNA (Figure 4B and Table 4). For the ctTFIIH XPB\_K392R complex we observed an activity of  $43.2 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ . The single activities of both Walker A variant complexes add up to  $66.8 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$  which is almost identical to our observed core ctTFIIH ATPase activity ( $73.8 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ ). These data indicate that no further activation of the ctXPB ATPase in the core ctTFIIH environment is achieved by the presence of the additional subunits. The ctXPD/ctp44/ctp62 complex displays an ATPase activity of  $59.5 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ , rendering XPD the major ATPase in core ctTFIIH. Finally, we investigated core ctTFIIH lacking ctp8. Here, we observed an activity of  $65.7 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$ . This corresponds to a reduction of around  $8 \mu\text{mol ATP}^*1^{-1}*\text{min}^{-1}$  when compared to core ctTFIIH containing ctp8. This reduction in activity is in line with the observed reduction in activation of XPB's ATPase, which emphasizes the importance of ctp8 within core ctTFIIH (43). Combined, these findings indicate that XPB's ATPase activity is mainly regulated by p52/p8 but not by other subunits within TFIIH.

#### ctXPB translocase activity is dependent on core ctTFIIH

After assessing how ctXPB's ATPase function is regulated, we investigated how this regulation affects the translocase function of ctXPB. XPB is the essential enzymatic component of the TFIIH 5'-3' translocase complex (25), therefore its ATPase activity should directly influence its ability to translocate on dsDNA. Since XPA was shown to stimulate TFIIH's translocase activity (28), we also investigated the role of XPA in this process. Therefore, we used a well established triplex disruption assay to monitor ctXPB translocase activity (25,28). Reconstituted core ctTFIIH exhibits a robust translocase activity ( $2236 \text{ u rel. FC}^*\text{min}^{-1}$ , nega-

**Table 3.** Kinetic parameters for ATP calculated from ATPase measurements in Figure 3C

Complex	$K_M$ ( $\mu\text{M ATP}$ )	$V_{\text{max}}$ ( $\mu\text{mol ATP} \cdot \text{l}^{-1} \cdot \text{min}^{-1}$ )	$k_{\text{cat}}$ ( $\text{l} \cdot \text{min}^{-1}$ )	$k_{\text{cat}}/K_M$ ( $\mu\text{M ATP} \cdot \text{min}^{-1}$ )
XPB/dsDNA	62.9	74.9	299.6	4.8
XPB/p52/p8	140	23.4	93.6	0.7
XPB/dsDNA/p52/p8	45	20.7	82.8	1.8



**Figure 4.** ATPase and translocase activities of ctXPB in core ctTFIIH. (A) ATPase activity profile of ctXPB activated by increasing amounts of dsDNA without ctp52 (open circles) and with fixed concentrations of ctp52/ctp34 (open triangles) and ctp52/ctp34/ctp8 (open rhombs). The x-axis shows the logarithmic concentration of dsDNA. The y-axis shows the NADH-consumption-derived ATP turnover rate in  $\mu\text{mol ATP}$  per liter per minute.  $N \geq 6$ . The data for the dsDNA curve (open circles) were taken from Figure 3A for better comparison. (B) ATPase activity of core ctTFIIH and subunit complexes in the presence of dsDNA. The mean values and standard deviations for each condition are listed in Table 4.  $N \geq 6$ . (C) DsDNA translocase activity of core ctTFIIH and different subunit complexes in the presence and absence of ctXPA. The mean values and standard deviations for each condition are listed in Table 4.  $N \geq 11$ .

**Table 4.** Measured mean values and standard deviations (SD) from ATPase measurements of TFIIH complexes in Figure 4B

Complex	ATPase activity ( $\mu\text{mol ATP} \cdot \text{l}^{-1} \cdot \text{min}^{-1}$ )	SD
XPB	66.1	10.2
XPB/p52/p34	17.6	3.0
XPB/p52/p34/p8	22.9	4.4
TFIIH	73.8	16.3
TFIIH - p8	65.7	16.4
XPD/p44/p62	59.5	14.6
TFIIH K48R	23.6	4.4
TFIIH XPB K392R	43.2	5.6
TFIIH XPB K392R XPD K48R	5.9	2.1
XPB K392R	0.4	1.2

tive control: 231 u rel.  $\text{FC} \cdot \text{min}^{-1}$ , Figure 4C and Table 5) which is further enhanced by the addition of a fourfold mo-

lar excess of ctXPA (5340 u rel.  $\text{FC} \cdot \text{min}^{-1}$ , Figure 4C and Table 5). Example time courses for the triplex displacement measurements are shown in Supplementary Figure S1B and the SDS PAGE analysis confirms the integrity and purity of the used complexes (Supplementary Figure S2). As our DNA triplex substrates would allow for more than one core ctTFIIH complex to bind, we investigated if cooperative interactions of multiple complexes on the same triplex might occur or even be necessary. We therefore performed titration experiments with either increasing amounts of ctTFIIH/ctXPA complex or increasing amounts of DNA. Both experiments indicate that translocase movement can still be observed at equimolar protein and DNA concentrations arguing against a cooperative behaviour (Supplementary Figure S6A, B). However, increasing the molar ratio of protein concentration relative to the DNA (Supplementary Figure S6A) leads to an increased activity that could be attributed to more than one core ctTFIIH being involved

**Table 5.** Measured mean values and standard deviations (SD) from translocase measurements of TFIIH complexes in Figure 4C

Complex	Translocase activity (relative fluorescence change*min <sup>-1</sup> )	SD
XPB	283	252
XPB/p52/p34/p8 + XPA	269	288
TFIIH	2236	756
TFIIH - p8 + XPA	1628	780
TFIIH XPD K48R + XPA	4742	1642
TFIIH XPB K392R + XPA	1071	740
TFIIH XPD K48R XPB K392R + XPA	721	852
TFIIH + XPA	5340	1248
no protein	231	175

in translocase activity. In addition, we show that a 112 bp triplex substrate results in the same activity measured with our assay settings (500 nM ctTFIIH, 150 nM DNA) as compared to the 52 bp triplex used in this study (Supplementary Figure S6C) suggesting that the length of the DNA substrate is not a decisive factor for the translocase activity.

Reconstituting the core ctTFIIH/ctXPA complex with the ctXPD Walker A variant (ctXPD\_K48R) has no effect on the translocase activity (4742 u rel. FC\*min<sup>-1</sup>, Figure 4C and Table 5) whereas reconstituting the core ctTFIIH/ctXPA complex with the ctXPB Walker A variant (ctXPB\_K392R) shows a strong decrease in activity (1071 u rel. FC\*min<sup>-1</sup>, Figure 4C and Table 5). When reconstituting a core ctTFIIH/ctXPA complex containing both ctXPB and ctXPD Walker A variants, a comparable decrease in activity was observed (721 u rel. FC\*min<sup>-1</sup>, Figure 4C and Table 5). We attribute the remaining minor activity observed for ctTFIIH/ctXPA XPB\_K392R as well as for ctTFIIH/ctXPA K392R/K48R to binding events, which destabilize the triplex, rather than to a real translocase movement. These results strongly suggest that ctXPB is the major motor for translocase activity within core ctTFIIH.

Interestingly, removal of ctp8 from the core ctTFIIH/ctXPA complex also decreases the translocase activity significantly to 1628 u rel. FC\*min<sup>-1</sup> (Figure 4C and Table 5). To our surprise ctXPB by itself does not exhibit any measurable translocase activity within our experimental conditions (283 u rel. FC\*min<sup>-1</sup>, Figure 4C and Table 5). Furthermore, a ctXPB/ctp52/ctp34/ctp8 complex with a 4-fold excess of ctXPA still yields no measurable translocase activity (269 u rel. FC\*min<sup>-1</sup>).

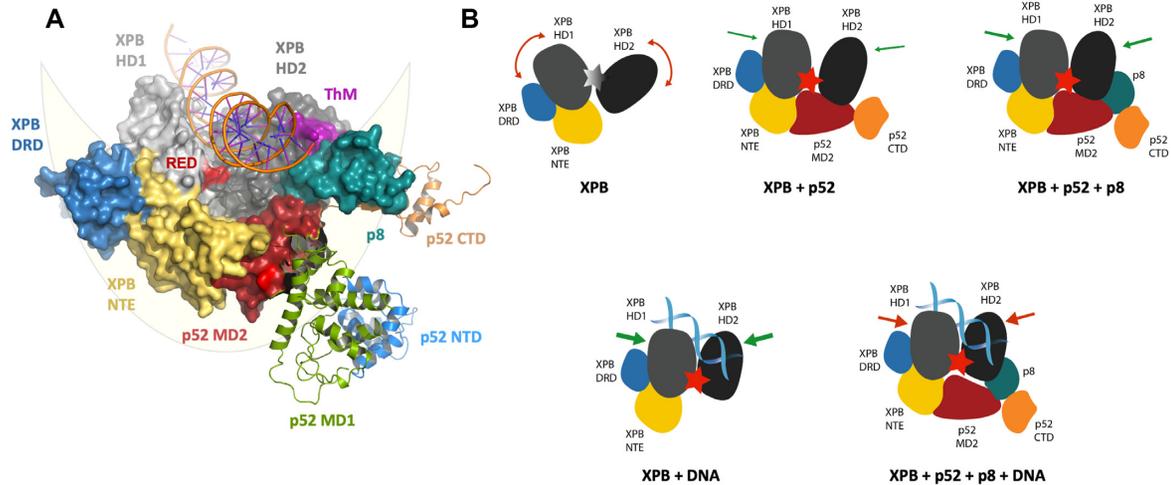
Finally, we addressed the question if the increase in translocase activity by ctXPA is due to the stimulation of XPB's ATPase activity. Interestingly, ctXPA does not seem to influence ctXPB's ATPase activity in any of the analysed conditions (Supplementary Figure S7A), thus excluding the possibility that the translocase effect mediated by XPA is due to further ATPase modulation. Taken together, our data suggest that the XPB translocase function is only operating in the context of core TFIIH and the enzymatic activity is solely provided by XPB.

## DISCUSSION

Recent structural data derived from cryo-EM studies have greatly improved our understanding towards the architecture of core TFIIH and its individual components (26–28). However, little is known so far about the intricate network within core TFIIH, i.e. how the individual subunits define and regulate the activity of the entire complex. In this work we focused on the XPB helicase and investigated the regulation of XPB (i) by its direct interaction partners p52 and p8; and (ii) in the context of the entire core TFIIH, to obtain insights how the ATPase is regulated, since this function is essential for transcription and DNA repair.

Our crystal structures of ctp52 show that p52 is a flexible protein that requires its binding partners XPB and p8 to assume a complete functional state. The high structural homology of the individual p52 domains but the different orientation of the domains relative to each other within our crystal structures compared to the cryo-EM TFIIH structures provides support for this hypothesis. We cannot exclude the possibility that the domain arrangement observed in our crystal structure is a crystallization artefact due to crystal packing. The modular composition of p52, however, may well be necessary to provide the flexibility to assume different conformational states thereby fulfilling TFIIH's functions in transcription and NER (26–28,44).

Importantly, the high structural homology between the *C. thermophilum* and the human proteins indicates a high functional homology of the two organisms that has been described previously (11,45). In this study, we extended this system towards the analysis of XPB regulation and the functional assembly of core TFIIH. We first analysed the interaction of ctXPB and ctp52 and the effect that ctp52 exerts on ctXPB. We observed that ctp52 is sufficient to enable ctXPB's ATPase function, which is well in line with previous work (6). To our surprise, this effect could be directly boosted by the addition of ctp8 without the presence of other core TFIIH proteins indicating a synergistic activation mechanism depending on p52/p8 only. Adding p8 to the XPB/p52 complex further enhanced the activity twofold (Figure 2A). One of the crucial players towards mediating the effect of p52 activation is the NTE of XPB. We analysed the interface formed by the NTE of XPB and MD2 of p52 through functional mutagenesis studies and showed that individual residues within this interface are critical for the p52 dependent ATPase activation. Based on the intricate network of interactions between XPB, p52, and p8 as observed in the cryo-EM structures, a sequential activation mechanism of XPB could be envisioned (Figure 5). In its apo form XPB is not able to hydrolyse ATP due to the high flexibility of the RecA like helicase domains (HD1 and HD2, Figure 5B), as observed in the crystal structures of the isolated XPB protein in which the helicase domains assume an orientation relative to each other which is not productive for ATP hydrolysis (4). Stabilisation and correct positioning is achieved by the combined action of p52 and p8. P52 interacts with the NTE and HD2 of XPB, thereby positioning HD1 and HD2 in an ATP hydrolysis competent state (26). Stability is further enhanced by p8, which also binds to HD2 of XPB. Remarkably, p8 also directly interacts with the ThM motif of XPB, clearly pointing towards a role for



**Figure 5.** Model for p52/p8 mediated activation and restriction of XPB's ATPase. (A) Structural model of the lunete-like ring that encircles the two XPB helicase domains. XPB comprises the following domains: NTE (yellow), DRD domain (blue), HD1 (light gray) and HD2 (dark gray). The ThM and RED motifs are highlighted in purple and red, respectively. P52 comprises NTD (blue), MD1 (green), MD2 (red), and CTD (orange). P8 is depicted in mint green, the DNA duplex in orange. The lunete-like ring that encircles the XPB helicase domains is schematically depicted in light yellow. Domains participating in the lunete-like ring are shown as solid surfaces. P52 domains not participating in the lunete-like ring are shown as cartoon. The structural model was derived from Schilbach, *et al.* (27) and our p52 structures. (B) Model of the differential XPB activation by DNA, p52/p8, p52 or p52/p8/DNA. If no activator is present, the XPB helicase domains HD1 and HD2 are too flexible and ATP hydrolysis cannot be performed. P52 binding fixes the helicase domains and thus brings them into close proximity to each other enabling ATP hydrolysis. Binding of p8 further extends the lunete-like ring and further restrains the flexibility of the two helicase domains resulting in an even higher activity. When DNA binds to XPB, the helicase domains are also restrained and brought into close proximity, leading to the highest ATPase activity. Simultaneous p52 or p52/p8 and DNA binding also results in an activation of XPB, but the activation level is limited due to the reduced flexibility of XPB.

the proper positioning of this crucial element within XPB (Figure 1B). Combined, the interaction network comprising the MD2 domain of p52, the NTE and DRD domains of XPB and p8 leads to the formation of a lunete ring which embeds the two helicase domains of XPB, thereby permitting an ATPase compatible conformation. A recent study indicates that binding of triptolide to Cys342 of XPB might disturb the interaction with p52/p8. Cys342 is located in HD1 of XPB and points towards the cleft between the two RecA like helicase domains (Supplementary Figure S7B). Therefore, triptolide binding might push HD1 and HD2 apart from each other forcing them into an open conformation (46). This hypothesis is in agreement with our data as this enforced open conformation might counteract p52 and p8 binding via the lunete-like ring, thereby impeding enzymatic activity.

To interrogate the effect of DNA on XPB we first analysed the dsDNA-mediated activation of XPB in the absence of any other subunit and observed that the ATPase activity is about three times higher than with p52/p8, indicating that higher ATPase rates of the XPB scaffold can be achieved. To our surprise, the combination of dsDNA and p52/p8 did not lead to additional activation. On the contrary, the presence of p52/p8 limits the activation of XPB to the state that was achieved just by the addition of p52/p8 (Figure 3A) which could be attributed to limiting the  $V_{\max}$  value of the ATPase function and not to DNA binding or changes of the apparent  $K_M$  for ATP (Figure 3B,C). Thus p52/p8 acts not only as a clutch (27) but also as a 'speed limiter' ensuring that XPB does not proceed too fast either in the presence or absence of dsDNA (Figure 5B). Remarkably, this speed limitation is also not released in the presence of

the other core TFIIH subunits (Figure 4B) indicating that the regulatory network formed by XPB, p52, and p8 (Figure 5A) is sufficient for the incorporation and regulation of XPB's ATPase within the TFIIH scaffold. Disrupting the contact points between the NTE of XPB and MD2 of p52 leads to the abrogation of the control of XPB by p52/p8 and restores the activation of XPB by dsDNA (Figures 2C and 3D). It thus seems that the presence of p52/p8 on the one hand is required for XPB to assume an ATPase competent conformation but on the other hand restricts that conformation so that higher ATPase rates cannot be achieved and thereby also assumes the role as a speed limiter.

The ATPase function of XPB is a prerogative for its subsequent actions in transcription and DNA repair (reviewed in (47)) and is highly associated with its translocase activity leading to promoter opening but may also be important for NER (25,28,48). We investigated the different ATPase competent XPB complexes towards their ability to translocate on dsDNA using a triplex displacement assay. To our surprise, the ATPase function of XPB alone is not sufficient to observe translocase activity (Figure 4C). Furthermore, the complex comprising XPB/p52/p34/p8/XPA led to the same result. Since the latter complex should include all necessary elements to stimulate XPB's translocase activity we conclude that only in a fully assembled core TFIIH translocase activity can be observed. Whether this is due to other TFIIH components aiding in productive XPB DNA binding or other TFIIH components are required for processivity remains currently unclear. However, TFIIH translocase activity is exclusively dependent on the ATPase activity of XPB, as shown by the comparison of core TFIIH assembled in the presence of the Walker A motif variants of XPD or

XPB, respectively. We observed that XPA was able to boost the translocase activity about 2.4-fold which is in line with the data from Kokic *et al.* (28). XPA interacts with dsDNA above the XPB subunit of core TFIIH like a clamp, thereby further restraining XPB (28). This conformation may stimulate the translocase activity of core TFIIH. Intriguingly, this stimulation is not caused by a modulation of the ATPase activity of XPB (Supplementary Figure S7A). We thus suggest that XPA increases the processivity of the complex. Despite its size, p8 plays a highly important role with respect to the translocase activity within core TFIIH. Removal of p8 from core TFIIH causes a significant decrease in XPA stimulated translocase activity that is even lower than the activity of core TFIIH without XPA (Figure 4C). Coin *et al.* reported that p8 plays a crucial role in the recruitment of XPA (49). Hence, the impact of p8 on XPB's translocase activity is twofold: its loss leads to a reduction in ATPase activity by affecting the overall stability of TFIIH but also impairs XPA recruitment. Combined, these observations emphasize the role of p8 as a vital factor for NER and transcription as reflected by its involvement in the disease TTD (43,50).

Our data show that XPB is a highly regulated protein. This regulation is embedded in several layers and could depend on the cellular process the protein is involved in. The maximum activation of XPB by DNA may only be important for other, less explored functions of XPB related to its localisation at the centrosome during mitosis or mRNA export (51,52) and thus requires different regulation or possibly even higher activity than in the context of TFIIH. When XPB fulfils its essential functions in transcription and NER, the main regulators are p52 and p8 that are able to activate XPB in a synergistic manner and at the same time act as a speed limiter, seemingly keeping the ATPase activity always at a constant level.

XPB's ATPase activity is crucial for NER and transcription initiation (53), however, the requirements for XPB in both processes differ. In transcription the absence of XPB is less detrimental than its presence in an inactive state (54). Since XPB exhibits instantaneously ATPase activity when combined with p52/p8, no additional trigger is required in the TFIIH environment, rendering XPB the only functional ATPase in a non-DNA bound TFIIH, whereas XPD activity is strictly DNA-dependent (11). This DNA independent ATPase activity could be instrumental to rearrangement/recruitment processes in the initial steps of transcription. This hypothesis is supported by the low translocase activity of core TFIIH in the absence of XPA and is well in line with the function that XPB adopts in transcription, where it acts like a molecular wrench by translocating on the DNA prior to the transcription start site thereby reeling the DNA into the active site cleft of the RNAPII. This leads to torsional tension and unwinds the DNA (27) but may require only a very low translocase activity to melt several nucleotides on the DNA before RNA polymerase takes over. In NER a DNA independent ATPase activity could aid towards the initial steps of TFIIH recruitment by XPC (9) and subsequent engagement with DNA. After these initial steps, the arrival of XPA would stimulate the translocase activity to enhance the unwinding of the repair bubble necessary for XPD to engage with the

DNA, yielding a repair competent complex that would also be in line with XPA's role in lesion verification (55).

In conclusion, our analysis permitted us to decipher the tight regulatory network controlling XPB's ATPase activity within TFIIH. We show in a stepwise approach how p52 and p8 activate XPB in an additive way. Importantly, this is independent from the activation of XPB by DNA. Despite being seen as an activator, the p52/p8 complex also limits the  $V_{max}$  of XPB's ATPase activation by DNA and thus assumes the function of a speed limiter, which clearly dominates XPB's regulation within TFIIH. Lastly, we show that despite its full ATPase capacity in isolation, XPB only acts as a translocase in the context of core TFIIH. Combined, our work redefines the regulatory network of the XPB ATPase and translocase shedding light on the inner regulatory mechanisms of core TFIIH in transcription and DNA repair.

## DATA AVAILABILITY

Atomic coordinates and structure factors have been deposited in the Protein Data Bank under the accession codes 6TRS (ctp52..121–514/ctp8) and 6TRU (ctp52..1–321).

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We would like to thank the staff from the beamline ID30A-3 at the European Synchrotron Radiation Facility and the staff at beamline 14.1 at BESSY for excellent support. Furthermore, we would like to thank Dr Maren Bleckmann and Michelle Endres from our insect cell expression core facility for expressing the Strep-tagged XPB in Hi5 cells.

## FUNDING

Deutsche Forschungsgemeinschaft [KI-562/11-1 to C.K.]. Funding for open access charge: Deutsche Forschungsgemeinschaft [KI-562/11-1].

*Conflict of interest statement.* None declared.

## REFERENCES

- Weeda,G., van Ham,R.C., Masurel,R., Westerveld,A., Odijk,H., de Wit,J., Bootsma,D., van der Eb,A.J. and Hoeijmakers,J.H. (1990) Molecular cloning and biological characterization of the human excision repair gene ERCC-3. *Mol. Cell. Biol.*, **10**, 2570–2581.
- Hwang,J.R., Moncollin,V., Vermeulen,W., Seroz,T., van Vuuren,H., Hoeijmakers,J.H. and Egly,J.-M. (1996) A 3' → 5' XPB helicase defect in repair/transcription factor TFIIH of xeroderma pigmentosum group B affects both DNA repair and transcription. *J. Biol. Chem.*, **271**, 15898–15904.
- Fan,L. and DuPrez,K.T. (2015) XPB: An unconventional SF2 DNA helicase. *Prog. Biophys. Mol. Biol.*, **117**, 174–181.
- Fan,L., Arvai,A.S., Cooper,P.K., Iwai,S., Hanaoka,F. and Tainer,J.A. (2006) Conserved XPB core structure and motifs for DNA unwinding: implications for pathway selection of transcription or excision repair. *Mol. Cell*, **22**, 27–37.
- Schaeffer,L., Moncollin,V., Roy,R., Staub,A., Mezzina,M., Sarasin,A., Weeda,G., Hoeijmakers,J.H. and Egly,J.-M. (1994) The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *EMBO J.*, **13**, 2388–2392.

6. Coin,F., Oksenysh,V. and Egly,J.-M. (2007) Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. *Mol. Cell*, **26**, 245–256.
7. Compe,E. and Egly,J.-M. (2012) TFIIH: when transcription met DNA repair. *Nat. Rev. Mol. Cell Biol.*, **13**, 343–354.
8. Tirode,F., Busso,D., Coin,F. and Egly,J.-M. (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. Oksenysh,V., de Jesus,B.B., Zhovmer,A., Egly,J.-M. and Coin,F. (2009) Molecular insights into the recruitment of TFIIH to sites of DNA damage. *EMBO J.*, **28**, 2971–2980.
10. Fuss,J.O. and Tainer,J.A. (2011) XPB and XPD helicases in TFIIH orchestrate DNA duplex opening and damage verification to coordinate repair with transcription and cell cycle via CAK kinase. *DNA Repair (Amst.)*, **10**, 697–713.
11. Kuper,J., Braun,C., Elias,A., Michels,G., Sauer,F., Schmitt,D.R., Poterszman,A., Egly,J.-M. and Kisker,C. (2014) In TFIIH, XPD helicase is exclusively devoted to DNA repair. *PLoS Biol.*, **12**, e1001954.
12. Lu,H., Zawal,L., Fisher,L., Egly,J.-M. and Reinberg,D. (1992) Human general transcription factor IIIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature*, **358**, 641–645.
13. Feaver,W.J., Svejstrup,J.Q., Henry,N.L. and Kornberg,R.D. (1994) Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. *Cell*, **79**, 1103–1109.
14. Roy,R., Adamczewski,J.P., Seroz,T., Vermeulen,W., Tassan,J.P., Schaeffler,L., Nigg,E.A., Hoeijmakers,J.H. and Egly,J.-M. (1994) The MO15 cell cycle kinase is associated with the TFIIH transcription-DNA repair factor. *Cell*, **79**, 1093–1101.
15. Glover-Cutter,K., Laroche,S., Erickson,B., Zhang,C., Shokat,K., Fisher,R.P. and Bentley,D.L. (2009) TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. *Mol. Cell Biol.*, **29**, 5455–5464.
16. Araújo,S.J., Tirode,F., Coin,F., Pospiech,H., Syväoja,J.E., Stucki,M., Hübscher,U., Egly,J.-M. and Wood,R.D. (2000) Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes Dev.*, **14**, 349–359.
17. Coin,F., Oksenysh,V., Mocquet,V., Groh,S., Blattner,C. and Egly,J.-M. (2008) Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Mol. Cell*, **31**, 9–20.
18. Peissert,S., Sauer,F., Grabarczyk,D.B., Braun,C., Sander,G., Poterszman,A., Egly,J.-M., Kuper,J. and Kisker,C. (2020) In TFIIH the Arch domain of XPD is mechanistically essential for transcription and DNA repair. *Nat. Commun.*, **11**, 1667–1613.
19. Weeda,G., van Ham,R.C., Vermeulen,W., Bootsma,D., van der Eb,A.J. and Hoeijmakers,J.H. (1990) A presumed DNA helicase encoded by ERCC-3 is involved in the human repair disorders xeroderma pigmentosum and Cockayne's syndrome. *Cell*, **62**, 777–791.
20. Weeda,G., Eveno,E., Donker,I., Vermeulen,W., Chevallier-Lagente,O., Taieb,A., Stary,A., Hoeijmakers,J.H., Mezzina,M. and Sarasin,A. (1997) A mutation in the XPB/ERCC3 DNA repair transcription gene, associated with trichothiodystrophy. *Am. J. Hum. Genet.*, **60**, 320–329.
21. Oh,K.-S., Khan,S.G., Jaspers,N.G.J., Raams,A., Ueda,T., Lehmann,A., Friedmann,P.S., Emmert,S., Gratchev,A., Lachlan,K. et al. (2006) Phenotypic heterogeneity in the XPB DNA helicase gene (ERCC3): xeroderma pigmentosum without and with Cockayne syndrome. *Hum. Mutat.*, **27**, 1092–1103.
22. Bukowska,B. and Karwowski,B.T. (2018) Actual state of knowledge in the field of diseases related with defective nucleotide excision repair. *Life Sci.*, **195**, 6–18.
23. Giglia-Mari,G., Coin,F., Ranish,J.A., Hoogstraten,D., Theil,A., Wijgers,N., Jaspers,N.G.J., Raams,A., Argenti,M., van der Spek,P.J. et al. (2004) A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nat. Genet.*, **36**, 714–719.
24. Grünberg,S., Warfield,L. and Hahn,S. (2012) Architecture of the RNA polymerase II preinitiation complex and mechanism of ATP-dependent promoter opening. *Nat. Struct. Mol. Biol.*, **19**, 788–796.
25. Fishburn,J., Tomko,E., Galburt,E. and Hahn,S. (2015) Double-stranded DNA translocase activity of transcription factor TFIIH and the mechanism of RNA polymerase II open complex formation. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 3961–3966.
26. Greber,B.J., Toso,D.B., Fang,J. and Nogales,E. (2019) The complete structure of the human TFIIH core complex. *Elife*, **8**, e44771.
27. Schilbach,S., Hantsche,M., Tegunov,D., Dienemann,C., Wigge,C., Urlaub,H. and Cramer,P. (2017) Structures of transcription pre-initiation complex with TFIIH and Mediator. *Nature*, **551**, 204–209.
28. Kokic,G., Chernev,A., Tegunov,D., Dienemann,C., Urlaub,H. and Cramer,P. (2019) Structural basis of TFIIH activation for nucleotide excision repair. *Nat. Commun.*, **10**, 2885.
29. Greber,B.J., Nguyen,T.H.D., Fang,J., Afonine,P.V., Adams,P.D. and Nogales,E. (2017) The cryo-electron microscopy structure of human transcription factor IIH. *Nature*, **549**, 414–417.
30. Li,M.Z. and Elledge,S.J. (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat. Methods*, **4**, 251–256.
31. Wang,W. and Malcolm,B.A. (1999) Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuickChange Site-Directed Mutagenesis. *BioTechniques*, **26**, 680–682.
32. Walden,H. (2010) Selenium incorporation using recombinant techniques. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 352–357.
33. Kabsch,W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
34. Tickle,I., Flensburg,C., Keller,P., Paciorek,W., Sharff,A., Vornrhein,C. and Bricogne,G. (2020) STARANISO. <http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi>.
35. Vornrhein,C., Blanc,E., Roversi,P. and Bricogne,G. (2007) Automated structure solution with autoSHARP. *Methods Mol. Biol.*, **364**, 215–230.
36. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
37. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
38. Murshudov,G.N., Skubák,P., Lebedev,A.A., Pannu,N.S., Steiner,R.A., Nicholls,R.A., Winn,M.D., Long,F. and Vagin,A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 355–367.
39. Bricogne,G., Blanc,E., Brandl,M., Flensburg,C., Keller,P., Paciorek,W., Pietro,Roversi, Sharff,A., Smart,O., Vornrhein,C. et al. (2017) Buster.
40. Kainov,D., Vitorino,M., Cavarelli,J., Poterszman,A. and Egly,J.-M. (2008) Structural basis for group A trichothiodystrophy. *Nat. Struct. Mol. Biol.*, **15**, 980–984.
41. Fregoso,M., Lainé,J.-P., Aguilar-Fuentes,J., Mocquet,V., Reynaud,E., Coin,F., Egly,J.-M. and Zurita,M. (2007) DNA repair and transcriptional deficiencies caused by mutations in the Drosophila p52 subunit of TFIIH generate developmental defects and chromosome fragility. *Mol. Cell Biol.*, **27**, 3640–3650.
42. Smurnyy,Y., Cai,M., Wu,H., McWhinnie,E., Tallarico,J.A., Yang,Y. and Feng,Y. (2014) DNA sequencing and CRISPR-Cas9 gene editing for target validation in mammalian cells. *Nat. Chem. Biol.*, **10**, 623–625.
43. Theil,A.F., Hoeijmakers,J.H.J. and Vermeulen,W. (2014) TTDA: big impact of a small protein. *Exp. Cell Res.*, **329**, 61–68.
44. Yan,C., Dodd,T., He,Y., Tainer,J.A., Tsutakawa,S.E. and Ivanov,I. (2019) Transcription preinitiation complex structure and dynamics provide insight into genetic diseases. *Nat. Struct. Mol. Biol.*, **26**, 397–406.
45. Radu,L., Schoenwetter,E., Braun,C., Marcoux,J., Koelmel,W., Schmitt,D.R., Kuper,J., Cianfèrani,S., Egly,J.M., Poterszman,A. et al. (2017) The intricate network between the p34 and p44 subunits is central to the activity of the transcription/DNA repair factor TFIIH. *Nucleic Acids Res.*, **45**, 10872–10883.
46. Uriostegui-Arcos,M., Aguayo-Ortiz,R., Valencia-Morales,M.D.P., Melchy-Pérez,E., Rosenstein,Y., Dominguez,L. and Zurita,M. (2020) Disruption of TFIIH activities generates a stress gene expression response and reveals possible new targets against cancer. *Open Biol.*, **10**, 200050.

47. Compe,E. and Egly,J.-M. (2016) Nucleotide excision repair and transcriptional regulation: TFIIH and beyond. *Annu. Rev. Biochem.*, **85**, 265–290.
48. Tomko,E.J., Fishburn,J., Hahn,S. and Galburt,E.A. (2017) TFIIH generates a six-base-pair open complex during RNAP II transcription initiation and start-site scanning. *Nat. Struct. Mol. Biol.*, **24**, 1139–1145.
49. Coin,F., Proietti De Santis,L., Nardo,T., Zlobinskaya,O., Stefanini,M. and Egly,J.-M. (2006) p8/TTD-A as a repair-specific TFIIH subunit. *Mol. Cell*, **21**, 215–226.
50. Theil,A.F., Nonnekens,J., Steurer,B., Mari,P.-O., de Wit,J., Lemaitre,C., Marteiijn,J.A., Raams,A., Maas,A., Vermeij,M. *et al.* (2013) Disruption of TTDA results in complete nucleotide excision repair deficiency and embryonic lethality. *PLoS Genet.*, **9**, e1003431.
51. Weber,A., Chung,H.-J., Springer,E., Heitzmann,D. and Warth,R. (2010) The TFIIH subunit p89 (XPB) localizes to the centrosome during mitosis. *Cell. Oncol.*, **32**, 121–130.
52. Mizuki,F., Namiki,T., Sato,H., Furukawa,H., Matsusaka,T., Ohshima,Y., Ishibashi,R., Andoh,T. and Tani,T. (2007) Participation of XPB/Ptr8p, a component of TFIIH, in nucleocytoplasmic transport of mRNA in fission yeast. *Genes Cells*, **12**, 35–47.
53. Egly,J.-M. and Coin,F. (2011) A history of TFIIH: two decades of molecular biology on a pivotal transcription/repair factor. *DNA Repair (Amst.)*, **10**, 714–721.
54. Sandoz,J. and Coin,F. (2017) Unified promoter opening steps in eukaryotic gene expression. *Oncotarget*, **8**, 84614–84615.
55. Li,C.-L., Golebiowski,F.M., Onishi,Y., Samara,N.L., Sugawara,K. and Yang,W. (2015) Tripartite DNA lesion recognition and verification by XPC, TFIIH, and XPA in nucleotide excision repair. *Mol. Cell*, **59**, 1025–1034.