Cross-linking mass spectrometric analysis of the endogenous TREX complex from *Saccharomyces cerevisiae*

CARINA KERN,[1,5] CHRISTIN RADON,[2,5] WOLFGANG WENDE,[1] ALEXANDER LEITNER,[3] and KATJA STRÄßER[1,4]

[1]Institute of Biochemistry, FB08, Justus Liebig University, 35392 Giessen, Germany
[2]Institute of Biochemistry and Biology, Department of Biochemistry, University of Potsdam, 14476 Potsdam-Golm, Germany
[3]Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, 8093 Zurich, Switzerland
[4]Cardio-Pulmonary Institute (CPI), EXC 2026, 35392 Giessen, Germany

ABSTRACT

The conserved TREX complex has multiple functions in gene expression such as transcription elongation, 3′ end processing, mRNP assembly and nuclear mRNA export as well as the maintenance of genomic stability. In *Saccharomyces cerevisiae*, TREX is composed of the pentameric THO complex, the DEAD-box RNA helicase Sub2, the nuclear mRNA export adaptor Yra1, and the SR-like proteins Gbp2 and Hrb1. Here, we present the structural analysis of the endogenous TREX complex of *S. cerevisiae* purified from its native environment. To this end, we used cross-linking mass spectrometry to gain structural information on regions of the complex that are not accessible to classical structural biology techniques. We also used negative-stain electron microscopy to investigate the organization of the cross-linked complex used for XL-MS by comparing our endogenous TREX complex with recently published structural models of recombinant THO–Sub2 complexes. According to our analysis, the endogenous yeast TREX complex preferentially assembles into a dimer.

Keywords: TREX; XL-MS; mass spectrometry; EM; nuclear mRNA export

INTRODUCTION

In eukaryotes, nuclear mRNP assembly is a key step of gene expression. Processing of the nascent mRNA occurs as it is synthesized by RNA polymerase II (RNAPII) and includes capping, splicing, and 3′ end formation. In parallel, the mRNA associates with multiple nuclear RNA-binding proteins (RBPs) that package it into a messenger ribonucleoprotein particle (mRNP) (Meinel and Strasser 2015; Singh et al. 2015; Wegener and Muller-McNicoll 2019; Wende et al. 2019; Khong and Parker 2020). Correct mRNP assembly is mutually dependent on mRNA processing and is a prerequisite for nuclear mRNA export. Furthermore, the composition of mRNPs can also affect later stages of the mRNA life cycle such as mRNA localization, translation, and degradation in the cytoplasm (Gehring et al. 2017).

A key player in nuclear mRNP assembly is the evolutionarily conserved TREX complex, which couples transcription to nuclear mRNA export (Jimeno et al. 2002; Strasser et al. 2002; Zenklusen et al. 2002). In *Saccharomyces cerevisiae*, the TREX complex is composed of the pentameric THO complex (Tho2, Hpr1, Mft1, Thp2, and Tex1), the ATP-dependent DEAD-box RNA helicase Sub2, the nuclear mRNA export adaptor Yra1 and the SR-like proteins Gbp2 and Hrb1 (Strasser et al. 2002; Hurt et al. 2004). In humans, the THO complex consists of THOC1 (Hpr1), THOC2 (Tho2), THOC3 (Tex1), THOC5 (structurally homologous to Thp2), THOC6 (not present in *S.c.*), and THOC7 (structurally homologous to Mft1) (Heath et al. 2016; Puhringer et al. 2020). In addition to THO, the human TREX complex contains UAP56/DDX39B (*S.c.* Sub2), ALYREF (*S.c.* Yra1), and several additional interaction partners such as SARNP (*S.c.* Tho1), ZC3H11A, IUF, LUZP4, POLDIP3, and CHTOP (Dufu et al. 2010; Heath et al. 2016). In addition to this high degree of conservation, the importance of TREX is also reflected by its implication in several human diseases (Heath et al. 2016).

TREX functions in transcription elongation, 3′ end processing, mRNP assembly and nuclear mRNA export

[5]These authors contributed equally to this work.

Corresponding authors: wolfgang.wende@chemie.bio.uni-giessen.de, leitner@imsb.biol.ethz.ch, katja.straesser@chemie.bio.uni-giessen.de

Article is online at http://www.rnajournal.org/cgi/doi/10.1261/rna.079758.123. Freely available online through the *RNA* Open Access option.

(Strasser et al. 2002; Zenklusen et al. 2002; Rougemaille et al. 2008). Furthermore, TREX—in part due to the DNA–RNA helicase activity of Sub2/UAP56 (Saguez et al. 2013; Perez-Calero et al. 2020)—prevents GC-rich mRNA elements from forming so-called R-loops, DNA–RNA hybrids of the nascent mRNA with the DNA template strand that cause hyperrecombination, DNA damage, and, ultimately, genome instability (Dominguez-Sanchez et al. 2011; Gomez-Gonzalez et al. 2011; Luna et al. 2019; San Martin-Alonso et al. 2021). In *S. cerevisiae*, the TREX complex is recruited to the site of transcription and the nascent mRNA by a number of interactions. THO is recruited by the carboxy-terminal domain of RNAPII, the Prp19 complex, Mud2, and RNA (Abruzzi et al. 2004; Chanarat et al. 2011; Meinel et al. 2013; Minocha et al. 2018). THO in turn recruits Sub2, which first adopts a semiclosed conformation and upon binding to mRNA and ATP changes to a closed state (Chen et al. 2021). Interaction of Sub2 with THO and Yra1 enhances its ATPase activity (Ren et al. 2017). Yra1 recruits the mRNA export receptor Mex67–Mtr2 to the mRNA, which exports the mRNA through the nuclear pore complex to the cytoplasm (Katahira et al. 1999; Strasser and Hurt 2000, 2001; Stutz et al. 2000; Zenklusen et al. 2001).

The molecular mechanism and the structural basis for these different TREX activities are still not completely understood. Ren et al. (2017) determined the first structural model of the hetero-pentameric THO complex with Sub2 at a resolution of 6.0 Å by X-ray crystallography. This structure is based on a truncated THO* core complex, in which potentially disordered regions were removed, and revealed an elongated architecture of THO with Sub2 bound in its center (Ren et al. 2017). More recently, high-resolution cyro-electron microscopy (EM) structures of human and yeast THO–Sub2 complexes with a resolution of up to 3.2 Å were reported (Puhringer et al. 2020; Schuller et al. 2020; Chen et al. 2021; Xie et al. 2021). The first cryo-EM reconstruction of a recombinant yeast THO–Sub2 complex was published by Schuller et al. (2020). This structure displays a dimeric architecture, consistent with the size-exclusion chromatogram. A similar cryo-EM study with heterologously expressed THO components and Sub2 was published by Chen et al. (2021). Structural insights into the interaction between THO and Gbp2 were recently obtained by combining cross-linking mass spectrometry (XL-MS) and cryo-EM (Xie et al. 2021). The obtained truncated THO*–Sub2 structure assembles in dimers as well as tetramers. The authors speculated that the tetramer is formed due to the truncations in THO* and focused on the dimeric structure for further analysis.

The arrangement of THO and Sub2 observed in the recombinant yeast structures closely resembles the cryo-EM structure of the recombinant human complex (Puhringer et al. 2020), although the human THO–UAP56 complex contains an additional protein, THOC6, which is not present in *S. cerevisiae*. Most recently, the structure of the endogenous human TREX complex was published (Pacheco-Fiallos et al. 2023), which overall is consistent with the structure of the recombinant complex (Puhringer et al. 2020). Both human complexes predominantly form a tetramer, although the recombinant THO–UAP56 complex also displays an octameric organization. Additionally, based on cryo-EM micrographs of TREX bound to the exon-junction complex and mRNAs, a globular architecture of human TREX–mRNP complexes was suggested (Pacheco-Fiallos et al. 2023).

The existing structures of the THO–Sub2 complex obtained by EM or crystallography suggest that the organization of the complex is conserved. However, all structures obtained from recombinantly expressed proteins, depending on complex preparation, exhibit different oligomeric states ranging from dimeric to octameric (Schuller et al. 2020; Chen et al. 2021; Xie et al. 2021).

Here, we report the structural analysis of the endogenous *S. cerevisiae* TREX complex by XL-MS, supported by negative-stain EM of the cross-linked complex, and compare it to the published recombinant yeast structures. Importantly, this is the first structural report on endogenously purified TREX complex from yeast with all nine subunits. This approach enables us to capture the TREX complex in different dynamic states and conformations that occur during transcription, mRNA processing and packaging as well as nuclear export. These dynamic TREX conformations are functionally relevant, but are not visible when analyzing the recombinant TREX complex. Chemical cross-linking experiments revealed that <40% of the cross-links occur in the conformationally rigid regions that are resolved in the recently reported structures of recombinant THO–Sub2 complexes. Instead, many cross-links are observed in flexible regions, in line with a recent study that also reported XL-MS combined with cryo-EM for the analysis of the recombinant THO–Sub2 complex (Xie et al. 2021). Our structural analysis of the cross-linked native TREX complex using negative-stain EM reveals that the endogenous *S. cerevisiae* TREX complex assembles into a dimer.

RESULTS

Purification of the endogenous TREX complex from *S. cerevisiae*

To obtain the *S. cerevisiae* TREX complex for structural analysis, we endogenously tagged its subunit Hpr1 with a carboxy-terminal FLAG-TEV-Protein A tag and purified the native complex by tandem affinity purification. A representative purification of TREX is shown in Figure 1A. As expected, the THO subunits Tho2, Mft1, Thp2, and Tex1 copurified with Hpr1 as well as the TREX components Sub2, Yra1, Gbp2, and Hrb1 (Fig. 1A,B). The identity of

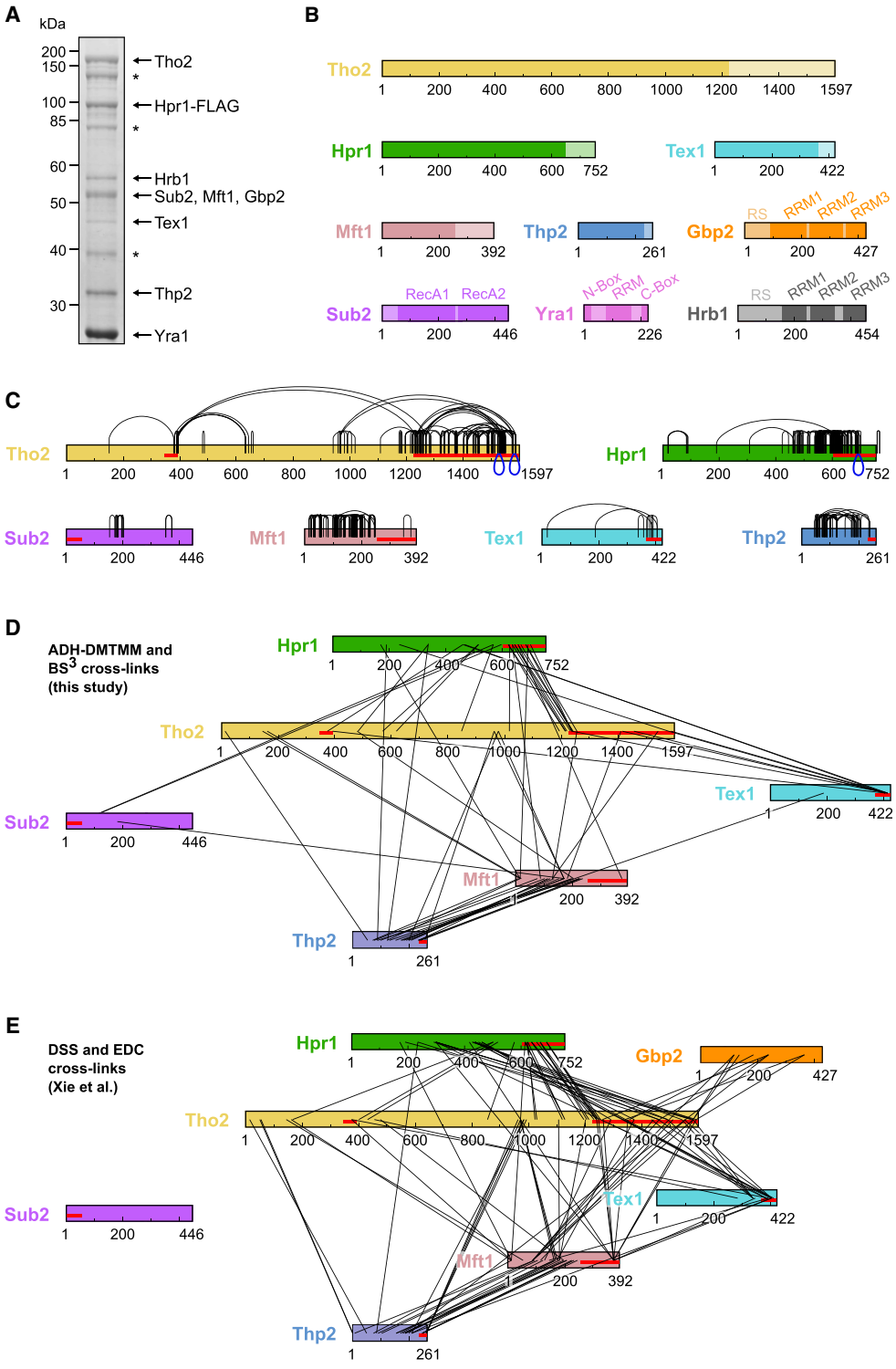

FIGURE 1. Visualization of cross-links on the TREX complex. (*A*) SDS–PAGE of the purified endogenous *S. cerevisiae* TREX complex. Subunits corresponding to Tho2, Hpr1-FLAG, Hrb1, Sub2, Mft1, Gbp2, Tex1, Thp2, and Yra1 were identified by mass spectrometry and are indicated accordingly. The asterisks indicate degradation bands. Molecular weight marker bands are indicated in kDa. (*B*) Scheme of the domain structures of TREX subunits. Flexible domains that are not observed in the solved structures are depicted in a lighter color. (*C–E*) Cross-links connecting residues within the same protein (*C*) and cross-links connecting residues of two different proteins (*D*) mapped on the sequence graphs of the corresponding proteins as black arcs and black lines, respectively. Segments longer than 25 residues that are missing in the PDB ID: 7apx structural model are shown as red lines on the sequence graph. Visualization was performed using xiNET (Combe et al. 2015). (*C*) Cross-linking data from this work. Cross-links that unambiguously correspond to intermolecular contacts between two molecules of the same protein are shown in blue as downward-facing loops. (*D*) Cross-links identified in this work. (*E*) Cross-links identified by Xie et al. (2021).

all these components was confirmed by mass spectrometric analysis. Interestingly, Yra1 is more abundant than the other TREX components (Fig. 1A). This overstoichiometric representation of Yra1 was recently also observed in TREX complexes consecutively purified via Hpr1 and Sub2, the direct interaction partner of Yra1 (Bonneau et al. 2023; published, while this manuscript was under review). Following purification, the complex was directly cross-linked either with BS^3, a homobifunctional cross-linker with amine-reactive *N*-hydroxysulfosuccinimide esters at each end of a 6-carbon-spacer-arm (11.4 Å), or with adipic dihydrazide (ADH-d_0/d_8, Sigma-Aldrich) and 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium (DMTMM) chloride. The endogenously purified, cross-linked TREX complex was then analyzed by XL-MS and negative-stain EM to gain insights into its structural arrangement.

Cross-linking mass spectrometry analysis of endogenously purified *S. cerevisiae* TREX complex

To investigate the molecular architecture of the endogenous yeast TREX complex, we first used XL-MS to map intra- and intermolecular contacts within the complex. Chemical cross-linking yields residue-specific amino acid contacts by forming covalent bonds between reactive side chains that are in spatial proximity (typically <30–35 Å apart). To obtain complementary information, we used two different cross-linking protocols. Cross-linking with BS^3, which predominantly reacts with primary amines on lysines and protein amino termini, yielded many cross-links in the flexible regions of TREX components. As a second cross-linking approach, we used a combination of ADH and DMTMM chloride. This cross-linking chemistry results in two types of cross-linking products, one incorporating the dihydrazide linker, connecting carboxylic groups (aspartic or glutamic acids), the other directly coupling amines to carboxyl groups activated by the reaction with DMTMM ("zero-length" cross-links) (Leitner et al. 2014a).

Using both cross-linking chemistries, we were able to obtain more than 550 cross-linked peptide pairs, corresponding to more than 445 nonredundant site pairs on the complex subunits (THO complex and Sub2), with the majority contributed by BS^3 and DMTMM links (Fig. 1; Supplemental Table S1). Although Yra1, Hrb1, and Gbp2 were present in the analyzed sample, no cross-links between these proteins and any of the core THO subunits were observed. This finding could reflect, for example, a suboptimal distribution of cross-linking sites at the respective binding interfaces, a reduced reactivity of such sites due to limited accessibility to the cross-linking reagents or the presence of these proteins at substoichiometric levels. Cross-links on individual subunits were heterogeneously distributed, with the largest numbers found on the carboxy-terminal regions of Tho2 and Hpr1, the first approximately 250 residues of Mft1 and on Thp2 (Fig. 1C). Note that cross-links between residues on the same protein may occur within one or between two copies of this protein. Cross-links connecting different subunits were most frequently observed between Tho2 and Hpr1 as well as between Mft1 and Thp2 (Fig. 1D).

Comparison of the cross-linking data with data from recombinant yeast TREX structures

Recently, high-resolution structures of reconstituted yeast THO–Sub2 complexes have been obtained in independent studies. To assess the compatibility of our cross-linking data of the endogenous TREX complex with the reported structures, we first mapped the cross-links depicted in Figure 1D onto the 3.4 Å cryo-EM structure from Schuller et al. (2020) (PDB ID: 7apx). Interestingly, for <40% of these cross-links (171 of 445), both of the linked residues were contained in the structure; for the remaining cross-links either one or both of the connected residues were not resolved in the structure. This is especially apparent in the carboxy-terminal part of Tho2, for which more than 350 residues are missing, and for Hpr1, for which the model lacks approximately 150 residues on the carboxy-terminal side (Fig. 1). For instance, although the unstructured carboxy-terminal region of Tho2 is not necessary for TREX complex formation, it plays a crucial role in promoting cell growth under reduced and elevated temperatures (Supplemental Fig. S1). Unfortunately, for both proteins the rich cross-linking data in the flexible regions do not allow modeling of the full-length protein with reasonable confidence, because the flexible regions have only relatively few contacts to the more ordered regions of the proteins or to other subunits.

For the subset of cross-links that map to the structure, ~78% of the contacts (133 of 171) were compatible with the monomeric unit with a maximal distance of 35 Å, and an additional three cross-links matched only to subunit contacts between two monomer units in the dimer. However, 35 of the 38 remaining site pairs also do not fit the dimer structure (PDB ID: 7aqo). The incompatible cross-links mostly reside on the subunits Thp2 and Mft1 or connect these two subunits. It should be pointed out that these cross-links were observed in multiple independent experiments and with different cross-linking chemistries. Moreover, the scores of most identifications were high so that we consider it unlikely that they correspond to false positive assignments. Therefore, we assume that these cross-links reflect different conformational or compositional states present in the native complex preparation, as we were not able to model any assembly state (monomeric, dimeric, or tetrameric) that would be compatible with all cross-linking data.

We furthermore compared our cross-linking data set (Fig. 1D) to the cross-linking results of Xie et al. (Fig. 1E). Xie et al.

(2021) also used XL-MS to confirm intra- and inter-molecular contacts in a reconstituted yeast THO–Sub2–Gbp2 complex, in their case using the amine-reactive disuccinimidyl suberate (DSS), a nonsulfonated analog of BS^3, and the zero-length coupling reagent *N*-(3-Dimethylaminopropyl)-*N*′-ethylcarbodiimide (EDC). Importantly, our analysis of the endogenous complex might reflect the capture of multiple conformations of the highly dynamic TREX complex. Compared to our data set, a larger fraction (243 of 314, ~75%) of the DSS/EDC links obtained by Xie et al. could be mapped to the structure determined by Schuller et al. (2020). In contrast to our data from the natively purified endogenous complex, the amino acid contacts that were covered in the structure were compatible with the 35 Å threshold. The higher fraction of compatible cross-links probably reflects the higher structural homogeneity for the recombinant complex. Figure 2 compares the inter-protein cross-links observed in both studies and shows that the most frequently covered subunit–subunit contacts are quite similar.

Structure determination of the cross-linked TREX complex by negative-stain EM

As the majority (over 60%) of the cross-links observed in our XL-MS analysis are located in regions that are not resolved in the recently published THO–Sub2 structures, for example, Schuller et al. (2020), we next examined whether the formation of the endogenous TREX complex differs from that of the recombinant THO–Sub2 complex. We were especially interested in the preferred oligomeric state of the endogenous TREX complex purified from its native environment, as the XL-MS data suggest different conformational or compositional states present in the sample. We therefore subjected a freshly BS^3 cross-linked complex to uranyl acetate negative-stain (UANS) EM and single-particle analysis (Supplemental Fig. S2). The sample showed a heterogeneous complex formation on the negative-stain grid (Fig. 3A); however, distinct dimer-like particles similar to previously published 2D classes were observed (Fig. 3B; Pena et al. 2012; Schuller et al. 2020; Xie et al. 2021). To estimate the fraction of dimer-like particles in the UANS data set, we compared reference-free 2D class averages to the published dimer map of the yeast THO–Sub2 complex by Schuller et al. using image cross-correlation techniques (Supplemental Fig. S2A). When opting for a cross-correlation coefficient cut-off value that includes distinct dimer views into the data set, the dimer-correlated fraction in the UANS data set is 50%, with another 27% at coefficients between 0.897 and 0.85. We thus conclude that the majority of particles in the data set have a dimer-like appearance. Subsequently, the dimer fraction was used to generate a 3D reconstruction without any symmetry applied, yielding an overall resolution of 26 Å (Fig. 3C). Comparison of the previously published yeast THO–Sub2 dimer (Schuller et al. 2020) with our 3D reconstruction or the reference-free class averages reveals a strong similarity between the two complexes (Fig. 3D,E; Supplemental Fig. S2A). At the resolution of the UANS data set, the endogenously purified *S. cerevisiae* TREX sample resembles the high-resolution structure of the in vitro assembled complex in terms of overall shape and order with small differences in domain orientations (Fig. 3E,F).

DISCUSSION

The TREX complex is essential for gene expression and plays crucial roles in transcription elongation, 3′ end processing, mRNP assembly, nuclear export, and genome stability. Recently, several high-resolution structures of the human and *S. cerevisiae* THO–Sub2 complex were published that elucidate TREX function from a structural perspective (Puhringer et al. 2020; Schuller et al. 2020; Chen et al. 2021; Xie et al. 2021). In most of these works, protein subunit purification and in vitro complex reconstitution were optimized to yield highly homogeneous complexes, for example, by truncation of flexible regions and a combination of prokaryotic and eukaryotic expression

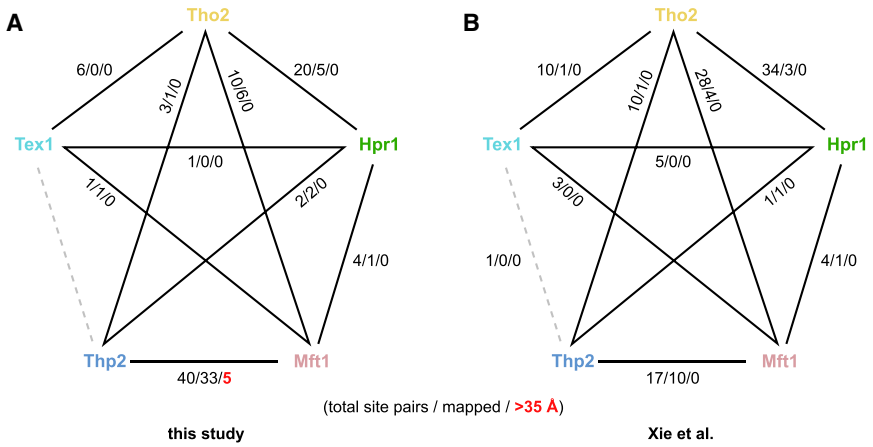

FIGURE 2. Comparison of intersubunit cross-links. (*A*) observed in our work, excluding contacts with Sub2, and (*B*) from cross-linking of the recombinant complex by Xie et al. (2021). Shown are the numbers of nonredundant cross-linking site pairs observed in total for each subunit pair, the number of contacts that could be mapped onto PDB ID: 7apx and therefore connect ordered regions resolved by cryo-EM, and (in red) those that exceed the distance of 35 Å chosen here as an upper threshold for cross-linkable distances. Proteins are color-coded as in Figure 1. A structural model of the dimeric TREX complex (PDB ID: 7aqo) is shown in Figure 3F.

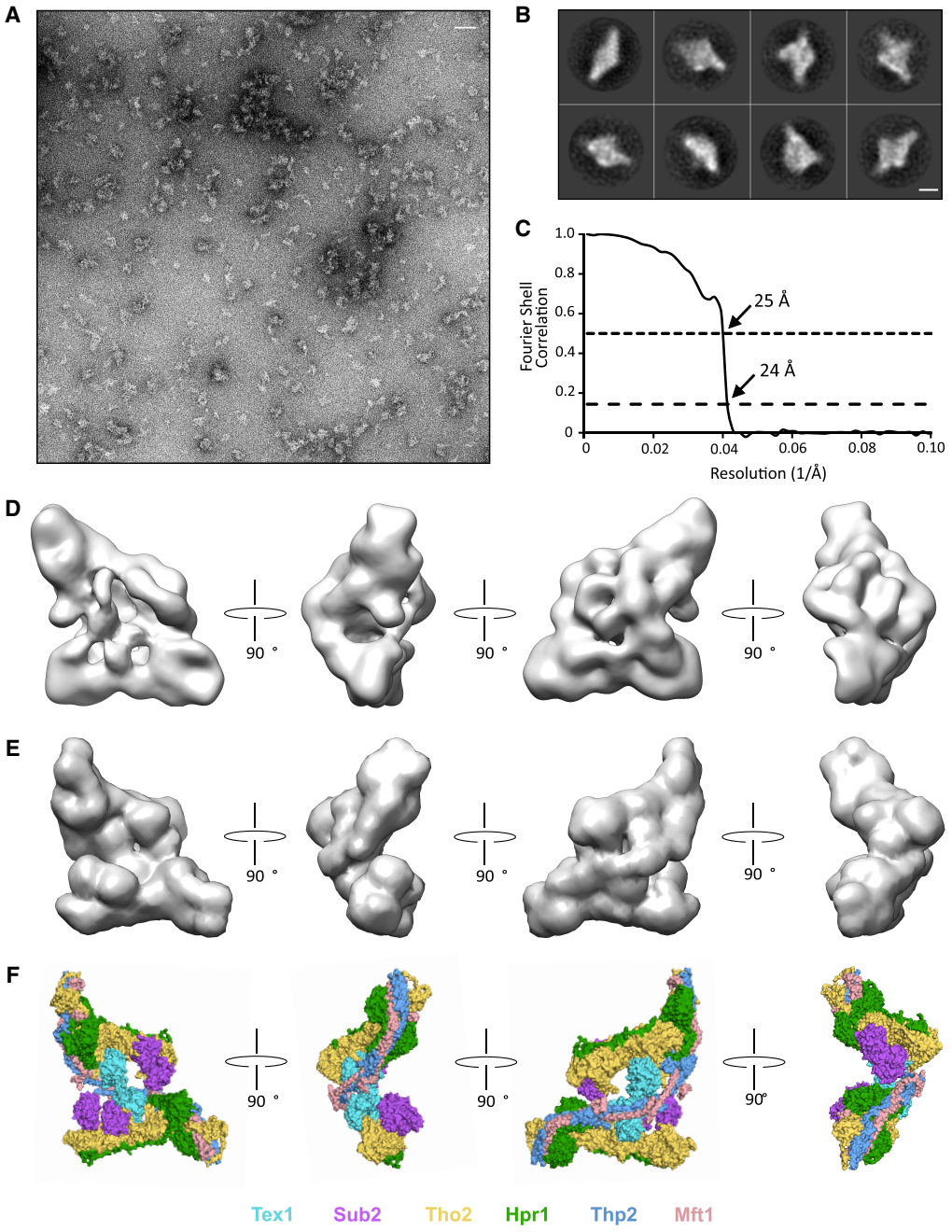

FIGURE 3. Negative-stain EM structure of the endogenous TREX complex from *S. cerevisiae*. (*A*) A representative negative-stain micrograph of the BS^3 cross-linked *S. cerevisiae* TREX complex. Scale bar represents 500 Å. (*B*) Representative reference-free 2D class averages of the *S.c.* TREX complex. Scale bar represents 100 Å. (*C*) Fourier shell correlation (FSC) plot for the negative-stain reconstruction of the *S.c.* TREX complex. The resolution at FSC = 0.143 (dashed line) and FSC = 0.5 (dotted line) are indicated. (*D*) Three-dimensional reconstruction of the *S.c.* TREX complex resolved to 26 Å. (*E*) Representative views of the Tho–Sub2 dimer structure (PDB ID: 7aqo) in surface representation lowpass-filtered to 25 Å. (*F*) Color-coded THO–Sub2 dimer PDB model depicted in the same views as in (*E*). Proteins are color-coded as in Figure 1.

systems. However, although the resulting structural models are highly similar with respect to the monomeric complexes, they exhibit different oligomeric states, raising the concern that the recombinant structures may not reflect complex organization in the native cellular environment of yeast. Overall, our XL data of the native TREX complex of *S. cerevisiae* support the structures of the recombinant TREX subcomplex and thus confirms the validity of the recent high-resolution structures for the interpretation of in vivo functional data.

Interestingly, the majority of our cross-linked residues (60%) are located in flexible regions that are not resolved

in the recently reported cryo-EM structures of the recombinant THO–Sub2/UAP56 complexes (Puhringer et al. 2020; Schuller et al. 2020; Chen et al. 2021; Xie et al. 2021), similar to a recent study that also used XL-MS for the analysis of the recombinant THO–Sub2 complex (Xie et al. 2021). In particular, this observation concerns the carboxy-terminal tails of Tho2, Hpr1, and—to a lesser extent—Tex1 (Fig. 1). While the presence of these flexible regions is not required for the formation of the TREX complex (Supplemental Fig. S1A; Xie et al. 2021), they are involved in the packaging and export of mRNPs. For instance, the carboxy-terminal domain of Hpr1 is ubiquitylated leading to faster degradation and interacts with the mRNA export receptor Mex67 (Gwizdek et al. 2005, 2006). Furthermore, our findings demonstrate that cells lacking the carboxyl terminus of Tho2 exhibit a significant growth impairment at both low (16°C) and high (37°C) temperatures (Supplemental Fig. S1B–D; Pena et al. 2012). These results highlight the critical role of the flexible Tho2 region in regulating mRNP biogenesis under stress conditions. This is supported by studies of the human Tho2 homolog, as carboxy-terminally truncated THOC2 mutants contribute to neurodevelopmental disorders (Kumar et al. 2018).

The observed cross-linking pattern is not untypical for flexible regions; however, according to our experience, the strong preference for these regions of the complex is unusual. Furthermore, we were unable to find conditions that resulted in a more balanced distribution of cross-links over the entire surface. XL-MS is frequently used to provide additional structural information on regions that are not accessible to classical structural biology techniques, as the distance restraints can be used as input for integrative/hybrid modeling approaches (Braitbard et al. 2019; Rout and Sali 2019; Koukos and Bonvin 2020). However, in this case, most of the intra-molecular cross-link contacts connect regions that are relatively close to each other in the primary sequence (i.e., within the flexible regions), and with relatively few contacts to more ordered regions of the complex, so that a restraint-based modeling approach is unlikely to be successful.

Among the cross-links covered by reported structures of the recombinant complex, close to 80% are compatible with a maximal distance of 35 Å between the involved residues. The remaining 20% could either be the result of cross-linking induced artifacts, or they could reflect other subunit arrangements in the natively purified complex. However, we argue against the former explanation for several reasons. First, cross-linking conditions were optimized to avoid excessive cross-linking. For example, the low reagent concentrations for the ADH + DMTMM reaction resulted in relatively few cross-links incorporating the ADH linker, as expected from recent observations (Mohammadi et al. 2021), and favored DMTMM-induced zero-length cross-links. Xie et al. used a DSS cross-linker concentration identical to our BS^3 conditions (0.5 mM), and their results show a better compatibility with the model from Schuller et al. (2020). Given that DSS and BS^3 are comparable in their reactivity, this also points to a limited influence of cross-linking conditions, and rather argues for a conformational difference of at least a subset of the respective purified complexes. Such differences in composition or conformation can be enhanced or even distorted in XL-MS experiments as contacts from lowly populated states may still result in high-confidence identifications if such cross-linked peptide pairs ionize and fragment well.

The negative-stain EM structure of the cross-linked endogenous *S. cerevisiae* TREX complex presented here corresponds in overall appearance reasonably well to the published cryo-EM structures of the recombinant THO–Sub2 complex (Schuller et al. 2020). Our endogenous TREX complex is predominantly dimeric and supports the findings for the recombinant THO–Sub2 complex, although the endogenously purified complex we model here additionally contains Hrb1, Gbp2, Yra1, and the carboxy-terminal 42 residues of Tex1 that were truncated in the recombinantly purified complexes. Small differences in domain orientation and extra densities between the native and recombinant structures are most likely due to the additional proteins, which unfortunately cannot be located without doubt at this resolution. However, the additional proteins present in our purification do not seem to cause the complex to adopt major alternative conformations.

Pacheco-Fiallos et al. (2023) showed that the endogenous human TREX complex is structurally consistent with the recombinant human THO–UAP56 complex (Puhringer et al. 2020). The only structure of endogenous, fully assembled yeast TREX complex so far shows an elongated, mainly monomeric assembly in negative-stain EM (Pena et al. 2012), which only partially fits our dimeric reconstruction as the dimer arrangement described could not be observed in our data set. Intriguingly, one protomer of the recombinant THO–Sub2 dimer complex (Schuller et al. 2020) fits well into our EM density, while the second protomer leaves some unfilled density. This observation could hint toward additional interactions with TREX subunits identified in our endogenously purified complex by MS (Yra1, Gbp2, Hrb1) that are not present in the structure of the heterologously expressed THO–Sub2 (Schuller et al. 2020; Xie et al. 2021). However, the unequal density distribution could also be due to differing stoichiometries of TREX components. Nevertheless, apart from these minor differences, we could not identify a major alternative conformation that would explain the 35 violating cross-links of our XL-MS analysis that were not compatible with the dimer structure from Schuller et al. (2020) or other assembly states like monomer or tetramer. We thus suggest that these cross-links could derive either from aggregated protein or from assembly intermediates of TREX complexes that are only present in the natively purified sample.

In summary, based on our XL-MS and negative-stain EM analyses, the structure of the native TREX complex of *S. cerevisiae* corresponds to the structures determined for the recombinant THO–Sub2 complex confirming the validity of the recent high-resolution structures for the interpretation of in vivo functional data.

MATERIALS AND METHODS

Purification of the native yeast TREX complex

A sequence coding for a FLAG-TEV-2xProtein-A (FTpA) tandem affinity tag was inserted carboxy-terminally into the endogenous *HPR1* locus of the *S. cerevisiae* strain RS453 (MATa) using standard yeast genetic techniques. Yeast was grown in 30 l YPD to $OD_{600} = 3.5$, harvested by centrifugation and frozen in liquid nitrogen. Frozen yeast pellets were lysed with a cryo-mill (SPEX) and stored at −80°C. The grindate was resuspended in buffer A (50 mM HEPES, pH 7.5; 200 mM KCl; 1.5 mM $MgCl_2$; 0.15% IGEPAL) containing protease inhibitor mix (2 µM pepstatin A, 0.6 µM leupeptin, 1 mM PMSF, and 2 mM benzamidine). The lysate was precleared at 4200*g* for 10 min at 4°C, followed by ultracentrifugation of the supernatant at 132,600*g* for 1 h at 4°C. The clear supernatant was transferred to four 50 mL conical centrifuge tubes. Each tube was rotated with 600 µL IgG beads (Cytiva) for 16–18 h at 4°C. Beads were separated from the lysate by centrifugation at 700*g* for 3 min at 4°C and washed twice with buffer A containing 1 mM DTT. For elution, beads were rotated with TEV protease in buffer A for 1 h at 16°C. Following centrifugation, the resulting supernatant (TEV eluate) was rotated with 450 µL FLAG beads (Sigma-Aldrich) in buffer B (20 mM HEPES, pH 8.4; 150 mM KCl) or buffer C (20 mM HEPES, pH 7.0; 150 mM KCl) for 1 h at 4°C. Beads were washed once with buffer B or C containing 0.15% IGEPAL. For elution, the sample was rotated with 100 µg/mL FLAG peptide (Sigma-Aldrich) for 15 min at room temperature. The purified *S.c.* TREX complex was concentrated and subsequently cross-linked with either bis(sulfosuccinimidyl)suberate (BS^3-d_0/d_{12}, Creative Molecules) in buffer B or with adipic dihydrazide (ADH-d_0/d_8, Sigma-Aldrich) and 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium (DMTMM) chloride (Sigma-Aldrich) in buffer C for 1 h at 37°C and 500 rpm using a ThermoMixer (Eppendorf). The BS^3 cross-linking reaction was stopped by incubation with 50 mM ammonium bicarbonate (NH_4HCO_3) for 20 min at 37°C at 500 rpm in a ThermoMixer. The ADH/DMTMM cross-link reaction was stopped by desalting with a Zeba Spin Desalting Column (ThermoFisher Scientific). The cross-linked samples were flash frozen in liquid nitrogen and stored at −80°C until further processing. For the XL-MS analyses, 40 µg purified *S.c.* TREX complex each were cross-linked in three independent experiments with (i) 0.5 mM BS^3, (ii) 1 mM ADH and 1 mM DMTMM, or (iii) 2.5 mM ADH and 2.5 mM DMTMM. For EM analysis, 20 µg *S.c.* TREX sample was cross-linked with BS^3.

Sample processing for mass spectrometry

After quenching or gel filtration, cross-linked samples were evaporated to dryness in a vacuum centrifuge and processed essentially as described previously (Leitner et al. 2014b). In brief, disulfide bonds were reduced with tris(2-carboxyethyl)phosphine, and free thiol groups in cysteines were alkylated with iodoacetamide. Proteins were digested in a two-step procedure, first with endoproteinase Lys-C (Wako; 1:100, 37°C for 2–3 h), then with trypsin (Promega; 1:50, 37°C overnight). Samples were purified by solid-phase extraction (SepPak tC18 cartidges, Waters) and fractionated by peptide-level size-exclusion chromatography (SEC; Superdex Peptide PC 3.2/300, GE) on an ÄKTA micro FPLC system (Leitner et al. 2012).

Liquid chromatography–tandem mass spectrometry

SEC fractions and additionally some unfractionated samples were analyzed by nanoflow LC–MS/MS. The BS^3 sample was initially analyzed on an Orbitrap Elite system and later reanalyzed on an Orbitrap Fusion Lumos instrument (both ThermoFisher Scientific), while the ADH/DMTMM samples were only analyzed on the latter instrument.

The Orbitrap Elite instrument was connected to an Easy nLC-1000 HPLC system (ThermoFisher Scientific) with a 150 mm × 75 µm Acclaim PepMap RSLC C_{18} column (2 µm particle size, ThermoFisher Scientific) and mobile phases A = water/acetonitrile/formic acid (98:2:0.15, v/v/v) and B = acetonitrile/water/formic acid (98:2:0.15, v/v/v). Peptides were separated at a flow rate of 300 nL/min with a gradient of 9%–35% B in 60 min. The mass spectrometer was operated in data-dependent acquisition mode with the detection of precursor ions in the Orbitrap at 120,000 resolution and the detection of fragment ions in the linear ion trap at normal resolution. For each sequencing cycle, the 10 most abundant precursor ions (top n mode) were selected for sequencing if they had a charge state of +3 or higher and were fragmented at 35% normalized collision energy. Dynamic exclusion was activated for 30 sec after one sequencing event.

The Orbitrap Fusion Lumos mass spectrometer was connected to an Easy nLC-1200 HPLC instrument (ThermoFisher Scientific) with a 250 mm × 75 µm Acclaim PepMap RSLC C_{18} column (2 µm particle size, Thermo Fisher Scientific) and mobile phases A = water/acetonitrile/formic acid (98:2:0.15, v/v/v) and B = acetonitrile/water/formic acid (80:20:0.15, v/v/v). Peptides were separated at a flow rate of 300 nL/min with a gradient of 11%–40% B in 60 min. The mass spectrometer was operated in data-dependent acquisition mode with detection of precursor ions in the Orbitrap at 120,000 resolution and detection of fragment ions in the linear ion trap at rapid resolution (low-resolution MS/MS) or in the Orbitrap at 30,000 resolution (high-resolution MS/MS). For each sequencing cycle, the most abundant precursor ions were selected for sequencing with a cycle time of 3 sec (top speed mode) if they had a charge state of +3 or higher and were fragmented at 35% normalized collision energy. Dynamic exclusion was activated for 30 sec after one sequencing event.

Identification of cross-linked peptides

MS/MS data acquired in the native Thermo.raw format were converted into the mzXML format by msconvert/ProteoWizard (Chambers et al. 2012) and searched using xQuest (Walzthoeni et al. 2012), version 2.1.4, against a custom database of target proteins (subunits of the TREX complex) and identified

contaminants. The false discovery rate (FDR) was controlled by a target/decoy search using reversed and shuffled sequences as decoys. xQuest search parameters included: enzyme = trypsin, maximum number of missed cleavages = 2, fixed modifications = carbamidomethylation on Cys, initial MS1 error tolerance = ±15 ppm, MS2 error tolerance = 0.2 Da for "common" fragment ions and 0.3 Da for "xlink" fragment ions (ion trap) or = ±20 ppm (Orbitrap), respectively. Cross-linking specificity was set to K and amino terminus for BS^3, D, or E for ADH, and K with D or E for DMTMM. Primary search results (target and decoy hits) were then further filtered according to the experimentally observed mass accuracy, a TIC value of >0.1 (for BS^3 or ADH) or >0.15 (for DMTMM) and a ΔS value of <0.9. All remaining candidate hits were further curated and only assignments for which a minimum of four bond cleavages overall or three consecutive bond cleavages were assigned for each peptide were retained. The FDR was adjusted to 5% at the cross-linked peptide pair level after the manual curation steps.

Structural validation of cross-links

Cross-linking data from our own experiments or retrieved from published work (Xie et al. 2021) were mapped on the structure of the THO–Sub2 complex from Schuller et al. (2020) (monomer: PDB ID: 7apx, dimer: PDB: 7aqo) using PyXlinkViewer (Schiffrin et al. 2020). An upper bound Euclidean distance of 35 Å (C_α–C_α) was selected for "compatible" cross-links. Visualization of cross-links for Figure 2 was performed using xiNET (Combe et al. 2015).

Negative-stain electron microscopy, data collection, and data processing

An amount of 3.5 µL of BS^3 cross-linked *S.c.* TREX complex (c = 60 µg/mL) was applied to a freshly glow-discharged carbon-coated copper mesh grid and stained two times with 2% uranyl acetate. Negative-stain EM images were collected on a Talos F200C TEM microscope (Thermo Fisher) operated at 200 kV equipped with a Ceta 16M camera (Thermo Fisher). Five hundred and ninety-nine micrographs were recorded at a magnification of 54 kx at a pixel size of 2.54 Å using EPU. The defocus range was set from 1.5 to 2.5 µm.

Image processing and 3D reconstruction were performed using Relion-3.0 (Zivanov et al. 2018). Contrast transfer functions were determined using CTFFIND4 (Rohou and Grigorieff 2015). All refinements used gold-standard FSC calculations and reported resolutions are based on the FSC = 0.143 criterion. Two-dimensional classification of 630 manually picked particles generated templates for semiautomated particle selection. The data set of 88,543 picked particles was subjected to several rounds of unsupervised 2D classification to remove bad particles resulting in a final data set of 42,784 particles. The 2D class averages of the 42,784 particle set and a Molmap of PDB ID: 7aqo lowpass-filtered to 25 Å were subjected to projection matching in IMAGIC (van Heel and Keegstra 1981) to analyze dimer representation in the data set. Only 2D class averages with a cross-correlation coefficient larger 0.897 were selected resulting in a final particle set of 22,024 particles. Selected particles were used for a 3D refinement in CryoSPARC v4 (Punjani et al. 2017). The EM map of the endogenous *S. cerevisiae* TREX complex was deposited under accession code EMD-17464.

DATA DEPOSITION

All XL-MS data were deposited to the PRIDE repository (Perez-Riverol et al. 2019) (PXD025843), and all identified cross-linked peptide pairs are listed in Supplemental Table S1. The EM map of the endogenous *S. cerevisiae* TREX complex was deposited under accession code EMD-17464.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank Manuel Koschitza for initial experiments for this project, Birte Keil for purification experiments of THO truncation mutants, and Cornelia Kilchert and Vera Bettenworth for critically reading the manuscript. A.L. would like to thank Ruedi Aebersold and Paola Picotti for access to laboratory infrastructure and instrumentation. EM data were recorded on the Talos F200 supported by the DFG under INST 336/148-1 FUGG. We thank Petra Wendler for the helpful discussions and for providing microscope data. C.R. thanks Andreas Heimann for IT support. This work was supported by the European Union (European Research Council Consolidator Grant, grant number 772049 to K.S. and Grant ULTRA-DD, grant number FP7-JTI 115766) and ETH Zurich (Scientific Equipment Grants).

Received June 23, 2023; accepted August 22, 2023.

REFERENCES

Abruzzi KC, Lacadie S, Rosbash M. 2004. Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes. *EMBO J* **23:** 2620–2631. doi:10.1038/sj.emboj.7600261

Bonneau F, Basquin J, Steigenberger B, Schafer T, Schafer IB, Conti E. 2023. Nuclear mRNPs are compact particles packaged with a network of proteins promoting RNA-RNA interactions. *Genes Dev* **37:** 505–517. doi:10.1101/gad.350630.123

Braitbard M, Schneidman-Duhovny D, Kalisman N. 2019. Integrative structure modeling: overview and assessment. *Annu Rev Biochem* **88:** 113–135. doi:10.1146/annurev-biochem-013118-111429

Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, et al. 2012. A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol* **30:** 918–920. doi:10.1038/nbt.2377

Chanarat S, Seizl M, Strasser K. 2011. The Prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes. *Genes Dev* **25:** 1147–1158. doi:10.1101/gad.623411

Chen C, Tan M, Wu Z, Zhang Y, He F, Lu Y, Li S, Cao M, Li G, Wu J, et al. 2021. Structural and functional insights into R-loop prevention and mRNA export by budding yeast THO-Sub2 complex. *Sci Bull* **66:** 2347–2352. doi:10.1016/j.scib.2021.08.004

Combe CW, Fischer L, Rappsilber J. 2015. xiNET: cross-link network maps with residue resolution. *Mol Cell Proteomics* **14:** 1137–1147. doi:10.1074/mcp.O114.042259

Dominguez-Sanchez MS, Barroso S, Gomez-Gonzalez B, Luna R, Aguilera A. 2011. Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet* **7:** e1002386. doi:10.1371/journal.pgen.1002386

Dufu K, Livingstone MJ, Seebacher J, Gygi SP, Wilson SA, Reed R. 2010. ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. *Genes Dev* **24:** 2043–2053. doi:10.1101/gad.1898610

Gehring NH, Wahle E, Fischer U. 2017. Deciphering the mRNP code: RNA-bound determinants of post-transcriptional gene regulation. *Trends Biochem Sci* **42:** 369–382. doi:10.1016/j.tibs.2017.02.004

Gomez-Gonzalez B, Garcia-Rubio M, Bermejo R, Gaillard H, Shirahige K, Marin A, Foiani M, Aguilera A. 2011. Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. *EMBO J* **30:** 3106–3119. doi:10.1038/emboj.2011.206

Gwizdek C, Hobeika M, Kus B, Ossareh-Nazari B, Dargemont C, Rodriguez MS. 2005. The mRNA nuclear export factor Hpr1 is regulated by Rsp5-mediated ubiquitylation. *J Biol Chem* **280:** 13401–13405. doi:10.1074/jbc.C500040200

Gwizdek C, Iglesias N, Rodriguez MS, Ossareh-Nazari B, Hobeika M, Divita G, Stutz F, Dargemont C. 2006. Ubiquitin-associated domain of Mex67 synchronizes recruitment of the mRNA export machinery with transcription. *Proc Natl Acad Sci* **103:** 16376–16381. doi:10.1073/pnas.0607941103

Heath CG, Viphakone N, Wilson SA. 2016. The role of TREX in gene expression and disease. *Biochem J* **473:** 2911–2935. doi:10.1042/BCJ20160010

Hurt E, Luo MJ, Rother S, Reed R, Strasser K. 2004. Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex. *Proc Natl Acad Sci* **101:** 1858–1862. doi:10.1073/pnas.0308663100

Jimeno S, Rondon AG, Luna R, Aguilera A. 2002. The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *EMBO J* **21:** 3526–3535. doi:10.1093/emboj/cdf335

Katahira J, Strasser K, Podtelejnikov A, Mann M, Jung JU, Hurt E. 1999. The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *EMBO J* **18:** 2593–2609. doi:10.1093/emboj/18.9.2593

Khong A, Parker R. 2020. The landscape of eukaryotic mRNPs. *RNA* **26:** 229–239. doi:10.1261/rna.073601.119

Koukos PI, Bonvin A. 2020. Integrative modelling of biomolecular complexes. *J Mol Biol* **432:** 2861–2881. doi:10.1016/j.jmb.2019.11.009

Kumar R, Gardner A, Homan CC, Douglas E, Mefford H, Wieczorek D, Ludecke HJ, Stark Z, Sadedin S, Broad CMG, et al. 2018. Severe neurocognitive and growth disorders due to variation in THOC2, an essential component of nuclear mRNA export machinery. *Hum Mutat* **39:** 1126–1138. doi:10.1002/humu.23557

Leitner A, Reischl R, Walzthoeni T, Herzog F, Bohn S, Forster F, Aebersold R. 2012. Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. *Mol Cell Proteomics* **11:** M111 014126. doi:10.1074/mcp.M111.014126

Leitner A, Joachimiak LA, Unverdorben P, Walzthoeni T, Frydman J, Forster F, Aebersold R. 2014a. Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. *Proc Natl Acad Sci* **111:** 9455–9460. doi:10.1073/pnas.1320298111

Leitner A, Walzthoeni T, Aebersold R. 2014b. Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline. *Nat Protoc* **9:** 120–137. doi:10.1038/nprot.2013.168

Luna R, Rondon AG, Perez-Calero C, Salas-Armenteros I, Aguilera A. 2019. The THO complex as a paradigm for the prevention of cotranscriptional R-loops. *Cold Spring Harb Symp Quant Biol* **84:** 105–114. doi:10.1101/sqb.2019.84.039594

Meinel DM, Strasser K. 2015. Co-transcriptional mRNP formation is coordinated within a molecular mRNP packaging station in *S. cerevisiae*. *Bioessays* **37:** 666–677. doi:10.1002/bies.201400220

Meinel DM, Burkert-Kautzsch C, Kieser A, O'Duibhir E, Siebert M, Mayer A, Cramer P, Soding J, Holstege FC, Strasser K. 2013. Recruitment of TREX to the transcription machinery by its direct binding to the phospho-CTD of RNA polymerase II. *PLoS Genet* **9:** e1003914. doi:10.1371/journal.pgen.1003914

Minocha R, Popova V, Kopytova D, Misiak D, Huttelmaier S, Georgieva S, Strasser K. 2018. Mud2 functions in transcription by recruiting the Prp19 and TREX complexes to transcribed genes. *Nucleic Acids Res* **46:** 9749–9763. doi:10.1093/nar/gky640

Mohammadi A, Tschanz A, Leitner A. 2021. Expanding the cross-link coverage of a carboxyl-group specific chemical cross-linking strategy for structural proteomics applications. *Anal Chem* **93:** 1944–1950. doi:10.1021/acs.analchem.0c03926

Pacheco-Fiallos B, Vorlander MK, Riabov-Bassat D, Fin L, O'Reilly FJ, Ayala FI, Schellhaas U, Rappsilber J, Plaschka C. 2023. mRNA recognition and packaging by the human transcription-export complex. *Nature* **616:** 828–835. doi:10.1038/s41586-023-05904-0

Pena A, Gewartowski K, Mroczek S, Cuellar J, Szykowska A, Prokop A, Czarnocki-Cieciura M, Piwowarski J, Tous C, Aguilera A, et al. 2012. Architecture and nucleic acids recognition mechanism of the THO complex, an mRNP assembly factor. *EMBO J* **31:** 1605–1616. doi:10.1038/emboj.2012.10

Perez-Calero C, Bayona-Feliu A, Xue X, Barroso SI, Munoz S, Gonzalez-Basallote VM, Sung P, Aguilera A. 2020. UAP56/DDX39B is a major cotranscriptional RNA-DNA helicase that unwinds harmful R loops genome-wide. *Genes Dev* **34:** 898–912. doi:10.1101/gad.336024.119

Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, et al. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47:** D442–D450. doi:10.1093/nar/gky1106

Puhringer T, Hohmann U, Fin L, Pacheco-Fiallos B, Schellhaas U, Brennecke J, Plaschka C. 2020. Structure of the human core transcription-export complex reveals a hub for multivalent interactions. *Elife* **9:** e61503. doi:10.7554/eLife.61503

Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14:** 290–296. doi:10.1038/nmeth.4169

Ren Y, Schmiege P, Blobel G. 2017. Structural and biochemical analyses of the DEAD-box ATPase Sub2 in association with THO or Yra1. *Elife* **6:** e20070. doi:10.7554/eLife.20070

Rohou A, Grigorieff N. 2015. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192:** 216–221. doi:10.1016/j.jsb.2015.08.008

Rougemaille M, Dieppois G, Kisseleva-Romanova E, Gudipati RK, Lemoine S, Blugeon C, Boulay J, Jensen TH, Stutz F, Devaux F, et al. 2008. THO/Sub2p functions to coordinate 3′-end processing with gene-nuclear pore association. *Cell* **135:** 308–321. doi:10.1016/j.cell.2008.08.005

Rout MP, Sali A. 2019. Principles for integrative structural biology studies. *Cell* **177:** 1384–1403. doi:10.1016/j.cell.2019.05.016

Saguez C, Gonzales FA, Schmid M, Boggild A, Latrick CM, Malagon F, Putnam A, Sanderson L, Jankowsky E, Brodersen DE, et al. 2013. Mutational analysis of the yeast RNA helicase Sub2p reveals conserved domains required for growth, mRNA export, and genomic stability. *RNA* **19:** 1363–1371. doi:10.1261/rna.040048.113

San Martin-Alonso M, Soler-Oliva ME, Garcia-Rubio M, Garcia-Muse T, Aguilera A. 2021. Harmful R-loops are prevented via different cell cycle-specific mechanisms. *Nat Commun* **12:** 4451. doi:10.1038/s41467-021-24737-x

Schiffrin B, Radford SE, Brockwell DJ, Calabrese AN. 2020. PyXlinkViewer: a flexible tool for visualization of protein chemical crosslinking data within the PyMOL molecular graphics system. *Protein Sci* **29:** 1851–1857. doi:10.1002/pro.3902

Schuller SK, Schuller JM, Prabu JR, Baumgartner M, Bonneau F, Basquin J, Conti E. 2020. Structural insights into the nucleic acid remodeling mechanisms of the yeast THO-Sub2 complex. *Elife* **9:** e61467. doi:10.7554/eLife.61467

Singh G, Pratt G, Yeo GW, Moore MJ. 2015. The clothes make the mRNA: past and present trends in mRNP fashion. *Annu Rev Biochem* **84:** 325–354. doi:10.1146/annurev-biochem-080111-092106

Strasser K, Hurt E. 2000. Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *EMBO J* **19:** 410–420. doi:10.1093/emboj/19.3.410

Strasser K, Hurt E. 2001. Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature* **413:** 648–652. doi:10.1038/35098113

Strasser K, Masuda S, Mason P, Pfannstiel J, Oppizzi M, Rodriguez-Navarro S, Rondon AG, Aguilera A, Struhl K, Reed R, et al. 2002. TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* **417:** 304–308. doi:10.1038/nature746

Stutz F, Bachi A, Doerks T, Braun IC, Seraphin B, Wilm M, Bork P, Izaurralde E. 2000. REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA* **6:** 638–650. doi:10.1017/S1355838200000078

van Heel M, Keegstra W. 1981. IMAGIC: a fast, flexible and friendly image analysis software system. *Ultramicroscopy* **7:** 113–130. doi:10.1016/0304-3991(81)90001-2

Walzthoeni T, Claassen M, Leitner A, Herzog F, Bohn S, Forster F, Beck M, Aebersold R. 2012. False discovery rate estimation for cross-linked peptides identified by mass spectrometry. *Nat Methods* **9:** 901–903. doi:10.1038/nmeth.2103

Wegener M, Muller-McNicoll M. 2019. View from an mRNP: the roles of SR proteins in assembly, maturation and turnover. *Adv Exp Med Biol* **1203:** 83–112. doi:10.1007/978-3-030-31434-7_3

Wende W, Friedhoff P, Strasser K. 2019. Mechanism and regulation of co-transcriptional mRNP assembly and nuclear mRNA export. *Adv Exp Med Biol* **1203:** 1–31. doi:10.1007/978-3-030-31434-7_1

Xie Y, Clarke BP, Kim YJ, Ivey AL, Hill PS, Shi Y, Ren Y. 2021. Cryo-EM structure of the yeast TREX complex and coordination with the SR-like protein Gbp2. *Elife* **10:** e65699. doi:10.7554/eLife.65699

Zenklusen D, Vinciguerra P, Strahm Y, Stutz F. 2001. The yeast hnRNP-Like proteins Yra1p and Yra2p participate in mRNA export through interaction with Mex67p. *Mol Cell Biol* **21:** 4219–4232. doi:10.1128/MCB.21.13.4219-4232.2001

Zenklusen D, Vinciguerra P, Wyss JC, Stutz F. 2002. Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. *Mol Cell Biol* **22:** 8241–8253. doi:10.1128/MCB.22.23.8241-8253.2002

Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7:** e42166. doi:10.7554/eLife.42166