Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Structural Characterization of the Timeless-Tipin-RPA Complex and its Interaction with ssDNA

Justine Martha Witosch
aus
Knurow, Polen

2014

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Frau Prof. Elena Conti, PhD betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 21.11.2014

.....

Justine Witosch

Dissertation eingereicht am	21.11.2014
1. Gutachterin	Prof. Elena Conti, PhD
2. Gutachterin	Prof. Dr. Eva Wolf
Mündliche Prüfung am	30.01.2015

Preface

Parts of this thesis have been published:

Justine Witosch, Eva Wolf, Naoko Mizuno. Architecture and ssDNA interaction of the Timeless-Tipin-RPA complex. *Nucleic Acids Research* (2014). doi: 10.1093/nar/gku960.

Contents

Summary

Timeless-Tipin (Tim-Tipin), also referred to as the fork protection complex, is an evolutionarily conserved protein complex coordinating DNA replication and guaranteeing replication fork stability. DNA replication is a highly dynamic and coordinated process controlled by the action of numerous proteins. Key regulatory enzymes, like DNA helicase and DNA polymerases, are accompanied by accessory factors ensuring the correct duplication of the genome. The unwound DNA is stabilized by the key factor Replication Protein A (RPA), which recruits important downstream proteins involved in unperturbed DNA replication or checkpoint signaling and DNA repair. The Tim-Tipin complex is recruited to replication forks via RPA and spatially bridges between the helicase and polymerase, thereby coupling the functions of these two machineries. However, details of the interactions between Tim-Tipin, RPA, and ssDNA, the underlying molecular mechanism of Tim-Tipin recruitment to RPA-ssDNA and/or ssDNA, and the three dimensional structure of the Tim-Tipin-RPA complex are unknown. In this thesis, the complex formation of nearly full-length Tim-Tipin and RPA was examined using biochemical methods. Tim-Tipin and RPA form a 258 kDa complex with a 1:1:1 stoichiometry. CryoEM analysis provided first structural insights into the Tim-Tipin-RPA interactions and revealed a globular architecture of the complex showing a ring-like and a U-shaped domain which is covered by a RPA lid. One exciting aspect was that RPA adopts a compact conformation in the Tim-Tipin-RPA complex similar to the structure of RPA bound to 30 nt ssDNA. This specific conformation was the basis for further biochemical studies investigating RPA's ssDNA length-dependent conformational switch and its impact on the recruitment of Tim-Tipin to ssDNA. Here, Tim-Tipin-RPA-ssDNA complex formation was shown to be modulated by RPA's conformation. The complex was only stable when RPA adopted the more compact 30 nt binding mode and dissociated when RPA was bound to ssDNA in the 8 nt binding mode. These dynamic interactions provide a mechanistic basis for RPA-based recruitment of Tim-Tipin to ssDNA and reveal an interesting aspect of coordination of the eukaryotic replication fork.

1. Introduction

1.1. DNA Replication in Eukaryotes

1.1.1. The DNA Replication Machinery

DNA replication is an essential process duplicating genomic information with high fidelity in the nucleus. The process of DNA synthesis relies on the coordinated action of multiple replication proteins. In eukaryotes, these proteins specifically assemble to a molecular machinery known as the replisome, which is composed of the helicase core, called mini-chromosome maintenance (MCM) 2-7 complex, associated with Cdc45 and the heterotetrameric GINS (Sld5-Psf1-Psf2-Psf3) complex, the polymerase α-primase, the replicative polymerases Pol ϵ and Pol δ, the sliding clamp PCNA, the PCNA loader replication factor C (RFC), and the replication protein A (RPA) [1].

For the initiation of DNA replication, the DNA has to be separated by a helicase at the origins of replication, which are sensed by the multi-protein origin recognition complex (ORC) together with the proteins Cdc6 and Cdt1 [2–4]. This complex also recruits the helicase, a process which is called licensing [5]. The activity of the loaded helicase depends on the additional accessory proteins Cdc45 and GINS [6, 7]. Once activated, the helicase opens the DNA double helix and produces single-stranded (ss) DNA, which is immediately coated by the heterotrimeric replication protein A, the most abundant single-strand binding protein in the cell [8, 9]. The RPA-ssDNA complex in turn mediates the stable attachment of the DNA polymerase α-primase (Pol α) to DNA, which *de novo* synthesizes the first RNA-DNA primer [10, 11].

Further, the bidirectional extension of the primer is performed by Pol ϵ and δ [12, 13], whose processivities are enhanced by the polymerase clamp PCNA (proliferating-cell-nuclear-antigen) preventing dissociation of the polymerase [14, 15]. The loading of the replicating polymerases requires a switch from Pol α-primase to Pol ϵ and Pol δ, which is mediated by the replication factor C specifically recognizing the 3′-hydroxyl end of the primer-template junction [16, 17]. After loading the replicative polymerases, the replisome is assembled and the replication fork progresses at the same time as

duplicating the genome.

In addition to the core replication proteins, other regulatory components such as the fork protection complex (FPC) ensure correct duplication of the genome [18]. The FPC consists of the proteins Timeless (Tim), the Timeless-interacting protein (Tipin), and Claspin and is thought to coordinate DNA unwinding and DNA synthesis by coupling the helicase and DNA polymerase activities [19]. In addition to interactions with the helicase and DNA polymerases, the FPC components Tim-Tipin also associate with RPA [19, 20]. However, how Tim-Tipin and Tim-Tipin in complex with RPA travel with the replisome and affect the DNA replication accuracy, has not been well understood (Figure 1.1 and Section 1.2).

1.1.2. DNA Synthesis under Replicative Stress

During synthesis, the replisome encounters regions that contain numerous obstacles like protein-DNA barriers or abnormal DNA topology which cause replicative stress and can reduce replisome activity [18]. For the protection of the genome integrity, it is crucial to sense replicative stress, to stabilize normal and stalled forks, and to restart DNA synthesis without any errors after recovery. Eukaryotic cells evolved surveillance factors, which regulate DNA replication and couple DNA synthesis to cell cycle progression and DNA damage repair. Thus, the ATR-Chk1 (Ataxia telangiectasia mutated (ATM) and Rad3-related, checkpoint kinase) intra-S phase checkpoint pathway delays or arrests the DNA synthesis progression in response to replication stress by modulating the activities of the cyclin-dependent kinase 2 (CDK2) [21]. Replicative stress inhibiting the replicative polymerases, but not affecting DNA helicase might result in accumulation of ssDNA, which is coated by RPA [22, 23]. RPA-coated ssDNA plays a positive role in the recruitment of the ATR-interacting protein (ATRIP) to the DNA lesions [24]. ATRIP in turn recruits its interaction partner ATR, a replication stress response kinase, which phosphorylates Chk1 [25] dependent on the presence of Claspin [26, 27]. In that cascade, Chk1 phosphorylates its downstream targets, which participate in the regulation of S phase progression [28, 29]. Interestingly, the phosphorylation of Chk1 is mediated by Tim-Tipin, which indirectly modulates the intra-S phase checkpoint response upon replicative stress (Figure 1.1 and Section 1.2). However, how Tim-Tipin influences the ATR-Chk1 checkpoint pathway remains elusive.

Furthermore, the induction of the checkpoint signaling also activates the appropriate DNA repair mechanism. RPA was shown to sense damaged DNA [30–32] and to recruit DNA repair proteins to sides of DNA damage by protein-protein interactions. Several

repair mechanisms are involved in DNA recovery.

In the nucleotide excision repair (NER), RPA senses damaged DNA and specifically interacts with XPA (xeroderma pigmentosum complementation group A protein). The complex recruits the corresponding endonucleases that remove bulky DNA lesions [32, 33]. Further, interaction and structural studies have revealed direct association of RPA with the uracil-DNA glycosylase 2 (UNG2) that initiates base excision repair (BER) by excising and replacing defective DNA bases [34, 35]. Finally, RPA is also required for double-strand break (DSB) repair by homologous recombination (HR), which is facilitated by direct RPA-Rad51 and RPA-Rad52 interactions [34, 36–38].

Interestingly, the interactions of RPA with the DNA repair proteins XPA and UNG2 are mediated by a similar interface as reported for the FPC member Tipin [20]. However, how the FPC affects the DNA repair mechanisms is not clear yet.

Taken together, the integrity of the genome is guaranteed by an intact replisome/ replication fork, which is stabilized by additional accessory components such as Tim-Tipin, RPA or Tim-Tipin-RPA.

1.2. Roles of Timeless and Timeless-Tipin in Circadian Clocks and DNA Replication

The function of the mammalian Tim-Tipin complex was first discussed on the basis of the function of the mammalian Timeless protein. Mammalian Tim (mTim) was first identified as an ortholog of the *Drosophila* Timeless protein (dTim), which is a key player in the circadian clock of flies. Publications of the past two decades are inconsistent regarding the role of mammalian Tim and its ambivalent functions are controversially discussed. Later studies indicate that mTim is directly involved in the molecular machinery of DNA replication, which apparently reflects the more prominent biological function of mTim.

1.2.1. Supporting Evidence for Timeless Function in Circadian Clocks

Mammalian Tim was identified based on its sequence similarity to *Drosophila* Tim, a key player in the circadian clock of flies [39–41]. Most organisms display biological rhythms in behavior like rest and activity or sleep and waking cycles. The periodicity of these rhythms averages 24 h and is a prerequisite for the adaptation of the organisms to the day and night cycles of the environment. The generation of these circadian rhythms (*circa diem*: Latin, 'approximately a day') is performed by a master clock located in

the suprachiasmatic nucleus (SCN) in the brain that perceives external Zeitgeber-stimuli like light and creates an output that regulates the biological rhythmical processes [42].

In general, intracellular clocks (master (SCN) and peripheral clocks (tissues like lung, liver, kidney)) consist of positive and negative feedback loops of transcription and translation that drive periodic rhythms in the mRNA and protein levels mediated by protein-protein and protein-DNA interactions, post-translational modifications, cellular localizations and nuclear translocations, expression rates, and degradation of clock proteins.

Drosophila Tim and its interacting partner *Drosophila* Period (dPer) are members of the negative feedback loop in the fly oscillator and inhibit their own transcription by binding to the transcription factors dClock and dCycle, which represent the positive limb of the feedback loop and activate the transcription of the *dPer* and *dTim* genes [43]. A comparison of the *Drosophila* clock with the mammalian clock emphasizes that many of the core clock functions of Tim in insects are carried out by the Cryptochrome proteins (mCry) in mammals, leaving the question what the function of mTim in the mammalian clockwork is?

There are a couple of indications supporting mTim's role in the mammalian circadian clock comparable to dTim's function in the fly clock. Since the formation of homo- or heteromultimeric complexes is a common feature of circadian clock proteins [41, 44–48], mTim should interact with other clock proteins. Indeed, mTim was found to associate with elements of the negative feedback loop, mCry1 and mCry2, by immunoprecipitation of *in vivo* expressed [49] or over-expressed proteins [50]. Moreover, as rhythmic expression and activity of clock proteins is mediated by clock protein-protein interactions, a knockdown of one of the component results in changed expression or activity of the other partner [44, 51, 52]. An alteration in the expression level of mPer and mCry proteins after mTim knockdown was reported, supporting mTim's role in the clock [52]. In addition, mTim was shown to repress the mClock-mBMAL1 (Brain Muscle Arnt-like 1) transcription activity [50] similar to dTim in the fly clock, where dTim inhibits its own transcription by binding to dClock-dCycle (homologs of mClock-mBMAL1) [53]. Further, mTim knockdown abolishes the circadian neuronal activity (firing rate) of the SCN, the site of the master clock, which demonstrates that mTim is required for SCN circadian rhythmicity [52]. Another profound characteristic of circadian clocks is the rhythmic expression of mRNAs and proteins [50, 54, 55]. mTim exists in two alternative transcripts [56] with only one isoform being expressed in a circadian manner. Adult animals exhibit a high and constitutive expression rate of a short Tim (Tim-s) isoform comprising the most C-terminal 475 residues. Full-length Tim expression appears to be

periodic and exists only in the embryonic stadium [52].

Despite these data partially supporting the role of mTim as a member of the negative feedback loop in the mammalian clock, the majority of data point to a minor role of mTim in the circadian clock.

1.2.2. Non-Supporting Evidence for Timeless Function in Circadian Clocks

Currently, several reports contradict the evidences for mTim function in circadian clocks. Interactions between the mPer proteins (mPer1, 2, 3) and mTim are ambiguously described in literature questioning mTim's role in the circadian clock. While on the one hand, mTim-mPer interactions were neither detected by co-immunoprecipitations of *in vivo* [49] or overexpressed proteins [50] nor by yeast-two-hybrid experiments [41], indirect association of mTim and mPer in the SCN was reported [52]. Protein-protein interactions also influence cellular localization of the clock proteins. In the fly clock, dPer and dTim associate and translocate into the nucleus, which is important for the repression of their own transcription [44]. In mammals, mPer proteins exhibit mostly cytoplasmic location and are shuttled into the nucleus upon mCry binding [50]. mTim was shown to be nuclear and did not translocate mPer into the nucleus in contrast to mCry.

In contrast to *bona fide* clock proteins, the deletion of mTim is embryonically lethal as described by Gotter *et al.* indicating a function of mTim in development [57]. Furthermore, phylogenetic analyses of Tim proteins identified a second gene cluster in the *Drosophlia* genome. In addition to the gene cluster containing the canonical, circadian clock *dTim* gene, a *dTim* paralog, *dTim-2/dTimeout*, was identified, which is required for chromosome stability [57–59]. Notably, sequence alignments revealed that mTim shares a greater similarity with the *Drosophlia* protein dTim-2/dTimeout than with dTim [57, 58] and therefore might be the true ortholog of dTim-2/dTimeout and not an ortholog of the clock-relevant dTim protein. In summary, these reports indicate that Tim's clock function still remains to be elucidated.

1.2.3. Tim-Tipin and its Role in DNA Replication

Collective observations point to other more prominent biological functions of Tim than the clock functions. To investigate the putative function of mTim, yeast-to-hybrid experiments were performed by Gotter *et al.* in 2003 and identified mTipin as a Timeless interacting protein, which was further confirmed by co-immunoprecipitation assays [60].

Tim and Tipin (hereafter mammalian proteins are meant, if not stated otherwise) form a complex via their N-terminal regions [61]. Moreover, the complex formation is important for their mutual stabilization and the loss of one interacting partner leads to the loss of the other partner [62] (Figure 1.2). Furthermore, evaluations of the subcellular location showed a mutual dependance. While Tim was mostly found in the nucleus [50], Tipin alone was equally distributed in both compartments of the cell. Interestingly, co-expression of Tipin with Tim promoted Tipin's nuclear location [60, 63].

Notably, other homologs of Tim and Tipin have been identified in *Metazoans* and fungi like Swi1 and Swi3 in fission yeast *Schizosaccharomyces pompe* [65, 66], Tof1 and Csm3 in budding yeast *Saccharomyces cerevisiae* [67, 68] as well as ce-TIM1 in *Caenorhabditis elegans* [69] indicating that these proteins are conserved in Eukaryotes.

Identification of Tim-Tipin binding to essential DNA replication proteins pointed to the assumption that Tim-Tipin might play a role in DNA replication as an replisome-associated protein (Figure 1.1 a). Thus, Tim-Tipin has been shown to bind to the MCM2-7 helicase subunits *in vivo* and *in vitro* in mammals [19, 70] and in yeast [71, 72]. The association of Tim-Tipin with very large replisome complexes consisting among others of the MCM helicase core, GINS and Cdc45 was reported in yeast and culminated in the middle S phase [73]. This interaction was recently also detected in human cells [74] suggesting a high conservation of these complexes among species. Tim-Tipin and Tim alone were also shown to bind to the replicating DNA-polymerases in cells and co-immunoprecipitation assays [19, 74]. Since Tipin alone did neither interact with the helicase nor with the DNA polymerases, these interactions are likely mediated by Tim. Tim-Tipin was also found to inhibit the helicase and ATPase activity of the CMG complex (Cdc45-MCM2-7-GINS) and to stimulate the activities of the DNA polymerases α, δ and ϵ [74, 75]. Collectively, these observations imply the direct involvement of Tim-Tipin in the catalytic activities of the replisome enzymes.

Depletion of Tim-Tipin leads to an accumulation of ssDNA observed in yeast and mammals [76–78] implying that Tim-Tipin might stabilize the DNA replisome by spatially bridging between the helicase and DNA polymerase and thereby coupling DNA unwinding with DNA synthesis [19] (Figure 1.1 b). Consistent with that study, an increased level of chromatin-bound RPA was observed in DNA replication-blocked and Tipin-depleted *Xenopus* extracts indicating raised presence of ssDNA [79]. Moreover, increased chromatid breaks, translocations, and sister-chromatid exchanges have been observed upon Tim or Tipin depletion [62] supporting the important role of Tim-Tipin throughout normal DNA replication [80].

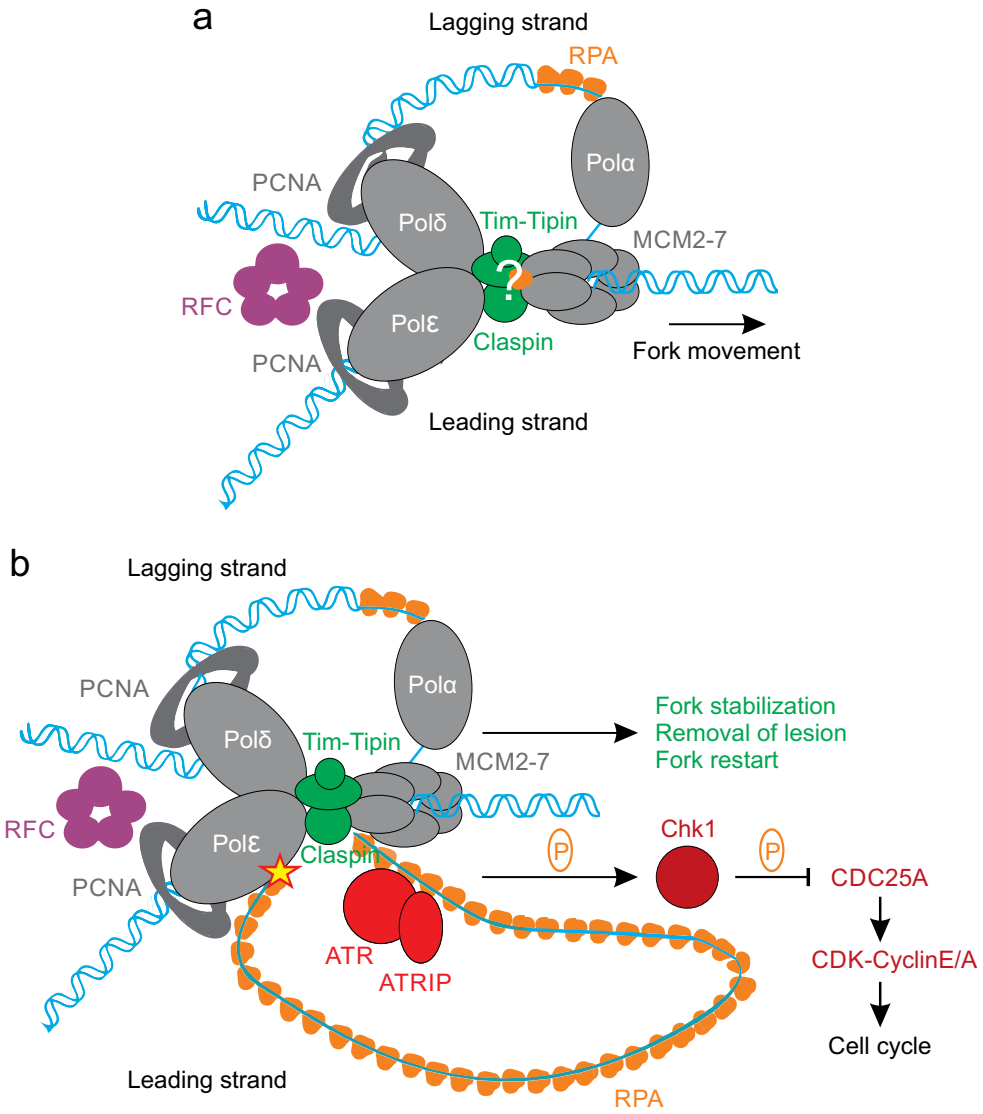


Figure 1.1.: Schematic representation of the DNA replication fork. (a) The replisome consists of the key proteins MCM2-7 helicase, the Pol α-primase, the replicative polymerases Pol ϵ and Pol δ, the sliding clamp PCNA, the PCNA loader RFC, and RPA. The FPC (Tim-Tipin and Claspin) couples DNA unwinding and synthesis by bridging between helicase and polymerase during unperturbed DNA-replication. The Tim-Tipin-RPA complex is mediated by a direct Tipin-RPA32 interaction. The assembly of the replisome harboring Tim-Tipin-RPA at the fork still needs to be elucidated (indicated by the question mark). *Figure legend continued on next page.*

In addition, the protein expression levels of Tim and Tipin exhibit an steady increase prior to and during the S phase [19, 81]. Reduced levels of Tim-Tipin have been reported to retard S phase progression, which is consistent with the hypothesis that Tim-Tipin is a key player in DNA replication [19, 61].

During unperturbed DNA replication RPA has been found to recruit additional accessory proteins, which are essential for correct processing and to help to coordinate these proteins at the replication fork [82]. Tipin was shown to co-localize with RPA32 and Tim with RPA70 in cells at regions undergoing DNA replication [19]. Complex formation of Tim-Tipin and RPA has been reported to be mediated by a direct interaction of Tipin and the C-terminal domain of RPA32 [19, 20] (Figure 1.2 b). Furthermore, the interaction of RPA and Tipin was also observed in the DNA-bound state [63, 83]. The direct interaction of Tim-Tipin and RPA raises the possibility that the Tim-Tipin-RPA complex is also located at ssDNA stretches in the replication fork.

However, DNA replication forks are vulnerable to DNA damage or might stall due to several obstacles [84]. RPA has been reported to be important for the delay of the cell cycle progression, the recruitment of repair proteins, and for the stabilization of the replication fork until the DNA damage or block is resolved [21]. Accumulated reports have shown that Tim-Tipin together with RPA might be involved in the DNA damage response and DNA repair. Indeed Tim-Tipin was reported to interact with components of the DNA replication checkpoint pathway (Figure 1.1 b). Tim-Tipin interactions with Chk1 and ATR-ATRIP were stimulated by UV- or hydroxyurea- (HU) induced replicative stress [81, 83]. The same reports also demonstrated that the phosphorylation of Chk1 in response to HU or UV was attenuated upon Tim and Tipin knockdown [61, 81, 83]. Additionally, Tim as well as Tipin were reported to directly bind to Claspin [27, 63], the mediator of the Chk1 phosphorylation by ATR-ATRIP. Moreover, knockdown of Tim-Tipin have lead to reduced nuclear accumulation of Claspin under replicative

Figure 1.1.: *(Previous page).* (b) UV-induced DNA damage (yellow star) stalls the DNA polymerase and leads to helicase-polymerase/leading and lagging strand uncoupling resulting in ssDNA accumulation (leading strand). RPA covers the locally increased ssDNA, which serves as a signal for the recruitment of ATR/ATRIP. The DNA damage sensors activate the intra-S phase checkpoint by Tim-Tipin-Claspin mediated Chk1 phosphorylation. Further effector proteins (CDC25A, CDK-CyclinE/A) arrest cell cycle. In addition, Tim-Tipin together with Claspin stabilize the stalled replication fork by spatially bridging between helicase and polymerases and thus preventing the replisome from disassembling. The FPC also promotes removal of DNA lesions and fork restart after the obstacle has been resolved. Other important DNA replication and checkpoint factors (e.g. GINS, Cdc45) are omitted for simplicity. Adopted from [64].

stress [61].

Like their mammalian counterpart, the yeast Tim-Tipin and Claspin homologs, Tof1, Csm3, Mrc1 or Swi1, Swi3, Mrc1, have been found to be also required for the activation of the effector kinases Rad53 or Cds1 [71, 77]. In addition, Tof1 and Csm3 (Tim and Tipin) have been reported to be important for the pausing and recovering of the replication fork at replication fork barriers [85].

Taken together, in addition to the role in unperturbed DNA replication, many reports support the involvement of Tim and Tipin in checkpoint signaling upon DNA damage and in DNA repair. Accumulated evidence therefore supports the importance of the Tim-Tipin complex as an evolutionary conserved player at the DNA replication fork maintaining genome integrity.

1.3. Replication Protein A and its Multiple Functions

The ubiquitous eukaryotic ssDNA binding protein RPA plays an essential role in all kinds of DNA metabolisms [9]. RPA consists of three very tightly associated subunits of approximately 70, 32, and 14 kDa, referred to as RPA70, RPA32, and RPA14 [86] and is only solubly expressed in bacteria as heterotrimer or RPA32/14 heterodimer [87, 88] (Figure 1.3). The ternary complex specifically recognizes ssDNA with a nano- to subnanomolar affinity in a cooperative manner, which is length-dependent [89–91]. Multiple DNA-binding domains have been identified in RPA and structurally characterized (for an overview see Section 1.4.3) [92–96]. Interestingly, the multi-module protein undergoes conformational changes and shows a gradual compaction upon DNA binding as the coverage of RPA by ssDNA progresses [97].

RPA was reported to play an important role in DNA replication initiation and elongation. Although RPA was found to be diffusely localized in the nucleus in G_1 and G_2 phase, it specifically becomes punctually located at replication foci prior to initiation and during DNA replication [98, 99]. Studies on the DNA replication of the simian virus 40 (SV40) have identified interactions of the viral replicative DNA helicase (SV40 T antigen, Tag) and RPA. This direct protein-protein interaction promotes the recruitment of RPA to the emerging ssDNA at the origin of replication. Binding of RPA to the polymerase α-primase was also reported in the SV40 system [10]. In the eukaryotic DNA replication, RPA is recruited to ssDNA by interactions with the MCM helicase subunit and thus stabilizes the MCM-unwinded DNA [100, 101].

In addition to proteins involved in DNA replication, RPA binds many factors in

processes of other DNA metabolisms including DNA repair and cell cycle checkpoint signaling. By sensing damaged DNA [30–32], RPA recruits DNA repair proteins to sides of DNA damage [32–38]. The interactions of these DNA repair factors with RPA are structurally very similar and are mediated by the RPA32 C-terminal domain (winged-helix, WH domain) [34]. Strikingly, Tipin has been reported to also interact with the RPA32 C-terminal WH domain [19, 20, 83]. Indeed, one study has recently shown that Tipin exhibits sequence similarity with XPA and UNG2 implying that these proteins share a similar binding interface on the RPA32-WH domain [20] (Figure 1.2). Moreover, the association of XPA with RPA was shown to be mutually exclusive with the association of Tipin with RPA, supporting a competitive binding to the same RPA32 surface [19].

Similar to ARTIP, Tipin also binds to DNA-bound RPA, as already elaborated in Section 1.2, and is required for the recruitment of Timeless and Claspin to the sides of DNA lesion [63]. Thus, the Tim-Tipin-RPA interaction might be important for the activation of the ATR-Chk1 checkpoint pathway.

Taken together, RPA couples its DNA binding activity to the recruitment of DNA processing factors by providing a binding platform for multiple proteins complexes (among FPC) involved in DNA replication, cell cycle checkpoint signaling, and DNA repair.

1.4. Molecular Architecture of Tim, Tipin, and RPA

1.4.1. Tim

The mouse Tim protein consists of 1197 amino acids (aa), which correspond to a molecular weight of 137.4 kDa. It does not contain any known domains that have been yet structurally or biochemically identified and therefore could not be associated with any structural super-family.

However, PSI-BLAST [102] search using the NCBI website (National Center for Biotechnology Information) revealed two conserved regions (Figure 1.2 a). The N-terminal region (aa 21-285) belongs to the Timeless protein family (pfam 04821). The C-terminal part harbors the Timeless protein C-terminal region (Timeless-C, pfam 05029, aa 722-1189). Protein secondary structure prediction analysis using PSIPRED [103] suggested that Tim's major regions are all α-helical (Appendix A.1). The three-dimensional (3D) prediction of Arm/HEAT repeats for Tim is controversially discussed (Arm/HEAT shortened for Armadillo/Huntingtin, Elongation factor 3, subunit of protein phosphatase 2A, yeast kinase TOR1 [104]). Unlike one study suggesting Tim to

be a member of the Arm/HEAT protein super-family by showing significant hits to a number of Arm/HEAT repeat proteins including importin α and β [105], a second contradicting study found no significant homology to Arm/HEAT proteins and indicated false positives and overpredictions in previous analysis [106].

A DDT domain (DNA-binding homeobox and different transcription factors) was identified in the *S. pompe* yeast Tim homolog, Swi1, and was shown to be conserved between mouse Tim (aa 332-385) and the yeast protein (aa 327-378) [107] (Figure 1.2 and Figure A.1). The DDT domain is a putative DNA-binding domain found in different transcription and chromosome remodeling factors [108]. The Swi1 DDT domain was reported to bind DNA *in vitro*, to associate with chromatin in cells, and to be important for the recruitment of Swi3 (Tipin homolog) to chromatin.

Further, four putative nuclear localization signals (NLS 1-4) were identified in Tim but only the C-terminal NLS 4 was reported to be likely relevant for the nuclear import of Tim [109]. In addition, Tim exhibits several acidic stretches within the protein sequence including a Glutamate (E)-repeat and five E-rich regions in the C-terminal part (Figure 1.2). The N-terminal region of human Tim (mouse Tim, aa 1-569) was reported to bind to human Tipin by co-immunoprecipitation assays [61].

1.4.2. Tipin

Mouse Tipin consists of 278 amino acids, which correspond to a molecular weight of 31.5 kDa. Like for Tim, no similarity to any known structural super-families was reported to date [60]. PSI-BLAST search revealed one conserved domain corresponding to the Swi3 super-family (*S. pompe* yeast Tipin homolog Swi3, pfam 07962, aa 63-146 (Figure 1.2 b)). Protein secondary structure prediction analysis suggested that Tipin is mostly α-helical and exhibits a Proline (P)-rich region at the N-terminus, which is likely to be disordered (Figure 1.2 b and Appendix A.2). Interestingly, Tipin possesses sequence similarity to XPA and UNG2 of approximately 48 amino acids (Figure 1.2 c). Both proteins play an important role in DNA damage repair and bind to the RPA32WH subunit via the Tipin-homolog region [34]. Further, structural analysis mapped the RPA-binding region within Tipin to the residues 182-215 [20] (Figure 1.2 d). The N-terminal region of human Tipin (mouse Tipin, aa 64-140) was reported to be significant for human Tim-binding and shows no overlap with the C-terminal RPA-binding site [61] (Figure 1.2 b). Strikingly, no interactions of Tim with RPA were reported supporting that Tipin harbors both binding sites (for RPA and Tim) [19, 83].

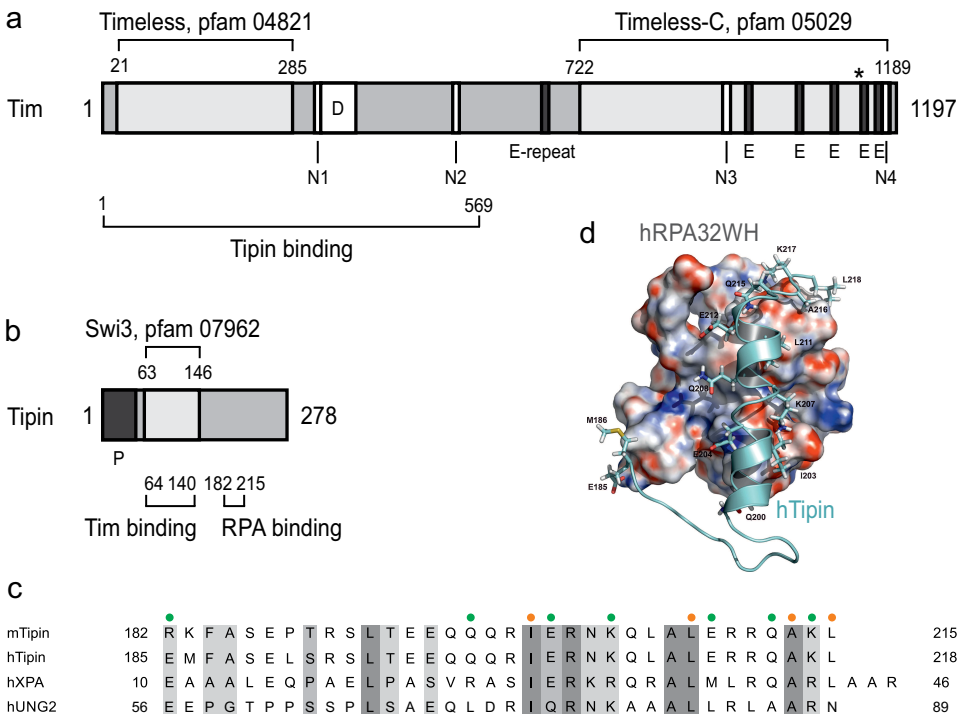


Figure 1.2.: Domain organization of Tim (a) and Tipin (b) and sequence similarities (c). (a) Light gray = N-terminal and C-terminal Timeless protein family regions. N1-N4 = nuclear localization signal (NLS); NLS 1, aa 316-332; NLS 2, 528-537; NLS 3, 935-946; NLS 4, 1175-1190. D = conserved DNA-binding domain DDT, which was suggested for the *S. pompe* homolog, Swi1. Dark gray = E-repeat, aa 662-672 and E-rich regions (E), aa 964-985, 1046-1055, 1099-1107, 1144-1152, 1167-1172. Amino acid boundaries are shown as numbers. The asterisk marks the end of the construct used in this thesis. Bottom brace = Tipin binding site. (b) Dark gray/P = Proline-rich region of Tipin, approx. aa 1-60. Light gray = Swi3 protein family region. Bottom braces = Tim and RPA binding sites. (c) Sequence similarities between mouse and human Tipin, XPA, and UNG2. Circles = hydrophobic (orange) and charge-charge (green) interactions between hTipin and hRPA32WH. (d) Electrostatic potential surface representation of human RPA32WH (aa 172-270, blue = positive, red = negative) and ribbon representation of human Tipin (aa 185-218). The binding interface is highlighted with sticks and residue numbers. Adopted from [20].

1.4.3. RPA

The ternary RPA complex consists of the proteins RPA70, RPA32, and RPA14. It is a modular protein complex and contains six domains adopting the OB-fold (Oligonucleotide/Oligosaccharide-binding fold), which is very common for ssDNA-binding proteins [110, 111]. Four OB-domains (N, A, B, C) are located in RPA70, which are tethered by flexible linkers and one each in RPA32 (D) and RPA14 (14) (Figure 1.3 a-c). Four

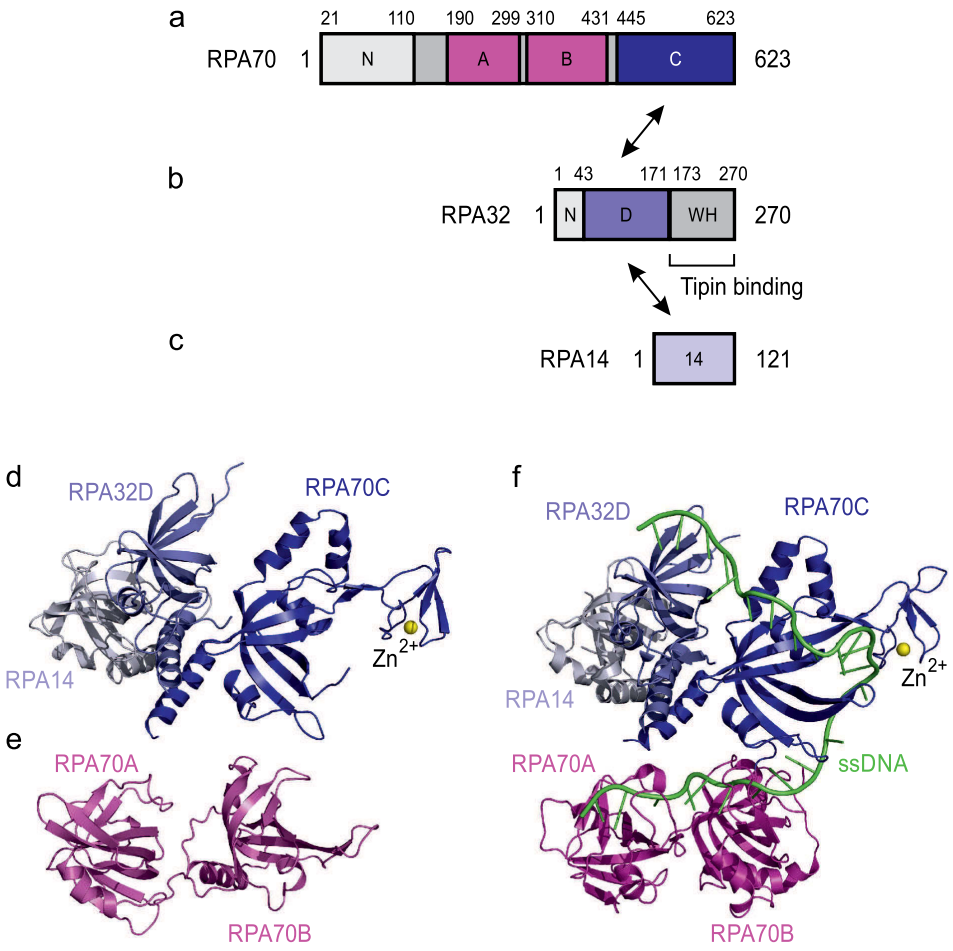


Figure 1.3.: Domain organization and crystal structures of RPA. (a) Schematic representation of RPA70 domain organization. N = N-terminal RPA70 domain. A, B, C = RPA70 DNA-binding domains (DBD). Amino acid boundaries of each domain are shown at the top. (b) RPA32 domain organization. N = Phosphoamino acid cluster. D = RPA32 DNA-binding domain. WH = RPA32 winged helix. (c) RPA14 domain organization. 14 = RPA14. (d) Crystal structure of RPA trimerization core consisting of RPA70C (dark blue), RPA32D (light blue), and RPA14 (gray) (PDB 1L1O). The bound zinc metal ion is represented as yellow sphere. (e) Crystal structure of RPA70A and RPA70B in the DNA-free state (PDB 1FUG). (f) Crystal structure of the RPA DNA-binding core consisting of RPA70ABC, RPA32D, and RPA14 bound to ssDNA (PDB 4GNX). Color code as in (d) and (e). The bound 32 nt ssDNA is represented in green.

of these OB-fold domains possess DNA-binding ability, namely DNA-binding domain (DBD) A, DBD-B, DBD-C in RPA70, and DBD-D in RPA32 and adopt different conformations upon DNA binding [92–96].

RPA binds to ssDNA in two discrete binding modes, which are referred to the 8 and 30 nucleotide (nt) binding mode and can coexist in a dynamic equilibrium in solution [112–115].

RPA70 DBD-A and DBD-B exhibit high inter-domain flexibility and adopt multiple conformations under DNA-free conditions [94] (Figure 1.3 e). In the 8 nt binding mode, RPA recognizes a minimum of 8 nucleotides by the RPA70 DBD-A and DBD-B only [92, 95] (Figure 1.3 a, magenta). Binding of ssDNA induces conformational changes and stabilizes the two domains in a tandem conformation [92]. Increasing length of bound ssDNA sequentially engages DBD-C and further stabilizes the linker tethering DBD-B and DBD-C [95]. In the 30 nt binding mode, the ssDNA is bound by RPA involving all four DBDs (A-D) [96] (Figure 1.3 f). Biochemical studies have shown that these two RPA binding modes exhibit different dissociation constants (K_D) of ∼50 nM for the 8 nt mode and ∼0.05 nM for the 30 nt binding mode [90, 116]. Further, a recent report from 2013 has shown a progressive compaction of RPA's overall architecture with increasing length of bound ssDNA [97].

The trimerization between the RPA subunits is mediated by RPA70 DBD-C, RPA32 DBD-D, and the entire RPA14, which form a three α-helix bundle interface [95]. Interestingly, DBD-C contains an zinc ribbon motif and the zinc metal ion impacts the stabilization and ssDNA-binding activity of the domain [117] (Figure 1.3 d).

Three additional modules have been identified in RPA, which are not incorporated into the structural core. The N-terminal RPA70N module has low DNA-binding affinity but was reported to mediate important protein-protein interactions with p53 and to stimulate Pol-α activity [118–120]. The N-terminal RPA32N module is predicted to be unstructured and becomes phosphorylated in a cell cycle-dependent manner and in response to DNA damage [121–124].

The RPA32 C-terminal module contains a winged helix (WH) domain, which was reported to be important for protein interactions with multiple DNA repair proteins and Tipin (Section 1.3) [20, 35, 125]. XPA, UNG2, and Tipin interact with the same interface of RPA32WH and are suggested to associate with a similar binding mode. All three proteins undergo conformational changes upon RPA32WH binding and adopt an α-helix conformation (Figure 1.2 d). Thus, RPA uses a single binding site to interact with several proteins involved in multiple processes of DNA metabolism and serves as an exchange point.

1.5. Aim of the Thesis

Since crystal structures of the RPA components, but no structural information about the Tim-Tipin-RPA complex have been available, little is known about the interactions between Tim-Tipin, RPA, and ssDNA, the underlying molecular mechanism of Tim-Tipin recruitment to ssDNA/RPA-ssDNA, and the three dimensional structure of Tim-Tipin-RPA. The first aim of the thesis was to solve the 3D structure of the Tim-Tipin-RPA complex. Therefore, the expression and purification protocol of the recombinant Tim-Tipin complex had to be established and the reported purification protocols of recombinant RPA had to be adapted to our experimental conditions. Second, to probe and to characterize the complex formation between Tim-Tipin and RPA, a combination of biochemical and biophysical approaches was used including pull-down assays, size exclusion chromatography, static light scattering, microscale thermophoresis, and cross-linking experiments in combination with mass spectrometry analysis. In addition, to gain structural insights into the Tim-Tipin-RPA interactions, the complex was analyzed by electron microscopy (EM). The 3D model obtained from negative stain and cryo-EM was validated by docking analysis of RPA crystal structures, antibody labeling, and gel-filtration analysis using RPA sub-complexes.

Based on the structural analysis of the Tim-Tipin-RPA complex, the following questions were to be answered: how ssDNA and/or RPA's binding modes affect the stability and ssDNA binding affinity of Tim-Tipin-RPA and how Tim-Tipin recognizes ssDNA in a RPA-independent manner? These questions were addressed by employing additional biochemical methods including size exclusion chromatography with ssDNA substrates, electrophoretic mobility shift assays, and fluorescence anisotropy experiments.

Finally, the results and conclusions were related to the molecular dynamics at the DNA replication fork, the DNA replication checkpoint, and DNA repair activities.

2. Results

2.1. Definition of the Working Constructs

2.1.1. Purification of Timeless and Crystallization Attempts

Prior to the initiation of the project, the cloning and expression of the first mouse Timeless (Tim) full-length protein and its constructs generated by rational design according to secondary structure prediction was performed by the Dortmund Protein Facility (DPF, Dortmund, Germany). The construct 997-1197 was solubly expressed and purified using IMAC (immobilized metal ion affinity chromatography) and SEC (size exclusion chromatography). Christiane Theiß (laboratory technician) conducted first limited proteolysis attempts to further define the best construct for structural studies. The tryptic digestion of Tim 997-1197 identified two new constructs with the boundaries 997-1134 and 997-1138.

At the beginning of this thesis, the purification of Tim 997-1138 was improved by an additional ion-exchange chromatography step (Figure 2.1) and octahedral-shaped crystals were obtained from a 72 mg/mL sample within 10 days in conditions containing different sized polyethylene glycol polymers (PEGs) (Figure 2.2 a).

Optimization of the initial condition did not give any hits, except two conditions containing crystals that were too small for data collection. Further crystals appeared in the initial screen after 1-2 months (Figure 2.2 b-c). While most of the crystals diffracted poorly from 9-5.5 Å, one crystal diffracted up to 4.4 Å (Figure 2.2 e), with which a complete diffraction data collection was performed. The crystal adopted the space group P422 with a = 93.42 Å and c = 201.03 Å. The Matthews probability analysis indicated 4-6 molecules per asymmetric unit [126, 127]. The crystals were shown to be comprised of a degradation product of Tim by silver-stained SDS-gels and MS/MS analysis (Figure 2.2 d).

To improve the diffraction of the crystals, limited proteolysis of Tim 997-1138 was performed and yielded a stable protein fragment of similar molecular weight (MW), when compared to the dissolved crystals analyzed by SDS-gel. The boundaries were

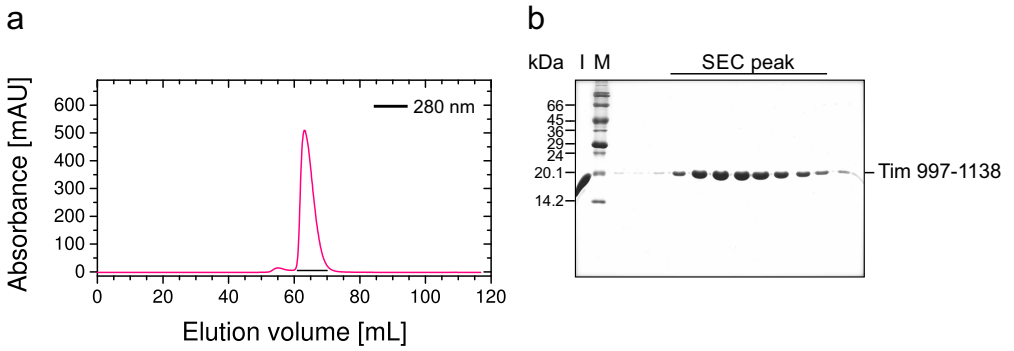


Figure 2.1.: SEC analysis of Tim 997-1138 using a Superdex 75 16/60 column. (a) Size exclusion chromatogram of purified Tim showing a single peak elution profile at 63 mL visualized by UV absorbance at 280 nm. (b) SDS-PAGE (sodium dodecyl sulphate-polyacrylamide gel electrophoresis) shows the peak fractions marked by a line. (I) Protein sample injected on gel-filtration column. (M) Protein marker in kDa.

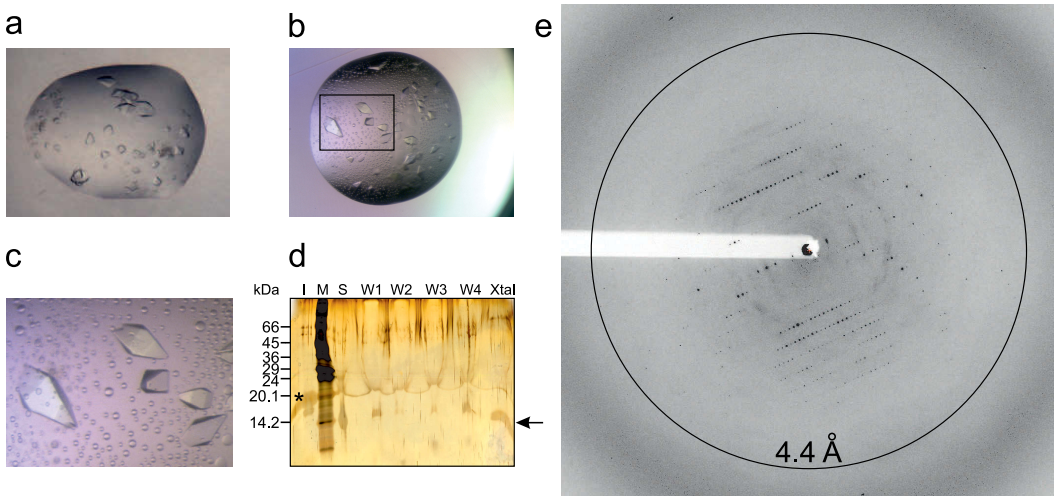


Figure 2.2.: Crystallization and X-ray diffraction of Tim 997-1138. (a) Initial crystals of Tim appeared after 10 days in 100 mM Sodium cacodylate pH 6.0, 12% PEG 20000. (b) Further crystals showed up in an additional condition after 1 month (100 mM Sodium cacodylate pH 6.2, 7% PEG 20000). The rectangle marks the enlarged area shown in (c). (d) SDS-PAGE analysis of dissolved crystals stained by silver nitrate. (I) Protein sample used for crystallization setup (marked by asterisk). (M) Protein marker in kDa. (S) Supernatant of the spun-down drop containing the analyzed crystals. (W1-W4) Content of washing steps. (Xtal) Washed and dissolved Tim 997-1138 crystals (marked by arrow). (e) Diffraction image obtained from crystals diffracting to 4.4 Å. The black circle shows the 4.4 Å resolution range.

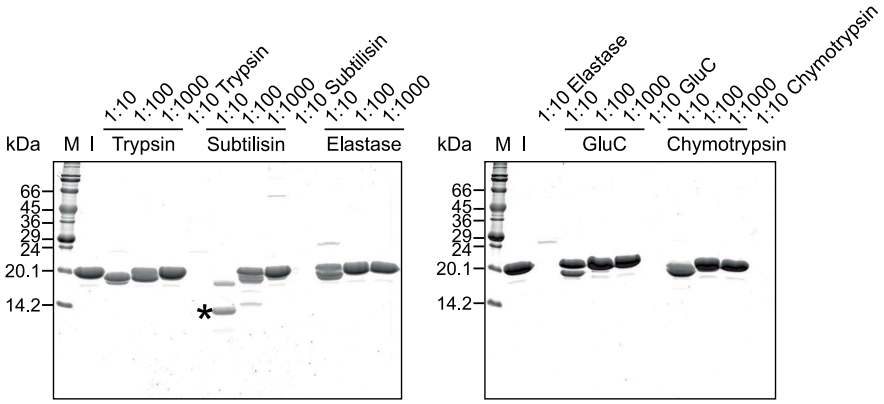


Figure 2.3.: Limited proteolysis of Tim 997-1138. The protein was incubated with five different proteases, each for 30 min on ice. The asterisk marks the stable product that was analyzed by ESI-MS and N-terminal sequencing and corresponds to Tim 1008-1125. The ratio above each lane illustrates the dilution of the 1 mg/mL protease stock solution. (M) Protein marker in kDa. (I) Protein sample used for the proteolytic analysis.

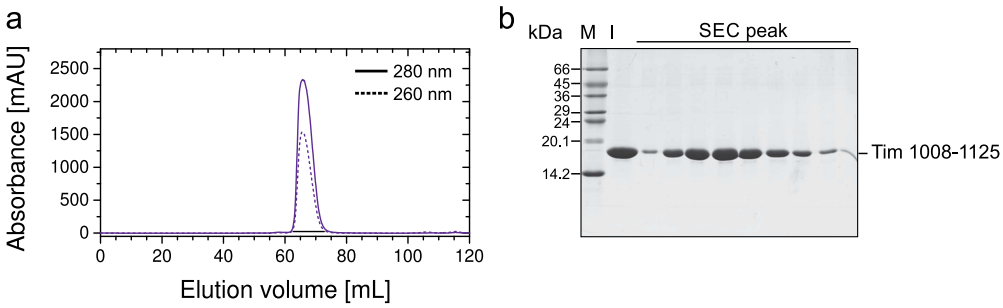


Figure 2.4.: SEC analysis of Tim 1008-1125 using a Superdex 75 16/60 column. (a) Size exclusion chromatogram of purified Tim showing a single peak elution profile at 66 mL visualized by UV absorbance at 280 nm and 260 nm. (b) SDS-PAGE shows the peak fractions marked by a line. (M) Protein marker in kDa. (I) Protein sample injected on gel-filtration column.

determined to be 1008-1125 (Figure 2.3).

The protein fragment (Tim 1008-1125) was purified (Figure 2.4) and set up for crystallization trials. Crystals showing an octahedral shape appeared after 21 days in different sized PEG polymers (Figure 2.5). The conditions were similar compared to those of Tim 997-1138. The crystals diffracted from 10-5.3 Å. The data was indexed to the tetragonal space group I422 with a = 88.68 Å and c = 274.47 Å. However, analysis of the dissolved crystals on silver-stained SDS-gel showed a degradation product as well (Figure 2.5 d). Further optimization did not yield reproducible crystals. The difficulty

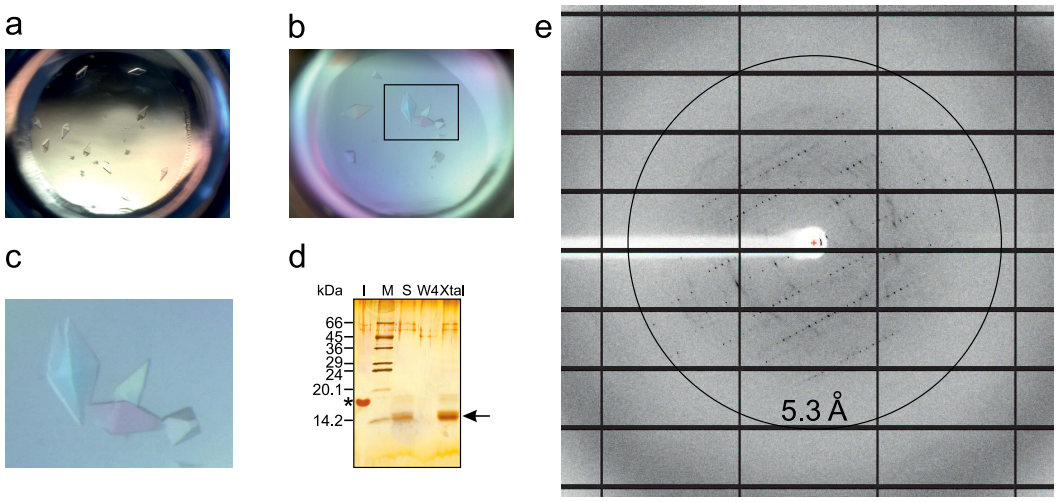


Figure 2.5.: Crystallization and X-ray diffraction of Tim 1008-1125. (a) Initial crystals of Tim appeared after 21 days in (a) 100 mM Tris pH 8.5, 18% PEG 4000 and (b) 100 mM MOPS pH 7.2, 12% PEG 20000. The rectangle marks the enlarged area shown in (c). (d) Silver-stained SDS-gel of dissolved Tim 1008-1125 crystals. (I) Protein sample used for crystallization setup (marked by asterisk). (M) Protein marker in kDa. (S) Supernatant of the spun-down drop containing the crystals. (W4) Content of the last washing step. (Xtal) Washed and dissolved Tim 1008-1125 crystals (marked by arrow). (e) Diffraction image obtained from crystals in (b) diffracting to 5.3 Å. The black circle shows the 5.3 Å resolution range.

of the reproduction of the crystallization is likely due to *in situ* degradation. Furthermore, low resolution of the collected data made the crystal structure determination very challenging.

To tackle the problem of protein degradation, the strategy was modified and longer Tim fragments were designed based on secondary structure prediction and sequence conservation. In addition, the expression in insect cells was attempted to solubilize full-length Tim.

Expression of Tim in insect cells was established by screening for the best expression conditions. After the optimization by varying the amount of the virus, the protein was purified using IMAC and SEC and yielded soluble protein. The protein was only stable in high sodium chloride concentrations (>300 mM) and was prone to degradation.

During the last step of the purification using SEC (Superdex 200 16/60), Tim (full-length) eluted at 47 mL corresponding to 9.5 MDa, while the theoretical molecular weight of Tim (full-length) is 162.3 kDa. This indicates that Tim (full-length) is in a soluble aggregated state (Figure 2.6 a, b). Nevertheless, to determine a stable protein core and to remove flexible parts which were prone to degradation, limited proteolysis

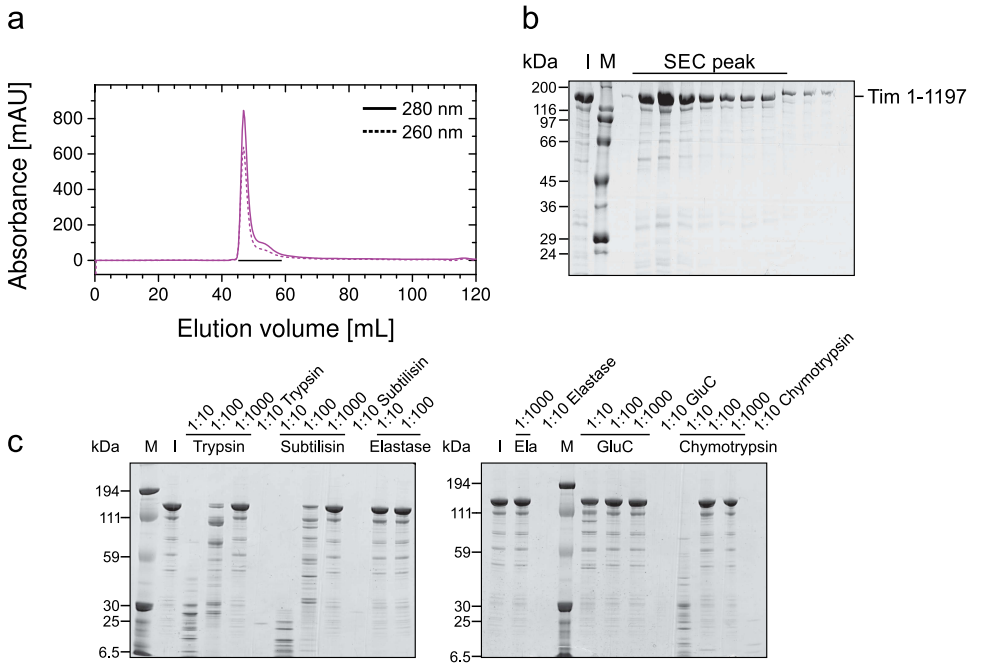


Figure 2.6.: SEC analysis of full-length Tim (aa 1-1197) using a Superdex 200 16/60 column. (a) Size exclusion chromatogram of purified Tim showing a peak at 47 mL (∼9.5 MDa) close to the void volume visualized by UV absorbance at 280 nm and 260 nm. (b) SDS-PAGE shows the peak fractions marked by a line. (I) Protein sample injected on gel-filtration column. (c) Limited proteolysis of Tim with several proteases. The ratio above each lane illustrates the dilution of the 1 mg/mL protease stock solution. (M) Protein marker in kDa. (I) Protein sample used for the proteolytic analysis.

was performed in the same way as described previously and the products were analyzed by SDS-PAGE (Fig. 2.6 c). Surprisingly, Tim was very stable in the presence of the tested proteases. Only by mixing high concentrations of Trypsin and Subtilisin, and at high concentrations of Chymotrypsin, Tim was susceptible to degradation but no pronounced fragments were observed.

In parallel, several constructs based on secondary structure prediction and sequence alignments were tested for expression in *E. coli* followed by pull-down experiments using cell lysates for the test of solubility. Besides the tested constructs as shown in Figure 2.1 and Figure 2.4 (Tim 997-1138 and Tim 1008-1125), only slightly longer C-terminal constructs of Tim yielded soluble protein despite using a solubility GST-tag, lower expression temperature and lower IPTG concentrations. All tested Tim fragments are listed in Figure 2.7. Although a longer construct (Tim 818-1138) could be purified, its limited proteolysis yielded the same proteolytic product of ∼13 kDa with

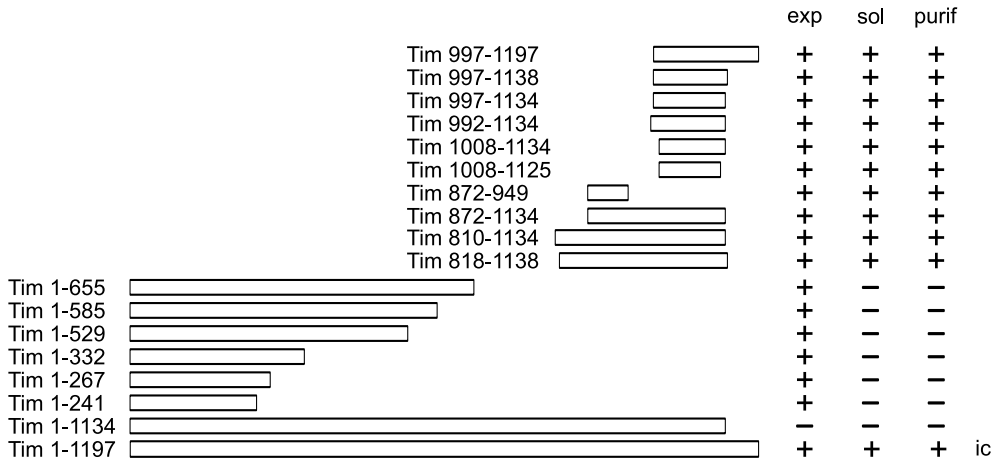


Figure 2.7.: Schematic representation of Tim constructs and boundaries tested for expression (exp), solubility (sol), and purification (purif). The success is indicated by plus (positive) or minus (negative). Only C-terminal fragments were soluble in *E. coli*. None of the N-terminal fragments yielded soluble protein. Tim 1-1134 was only soluble when co-expressed with Tipin (see Section 2.1.3 for further details). Full-length Tim (1-1197) was solubilized in insect cells (ic), but eluted at void volume in SEC (Figure 2.6 a).

Subtilisin (marked by asterisk, Figure 2.8 and Figure 2.3 for comparison). A slightly longer construct was identified with Elastase and the boundaries were identified to be aa 950-1129 (marked by plus, Figure 2.8). This construct was not followed up, because it contained an E-rich region at the N-terminus (aa 964-985). The construct 992-1134 was cloned instead.

The soluble Tim fragments were purified and set up for crystallization. None of the crystallization experiments of the rationally designed constructs or their proteolytic cores yielded better diffracting crystal than the crystals shown in Figure 2.2 and Figure 2.5.

2.1.2. Purification of Tipin and Crystallization Attempts

To examine the structure of the Timeless-interacting protein, full-length Tipin (aa 1-278) was first tested for expression in *E.coli*. The protein was well expressed using a His-GST solubility tag and eluted as a single peak with a small shoulder at 65 mL (∼860 kDa, theoretical MW 31.9 kDa) indicating formation of higher oligomers (Figure 2.9 a, b). To define the core suitable for crystallization, the protein was subjected to proteases (Figure 2.9 c). Three proteolytic products were observed by SDS-PAGE and

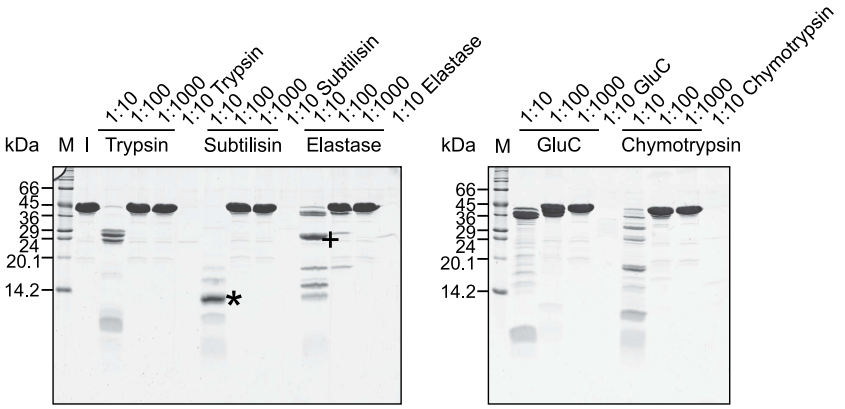


Figure 2.8.: Limited proteolysis of Tim 818-1138. The protein was incubated with five different proteases, each for 30 min on ice. The asterisk marks the stable product that exhibits the same size as the proteolytic fragment from Tim 997-1138 (Figure 2.3). The plus marks a construct (aa 950-1129) that contains an N-terminal E-rich unstructured region, which was not followed up. Instead, 992-1134 was tested for expression (see Figure 2.7). (M) Protein marker in kDa. (I) Protein sample used for the proteolytic analysis.

further examined by N-terminal sequencing and ESI-MS analysis to define the boundaries. Fragments (1), (2) and (3) correspond to aa 1-125, 1-136 and 1-169. In addition, based on secondary structure prediction and sequence alignments further constructs were designed as depicted in Figure 2.10 and tested for expression and solubility by pull-down experiments.

The constructs deriving from limited proteolysis (1)-(3) and three rationally designed constructs (55-150, -169, -220) were soluble but were highly prone to degradation during expression and purification (Figure 2.10). Therefore only full-length Tipin was submitted for crystallization experiments but did no yield any crystals.

Since individual full-length Tim and Tipin eluted as high oligomers and limited proteolysis did not result in more stable and better behaving protein fragments, co-expression of Tim and Tipin was tested (Section 2.1.3).

2.1.3. Solubilization of Timeless by Co-Expression with Tipin

Full-length Tim is not soluble in *E. coli* as proven by expression tests in normal (BL21) and codon-optimized (BL21 pRARE) *E. coli* cells followed by pull-down on Ni-beads (Figure 2.11 a, lane 4-7). Only C-terminal fragments of Tim can be expressed and purified as described in Section 2.1.1.

To solubilize Tim in the bacterial system, co-expression with its interacting partner

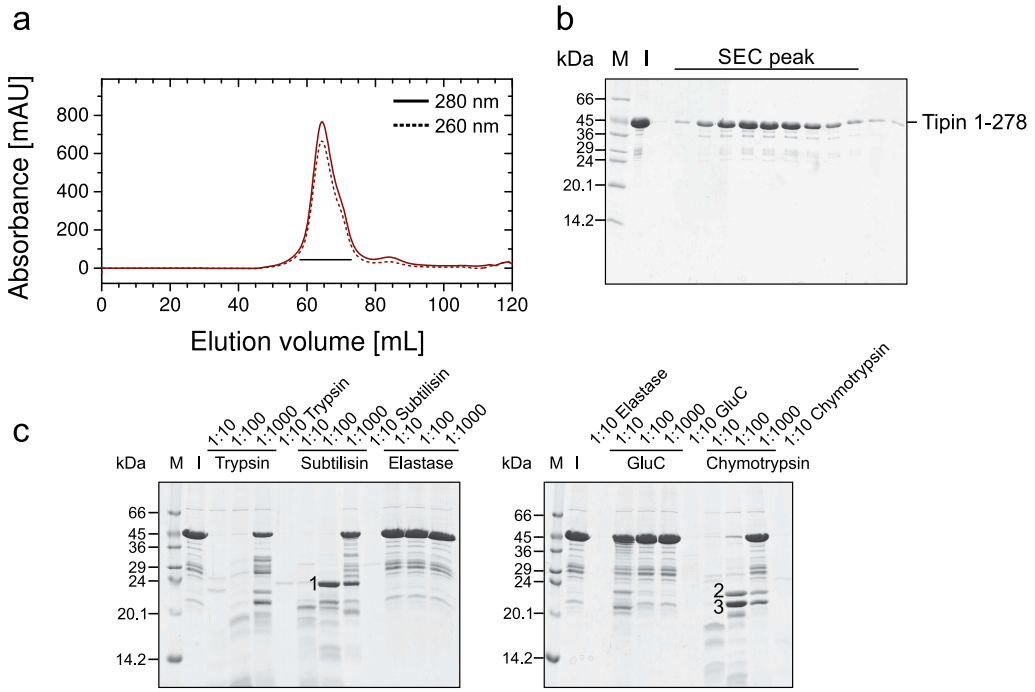


Figure 2.9.: SEC analysis of full-length Tipin (aa 1-278) using a Superdex 200 16/60 column. (a) Size exclusion chromatogram of purified Tipin showing a peak at 65 mL (∼860 kDa) indicating formation of higher oligomers visualized by UV absorbance at 280 nm and 260 nm. (b) SDS-PAGE shows the peak fractions marked by a line. (I) Protein sample injected on gel-filtration column. (c) Limited proteolysis of Tipin with several proteases. (M) Protein marker in kDa. (I) Protein sample used for the proteolytic analysis. (1),(2) and (3) indicate stable Tipin fragments.

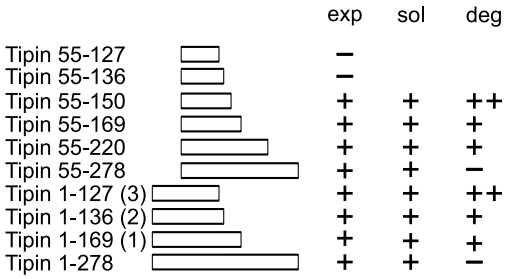


Figure 2.10.: Schematic representation of Tipin constructs and boundaries tested for expression (exp), solubility (sol), and degradation (deg) in *E. coli*. The success is indicated by plus (positive) and minus (negative). The proteolytic products are numbered (1), (2), and (3). Most of the designed constructs were soluble, but were prone to degradation.

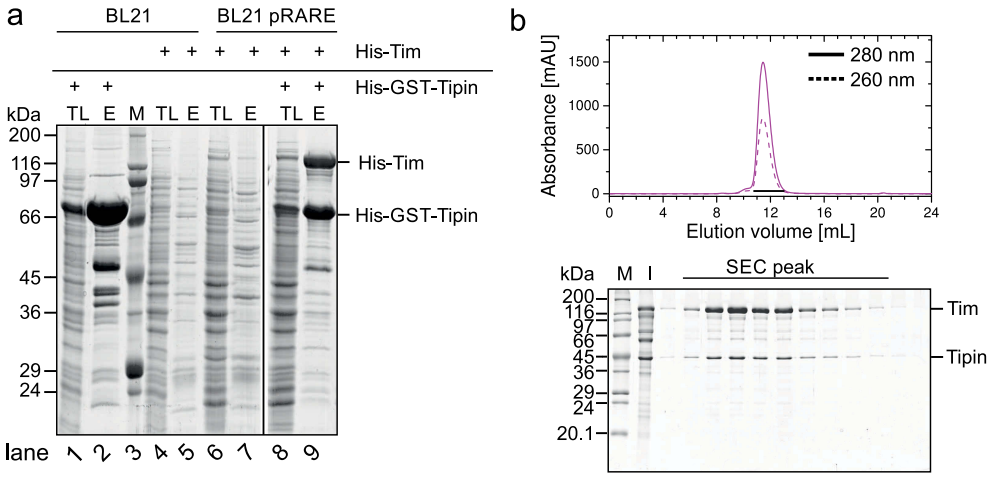


Figure 2.11.: Solubilization of Tim by co-expression with Tipin. (a) Recombinant expression in *E. coli* (TL, total lysate) and *in vitro* pull-down assays (E, Eluate) of His-Tim and His-GST-Tipin using GSH beads (lane 1, 2 and 8, 9) and Ni beads (lane 4-7). Coomassie stained SDS-gels are shown. Tim alone is not expressed (lane 4-7). In contrast, co-expression with Tipin solubilized Tim (lane 8, 9). (b) SEC of purified Tim-Tipin complex showing a single peak at 11.5 mL corresponding to 280 kDa and visualized by UV absorbance at 280 nm and 260 nm (top). SDS-PAGE analysis of the SEC peak marked by a line (bottom). (M) Protein marker in kDa. (I) Protein sample injected on SEC column. Adopted from [128].

Tipin was tested. First, almost full-length Tim was cloned containing aa 1-1134. These boundaries are derived from limited proteolysis attempts described in Section 2.1.1.

Co-expression of His-tagged Tim (aa 1-1134) with full-length His-GST-tagged Tipin (aa 1-278) solubilized Tim. These co-expressed proteins formed a stable complex on GSH beads in pull-down experiments (Figure 2.11 a, lane 8, 9). The Tim-Tipin complex could be further purified using IMAC, anion-exchange chromatography (AEX) and SEC (see Section Materials and Methods) resulting in 1.4 mg of pure protein complex per 1 L TB culture (23 g cell pellet) (Figure 2.11 b).

Next, limited proteolysis (Figure 2.12) was performed on the Tim-Tipin complex to identify the interaction core. Tim-Tipin was highly stable in presence of the tested proteases, except for high concentrations of Subtilisin and Chymotrypsin, where the complex was completely degraded. The 1:100 Trypsin dilution showed several stable bands in SDS-PAGE in the preliminary (Figure 2.12 a) and time-optimized experiment (Figure 2.12 b). The 1:100 Trypsin dilution with 15 min and 200 min incubation time were chosen for large scale limited proteolysis. The proteins were digested into a number of fragments but these fragments co-eluted together in one peak (15 min, Figure 2.12 c

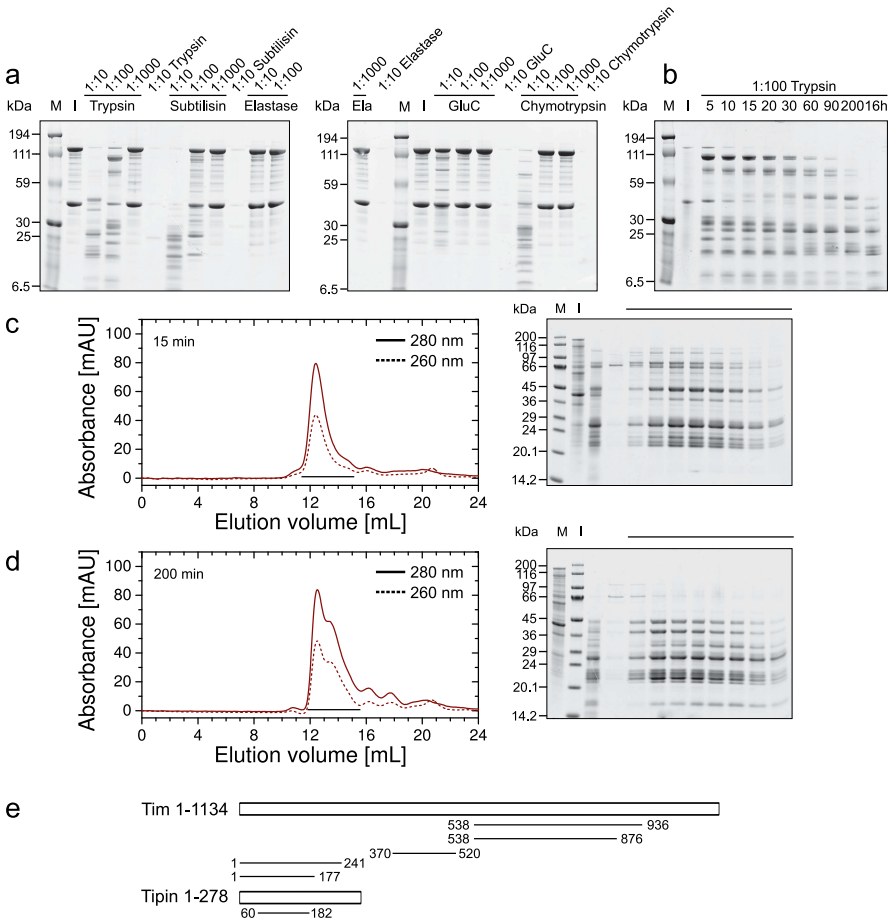


Figure 2.12.: Limited proteolysis of the Tim-Tipin complex. (a) The complex was incubated with five different proteases, each for 30 min on ice. The ratio above each lane illustrates the dilution of the 1 mg/mL protease stock solution. (b) Time course of the 1:100 Trypsin condition. The numbers correspond to time in minutes. (h) hour. (c) Large scale proteolysis using the same Trypsin dilution as in (b). The protein was subjected to the protease for 15 min. Before SEC analysis using a Superdex 200 10/300 GL column the reaction was stopped by adding a protease inhibitor. (d) Large scale proteolysis as in (c) with 200 min incubation time. (e) Proteolytic stable constructs of Tim-Tipin determined by N-terminal sequencing and ESI-MS. (M) Protein marker in kDa. (I) Protein sample used for the proteolytic analysis.

and 200 min, Figure 2.12 d). The results suggest that proteolysis occurs at flexible regions without affecting the entire architecture of the complex. Subsequent N-terminal sequencing and ESI-MS (Fig. 2.12 e) identified the amino acids recognized by the proteolysis. These identified residues correspond to loop regions between helices based on the secondary structure predictions (Section 1.4) and did not show any sign of flexible

domains that are separable from the core.

Taken together, Tim and Tipin were only stable when they were co-expressed and co-purified and formed a stable protein core. For the following experiments this stable complex (Tim 1-1134 and Tipin 1-278) was chosen to examine the interaction with RPA (Section 2.2.1).

2.1.4. Purification of Replication Protein A

To carry out the complex formation of Tim-Tipin and RPA, individual RPA had to be purified. RPA is a trimeric complex consisting of three RPA proteins (RPA70, RPA32, RPA14) (Section 1.4.3). In this thesis RPA's DNA-binding core (RPA70ABC, RPA32D) together with the RPA14 and RPA32 WH-domain was used. On the basis of published crystal structures of human RPA, sequence alignments, and secondary structure predictions (Section 1.4.3), homologous constructs of mouse RPA were designed for recombinant expression in *E. coli*, namely RPA70 DBD-A, -B, -C (aa 190-623), RPA32 DBD-D and WH-domain (aa 43-270), and full-length RPA14 (aa 1-121). All three RPA subunits were cloned as N-terminal His-fusions and expressed in BL21 (DE3) gold by co-transformation of the cells with three plasmids. As a DNA binding protein, the complex had the tendency to be contaminated with nucleic acids. To remove the DNA, DNaseI and Urea were added to the cell lysate. In addition a high salt wash was performed during IMAC. Finally, the RPA complex was purified to homogeneity by AEX and SEC (Figure 2.13) as described in Materials and Methods. RPA eluted

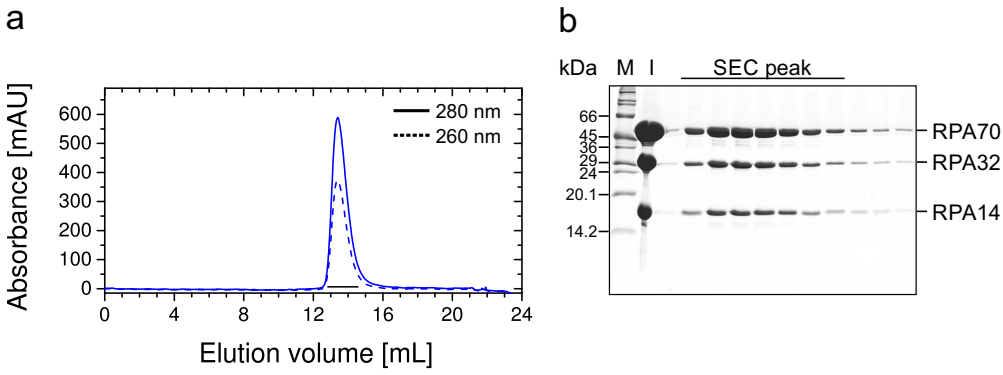


Figure 2.13.: Purification of RPA. (a) SEC analysis of RPA using a Superdex 200 10/300 GL column. The elution profile shows a single peak at 13.4 mL corresponding to 105 kDa and was visualized by UV absorbance at 280 nm and 260 nm. (b) SDS-PAGE of the peak fractions marked by line. (M) Protein marker in kDa. (I) Protein sample injected on SEC column. Adopted from [128].

as a single peak in SEC at elution volumes corresponding to 105 kDa (estimated based on molecular weight standard proteins) and contained all three subunits with a 1:1:1 stoichiometry. The yield of the purified complex was 0.3 mg protein/24 g cell pellet.

2.2. Reconstitution of the Tim-Tipin-RPA Complex

2.2.1. Timeless-Tipin in Complex with Replication Protein A

To analyze the interaction of Tim-Tipin and RPA, the individual components were purified (Section 2.1.3 and 2.1.4). To test if the purified RPA complex interacts with Tim-Tipin, *in vitro* GST-pull-down experiments using purified His-GST-tagged Tim-Tipin and purified His-tagged RPA were performed (Figure 2.14 a). RPA does not unspecifically bind to GSH-beads or GST ((a) lane 2, 3). His-GST-Tim-Tipin was able to precipitate His-RPA ((a) lane 5) proving the complex formation. The homogeneity of the complexes was tested by native PAGE (Figure 2.14 b, top). RPA and Tim-Tipin

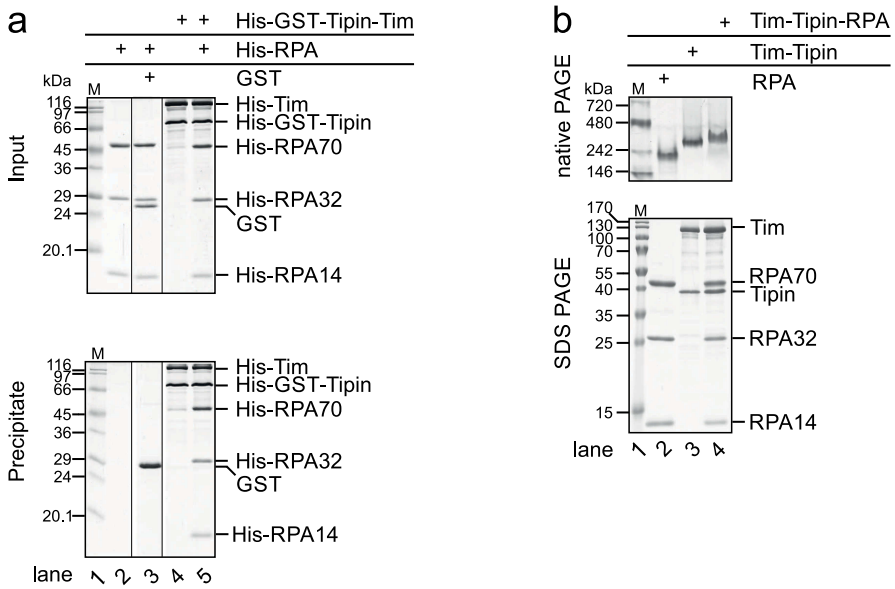


Figure 2.14.: Qualitative analysis of the Tim-Tipin-RPA complex formation. (a) GST-pull-down experiment showing binding of His-RPA to His-GST-Tim-Tipin. The Coomassie stained SDS-gels show the input (top) and precipitate (bottom). (M) Molecular weight marker in kDa. (b) Native PAGE analysis demonstrates a homogenous complex formation of RPA, Tim-Tipin, and Tim-Tipin-RPA (top). SDS-PAGE of the protein samples used in native PAGE (bottom). (M) Native molecular weight marker in kDa. Adopted from [128].

ran as a single band ((b) lane 2, 3). A shift in mobility to higher molecular weight was observed for Tim-Tipin in presence of RPA, when both complexes were mixed in a 1:1 ratio ((b) lane 4). SDS-PAGE analysis demonstrated the presence of all proteins used for native PAGE analysis ((b) bottom).

To examine the Tim-Tipin-RPA complex formation quantitatively, analytical SEC and static light scattering (SLS) were performed. Individual Tim-Tipin or RPA eluted as a single peak (RPA, blue, 110 kDa; Tim-Tipin, magenta, 321 kDa; Figure 2.15 and Table 2.1). The 1:1 molar ratio mixture of RPA and Tim-Tipin shifted to higher molecular weight (Tim-Tipin-RPA, green, 388 kDa). The mixture with one molar excess of RPA (2:1) showed a peak for Tim-Tipin-RPA and a peak from the excess of RPA, which was not incorporated (cyan). The molecular weight of the proteins was estimated by means of the calibration of the gel-filtration column by molecular weight standard proteins and the elution volumes of the complexes (Table 2.1, second and third column).

Since the experimental molecular weights of Tim-Tipin and Tim-Tipin-RPA determined by SEC were higher than the theoretical molecular weights (Table 2.1, sixth column), and the elution volumes of the protein complexes were distant from that predicted by the calibration of the column, static light scattering was performed to determine molecular weights more precisely. SLS measurements were conducted either by Dr. Claire Basquin, Dr. Jörg Tittor at MPIB or by myself. The experimental (SLS) molecular weight for RPA is 106 ± 6 kDa, for Tim-Tipin 164 ± 12 kDa and for Tim-Tipin-RPA 273 ± 31 kDa (Figure 2.16), which is in line with the theoretical molecular weights and corresponds to a 1:1:1 RPA complex, 1:1 Tim-Tipin complex and 1:1:1 Tim-Tipin-RPA complex (Tim:Tipin:RPA). Notably, the Tim-Tipin-RPA complex showed faint dissociation in the SLS analysis.

Table 2.1.: Elution volumes, experimental, and theoretical molecular weights of RPA, Tim-Tipin, and Tim-Tipin-RPA. Adopted from [128].

Protein complex	Elution volume [mL]	Experimental$^{SEC\delta}$ molecular weight [kDa]	Experimental$^{SLS\%}$ molecular weight [kDa]	Experimental^{SG*} molecular weight [kDa]	Theoretical$^{\#}$ molecular weight [kDa]
RPA	1.63 ± 0.02	110	106 ± 6	n.d.	96.2
Tim-Tipin	1.46 ± 0.01	321	164 ± 12	158	162.3
Tim-Tipin-RPA	1.43 ± 0.02	388	273 ± 31	n.d.	258.5

$^{\delta}$ Molecular weights derived from size exclusion chromatography (SEC) were estimated based on the comparison with molecular mass standard proteins.
$^{\%}$ Molecular weights determined by static light scattering (SLS).
* Molecular weights determined by sucrose-gradient sedimentation assay (SG).
$^{\#}$ Molecular weights calculated based on the amino acid sequence.
n.d. = not determined.

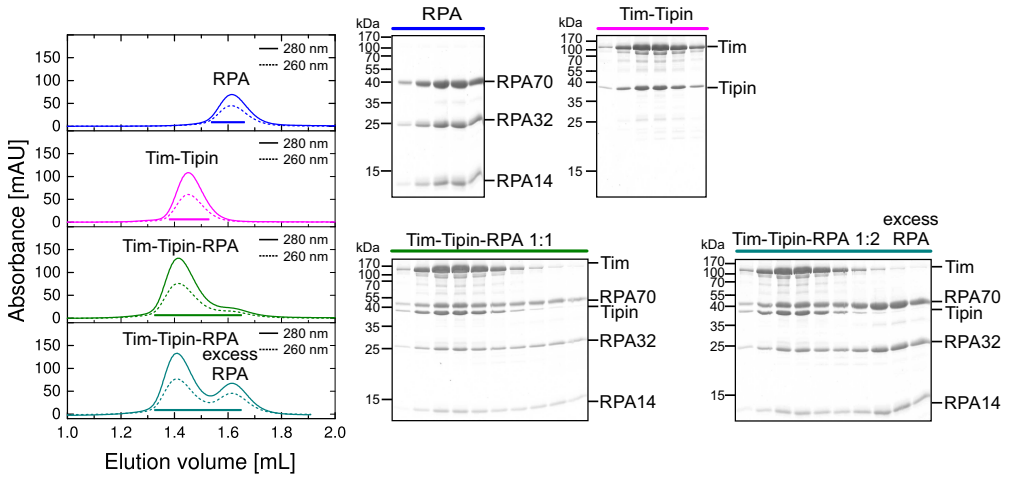


Figure 2.15.: Quantitative analysis of the Tim-Tipin-RPA complex formation. SEC analysis of RPA (blue), Tim-Tipin (magenta), Tim-Tipin-RPA at a 1:1 (green) and at a 1:2 molar ratio (Tim-Tipin:RPA). The Tim-Tipin-RPA complex shifted to higher molecular weights demonstrating complex formation. The elution profile was visualized by UV absorbance at 280 nm (solid) and 260 nm (dashed). Peak fractions marked by lines were analyzed by Coomassie stained SDS-gels (right). Adopted from [128].

Finally, to estimate the size of the Tim-Tipin complex with an independent method from SEC and SLS, sucrose-gradient centrifugation (SG) was performed. Tim-Tipin (162 kDa) migrated at fractions 7-9, with a maximum population in fraction 8 and showed the same sedimentation behavior as γ-globulin (158 kDa), which also migrated with a maximum population in fraction 8 (Figure 2.17). The analysis by SG further confirmed the molecular weight of Tim-Tipin.

2.2.2. Binding Affinity of Tim-Tipin to RPA

To determine the binding affinity (K_D) of Tim-Tipin and RPA, microscale thermophoresis (MST) was performed. MST uses the thermophoretic movement of molecules within an infrared laser induced temperature gradient and measures the change of that movement due to alterations of the hydration shell and charge or size of the molecule upon binding of the interaction partner. The motion is monitored by fluorescence of the molecule. Before MST measurement, RPA was fluorescently labeled with Cy3 as described in Section 4.2.5. The average number of lysines labeled per RPA was calculated by Equation 4.2 to be 4.4. The integrity and homogeneity of labeled RPA was tested by native PAGE and analytical SEC (Figure 2.18).

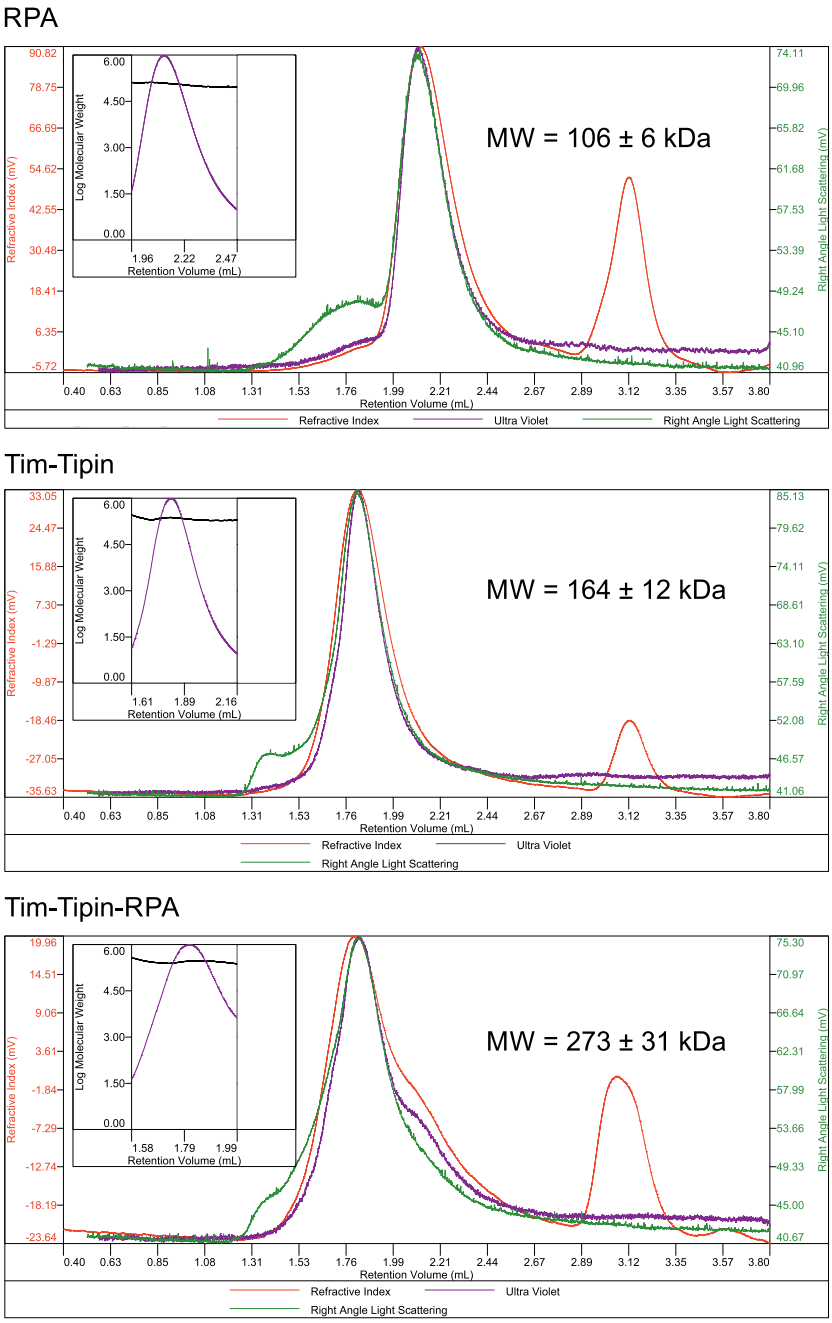


Figure 2.16.: Molecular weight determination of RPA, Tim-Tipin, and Tim-Tipin-RPA by static light scattering (SLS). Representative profiles are shown. The curves are UV absorbance (violet), light scattering (green), and refractive index (red) for the indicated protein complex solution. The top left panels are the average molecular weight values in the population of the peaks shown in logarithmic scale. The refractive index peaks at ∼3mL show no UV absorbance and no light scattering indicating small molecules (e.g. glycerol). Adopted from [128].

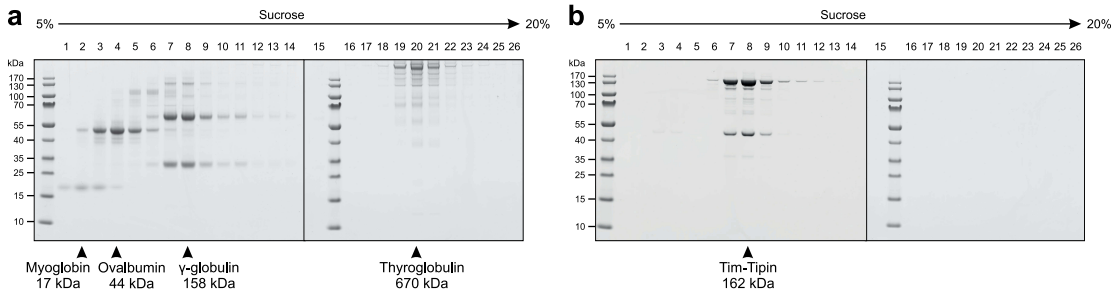


Figure 2.17.: Molecular weight determination of Tim-Tipin by sucrose-gradient centrifugation analysis (SG) and comparison to standard proteins. Coomassie stained SDS-gels of approximately 500 µL fractions from a 5-20% (w/v) sucrose-gradient with sedimented (a) standard proteins and (b) Tim-Tipin complex are shown. The standard proteins were Myoglobin (17 kDa), Ovalbumin (44 kDa), γ-globulin (158 kDa for the entire molecule of two light and two heavy chains), and Thyroglobulin (670 kDa for entire homo-dimeric molecule). The Tim-Tipin complex (162 kDa) migrated at the same position within the gradient as γ-globulin (158 kDa). Adopted from [128].

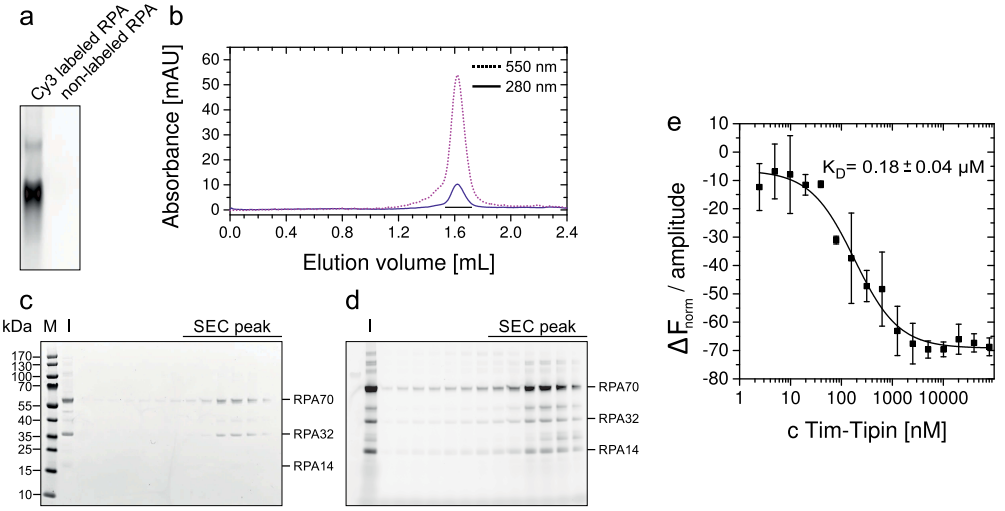


Figure 2.18.: Fluorescence labeling of RPA and binding of Tim-Tipin to RPA analyzed by microscale thermophoresis (MST). (a) Fluorescence image of native PAGE analysis of Cy3 labeled RPA. Non-labeled RPA as control. (b) Analytical SEC analysis of RPA using a Superose 6 3.2/PC column. The elution profile was visualized by UV absorbance at 280 nm (solid) and 550 nm (dashed). (c) Peak fractions marked by a line in (b) were analyzed by Imperial stained SDS-gel. (d) Fluorescence image of SDS-PAGE shown in (c). Note, RPA14 was only visible in the fluorescence image of the SDS-gel, likely due to the detection limit of the Imperial stain. (M) Protein marker in kDa. (I) Protein sample injected on SEC. (e) MST measurement. The concentration of Cy3 labeled RPA was kept constant at 100 nM while the Tim-Tipin concentration varied from 2.5 nM to 80550 nM. The resulting binding curve yields a K_D of 0.18 ± 0.04 µM. Adopted from [128].

The labeled RPA complex showed a homogenous behavior in native PAGE (Figure 2.18 a) and eluted as a single peak in analytical SEC showing the specific Cy3 absorbance at 550 nm (Figure 2.18 b). SDS-PAGE analysis displayed the presence and labeling of all RPA subunits (Figure 2.18 c, d). Finally, the determination of the K_D was performed using Cy3 labeled RPA, as previously described, and non-labeled Tim-Tipin in the MST measurement. The binding affinity of Tim-Tipin to RPA was measured to be 0.18 ± 0.04 µM (Figure 2.18 e).

2.3. Structural Studies of the Tim-Tipin-RPA Complex Using EM

2.3.1. Sample Preparation of Tim-Tipin-RPA Using GraFix

To obtain insights into the three dimensional (3D) architecture of Tim-Tipin-RPA, the complex was reconstituted using SEC (Section 2.2.1) and directly observed by negative stain electron microscopy (EM) (Figure 2.19). The EM micrographs showed low image contrast and the sample displayed slight heterogeneity, aggregates, and additional small densities. Despite dissociation of the complex particles (Figure 2.19 a), 1914 intact particles were selected and classification and averaging of those particles was performed. The 2D class averages showed an overall globular architecture with a ring- and U-like shape (Figure 2.19 a, A).

To reduce the problem of heterogeneity, likely due to particle dissociation during EM grid preparation, and to increase the image contrast, the GraFix method [129] was employed (Figure 2.19 b, c). During GraFix preparation, the complexes are gradually and chemically fixed by glutaraldehyde and are concurrently purified by sedimentation in a sucrose gradient. After cross-linking, the content of each fraction of the gradient containing the cross-linked particle sedimented according to size and the sample without glutaraldehyde as control were analyzed by SDS-PAGE (Figure 2.19 c, d). Three fractions of the cross-linked sample (C8, C10 and C12) displayed sharp and well-defined bands on the SDS-gel. Particles in fraction C8 migrated with higher molecular mass (300 ± 100 kDa), likely containing either large aggregates, different cross-linking levels or heterogeneous particle composition. Particles in fraction C12 migrated close to the 170 kDa protein marker band indicating complex dissociation (170 ± 70 kDa). Fraction C10 contained particles migrating close to the expected molecular weight of Tim-Tipin-RPA (theoretical MW 258 kDa, experimental MW $230, \pm 80$ kDa).

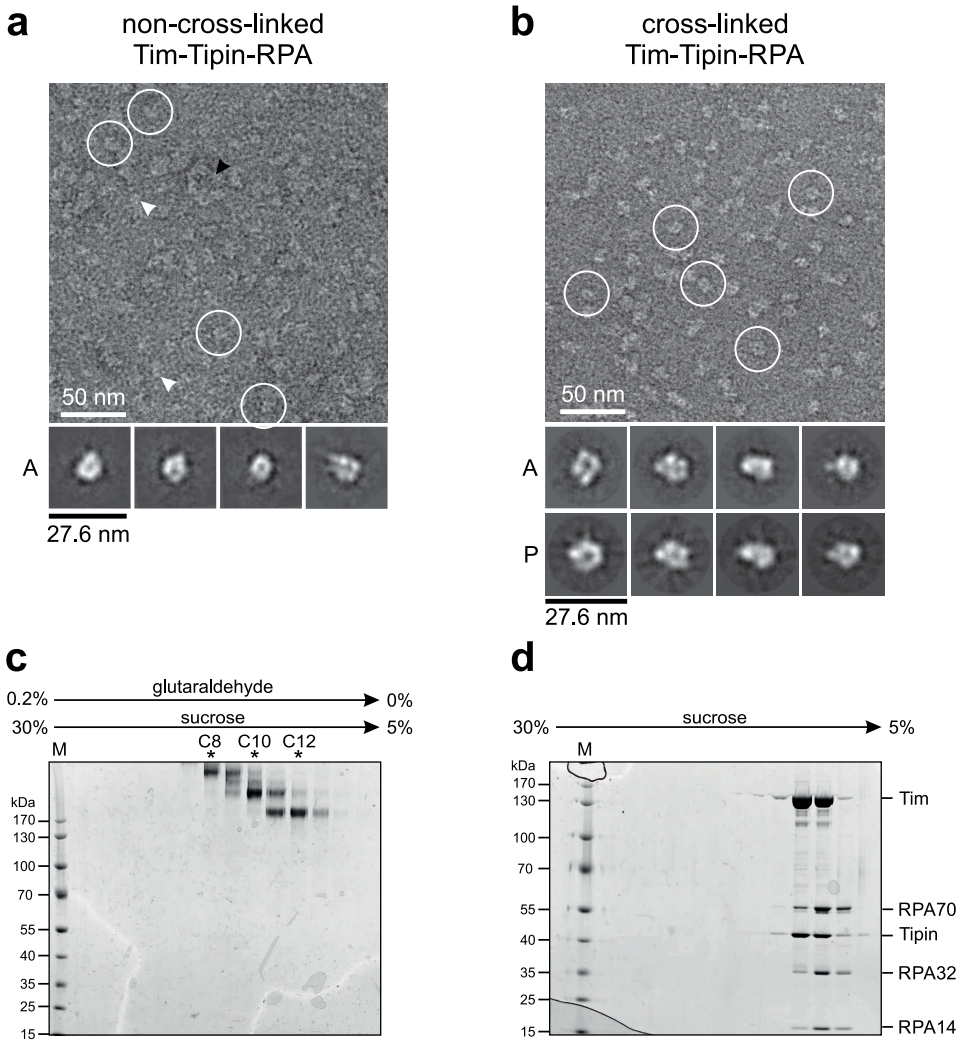


Figure 2.19.: Negative stain EM analysis and GraFix of Tim-Tipin-RPA. (a) Negative stain image of non-cross-linked Tim-Tipin-RPA. The white circles show examples of the Tim-Tipin-RPA complex. An aggregate and partial dissociation of the non-cross-linked complex are marked by a black and white arrowhead. (A) Representative 2D class averages. (b) Negative stain image of cross-linked Tim-Tipin-RPA. (A) Representative 2D class averages. (P) Reprojections of the 3D reconstruction. Fractions of (c) cross-linked Tim-Tipin-RPA using the GraFix method and (d) non-cross-linked Tim-Tipin-RPA analyzed by SDS-PAGE (4-12% Bis-Tris, Invitrogen). Fractions lacking glutaraldehyde were used as control (d). Fractions indicated by asterisk were selected for mass spectrometry. The stoichiometry of cross-linked Tim-Tipin-RPA in fraction C10 was calculated to be 1:1 (Tim-Tipin:RPA) using quantitative mass spectrometry analysis (iBAQ) as described in the Methods Section. Fraction C10 was used for subsequent EM analysis in (b). (M) Molecular weight marker. Adopted from [128].

Table 2.2.: Mass spectrometry (MS) analysis and intensity based absolute quantification (iBAQ). MS analysis identified all five proteins within the cross-linked samples. The stoichiometry of the protein complexes was determined by dividing the sum of all peptide peak intensities by the number of theoretically observable tryptic peptides. The quotient is specified in the table. Error bars for all log intensities were estimated to be ± 0.05. Adopted from [128].

	Fraction C8	Fraction C10	Fraction C12
Protein	log Intensity	log Intensity	log Intensity
RPA14	6.67	7.14	6.13
RPA32	6.41	7.09	6.02
RPA70	6.68	7.18	6.19
Tim	7.06	7.35	7.53
Tipin	6.74	7.00	7.19
average			
RPA	6.6 ± 0.2	7.1 ± 0.1	6.1 ± 0.1
Tim-Tipin	6.9 ± 0.2	7.2 ± 0.3	7.4 ± 0.2

To identify the presence of all components in the cross-linked samples and to determine their stoichiometry, a combination of trypsin digestion and mass spectrometry (MS) was performed (Table 2.2). MS analysis identified RPA14, RPA32, RPA70, Tim, and Tipin in all three fractions (C8, C10, C12). The stoichiometry of the Tim-Tipin and RPA subcomplexes was quantified using MaxQuant and the integrated iBAQ package (intensity based absolute quantification (iBAQ)) [130]. iBAQ revealed a 1:1 stoichiometry of Tim-Tipin to RPA in fraction C10 (log Intensity RPA 7.1 ± 0.1 and Tim-Tipin 7.2 ± 0.3, Table 2.2). Fraction C8 and C12 contained non-stoichiometric complexes with excess of Tim in C8 (log Intensity Tim 7.06, other proteins 6.41-6.74) and excess Tim-Tipin in C12 (log Intensity RPA 6.1 ± 0.1 and Tim-Tipin 7.4 ± 0.2, Table 2.2).

The negative stain EM analysis of the cross-linked sample was performed using fraction C10 (Figure 2.19 d). The resulting specimens were homogeneous and the electron micrographs showed a stable, monodispers complex of Tim-Tipin-RPA without any detectable aggregates for the cross-linked sample. In addition, the image contrast was increased with chemical fixation compared to the non-cross-linked complex (Figure 2.19 a). The 2D class averages of 13311 selected particles of the GraFix sample displayed similar features (Figure 2.19 b, A) in comparison to the 2D class averages of the non-cross-linked sample (Figure 2.19 a, A). The cross-linked Tim-Tipin-RPA 2D classes revealed sharp features in various views displaying a core ring-like shape, a U-shape and particles with four different globular domains (Figure 2.19 b, A). The complex showed a uniform size of ∼150 Å suggesting an overall globular architecture.

2.3.2. Initial Model Generation by Random-Conical Tilt and Negative Stain EM Reconstruction of the Tim-Tipin-RPA Complex

Single-particle EM analysis helps to determine and to understand the structure of biological macromolecules from micrographs. One of the key points in this analysis is the generation of a 3D map out of a series of 2D images. To obtain a starting 3D model, the angles positioning the 2D images into the right orientation in 3D space have to be determined. The random conical tilt (RCT) method [131] was used for the angular assignment. The principle of RCT is to collect untilted and tilted micrographs pairs of the same particles, with known angles of the tilted images and the tilt axis direction. Coordinating the images of the untilted and tilted particles relative to each other in 3D space by the known angles and tilt axis, results in an *ab initio* 3D model. Here, 857 pairs of untilted and tilted particles (Figure 2.20 a, b) were interactively selected, the corresponding angles were assigned, and eight initial models were generated as described in Radermacher et al. [131]. Two most similar models (Figure 2.20 c, d) with prominent features were merged and used as the initial model (Figure 2.20 e).

For the subsequent 3D reconstruction the projection matching refinement method was employed by comparing the reprojections of the RCT model to the 2D class averages as well as raw data. The 3D reconstruction was improved by iterative reprojection matching of the merged RCT reconstruction as a reference and the 13311 selected untilted particles. The refined reconstruction at ∼23 Å resolution (according to the Fourier Shell Correlation (FSC) = 0.5 criterion) displayed an overall globular shape with the size of ∼150 Å x ∼120 Å x ∼120 Å (Figure 2.21). The most prominent feature of the model displays a ring-like density (Figure 2.21, front) with a ∼30 Å-wide channel (channel 1, dashed circle). This channel is closed at the bottom by a lid density (lid). The view rotated by 90°revealed a connection between the lid and the ring-like structure forming a U-like feature (side, U-shape). This U-shape feature incloses a second channel (channel 2, dashed circle), which is formed by the lid and the protruding density of the ring-like core to the back. The back view unveiled four different domains (marked by asterisks).

2.3.3. Cryo-EM 3D Reconstruction of the Tim-Tipin-RPA Complex

For cryo-EM reconstruction of the cross-linked Tim-Tipin-RPA complex, the specimen was frozen in vitreous ice (Figure 2.22 a). 39679 particles were chosen for the cryo-EM analysis. The 3D reconstruction of the Tim-Tipin-RPA complex was obtained by using the 3D model of the negative-stained samples as a reference (Figure 2.21). The

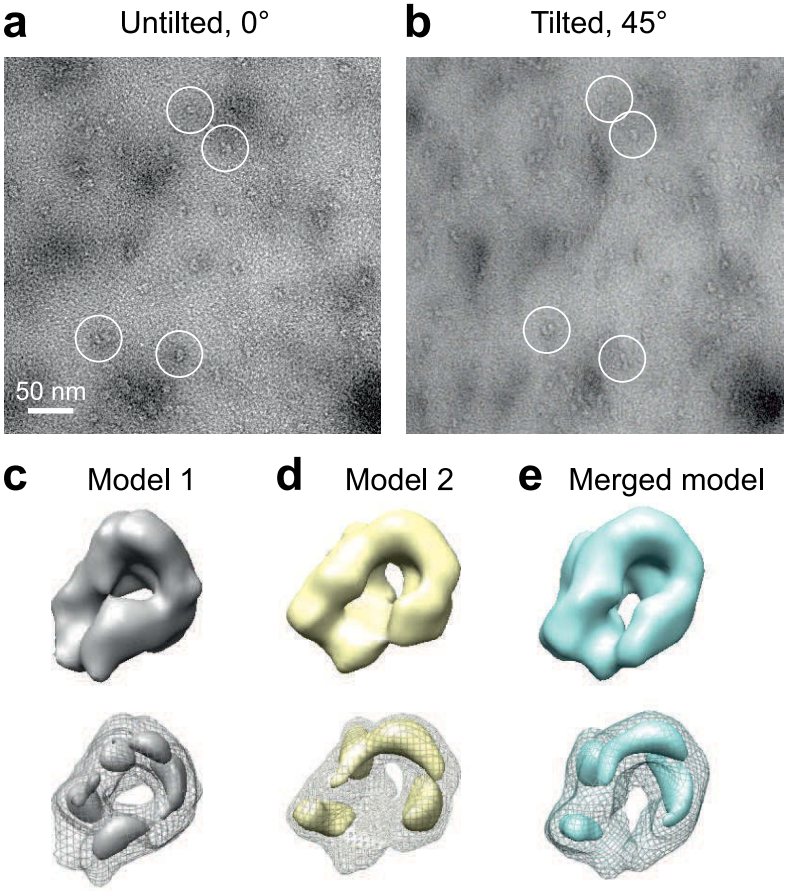


Figure 2.20.: Random conical tilt (RCT) reconstruction of Tim-Tipin-RPA. Corresponding pair of (a) untilted- and (b) tilted-specimen micrographs. Examples of the Tim-Tipin-RPA complex are marked with white circles. (c, d) Two similar RCT models showing prominent characteristics (top) and their mesh surface representation (bottom) showing the core density (solid surface). (e) Merged 3D density map used as initial RCT model. Representation as in (c) and (d). Adopted from [128].

3D model was refined by iterative projection matching resulting in a cryo-EM map at 17 Å resolution (Figure 2.22 b, c). The final Tim-Tipin-RPA model shows a ring-like structure (Figure 2.22 d, front) forming a ∼30 Å channel (channel 1) that is closed by a lid density at the bottom (lid). Further, a second channel (channel 2) is formed by the lid and the ring-like density. The back view of the 3D reconstruction displays four distinctive domains (back, marked by asterisks). The reprojections of the 3D cryo-EM model of the Tim-Tipin-RPA complex (Figure 2.22 a, P) showed good consistency when compared to the 2D class averages (Figure 2.22 a, A) for verification.

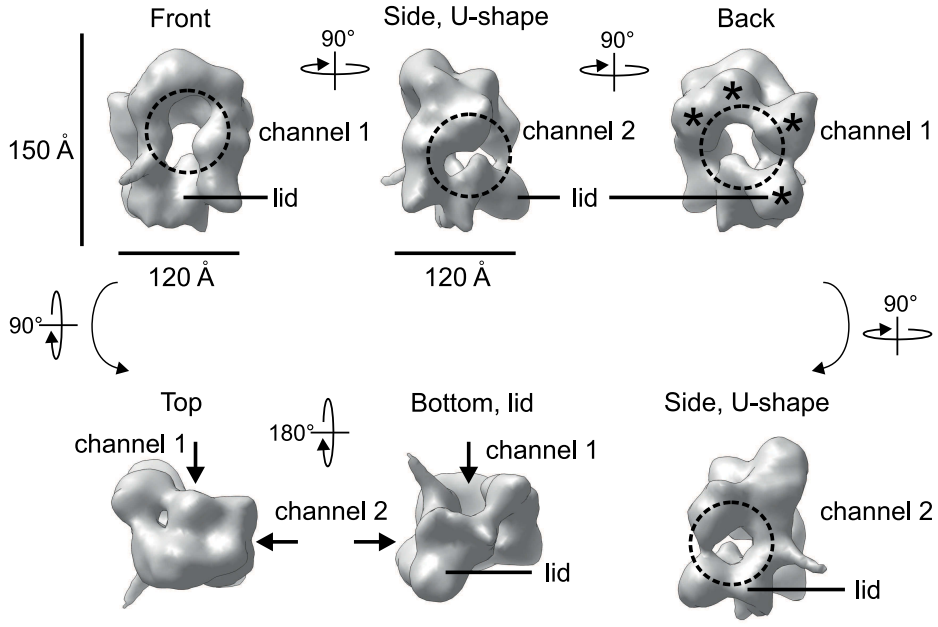


Figure 2.21.: Negative stain EM 3D reconstruction of Tim-Tipin-RPA in different views (front, side, back, top, and bottom) related by rotation around the y- and x-axis as indicated. The lid volume and the scale bar are marked. For more details see main text. Adopted from [128].

2.3.4. Negative Stain Analysis of the Individual Tim-Tipin Complex and RPA

To further elucidate the structural composition of the Tim-Tipin-RPA complex, negative stain EM of individual Tim-Tipin and RPA was performed. Tim-Tipin and RPA were cross-linked using the GraFix method and each fraction of the gradient was analyzed on SDS-PAGE after cross-linking as described for the Tim-Tipin-RPA complex (Section 2.3.1). Four fractions of the cross-linked Tim-Tipin complex showed distinct bands on the SDS-gel (Figure 2.23 a, b, asterisks). One fractions migrated close to the 170 kDa marker band, which is consistent with the theoretical molecular weight of the Tim-Tipin complex (162 kDa). Fraction G9 was chosen for negative stain EM analysis. The cross-linking of RPA showed less defined bands on SDS-PAGE. The fraction G9 migrated close to the expected molecular weight of RPA (96 kDa) (Figure 2.23 c, d) and was therefore chosen for negative stain EM analysis. The electron micrographs of the GraFix samples of Tim-Tipin and RPA showed good image contrast and a stable complex formation with no aggregates (Figure 2.23 e, f).

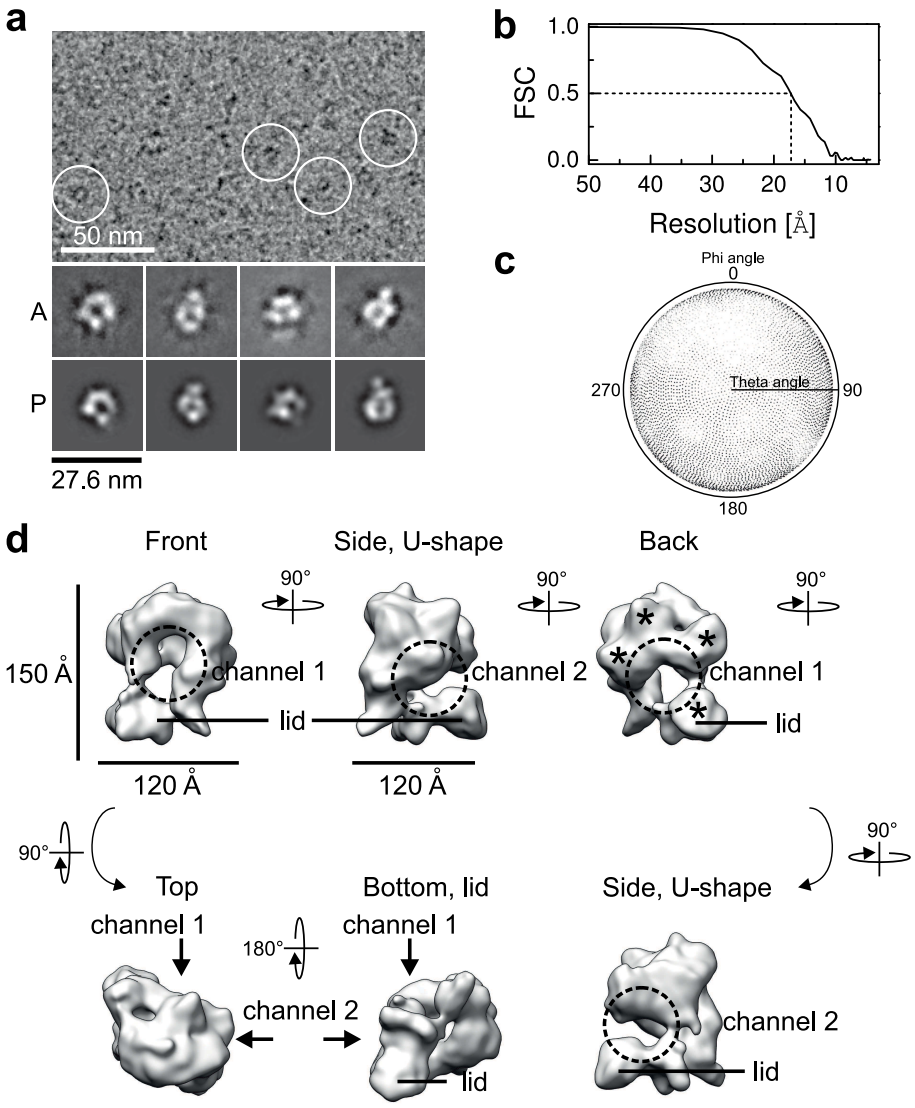


Figure 2.22.: Cryo-EM analysis of the Tim-Tipin-RPA complex. (a) Micrograph of Tim-Tipin-RPA under cryo-conditions. Examples of individual particles are circled in white. (A) Representative 2D class averages. (P) Projections of the 3D reconstruction. (b) Fourier shell correlation (FSC) plot of Tim-Tipin-RPA. The measured resolution is 17 Å using the 0.5 cut-off criterion. Selected particles were divided into two and two reconstructions were individually calculated for the estimation of the resolution by FSC. (c) Angular distribution plot showing the particles relative to the 3D volume in the final step of refinement. The intensity of the dots is proportional to the number of particles assigned to a certain reference reprojection by projection matching. (d) Cryo-EM reconstruction of the Tim-Tipin-RPA complex. The negative stain 3D model (Figure 2.21) was used as initial reference. Different views of the reconstruction are shown (front, side, back, top, and bottom) which are related by rotation around the y- and x-axis as indicated. The lid volume and the scale bar are marked. For more details see main text. Adopted from [128].

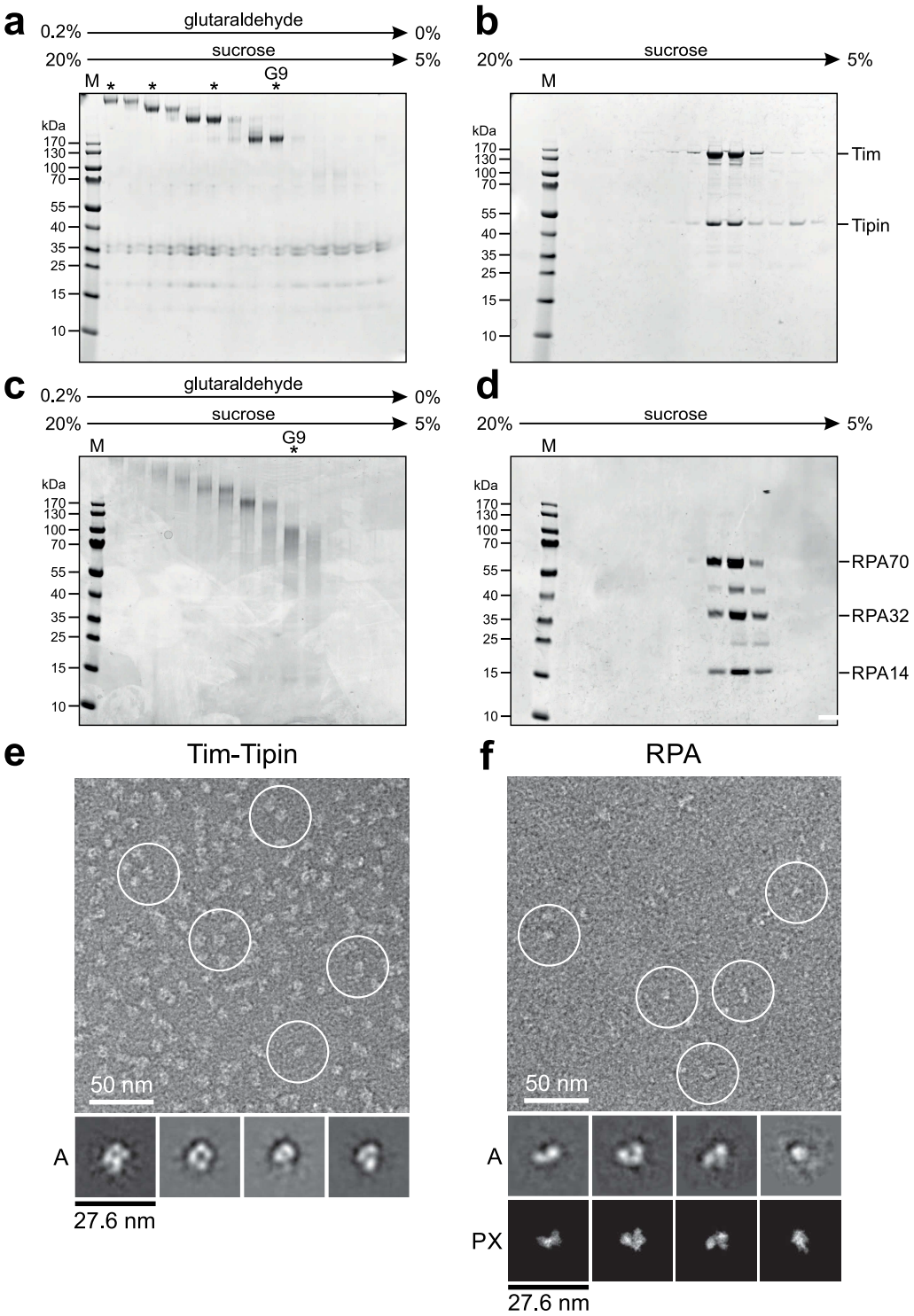


Figure 2.23.: *See next page for figure legend.*

Figure 2.23.: *(Previous page).* GraFix and negative stain EM analysis of individual Tim-Tipin and RPA. Fractions of (a) cross-linked Tim-Tipin, (b) non-cross-linked Tim-Tipin, (c) cross-linked RPA, (d) non-cross-linked RPA analyzed by SDS-PAGE (4-12% Bis-Tris, Invitrogen). Fractions lacking glutaraldehyde were used as control. Fractions indicated by G9 were selected for EM. (M) Molecular weight marker. (e) Negative stain images of the cross-linked Tim-Tipin complex indicate that the Tim-Tipin shape predominates the structure of the Tim-Tipin-RPA complex (Figure 2.22 a). (f) Negative stain images of cross-linked RPA show that RPA adopts a rod and horse shoe-shape structure. The comparison of the 2D class averages of RPA (A) and the reprojections (PX) of the crystal structure of RPA bound to 32 nt ssDNA [96] (PDB 4GNX) revealed similar features. White circles mark examples of the Tim-Tipin and RPA complex. (A) Representative 2D class averages. (PX) Reprojections of the crystal structure. Adopted from [128].

10872 particles for Tim-Tipin and 1673 particles for RPA were chosen for negative stain EM analysis. The 2D class averages of Tim-Tipin and RPA revealed specific structures. The Tim-Tipin complex showed a ring-like structure with ∼100 Å x ∼100 Å dimensions (Figure 2.23 e, A). RPA displayed a rod-like shape (∼90 Å x ∼55 Å) and a smaller horse shoe-shape (U-shape, ∼90 Å x ∼80 Å) with 2 or 3 globular densities (Figure 2.23 f, A). The comparison of the RPA 2D class averages to the reprojections of the crystal structure containing four DBDs-A-D of RPA bound to 32 nt ssDNA [96] displayed good agreement with the dimensions and structural features (Figure 2.23 f, A, PX). The reprojections showed similar shapes and displayed a rod-like and U-shape structure similar to the 2D class averages.

The comparison of the 2D class averages of Tim-Tipin-RPA (Figure 2.22 a), Tim-Tipin, and RPA only (Figure 2.23 e, f) shows that Tim-Tipin is the component responsible for the characteristic features of the Tim-Tipin-RPA complex showing a ring-like shape with comparable dimensions.

2.3.5. Antibody Labeling of RPA and Tipin within the Tim-Tipin-RPA Complex

To map the location of RPA and Tipin in the structure of the Tim-Tipin-RPA complex, antibody labeling of RPA and Tipin was performed. A commercially available polyclonal antibody against the N-terminus of Tipin and a monoclonal antibody against the RPA DBD-A (Figure 2.24 a) were chosen on the basis that their epitopes were different to RPA's binding site to Tim-Tipin or Tipin's binding site to RPA, respectively and therefore do not interfere with complex formation of Tim-Tipin and RPA (Figure 2.24). The binding of RPA DBD-A (Figure 2.24 a) to Tim-Tipin was tested and showed no signs of interaction according to the SEC analysis (Figure 2.34 b).

Cross-linked Tim-Tipin-RPA was incubated with the RPA antibody and negatively stained. For Tipin antibody labeling cross-linked Tim-Tipin-RPA was incubated with the Tipin antibody and the immune complex was further purified by SEC to better distinguish between labeled and non-labeled complexes. Images of the labeled Tim-Tipin-RPA complex showed a similar U-shape architecture as non-labeled Tim-Tipin-RPA (Figure 2.24 b). For the RPA-antibody labeled Tim-Tipin-RPA complex, the 2D class averages and raw images showed an additional density at the side of the U-like shape (Figure 2.24 b, arrowhead). The additional density appeared in the corresponding position of 3D reconstruction at channel 2 at the left-bottom side of the ring-like feature of Tim-Tipin-RPA (Figure 2.24 c). The lid domain likely corresponds to RPA accommodating RPA70 DBD-A and -B. Further, antibody labeling with the Tipin antibody revealed an additional density at the top right of the U-shape density (Figure 2.24 b, arrowhead). Therefore, the top right of the ring-like density likely accommodates the RPA32WH-Tipin-C-terminus interface. This result is consistent with previous reported protein-protein interaction studies [19, 20]. Finally, this result also suggests, that Timless is most likely located in the right part of the ring-like density in the 3D reconstruction.

2.3.6. Docking Analysis of RPA into the Tim-Tipin-RPA 3D Reconstruction

RPA is known to adopt different conformations dependent on the length of ssDNA and to undergo two major transitions upon ssDNA binding. Several studies have provided insights into RPA's architecture. To find out, which of these RPA conformations is most consistent with our EM model and to gain molecular insight into the structure of the Tim-Tipin-RPA complex (Figure 2.22), published crystal structures and SAXS models of RPA were docked into the EM model (Figure 2.25).

A small angle X-ray scattering (SAXS) analysis of RPA bound to ssDNA was published investigating RPA's DNA binding modes [97]. DNA-free RPA displays an extended conformation with a maximal diameter of 171 Å and very flexible relative DBD arrangement. Upon ssDNA binding, RPA undergoes a successive compaction as the coverage by ssDNA progresses. The first transition in the presence of short ssDNA ($\sim$10 nt) results in the compaction of DBD-A and -B, while the trimerization core of RPA (DBD-C, -D and RPA14) still remains flexibly tethered. The second transition upon binding to longer ssDNA substrates (>20 nt) leads to a more compact architecture by closer positioning the trimerization core to RPA70 DBD-A and -B. The crystal structure of RPA bound to 32 nt ssDNA revealed an even more compact architecture showing a horse

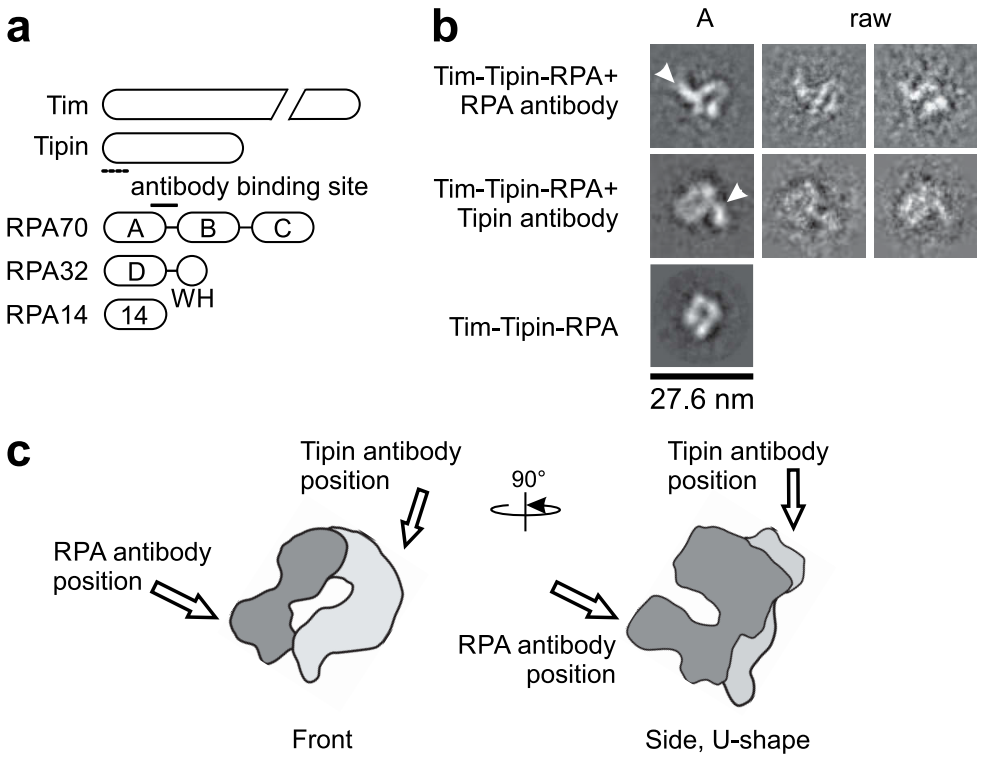


Figure 2.24.: Antibody labeling of Tim-Tipin-RPA using negative stain EM. (a) Domain architecture of the Tim-Tipin-RPA complex showing the antibody binding site within RPA70 DBD-A (black line, aa 268-304) and within the first 50 N-terminal amino acids of Tipin (dashed line). (A, B, C, D) DNA binging domains of RPA. (WH) Winged helix domain of RPA32. (14) RPA14. (b) Representative 2D class average (A) and raw images of negative stained Tim-Tipin-RPA with an antibody against the RPA70 subunit (first row), with an antibody against the N-terminus of Tipin (second row), and Tim-Tipin-RPA without antibody (third row). The RPA antibody binds to the left-bottom and the Tipin antibody to the top right of the ring-like density (marked by an arrowhead). (c) Schematic representation of the 3D model of Tim-Tipin-RPA and the antibody binding site in two different views. The putative localization of RPA within the model is represented as dark gray area and corresponds to a volume of approximately 96 kDa. The position of the RPA and Tipin antibody is indicated by an arrow. Adopted from [128].

shoe-like RPA conformation (PDB 4GNX) [96]. The published RPA structures were compared to the RPA structure in the Tim-Tipin-RPA 3D reconstruction as suggested by antibody labeling (Section 2.3.5). The SAXS model of DNA-free RPA [97] (BioIsis ID RPADCP) could not be fitted into the Tim-Tipin-RPA 3D model, leaving nearly half of the model outside of the envelope (Figure 2.25 a). Docking of the crystal structure of RPA bound to 32 nt ssDNA [96] and docking the corresponding SAXS model [97] fitted the cryoEM reconstruction (Figure 2.25 b, c). Finally, crystal structures of DNA-free

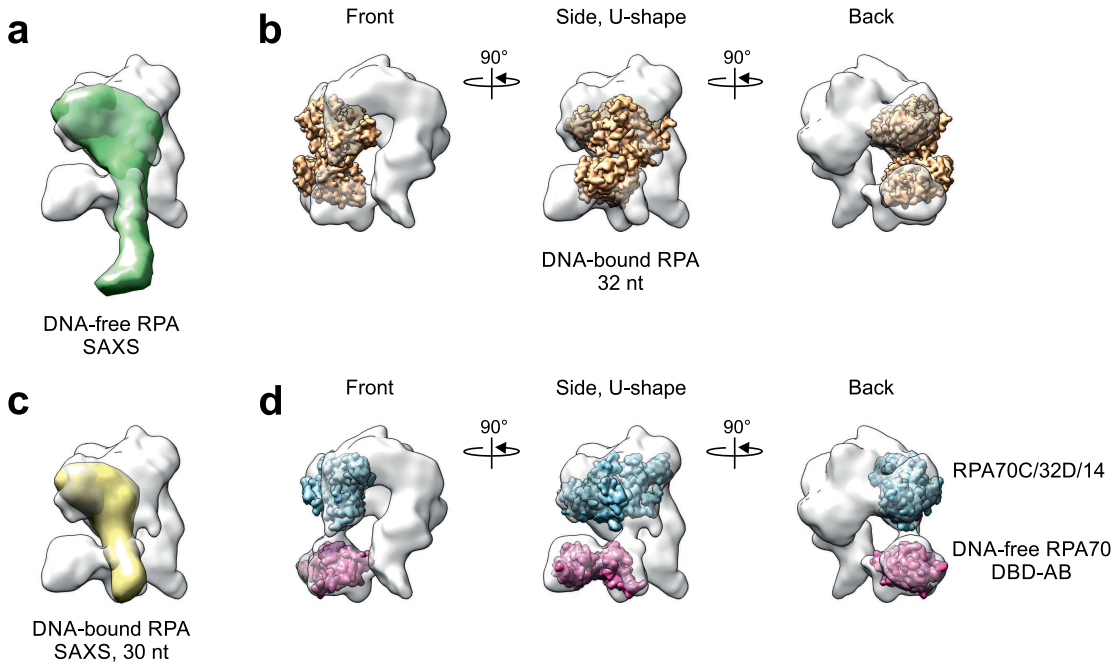


Figure 2.25.: Docking analysis. RPA models were fitted into the Tim-Tipin-RPA 3D reconstruction shown as transparent surface representation. RPA models: (a) SAXS model of DNA-free RPA (BioIsis ID RPADCP). (b) Crystal structure of RPA bound to 32 nt ssDNA (PDB 4GNX). (c) SAXS model of 30 nt DNA-bound RPA (BioIsis DBC30Y). (d) Crystal structure of the RPA trimerization core (blue, PDB 1L1O) and crystal structure of DNA-free RPA70 DBD-A and -B (magenta, PDB 1FGU chainA). Adopted from [128].

RPA70 DBD-A and -B and the trimerization core were individually placed in the 3D reconstruction. This approach takes the relative flexibility of the RPA domains into consideration and allows an optimized positioning of the subunits into the 3D density. The docking of the individual crystal structure of DNA-free RPA70 DBD-AB (PDB 1FGU chainA) and the RPA trimerization core (RPA70C/32D/14, PDB 1L1O) into the Tim-Tipin-RPA 3D reconstruction gave the best fitting (Figure 2.25 d). The docking analysis clearly shows that the rather compact conformation of RPA showing a horse shoe-like conformation as reported by Fan *et al.* 2012 and Brosey *et al.* 2013 is most consistent with the 3D reconstruction, while the extended conformation of RPA (171 Å in length) could not be reasonable accommodated in the EM model (Figure 2.25 a).

To verify that the compact RPA conformation in the 3D reconstruction occurs due to binding to Tim-Tipin and not due to the chemical fixation, the 2D class averages of the negative stained non-cross-linked sample were compared to the class averages and

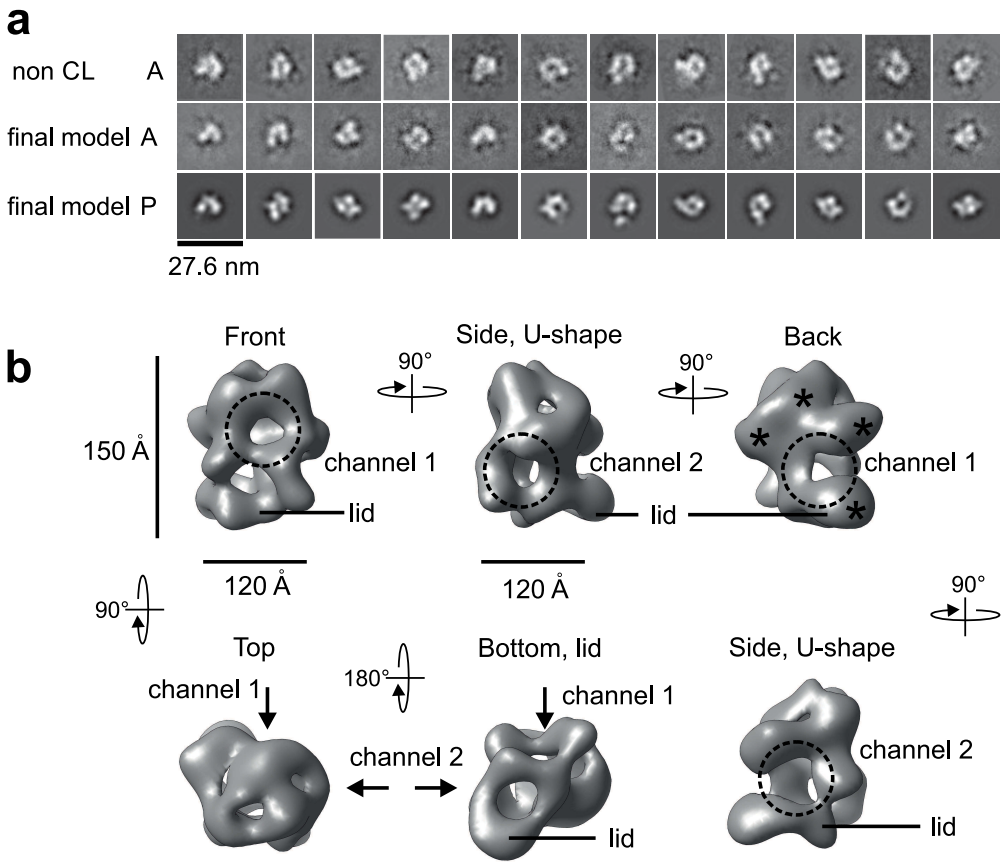


Figure 2.26.: EM analysis of non-cross-linked Tim-Tipin-RPA and comparison to the cryo-EM specimen. (a) Reference-free 2D class averages (A) of non-cross-linked Tim-Tipin-RPA (first row, non CL). Reference-free 2D class averages of cross-linked and vitrified Tim-Tipin-RPA (second row). Reprojections (P) of the final cryo-EM model (third row). (b) 3D reconstruction of non-cross-linked Tim-Tipin-RPA in various view as in Figure 2.21. The comparison of the non-cross-linked and the cross-linked specimens revealed that RPA adopts a compact conformation due to Tim-Tipin binding and is independent of the chemical fixation. Adopted from [128].

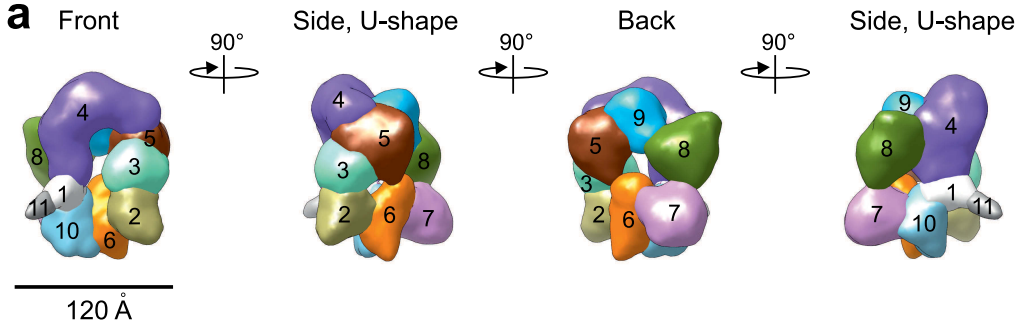
reprojections of the negative stain 3D reconstruction of cross-linked Tim-Tipin-RPA. The class averages of the untreated sample showed comparable features to the GraFix sample. (Figure 2.26 a). The 3D reconstruction of the non-cross-linked sample was performed using the negative stain GraFix EM model as reference. The gained overall structure revealed no difference on the structural arrangement of RPA (Figure 2.26 b) indicating that the binding of Tim-Tipin to RPA clamps RPA's conformation in the Tim-Tipin-RPA complex.

2.3.7. Tim-Tipin Assignment Using Difference Mapping

To assign the electron density of the Tim-Tipin complex in the 3D reconstruction of Tim-Tipin-RPA, 2D class averages of Tim-Tipin were aligned to the reprojections of various 3D Tim-Tipin-RPA models missing the putative RPA density (Tim-Tipin-ΔRPA).

First, fourteen 3D models displaying the electron density of Tim-Tipin only and missing the RPA region were generated by deleting the putative RPA density within the 3D Tim-Tipin-RPA map (Appendix B.1). These Tim-Tipin-ΔRPA 3D models were obtained by subtracting the RPA density enclosing the expected volume for RPA. For this analysis, the segmentation map function in Chimera [132] was used as described in Section 4.2.6. Briefly, the 3D model was sub-divided into eleven segments (Figure 2.27 a). Specific segments giving the final putative RPA volume were merged (Figure 2.27 b, merged segments) and subtracted from the 3D Tim-Tipin-RPA map resulting in fourteen models (Figure 2.27 b, difference map number 1-14 and Appendix B.1). Difference map number 15 was derived from the antibody labeling and docking analysis described in Section 2.3.5 and Section 2.3.6. Briefly, the volume recognized by the RPA antibody and accommodating the crystal structure of the RPA trimerization core and the crystal structure of DNA-free RPA70 DBD-A and -B was subtracted (Figure 2.24 c and Figure 2.25 d).

Tim-Tipin 2D class averages (Figure 2.23 e and Appendix B.2) were aligned to the reprojections of each difference map. The total number of the averages was plotted as bar chart against the reference to which the 2D class average was best aligned (difference map number, Figure 2.27 c). This approach assumes that the more Tim-Tipin 2D class averages are aligned to the reprojection of a specific reference/difference map, the better this reference represents the Tim-Tipin electron density. Reprojections of difference maps missing the lid region and parts of the U-shape electron density (number 14, 15) showed the highest number of aligned 2D class averages (14 and 17) indicating that these reprojections resemble the 2D averages the best (Figure 2.27 c). This result is in agreement with the antibody labeling and the docking analysis (Section 2.3.5 and Section 2.3.6) showing RPA being located in a horse shoe-like conformation in the 3D map.



b

Difference map number	Merged segments
1	6, 7, 10
2	7, 8
3	2, 3, 5
4	4
5	3, 5, 9
6	5, 9
7	9, 8
8	2, 6, 7
9	2, 3, 6
10	1, 4, 11
11	1, 7, 10, 11
12	2, 6, 10
13	1, 2, 6, 10, 11
14	7, 8, 10
15	fit

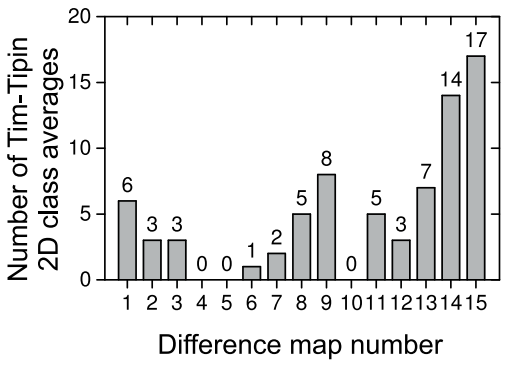


Figure 2.27.: Difference mapping of Tim-Tipin. (a) Negative stain 3D reconstruction of Tim-Tipin-RPA sub-divided into 11 segments. (b) Segments enclosing the RPA volume were merged as defined in the table and subtracted from the 3D map. (c) Bar chart showing the total number of the Tim-Tipin 2D class averages, which aligned the best with the reprojections of the 3D maps number 1 to 15.

2.4. Biochemical Analysis of the Tim-Tipin-RPA Complex

2.4.1. SEC and FA Analysis of ssDNA Binding

The EM structure revealed that RPA adopts a compact, horse shoe-like conformation in the Tim-Tipin-RPA complex. Several studies have shown that RPA undergoes conformational changes dependent on the length of the ssDNA substrate and the number of RPA molecules bound to ssDNA. To examine if the conformational changes of RPA affect the stability and ssDNA binding activity of the Tim-Tipin-RPA complex, SEC analysis using ssDNA substrates of various length (60, 31, 14 nt ssDNA) at different

protein:ssDNA stoichiometric ratios was performed. For comparison, the binding of individual RPA and the Tim-Tipin complex to various ssDNA substrates under our experimental conditions was examined.

Pre-incubation of RPA with excess amounts of 60 or 31 nt ssDNA (protein:DNA, 16 µM:32 µM), shifted the SEC peak to higher molecular weight (Table 2.3 and Figure 2.28 a, b, left) compared to DNA-free RPA (Figure 2.15). The chromatogram showed increased 260/280 ratio at the peak fraction (1.57 ± 0.05, 60 nt; 1.46 ± 0.01, 31 nt ssDNA) indicating DNA-bound RPA. The change of the molecular weight from 110 kDa to approximately 180 kDa indicates that one RPA is bound to ssDNA, likely in its 30 nt binding mode. When RPA was incubated with sub-stoichiometric amounts of DNA (protein:DNA, 32 µM:8 µM or 16 µM:8 µM) the SEC peak further shifted to higher molecular weight (Table 2.3 and Figure 2.28 a, b, right) and still showed increased 260/280 ratio (1.32 ± 0.03, 60 nt; 1.28 ± 0.05, 31 nt ssDNA). The shift to even higher molecular weight indicates an accumulation of RPA on ssDNA and likely a conformational change from the 30 nt to the 8 nt binding mode. In addition, the mixture of RPA with short ssDNA (14 nt) in all tested protein:DNA ratios showed an infinitesimal shift of the RPA peak and increased 260/280 ratio demonstrating that only one RPA molecule is bound to ssDNA, presumably in the 8 nt binding mode ($260/280 = 1.47 \pm 0.24$ and 1.27 ± 0.06; Table 2.3 and Figure 2.28 c, left, right). These results are consistent with published data, which have shown that under conditions where either the ssDNA is short (here 14 nt), or RPA oligomers are bound to ssDNA (here sub-stoichiometirc DNA amounts), RPA employs the 8 nt binding mode [82, 92, 95, 113, 115].

Next, the ssDNA binding characteristics of Tim-Tipin were examined. The Tim-Tipin complex exhibited only a small shift to higher molecular weight with excess ssDNA (Table 2.3 and Figure 2.28 a, b, left) compared to DNA-free Tim-Tipin (Figure 2.15). The small shift of the peak is presumably based on the detection limit of the experimental system. Further, the chromatogram showed a slightly increased 260/280 ratio at the peak fraction (0.98 ± 0.14, 60 nt; 0.74 ± 0.01, 31 nt ssDNA; Table 2.3) compared to DNA-free Tim-Tipin ($260/280 = 0.55 \pm 0.01$; Figure 2.15), indicating that Tim-Tipin is only partially occupied by ssDNA.

Additionally, to quantify Tim-Tipin binding to ssDNA, fluorescence anisotropy (FA) was performed. Binding isotherms showed that Tim-Tipin binds to 30 nt ssDNA with a K_D of 1.7 ± 0.2 µM and to 60 nt ssDNA with a K_D of 0.29 ± 0.01 µM (Figure 2.29).

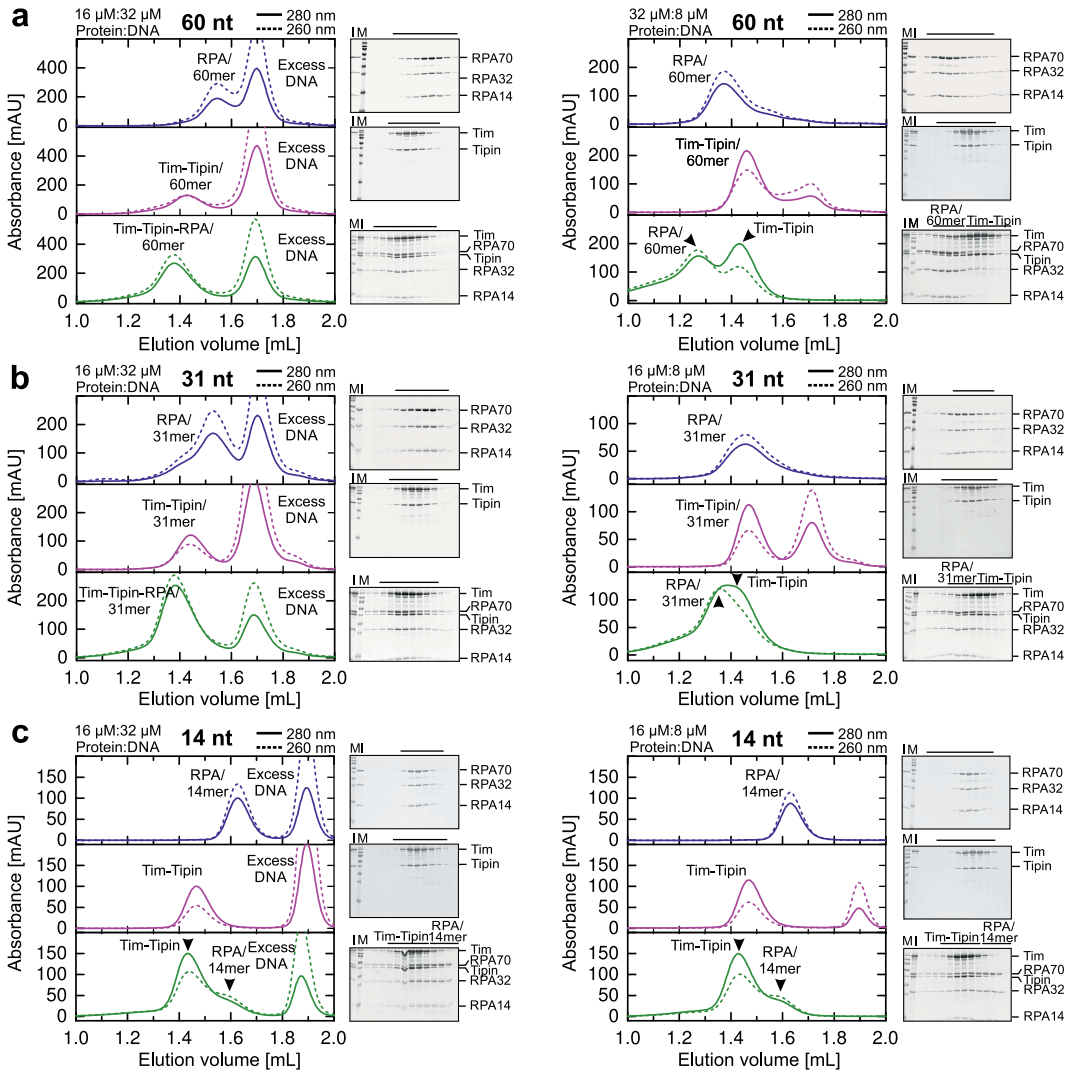


Figure 2.28.: Size exclusion chromatograms of RPA (blue), Tim-Tipin (magenta), and Tim-Tipin-RPA (green) with 60 nt (a), 31 nt (b), and 14 nt (c) ssDNA at different protein:ssDNA stoichiometric ratios. Coomassie blue stained SDS-gels show the SEC protein peak fractions (black line). (M) Protein marker. (I) Protein mixture injected on SEC. Solid line: UV absorbance at 280 nm. Dashed line: UV absorbance at 260 nm. Excess amounts of 60 and 30 nt ssDNA resulted in an association of Tim-Tipin-RPA with ssDNA. Sub-stoichiometric amounts of ssDNA and short ssDNA (here 14 nt) resulted in a breaking of the complex into DNA-free Tim-Tipin and DNA-bound RPA presumably because of the conformational change of RPA from the 30 nt to the 8 nt binding mode. Adopted from [128].

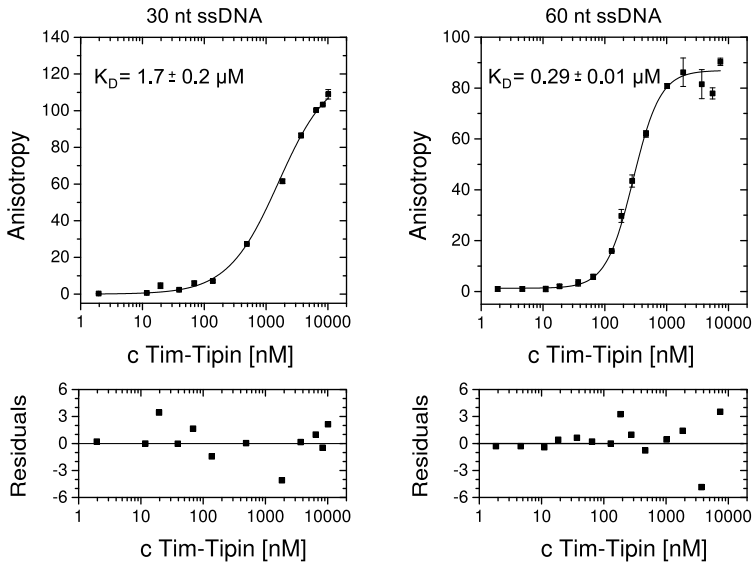


Figure 2.29.: Representative binding isotherms for binding of Tim-Tipin to 30 nt (left) and 60 nt (right) using 5′-FAM labeled ssDNA analyzed by fluorescence anisotropy. The anisotropy is shown as a function of the protein concentration in logarithmic scale. The K_D was calculated to be 1.7 ± 0.2 µM to 30 nt and 0.29 ± 0.01 µM to 60 nt ssDNA. The fit of the binding isotherms is shown as solid line and was calculated by using a nonlinear fit to the Hill equation with a Hill coefficient of n = 1.2 ± 0.2 (30 nt) and 2.2 ± 0.1 (60 nt) as described in Section 4.2.5. The error bars represent the standard deviation of three independent measurements. (Bottom) Residual plots for the curve fit (difference between the calculated anisotropy of the curve fit and the measured anisotropy) show a good quality of the regressions and no bias. Adopted from [128].

When Tim-Tipin was incubated with sub-stoichiometric amounts of ssDNA as described for RPA, the SEC peak was not shifted to higher molecular weight (Figure 2.28 a, b, right) and the 260/280 ratio was infinitesimal increased (0.74±0.06, 60 nt; 0.57±0.02, 31 nt ssDNA; Table 2.3), which is in agreement with the low affinity of Tim-Tipin to ssDNA as determined by FA (Figure 2.29) and with the partial binding of ssDNA to Tim-Tipin in SEC. Notably, no binding of Tim-Tipin to 14 nt ssDNA was detected (Figure 2.28 c, $260/280 = 0.54 \pm 0.02$ (left) 0.54 ± 0.00 (right)).

Finally the ssDNA binding characteristics of Tim-Tipin-RPA to various length and concentrations of ssDNAs using SEC were analyzed. The Tim-Tipin-RPA complex eluted at higher molecular weight when incubated with excess amounts of 60 or 31 nt ssDNA (Figure 2.28 a, b, left) compared to the DNA-free complex (Figure 2.15) and displayed increased 260/280 ratio (1.20±0.04, 60 nt; 1.11±0.04, 31 nt ssDNA; Table 2.3). This implies that ssDNA is bound to the intact Tim-Tipin-RPA complex with a RPA

Table 2.3.: Elution volumes, 260/280 ratio, experimental, and theoretical molecular weights of RPA, Tim-Tipin (TTP), and Tim-Tipin-RPA. Adopted from [128].

DNA substrate/ Protein:ssDNA (μM)	Protein mixture/ DNA substrate	Elution volume* [mL]	260/280*	Experimental$^{SEC\delta}$ MW [kDa]	Theoretical# MW [kDa]
	RPA	1.63 ± 0.02	0.63 ± 0.02	110	96.2
	Tim-Tipin	1.46 ± 0.01	0.55 ± 0.01	321	162.3
	Tim-Tipin-RPA	1.43 ± 0.02	0.56 ± 0.01	388	258.5
60 nt	RPA	1.54 ± 0.00	1.57 ± 0.05	194	114.6
16:32	Tim-Tipin	1.44 ± 0.01	0.98 ± 0.14	364	180.7
	Tim-Tipin-RPA	1.39 ± 0.01	1.20 ± 0.04	498	276.9
60 nt	RPA	1.38 ± 0.01	1.32 ± 0.03	531	403.2
32:8	Tim-Tipin	1.47 ± 0.01	0.74 ± 0.06	301	180.7
	Tim-Tipin-RPA	1.44 ± 0.01 (TTP)	0.66 ± 0.05	364	
		1.33 ± 0.05 (RPA)	1.13 ± 0.04	726	
31 nt	RPA	1.55 ± 0.03	1.46 ± 0.01	182	105.7
16:32	Tim-Tipin	1.46 ± 0.02	0.74 ± 0.01	321	171.8
	Tim-Tipin-RPA	1.40 ± 0.03	1.11 ± 0.04	468	268.0
31 nt	RPA	1.48 ± 0.02	1.28 ± 0.05	283	201.9
16:8	Tim-Tipin	1.47 ± 0.01	0.57 ± 0.02	301	171.8
	Tim-Tipin-RPA	1.41 ± 0.03 (TTP)	0.87 ± 0.00	439	
		1.38 ± 0.03 (RPA)	0.98 ± 0.01	531	
14 nt	RPA	1.63 ± 0.00	1.47 ± 0.24	110	100.5
16:32	Tim-Tipin	1.48 ± 0.01	0.54 ± 0.02	283	166.6
	Tim-Tipin-RPA	1.45 ± 0.02 (TTP)	0.66 ± 0.08	342	
		1.62 ± 0.01 (RPA)	1.27 ± 0.05	117	
14 nt	RPA	1.63 ± 0.00	1.27 ± 0.06	110	100.5
16:8	Tim-Tipin	1.48 ± 0.01	0.54 ± 0.00	283	166.6
	Tim-Tipin-RPA	1.45 ± 0.02 (TTP)	0.64 ± 0.05	342	
		1.62 ± 0.01 (RPA)	1.22 ± 0.04	117	
ssDNA only	60 nt	1.70	1.82	64	18.4
0:16	31 nt	1.70	1.76	64	9.5
	14 nt	1.88	2.34	21	4.3
	Thyroglobulin	1.32 ± 0.02*			670
	γ-globulin	1.63 ± 0.03*			158
	Ovalbumin	1.76 ± 0.01*			44
	Myoglobin	1.91 ± 0.02*			17
	Vitamin B_{12}	2.15 ± 0.02*			1.35

* Average numbers and standard deviations of at least two independent measurements are shown.
$^{\delta}$ Molecular weights derived from size exclusion chromatography (SEC) were estimated based on the comparison with molecular mass standard proteins.
\# Molecular weights calculated based on the amino acid sequence.

conformation likely resembling its compact 30 nt binding mode. When the complex was incubated with sub-stoichiometric amounts of ssDNA, dissociation of Tim-Tipin-RPA to DNA-free Tim-Tipin and DNA-bound RPA was observed (Figure 2.28 a, b, right) with a RPA conformation likely resembling its 8 nt binding mode. Interestingly, when Tim-Tipin-RPA was mixed with short ssDNA (14 nt), the complex showed dissociation

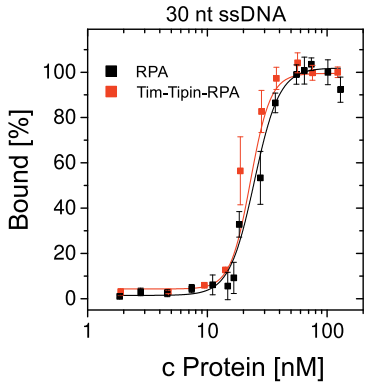


Figure 2.30.: Representative binding isotherms for binding of Tim-Tipin-RPA and RPA to 5′-FAM labeled 30 nt ssDNA analyzed by fluorescence anisotropy. The maximum anisotropy of each protein complex was set to 100% ssDNA-bound protein complex for normalization and is shown as a function of the protein concentration in logarithmic scale. The ssDNA concentration was kept constant at 10 nM. The K_D was calculated to be 24.7 ± 1.3 nM for RPA alone and 23.5 ± 1.5 nM for Tim-Tipin-RPA. Notably, the Tim-Tipin-RPA complex dissociates with sub-stoichiometric ssDNA amounts. Therefore the Tim-Tipin-RPA binding isotherm represents RPA binding to ssDNA at protein concentrations higher than 10 nM. The solid line represents the fit of the binding isotherms using a nonlinear fit to the Hill equation with a Hill coefficient of n $= 5.0 \pm 0.7$ (Tim-Tipin-RPA) and 4.2 ± 0.5 (RPA). The error bars show the standard deviation of three independent measurements.

to DNA-free Tim-Tipin and DNA-bound RPA in all tested conditions (Figure 2.28 c, left, right). The SEC analysis indicates that upon change of the RPA conformation to the 8 nt mode (at sub-stoichiometirc ssDNA amounts or short ssDNA), Tim-Tipin dissociates from DNA-bound RPA.

Although dissociation of Tim-Tipin-RPA was observed using SEC with sub-stoichiometric ssDNA amounts (excess amounts of protein complex), the quantification of Tim-Tipin-RPA binding to 30 nt ssDNA was attempted using fluorescence anisotropy and compared to RPA alone. Binding isotherms of Tim-Tipin-RPA showed similar binding curves to the curves of RPA alone, indicating that after Tim-Tipin-RPA dissociation, RPA likely takes over ssDNA binding (Figure 2.30). This interpretation is supported by the SEC results, where RPA eluted as ssDNA-protein complex and Tim-Tipin as DNA-free protein at sub-stoichiometric ssDNA amounts (Figure 2.28 a, b, c, right).

2.4.2. EMSA of RPA and Tim-Tipin Using ssDNA

The SEC analysis revealed that the change of RPA from the 30 nt to the 8 nt mode coincides with the dissociation of the Tim-Tipin-RPA complex. To gain further insight

into the accumulation of RPA and Tim-Tipin and the association of Tim-Tipin-RPA with ssDNA, EMSA analysis was performed using ssDNA substrates of various lengths (60, 30, 14 nt ssDNA) at the same protein:ssDNA stoichiometric ratio as applied in the SEC analysis (Section 2.4.1).

RPA incubation with excess amounts of 60 or 30 nt ssDNA resulted in one molecule of RPA bound to one molecule of 60 or 30 nt ssDNA in the EMSA (Figure 2.31 a, lane 3 and 9). This is consistent with the shift of RPA to higher molecular weight in SEC (from 110 kDa to ∼180 kDa) bound to one molecule of 60 nt or 31 nt ssDNA (Figure 2.28 a, b, left).

As the ssDNA concentration was reduced to sub-stoichiometric amounts the EMSA showed RPA-ssDNA complexes with stepwise reduced mobility indicating the accumulation of excess RPA on ssDNA. The degree of accumulation was dependent on the length of the ssDNA. The stoichiometry of RPA to ssDNAs at the saturation level was determined to be two (RPA II) for 30 nt (Figure 2.31 a, lane 11) and four (RPA IV) for 60 nt (Figure 2.31 a, lane 5). In contrast, one RPA molecule (RPA I) was detected for the shortest ssDNA substrate (14 nt) (Figure 2.31 a, lane 13-18).

Notably, in the EMSA of RPA using 14 nt ssDNA, a faint band below (marked by asterisk, lane 14-16) and above the major shifted band (marked by plus, lane 17) appeared. To further investigate these bands and for a better comparison of the migration of the RPA-ssDNA complexes, an EMSA additionally using 8 nt ssDNA was performed on the same gel. In this experiment the lower bands were not observed in the 8 nt ssDNA condition (Figure 2.31 b, lane 11-13), in contrast to the conditions using 14 nt ssDNA (Figure 2.31 b, lane 7-9, asterisk). The result indicates that the lower bands are most probably derived from a certain population of RPA adopting the compact 30 nt binding mode, which cannot be achieved in presence of the very short 8 nt ssDNA. Interestingly, the migration of RPA is slightly different between the RPA-8 nt and RPA-14 nt ssDNA complexes, which could derive from a slightly different RPA conformation.

The upper band in Figure 2.31 a (lane 17, marked by plus) could not be reproduced in the follow-up EMSA using 14 nt ssDNA, but a similar migrating band was detected using 8 nt ssDNA (Figure 2.31 b, lane 11, marked by plus). Hence, the upper band is likely an artifact, which could derive from a conformational heterogeneity of RPA.

In summary, the EMSA demonstrates that RPA undergoes a conformational change from the 30 nt (long and excess amount of ssDNA) to the 8 nt binding mode under our experimental conditions (short or sub-stoichiometric amount of ssDNA), which is in agreement with published data [82, 92, 95, 113, 115].

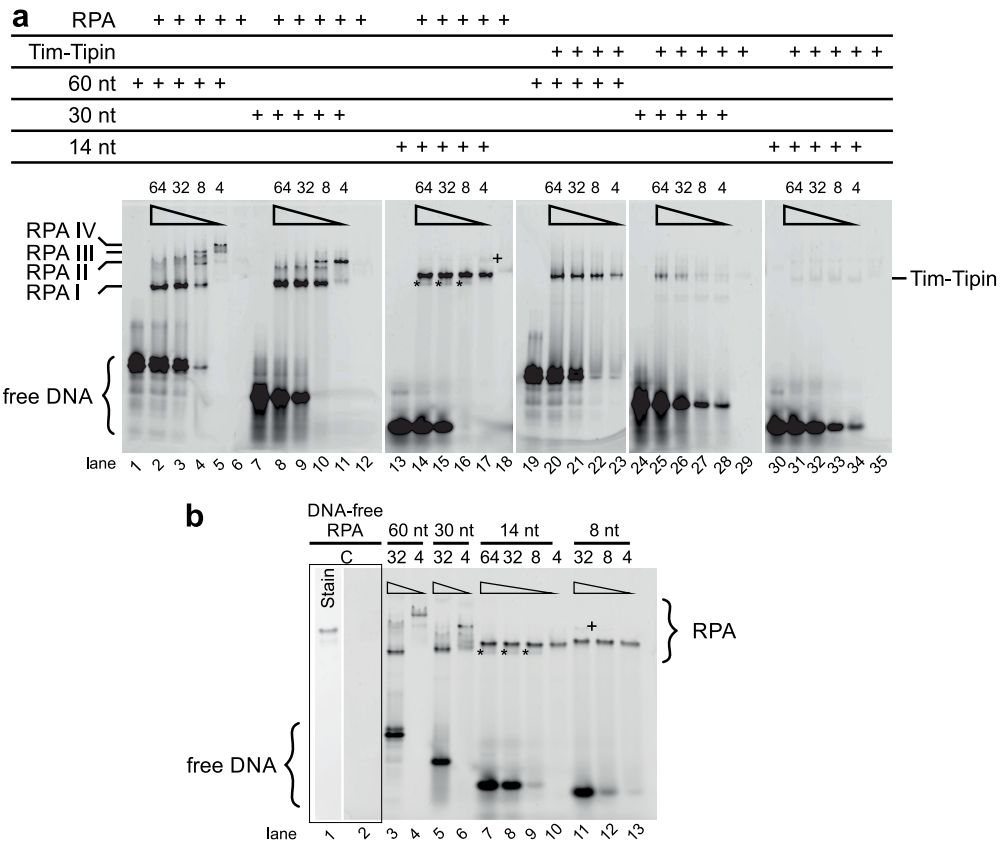


Figure 2.31.: Electrophoretic mobility shift assay (EMSA) of RPA and Tim-Tipin. 60, 30, and 14 nt 5′-FAM labeled ssDNA substrates were used in (a). Decreasing amounts of the added ssDNA substrate are indicated by a triangle (total concentration of 64, 32, 8, and 4 µM). The protein concentration was kept constant at 16 µM. Four RPAs were identified on 60 nt ssDNA (RPA I-IV, lane 1-6), two RPAs on 30 nt (RPA I-II, lane 7-12), and one RPA on 14 nt ssDNA (RPA I, lane 13-18). For Tim-Tipin, only one molecule was bound to ssDNA, indicating a 1:1 binding stoichiometry to the substrate. The complex shows the highest preference for 60 nt ssDNA (Tim-Tipin, lane 19-23) followed by 30 nt ssDNA (lane 24-29) and no binding to 14 nt ssDNA (lane 30-35). Unbound DNA is marked as free DNA. (b) EMSA of RPA using a 60, 30, 14, and 8 nt 5′-FAM labeled ssDNA substrate analyzed on the same native gel for better comparison. ssDNA and protein total concentrations as in (a). Frame = Control (C), DNA-free RPA. Stain = Coomassie stained native PAGE. Lane 2-13 = fluorescence image of the native gel. The asterisks (lane 7-9) indicate a small RPA population in the 30 nt binding mode. The cross (lane 11) indicates an artifact probably from a conformational heterogeneity of RPA. Adopted from [128].

Further, in the EMSA of Tim-Tipin with 60 and 30 nt ssDNA a band-shift to protein-DNA complexes with reduced mobility was observed (Figure 2.31, lane 19-23 for 60 nt; lane 24-29 for 30 nt), which was more pronounced with 60 nt ssDNA than with 30 nt

ssDNA. Notably, no binding of Tim-Tipin to 14 nt ssDNA was detected (Figure 2.31, lane 30-35). The EMSA results are in agreement with the SEC and FA analyses (Figure 2.28 and Figure 2.29) showing that the affinity of Tim-Tipin to ssDNA increases with the length of the DNA substrate.

Finally the ssDNA binding characteristics of Tim-Tipin-RPA were examined by EMSA. The detection of the different protein-ssDNA was supported by immunoblotting. Excess amounts of long (here 60 nt) ssDNA yielded the association of ssDNA with Tim-Tipin-RPA (Figure 2.32, lane 2 and 3, arrowhead P). Notably, this super-shift was not observed for RPA alone with long ssDNA (Figure 2.31 lane 2 and 3). Nevertheless, a partial dissociation of the Tim-Tipin-RPA-ssDNA complex was detected (Figure 2.32, lane 2 and 3, marked by R). Further, when Tim-Tipin-RPA was pre-incubated with sub-stoichiometric amounts of ssDNA (Tim-Tipin-RPA in excess), a complete dissociation of the complex into DNA-bound RPA and DNA-free Tim-Tipin was observed. RPA displayed accumulation on ssDNA with up to 4 RPAs for 60 nt ssDNA (Figure 2.32, lane 4 and 5, R), which is consistent with the EMSA of RPA-ssDNA alone (Figure 2.31, lane 4 and 5).

When the Tim-Tipin-RPA complex was incubated with short ssDNA (here 14 nt), the dissociation of the complex into DNA-bound RPA (marked by R) and DNA-free Tim-Tipin (marked by T) was observed with excess as well as at sub-stoichiometric amounts of ssDNA (Figure 2.32, lane 6-10), which is in line with our SEC analysis (Figure 2.28 c, left, right).

Finally, to stabilize Tim-Tipin-RPA-ssDNA complex formation, the protein-ssDNA mixture was cross-linked after incubation with 60 nt ssDNA and analyzed by EMSA. As under non-cross-linked conditions, the stabilized Tim-Tipin-RPA complex showed significant binding to ssDNA (Figure 2.32 b, lane 16, arrowhead P) as a super-shifted band, which was not observed with cross-linked RPA and Tim-Tipin (Figure 2.32 b, lane 14 and 15, R and T).

2.4.3. EMSA of RPA and Tim-Tipin Using Replication Fork-Like DNA

In the previous analyses, the binding characteristics of Tim-Tipin to ssDNA were examined using SEC and FA (Section 2.4.1), and EMSA (Section 2.4.2). The results showed a higher affinity of Tim-Tipin to ssDNA with increasing length of the DNA substrate. Interestingly, several studies have demonstrated the association of Tim-Tipin with components of the DNA replication fork. To gain additional insights into the binding characteristics of Tim-Tipin to fork-like DNA, EMSA was performed using

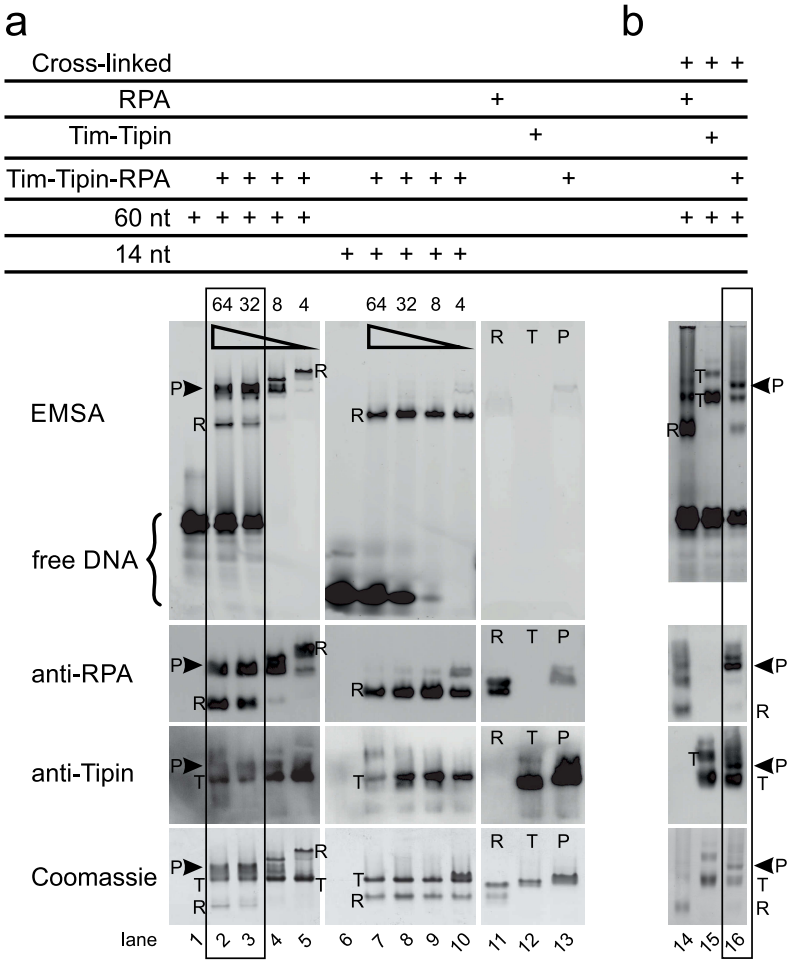


Figure 2.32.: Electrophoretic mobility shift assay (EMSA) of Tim-Tipin-RPA using long (60 nt) or short (14 nt) $5'$-FAM labeled ssDNA substrate. (a) Decreasing amounts of the added ssDNA substrate are indicated by a triangle (total concentration of 64, 32, 8, and 4 µM). The protein concentration was kept constant at 16 µM. The ssDNA-protein complexes were visualized by fluorescence imaging of the labeled ssDNA (EMSA), immunoblot analysis (anti-RPA, anti-Tipin), and by Coomassie-staining of the native gels (Coomassie). R = RPA; T = Tim-Tipin; P = pentameric Tim-Tipin-RPA complex. Lane 11-13 show the DNA-free proteins. While Tim-Tipin-RPA binds to long ssDNA under excess ssDNA conditions (lane 2, 3), sub-stoichiometric amounts of the 60 nt ssDNA substrate (lane 4, 5) or short ssDNA (14 nt; lane 6-10) lead to a complete dissociation of the complex to DNA-free Tim-Tipin and DNA-bound-RPA. (b) EMSA of glutaraldehyde cross-linked protein complexes using long (60 nt) $5'$-FAM labeled ssDNA. The protein-ssDNA complexes (0.6 µM Tim-Tipin-RPA and 1.2 µM ssDNA; 0.92 µM Tim-Tipin and 1.84 µM ssDNA; 1.56 µM RPA and 3.12 µM ssDNA) were cross-linked after incubation with ssDNA, analyzed by EMSA, and visualized as in (a). A significant amount of cross-linked Tim-Tipin-RPA is bound to ssDNA (lane 16) and shows a super-shift, which is not observed with Tim-Tipin and RPA alone (lane 14, 15). Adopted from [128].

various DNA substrates (Figure 2.33 a-c). Two DNA substrates showed fork-like structures with two kinds of junctions. The first DNA substrate displayed features of a splayed double-stranded DNA with 30 nt long ssDNA overhangs (splayed (sp)DNA; Figure 2.33 a). The second substrate displayed features of a splayed double-stranded DNA with overhangs supplemented with complementary DNA strands (replication fork (RF)DNA; Figure 2.33 b). To better discriminate whether Tim-Tipin is binding to the ssDNA-, dsDNA-part or to the junction, dsDNA was tested as third substrate (dsDNA; Figure 2.33 c). In addition, the DNA binding studies were conducted with RPA as control.

When RPA was incubated with the DNA substrates, a shift to slower migrating protein-DNA complexes was only observed with splayed DNA (Figure 2.33 d, lane 14, 15), which is in agreement with RPA's characteristic binding to ssDNA. Tim-Tipin did not show any preference for any of the tested DNAs and formed protein-DNA complexes in all tested EMSA conditions (Figure 2.33 d, lane 18, 19, dsDNA; lane 22, 23, RFDNA; lane 26, 27 spDNA).

From these results it is not possible to conclude, whether Tim-Tipin prefers binding to any junctions, to ssDNA or dsDNA. Rather the EMSA indicates a broad specificity of Tim-Tipin in DNA recognition.

2.4.4. RPA Domains Identification for Tim-Tipin-RPA Formation

The Tim-Tipin-RPA complex was successfully reconstituted on SEC as described in Section 2.2.1 by incubation of the Tim-Tipin complex with the ternary RPA. Interestingly, Tipin has been reported to interact with the RPA32 C-terminal WH domain [19, 20, 83]. To investigate which components of the trimeric RPA complex facilitate intact Tim-Tipin-RPA formation, various RPA sub-complexes were cloned, expressed and purified, and tested for the interaction with the Tim-Tipin complex using SEC. On the basis of published crystal structures of human RPA, sequence alignments and secondary structure predictions (Section 1.4.3), four homologous constructs of mouse RPA were designed for recombinant expression in *E. coli*.

First, the RPA32 C-terminal WH domain was cloned (RPA32, aa 173-270) and purified to homogeneity (Figure 2.34 a, blue). The SEC peak of the WH domain showed very low absorbance at 280 nm, because of low contents of tyrosines or tryptophans. Secondly, the major ssDNA binding domains of RPA70, DBD-A and -B, were cloned (RPA70, aa 190-431) and purified. The SEC elution profile showed a single peak and pure protein on SDS-PAGE (Figure 2.34 b, blue). Thirdly, RPA32 DBD-D together with

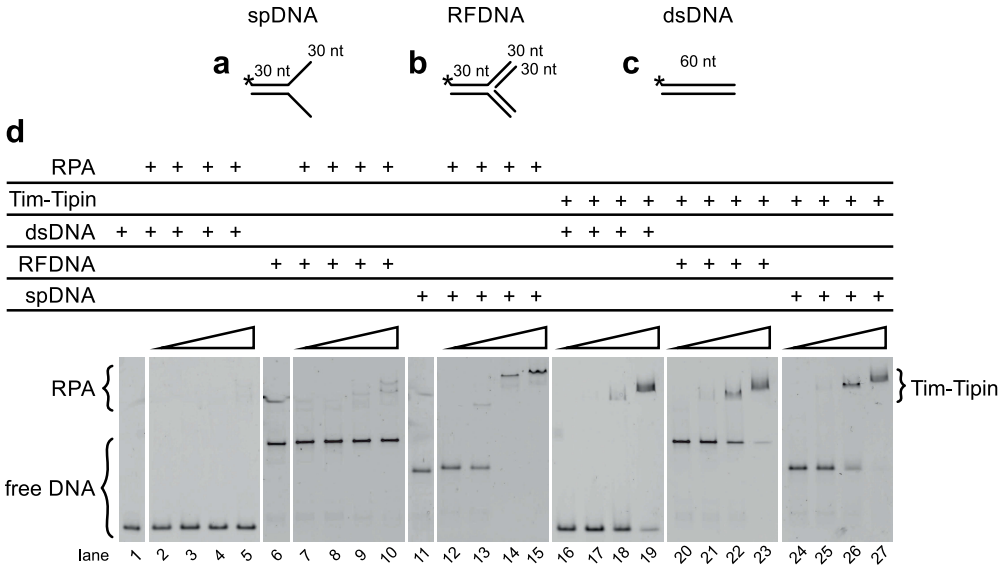


Figure 2.33.: Electrophoretic mobility shift assay (EMSA) using splayed DNA (spDNA, a), replication-fork DNA (RFDNA, b), and double-stranded DNA (dsDNA, d). The DNA substrates were 5′-FAM labeled (marked by asterisk). The triangle in the EMSA (d) indicates increasing amounts of the added protein complex (total concentration of 0.08, 0.24, 0.8, and 4.8 µM). The DNA substrate concentrations were kept constant at 80 nM. RPA binds only to DNA substrates with present ssDNA-parts. The Tim-Tipin complex shows a broad preference and binds to all tested DNA substrates. Unbound DNA is marked as free DNA. Lane 1, 6, and 11 show the negative control with DNA only.

the C-terminal WH-domain (RPA32, aa 43-270) and RPA14 (aa 1-121, full-length) were chosen for the subsequent binding study. The dimeric complex was purified to high purity and homogeneity as demonstrated by SDS-PAGE and SEC analysis (Figure 2.34 c, blue). Finally, all three RPA subunits lacking the major DBD-A and -B (RPA70 DBD-C, aa 445-623; RPA32 DBD-D and WH-domain, aa 43-270; RPA14 fl, aa 1-121) were cloned and showed pure complex after purification (Figure 2.34 d, blue).

For the subsequent interaction study of Tim-Tipin with RPA, the reported interaction of Tipin with RPA32 WH domain was first taken into consideration. When the purified RPA32 WH domain was incubated with the Tim-Tipin complex, only a weak interaction was detected (Figure 2.34 a). Further, the mixture of Tim-Tipin and RPA70 DBD-A and -B or RPA32/14 (Figure 2.34 b, c) showed no interactions and the individual components eluted in distinct peaks. Finally, when Tim-Tipin was incubated with the ternary RPA70DBD-C/32/14 complex (lacking RPA70 DBD-A and -B) an interaction with Tim-Tipin was detected (Figure 2.34 d). This interaction was weaker compared to the RPA70/32/14 complex containing all DBD-A, -B, -C, and -D (Figure 2.15).

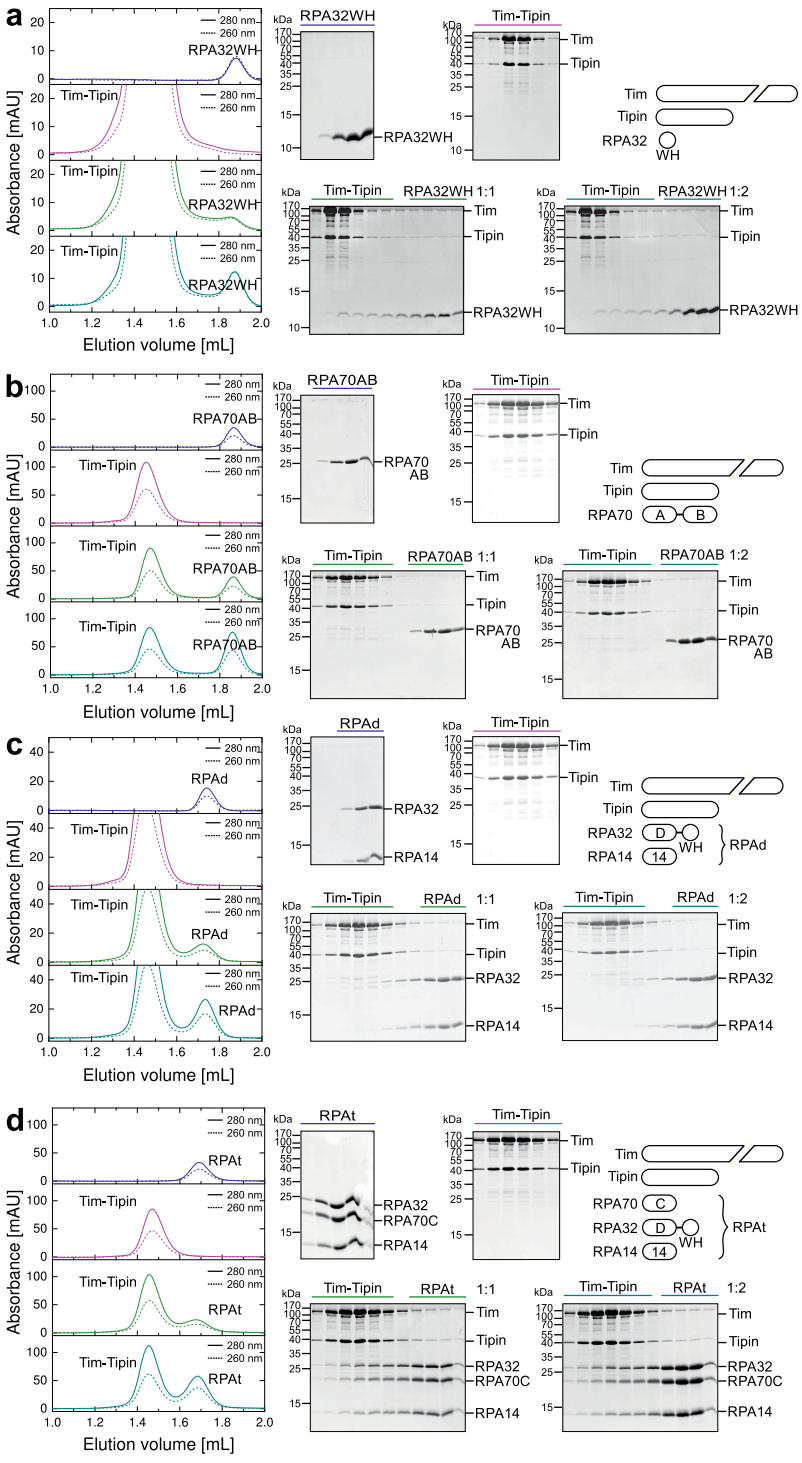


Figure 2.34.: *See next page for figure legend.*

Figure 2.34.: *(Previous page).* Size exclusion chromatography (SEC) analysis of various RPA sub-complexes binding to Tim-Tipin. Different RPA sub-complexes (blue, (a) RPA32 WH domain, (b) RPA70 DBD-A and -B, (c) RPA32/14 (RPAd), (d) RPA70C/32/14 (RPAt)) were mixed with Tim-Tipin at a molar ratio of 1:1 (green) and 2:1 (cyan). RPA32WH and RPA32/14 did not significantly bind to Tim-Tipin. RPA70C/32/14 formed a weaker complex compared to RPA70/32/14 (Figure 2.15) and RPA70AB showed no binding to Tim-Tipin. Solid line: UV absorbance at 280 nm. Dashed line: UV absorbance at 260 nm. The peak fractions for each sample mixture were analyzed on Coomassie stained SDS-gels (right). For the SDS-PAGE analysis in (a) only one fourth of the sample was used as input for the SDS-gel when compared to (b)-(d). The same SEC run and SDS-gel for Tim-Tipin is illustrated in (b) and (c). Adopted from [128].

In summary, although RPA70 DBD-A and DBD-B do not directly interact with Tim-Tipin, this subunit plays an indirect role in stabilizing the Tim-Tipin-RPA complex. Hence, ternary RPA containing all four DBDs A-D is required and individual RPA subunits are not sufficient for the Tim-Tipin-RPA complex formation.

2.4.5. Timeless-Tipin in Complex with Full-Length Replication Protein A

The RPA complex used in this thesis is lacking the RPA70 N-terminal domain (RPA70N, amino acids 1-189) and the first 42 amino acids of RPA32 (RPA32N) (Figure 1.3). To investigate whether the Tim-Tipin-RPA complex formation and its biochemical activity is affected by the lack of RPA70N and RPA32N, full-length RPA (RPAFL) was cloned, expressed and purified. The SEC elution profile showed two peaks corresponding to the RPAFL protein and an excess of the RPA32FL/RPA14 dimer (Figure 2.35 a). Although RPAFL was accessed in high purity as shown by SDS-PAGE, a degradation of RPA32FL occurred during purification (arrowhead, black). RPA32FL was truncated to a construct lacking the cloning overhang and the first two residues as identified by Edman sequencing.

Further the purified RPAFL complex was tested for the interaction with the Tim-Tipin complex using SEC. The mixture of RPAFL and Tim-Tipin shifted to higher molecular weight compared to the truncated Tim-Tipin-RPA complex (Figure 2.35 b). SDS-PAGE confirmed the successful reconstitution of Tim-Tipin-RPAFL. This result indicates, that the additional domains do not interfere with complex formation.

Finally, the stability and ssDNA binding activity of the Tim-Tipin-RPAFL complex was tested by SEC using ssDNA substrates of various length (60, 31, 14 nt ssDNA) at different protein:ssDNA stoichiometric ratios. When Tim-Tipin-RPAFL was incubated with excess amounts of 60 and 31 nt ssDNA the complex eluted in a single peak and showed increased 260/280 ratio at the peak fraction (260/280 = 1.24, 60 nt;

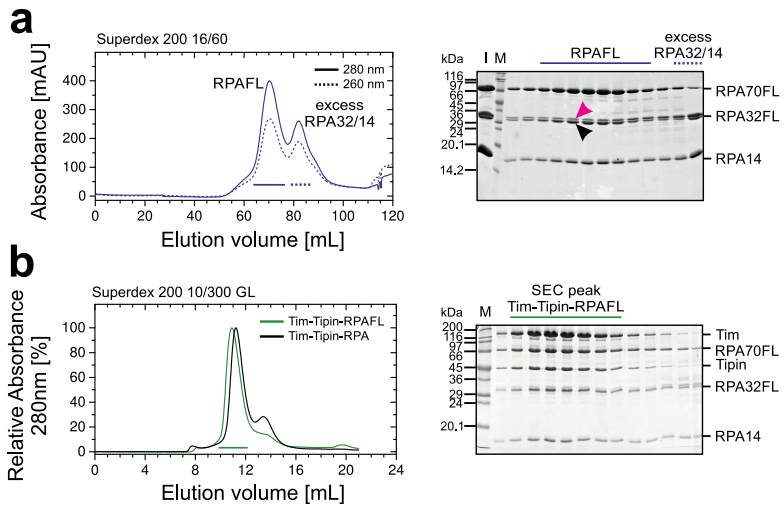


Figure 2.35.: Purification of RPAFL and Tim-Tipin-RPAFL reconstitution. (a) SEC analysis of RPAFL using a Superdex a 200 16/60 column. The elution profile shows two peaks corresponding to RPAFL and excess RPA32/14 as visualized by UV absorbance at 280 nm and 260 nm. SDS-PAGE shows the peak fractions marked by solid and dashed line. (M) Protein marker in kDa. (I) Protein sample injected on SEC column. The observed degradation of RPA32FL to a truncated construct lacking the cloning overhang and the first two residues is marked by a black arrowhead. RPA32FL is marked by magenta arrowhead. (b) SEC analysis of the Tim-Tipin-RPAFL complex formation. Tim-Tipin and RPAFL (green) shifted to higher molecular weights compared to truncated Tim-Tipin-RPA lacking RPA70N and RPA32N (black). The protein mixture was loaded on a Superdex 200 10/300 GL column and the elution profile was visualized by UV absorbance at 280 nm. Peak fractions marked by line were analyzed by Coomassie stained SDS-gels (right) and demonstrated complex formation. Adopted from [128].

1.05, 31 nt ssDNA) indicating ssDNA binding of the intact Tim-Tipin-RPAFL with RPA likely in the 30 nt binding mode (Figure 2.36 a left and b). The slightly lower 260/280 ratio (260/280 = 1.05) for Tim-Tipin-RPAFL using the 31 nt ssDNA in comparison to Tim-Tipin-RPA (260/280 = 1.11 $\pm$ 0.04) might result from oligomerization of RPAFL (0.9 mL, marked by asterisk). Further, the mixture of the complex with sub-stoichiometric amounts of 60 nt ssDNA and short ssDNA (14 nt) resulted in the dissociation of the complex to DNA-free Tim-Tipin and DNA-bound RPA probably adopting its 8 nt binding mode (Figure 2.36 a, right and c).

Notably, oligomers of RPAFL have been observed in the tested conditions eluting at 0.9 mL and between 1.0 and 1.2 mL (Figure 2.36, marked by asterisk), which have not been observed with truncated RPA and Tim-Tipin-RPA. The oligomerization was increased, when the RPA concentration was raised to 32 µM (Figure 2.36 a, right).

The SEC analysis showed an intact complex formation and a similar behavior of Tim-

Tipin-RPAFL as the truncated complex without RPA70N and RPA32N in terms of the dependence on the ssDNA length and the protein:ssDNA stoichiometry suggesting that RPA70N and RPA32N do not affect Tim-Tipin-RPAFL complex formation and ssDNA binding.

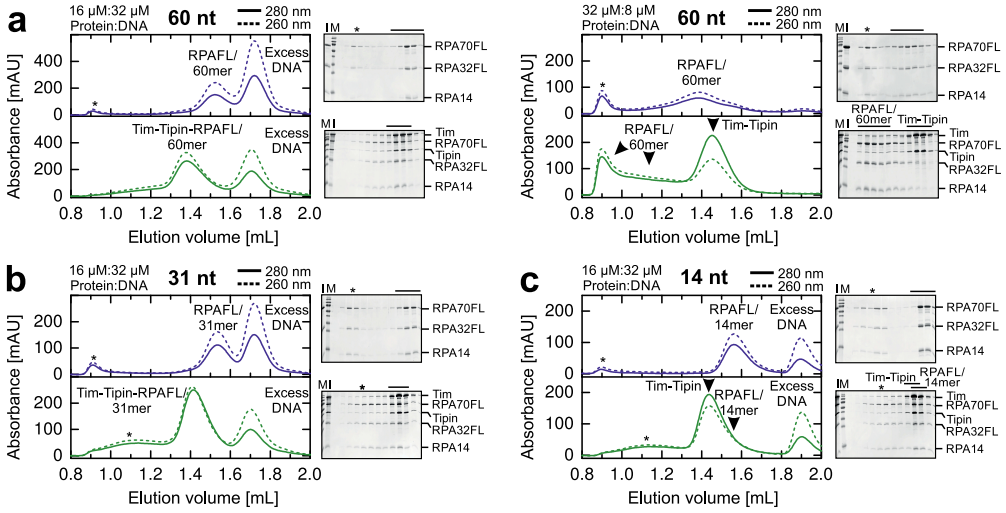


Figure 2.36.: Size exclusion chromatograms of RPAFL (blue) and Tim-Tipin-RPAFL (green) with 60 nt (a), 31 nt (b), and 14 nt (c) ssDNA at different protein:ssDNA ratios and SDS-gels showing the SEC protein peak fractions (black line). (I) Protein mixture injected on SEC. (M) Protein marker. Solid line: UV absorbance at 280 nm. Dashed line: UV absorbance at 260 nm. RPAFL displayed oligomers (marked by asterisk), which have not been observed with the truncated RPA. Tim-Tipin-RPAFL shows a similar behavior like Tim-Tipin-RPA lacking RPA70N and RPA32N and forms an intact ssDNA-protein complex when RPA is in the 30 nt binding mode (long and excess amount of ssDNA) and shows dissociation when RPA adopts its 8 nt binding mode (sub-stoichiometric amounts and short ssDNA). Adopted from [128].

3. Discussion

The study carried out in this thesis provides a structural and biochemical basis of the Tim-Tipin interaction with RPA. The 3D model of the complex shows a first view of the architecture of the almost full-length Tim-Tipin-RPA complex with RPA employing a compact conformation similar to the recently reported structure of RPA bound to 30 nt ssDNA [96, 97]. In addition, the results help to understand the RPA-controlled Tim-Tipin recruitment to ssDNA, which is strongly dependent on RPA's binding mode. Specifically, RPA binding to ssDNA in its 8 nt binding mode reduces the stability of the Tim-Tipin-RPA complex resulting in the dissociation of Tim-Tipin from the complex. Finally, the RPA-controlled Tim-Tipin recruitment to ssDNA provides interesting implications for the organization of DNA replication, the DNA replication checkpoint, and DNA repair.

This Section will discuss the reconstitution (Section 3.1 and 3.2) and the structural characterization of the Tim-Tipin-RPA complex (Section 3.3), as well as the dispensability of the N-terminal RPA70 and RPA32 and C-terminal Tim regions for complex formation and function (Section 3.4). Further, the Tim-Tipin-RPA dissociation upon change of the RPA conformation in presence of short or sub-stoichiometric ratios of ssDNA and the putative ssDNA path in the stable Tim-Tipin-RPA-ssDNA complex will be taken into consideration (Section 3.5 and 3.6). Moreover, the structural and biochemical characterization provides implications for the role of Tim-Tipin-RPA and Tim-Tipin in the effective organization of DNA replication (Section 3.5 and 3.6), the DNA replication checkpoint, and DNA repair (Section 3.7) and finally hints to a mechanistic basis for replication fork organization using dynamic interactions (Section 3.8). The last Section will provide a short overview of open questions and an outlook (Section 3.9).

3.1. Reconstitution of the Tim-Tipin-RPA Complex In Vitro

The biochemical and biophysical analyses in this thesis report the *in vitro* assembly of the Tim-Tipin and Tim-Tipin-RPA complex at stoichiometric levels. This study shows that individual Tim and Tipin form large oligomers and are only stable when they are

co-expressed and co-purified in agreement with previous reports [74, 83]. Other studies using immunoprecipitation assays of *in vitro* translated or over-expressed proteins also report that Tim forms homo-multimeric complexes, which are disrupted by Tipin [60, 83]. The SEC analysis in this study indicates an experimental molecular weight of 321 kDa for Tim-Tipin, which almost corresponds to the double value of the theoretical molecular weight (162 kDa) calculated based on the amino acid sequence. This discrepancy might arise from the fact that SEC determines the hydrodynamic radius of the protein and its calibration is based on elution volumes of proteins of globular shape. Elongated or unstructured proteins can elute at positions corresponding to higher molecular weights than a globular protein with similar size [133]. Although controversially discussed, Tim was suggested to be a member of the Arm/HEAT protein super-family [105], which could hint at an elongated shape. Nevertheless, the EM analysis in this thesis of individual Tim-Tipin revealed a globular shape of the complex (Section 2.3.4). The apparent disagreement in the elution volume might derive from a certain interaction of the Tim-Tipin complex with the SEC resin. SLS analysis, representing a more reliable technique of molecular weight determination, and sucrose gradient sedimentation assays in this study show a monomeric complex of Tim-Tipin with a 1:1 stoichiometry corresponding to a molecular weight of approximately 160 kDa.

Moreover, this study examines the *in vitro* reconstitution of the Tim-Tipin-RPA complex using to my knowledge the longest Tim-Tipin and RPA fragments (almost full-length). The mixture of Tim-Tipin and RPA elutes basically in one shifted peak, but shows an exaggerated molecular weight as observed for Tim-Tipin alone. However, SLS analysis determines a molecular weight of approximately 260 kDa, which clearly demonstrates a 1:1:1 (Tim:Tipin:RPA) stoichiometry of the SEC-reconstituted Tim-Tipin-RPA complex.

In this study, micro scale thermophoresis experiments using almost full-length proteins determine a K_D in the low micro-molar range ($K_D = 0.18 \pm 0.04$ µM) between RPA and the Tim-Tipin complex. One limitation of the MST measurement is that this method uses fluorescently labeled proteins (here Cy3-labeled RPA). The labeling takes place at primary amino groups (lysines, K) and may in this way insulate this residues from interactions with the binding partner. The used RPA constructs possess 54 lysines with approximately half of the residues (27 K) being exposed to the solvent. The degree of labeling was determined to be 4.4 lysines/858 aa ($\hat{=}$ 1 lysine/200 aa) for RPA indicating that four solvent-exposed lysines were modified by the fluorophore and may be hindered to bind to Tim-Tipin. In this case, the measured K_D could be underestimated.

If a label-free method was used, we might obtain a higher affinity of Tim-Tipin to RPA. To conclusively determine the K_D, a label-free method (e.g. isothermal titration calorimetry) could be applied.

However, one study from 2010 determined the K_D from fitting of the chemical shift perturbation of amino acid E252 in human RPA32WH domain upon hTipin (amino acids 185-218) binding to be 0.5 ± 0.2 µM [20]. In contrast to that study, nearly full-length proteins were used in this thesis. Here, the data shows a similar range as previously reported but reflects the affinity of the entire complex.

3.2. Requirement of Three RPA Subunits in the Tim-Tipin-RPA Complex

The SEC analysis performed in this thesis using various RPA truncations shows that the Tim-Tipin-RPA complex formation requires all three RPA subunits. In this study only a weak interaction of Tim-Tipin and RPA32WH is observed and no interaction of RPA70AB and RPA32/14 is detected. In particular, only the DNA-binding core of RPA (RPA70ABC, RPA32D, RPA14) together with the RPA32 WH-domain show relevant complex formation (Section 2.4.4). These results go beyond the study of Ali *et al.* and Gotter *et al.* only focusing on RPA32 and Tipin fragments and reporting that the complex formation of Tim-Tipin and RPA is mediated by a direct interaction of Tipin and the C-terminal RPA32 WH-domain [19, 20]. In summary, this study propose that the involvement of all RPA subunits is important for the relevant Tim-Tipin and RPA interaction under DNA-free conditions.

3.3. Adoption of a Compact Conformation of RPA in the Tim-Tipin-RPA Complex

The structural study of the Tim-Tipin and the Tim-Tipin-RPA complex in this thesis using negative stain and cryo-EM analysis provides the first view of these fully assembled complexes. Despite the number of reports investigating the structure of RPA subunits and RPA32WH in complex with a Tipin peptide, no structural analysis on Tim-Tipin-RPA has been published so far [20, 92–96].

In this study, the sample for the EM analysis displays faint dissociation and low image contrast as determined by SLS analysis, density gradient sedimentation, and negative stain EM of the Tim-Tipin-RPA complex (Section 2.2.1 and 2.3.1). To circumvent

this issues, the GraFix method was applied. This study and several other reports successfully applied this method of mild fixation and purification in a density gradient to stabilize macromolecular complexes for EM analysis [129, 134]. The quantitative mass spectrometry analysis identified all subunits in the cross-linked Tim-Tipin-RPA complex with a 1:1 stoichiometry (Tim-Tipin:RPA).

The EM analysis using negatively stained and vitrified specimens shows comparable overall architectures with a ring-like density at the front view covered by a lid at the bottom and a horse shoe-like density at the side view. The antibody labeling against the RPA70 DBD-A suggests that this RPA subunit is situated at the lid and the lower part of the U-shape density of the complex, respectively. For RPA, the crystal structures of the RPA70AB together with the RPA trimerization core (RPA70C, RPA32D, RPA14) and RPA's DNA-binding core (RPA70ABC, RPA32D, RPA14) bound to 30 nt ssDNA fit well into the EM model and support the good consistency between the 3D reconstruction and known RPA structures [94–96] (Figure 3.1 a, b). The EM analysis accompanied by the validation exercises (Section 2.3) show that RPA adopts a compact conformation in the DNA-free Tim-Tipin-RPA complex. This conformation resembles RPA's conformation in the 30 nt ssDNA binding mode and shows RPA70ABC, RPA32D and RPA14 in a horse shoe-like shape, which was previously reported [96, 97].

The relative orientation of Tim-Tipin to RPA was also mapped in this study using an antibody against Tipin and revealed the localization of Tipin at the upper part of the horse shoe-shape density. Previous studies reporting that Tipin directly interacts with the RPA32WH domain [20] motivated the placement of the RPA32WH-Tipin-C-terminus interface at the top of the ring-like density. The fact, that no interactions of Tim with RPA were reported so far [19, 83] together with the results of the mapping and docking analysis in this thesis suggest that Tim is most likely located at the right part of the ring-like density in the 3D reconstruction.

The antibody labeling, docking analysis, and difference mapping tempt to assign an envelope for Tim-Tipin (Figure 3.1 c). The shape of Tim-Tipin may be described as fingers, palm and thumb of a hand with the RPA32-WH-Tipin-C-Terminus interface being located at the top of the fingers and Tipin in the region of the finger joints. The palm could harbor Tim with its C-terminus extending to the thumb. However, this view is hypothetical and only limited conclusion can be drawn form the structural analysis in the present study. Thus, the Tim-Tipin structural model clearly requires further investigations.

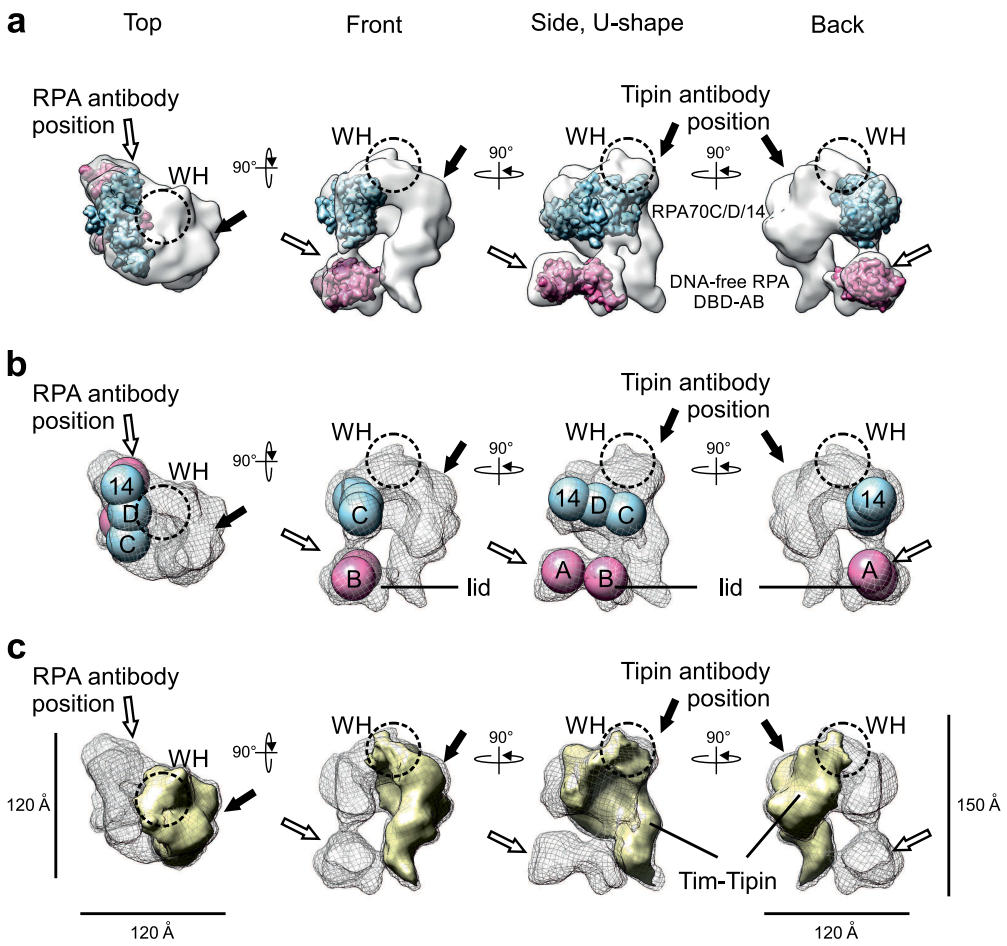


Figure 3.1.: Summary of the EM results and putative Tim-Tipin envelope. (a) Based on antibody labeling using an antibody against RPA70A, the crystal structure of DNA-free RPA70 DBD-A and -B (magenta, PDB 1FGU chainA) was fitted into the lower part of the U-shape domain (lid domain) in the Tim-Tipin-RPA 3D reconstruction. The RPA antibody position is indicated by a white arrow. Antibody labeling using an antibody against the N-terminus of Tipin suggested that the top right of the ring-like density (front) likely accommodates the RPA32WH-Tipin-C-terminus interface. The Tipin antibody position is indicated by a black arrow and the putative WH-Tipin location by a dashed circle (WH). Docking analysis fitted the crystal structure of the RPA trimerization core (blue, PDB 1L1O) into upper part of the U-shape density. (b) Schematic representation of the summary in (a). The densities show a mesh representation of the Tim-Tipin-RPA 3D reconstruction and the RPA subunits are depicted as spheres. A = RPA70 DBD-A, B = RPA70 DBD-B, C = RPA70 DBD-C. D = RPA32 DBD-D. 14 = RPA14. (c) Putative Tim-Tipin envelope (solid yellow density) and its location in the 3D map (mesh representation). It should be noted that only RPA's DNA binding core (RPA70ABC/RPA32D/RPA14) was deleted from the 3D map, thereby the yellow envelope harbors a small contribution from RPA32WH (MW ~10 kDa). Four different views of the reconstruction are shown (top, front, side, back), which are related by rotation around the y- and x-axis as indicated.

The comparison of the 2D class averages of the GraFix cross-linked and the SEC-reconstituted specimens shows no detectable differences on the structural arrangement of the RPA conformation (Figure 2.26), further supporting the notion that the compact conformation of RPA in the Tim-Tipin-RPA complex derives from binding to Tim-Tipin rather than from chemical fixation by glutaraldehyde. Despite this analysis, the question if direct interactions of Tim-Tipin with RPA are driving RPA's structure in a compact conformation remains unanswered. One the one hand, the SEC results using truncated RPA proteins show no interactions between RPA70AB and Tim-Tipin (Section 2.4.4) indicating that this subunit may play an indirect role in stabilizing Tim-Tipin-RPA and that the overall architecture of RPA is necessary for a relevant interaction. However, another possible explanation for RPA's specific structural arrangement could be weak contacts between Tim-Tipin/Tim and RPA70AB, which are not detectable by SEC. The weaker interactions between the ternary RPA70DBD-C/32/14 complex (lacking RPA70 DBD-A and -B) and Tim-Tipin compared to the RPA70/32/14 complex containing all DBD-A, -B, -C, and -D may support this hypothesis. Further, the close proximity of the putative Tim C-terminus to the lid domain accommodating RPA70AB in the EM model may hint at an additional interface between RPA and Tim-Tipin/Tim (Figure 3.1 c). Hence, further experiments are required to conclusively verify a second interface and to answer the question what is the driving force for RPA's specific compact architecture.

3.4. Dispensability of RPA70N, RPA32N, and Tim C-terminus for Tim-Tipin-RPA Complex Function

In this study, RPA's DNA binding core together with the RPA32 C-terminal WH domain, which was reported to mediate Tipin interactions, were used. The N-terminal region of RPA70 (aa 1-189, RPA70N), the first 42 residues of RPA32 (RPA32N), and the C-terminus of Tim (aa 1135-1197) were excluded for the structural and biochemical analysis. These truncations are justified by several observations emphasizing that the deleted segments do not affect the conclusions drawn from this study. First, RPA70N and RPA32N are deleted because they are not included into the DNA-binding core and thus do not contribute to RPA's ssDNA binding affinity in the 30 nt binding mode [82, 135]. Second, this two subunits are also structurally independent of RPA's DNA-binding core. For instance, RPA70N is connected by a long, unstructured, and poorly conserved linker and RPA32N is proven to be unstructured [136–138]. Third, Tim is truncated at the C-terminus because it shows severe degradation in limited proteolysis

(Section 2.1.1 and 2.1.3). Finally, truncated Tim-Tipin was fully able to bind to RPA highlighting the dispensability of this C-terminal region for Tim-Tipin-RPA complex formation.

Despite these observations, the reconstitution of Tim-Tipin-RPA using full-length RPA proteins was tested and could be successfully carried out. Further, the Tim-Tipin-RPA(FL) complex shows similar behavior to Tim-Tipin-RPA including truncated RPA proteins in terms of the dependence on ssDNA length and protein:ssDNA stoichiometry (Section 2.4.5). Thus, the truncated N-terminal RPA70 and RPA32 regions do not affect the complex formation and the biochemical function of the complex based on the SEC analysis.

However, the deleted RPA and Tim fragments may have further relevant biological functions. The C-terminal part of Tim comprises a nuclear localization signal, which was reported to be likely relevant for the nuclear import of Tim and/or Tim-Tipin and may be important for its *in vivo* function [109]. RPA32N becomes phosphorylated at the first 33 amino acids in a cell-cycle dependent manner and also in response to DNA damage [139, 140]. The phosphorylation introduces negative charge to the N-terminus and leads to a conformational change of RPA. Reportedly, this phosphorylation-induced inter-subunit rearrangement promotes the interaction of RPA32N with RPA70N [139]. In addition, RPA70N represents an important RPA-protein interaction domain and was reported to bind the p53 tumor suppressor, Rad9 (subunit of the Ras9-Hus1-Rad1 (9-1-1) complex), ATRIP (S phase checkpoint) and Mre11 (double-strand break/recombinational repair) [118, 141]. Thus, the phosphorylation-induced conformational change of RPA may modulate RPA's interaction with other proteins and/or ssDNA and may regulate RPA's function in that sense *in vivo*.

Potentially, to further characterize the role of RPA70N and RPA32N in recruitment and release of Tim-Tipin, truncated versions of RPA lacking the RPA70N and RPA32N regions could be used for *in vivo* experiments. In parallel, *in vitro* studies using phosphomimetic proteins could be used to characterize the influence of phosphorylation on Tim-Tipin-RPA complex formation and ssDNA-binding activity.

3.5. Tim-Tipin-RPA Dissociation and Resulting Implications for DNA Replication

Further, the SEC and EMSA analyses using various ssDNA substrates suggest that the stability and the ssDNA binding ability of Tim-Tipin-RPA also highly depends on RPA's

conformation in the presence of ssDNA and RPA's ssDNA binding mode, respectively. This study shows that the Tim-Tipin-RPA-ssDNA complex stays intact when RPA adopts the 30 nt binding mode at excess amounts of ssDNA and with ssDNA $\geq$ 30 nt engaging all four DBDs A-D [96, 97]. When the ssDNA length is short (<15 nt) or at sub-stoichiometric ssDNA amounts (excess amount of protein), the Tim-Tipin-RPA-ssDNA complex dissociates into DNA-free Tim-Tipin and DNA-bound RPA (Figure 2.28 and Figure 2.32, summarized in Figure 3.2). At this conditions RPA binds ssDNA in the 8 nt binding mode with DBD-A and DBD-B only contacting the ssDNA. In summary, these observations support the notions of a link between RPA's ssDNA-dependent conformation and RPA-protein interactions.

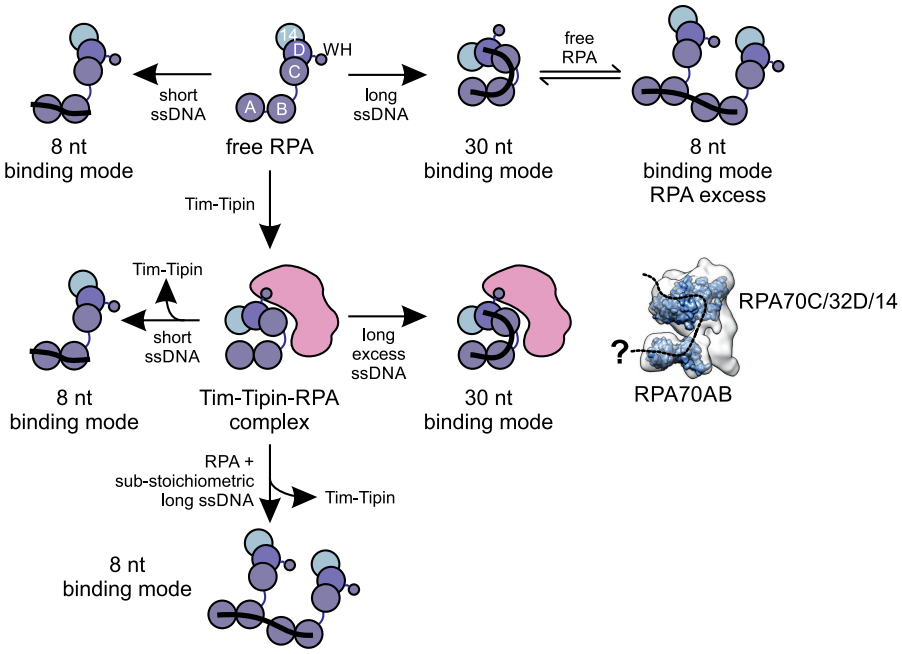


Figure 3.2.: Graphical summary of Tim-Tipin-RPA formation relying on the ssDNA length-dependent RPA conformation. RPA possesses two DNA binding modes (8 nt and 30 nt binding mode), which can coexist in a dynamic equilibrium (top). In the DNA-free state, Tim-Tipin and RPA form a complex and RPA's conformation is fixed into a compact mode resembling RPA's conformation with long ssDNA. An intact complex formation is also observed with long and excess amounts of ssDNA. In the presence of short or at sub-stoichiometric amounts of long ssDNA, thus upon the conformational change of RPA to the 8 nt binding mode, the complex dissociates into DNA-free Tim-Tipin and DNA-bound RPA. The RPA conformation-based association and dissociation of Tim-Tipin-RPA provides a molecular basis for the recruitment and correct arrangement of involved proteins at the replication fork. The hypothetical ssDNA path in the Tim-Tipin-RPA EM reconstruction is indicated as dashed line (ssDNA taken from RPA-core-32 nt complex structure, PDB 4GNX). Adopted from [128].

Several reports suggest that the Tim-Tipin complex is located between the helicase and polymerase and plays a role in coupling DNA-unwinding and DNA-synthesis by directly affecting the catalytic activities of these enzymes. These observations together with the results of this thesis lead to the speculation that the recruitment of the Tim-Tipin complex to replications forks and its release might be modulated by RPA's ssDNA-dependent conformation. While the 30 nt binding mode of RPA might be important for the recruitment of the Tim-Tipin complex to the replication fork, the conformational switch of RPA from the 30 nt to the 8 nt binding mode might be crucial for the release of the Tim-Tipin complex, which is then placed at its optimal location between the helicase and polymerase to couple the functions between these two machineries [19, 70, 74, 75, 79]. The conformational switch could be triggered by the oligomerization of RPA at opened ssDNA, which probably harbors RPA in its 8 nt binding mode. Moreover, RPA's conformation was also suggested to be modulated by binding of other proteins, for example by the SV40 Tag helicase [9, 142]. RPA's recruitment to the replication fork is proposed to be dependent on DNA unwinding by the helicase and specific RPA-helicase interaction. The RPA-helicase interaction was only stable when RPA adopted the 8 nt binding mode on ssDNA thus forming a ternary ssDNA-RPA-Tag helicase complex. These findings would further support the release of Tim-Tipin by RPA's conformational switch to the 8 nt binding mode upon recruitment, which could be induced by RPA-protein interactions (Figure 3.3).

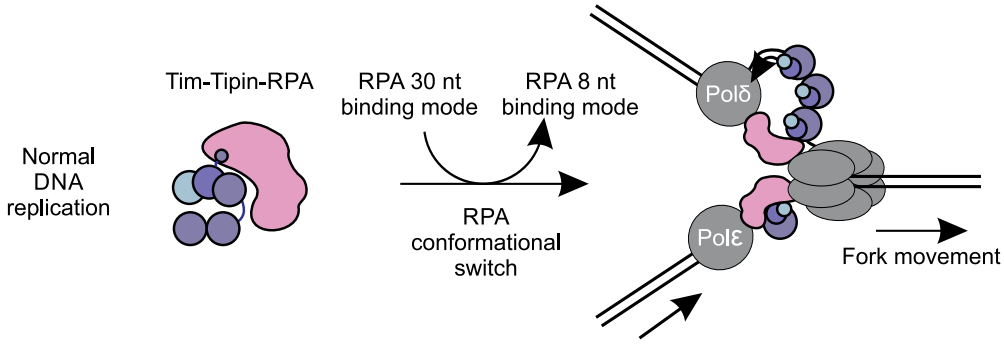


Figure 3.3.: Hypothetical role of Tim-Tipin-RPA at the replication fork under normal conditions. Tim-Tipin-RPA could form a complex with RPA in its 30 nt binding mode. After recruitment to the replication sites, Tim-Tipin could be 'handed off' to its optimal position coupling helicase and polymerase function. During this process, the release of Tim-Tipin could be induced by a conformational switch of RPA to the 8 nt binding mode by interactions with other proteins. Adopted from [128].

Interestingly, the DNA-binding studies reveal a different behavior of the Tim-Tipin complex in comparison to other RPA-binding proteins, which is profoundly discussed in Section 3.7. Briefly, many proteins destabilize RPA's 30 nt binding mode by binding to RPA's protein interaction domains (RPA70N, RPA70AB, RPA32WH) [34, 116, 142]. Protein-RPA interactions are coupled to the conformational shift of RPA from the 30 nt to the 8 nt binding mode and may lead to a protein-mediated displacement of RPA from ssDNA. In contrast, our DNA-binding studies show that the Tim-Tipin-RPA complex formation is stabilized when RPA adopts its 30 nt binding mode indicating a different behavior.

3.6. Biochemical Implications for the ssDNA Path in the Tim-Tipin-RPA Complex

The SEC, EMSA and FA experiments in this study show that Tim-Tipin binds ssDNA in a length-dependent manner with sub-micromolar affinity to long (60 nt) and no detectable affinity to short ssDNA substrates (14 nt) (Section 2.4.1 and 2.4.2). Here, the FA analysis also points out that RPA's ssDNA binding affinity is significantly higher compared to Tim-Tipin (summarized in Figure 3.4) and that the ssDNA binding in the Tim-Tipin-RPA complex is dominated by RPA as examined by FA and SEC analysis (Figure 2.28 and 2.30).

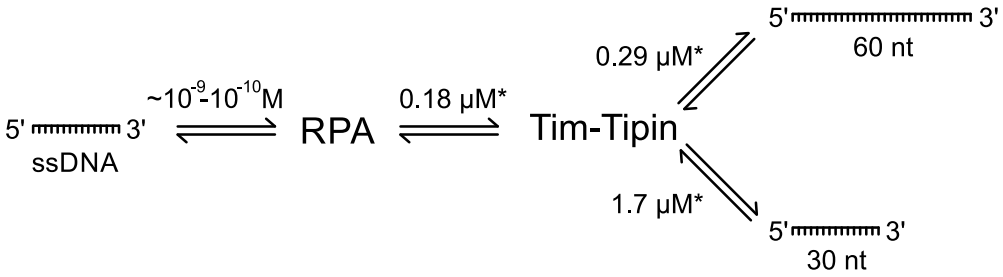


Figure 3.4.: Interaction scheme of RPA with ssDNA, RPA with Tim-Tipin, and Tim-Tipin with ssDNA. While RPA binds ssDNA with high (nano- to sub-nanomolar) affinity, Tim-Tipin shows micromolar affinity, which is length-dependent suggesting that Tim-Tipin binding to ssDNA plays only a minor role in the Tim-Tipin-RPA complex and is taken over by RPA. The asterisk indicate the K_D determined in this thesis. Adopted from [128].

Based on these results together with the fact that Tim-Tipin-RPA complex formation depends on RPA's conformation and ssDNA, this study suggests that Tim-Tipin plays only a minor role in ssDNA binding in the Tim-Tipin-RPA complex, while RPA takes

over DNA binding. The hypothetical path of the ssDNA could pass RPA's ssDNA binding core (Figure 3.2).

The weak DNA binding affinity may be advantageous for Tim-Tipin to stay proximal to the key replisome proteins. This hypothesis is in line with the reports suggesting the involvement of Tim-Tipin as an accessory protein in the DNA replication machinery coupling the function of the helicase and polymerase [19, 70, 74, 75, 79]. In this scenario, RPA may recruit Tim-Tipin to the replication fork and may 'hand off' the complex to the key replicative enzymes (Figure 3.3).

Interestingly, the structural characterization and the docking and mapping analysis of the Tim-Tipin-RPA complex may provide an explanation for the length-dependent affinity of Tim-Tipin to ssDNA and hint at the hypothesis of the presence of one long or multiple ssDNA binding sites all over Tim-Tipin. Considering the rise per base pair in B-DNA of 3.4 Å and of 3.7 Å in the elongated Z-DNA, the rise per base in ssDNA could probably correspond to be 3.5-4 Å. Further, the 3D reconstruction reveals that the Tim-Tipin complex is likely located in the right part of the ring-like density and spans about ∼150 Å (Figure 3.1). Taken together, Tim-Tipin could be spanned by 40-50 nucleotides all across its surface/DNA binding site when taking the bended form of Tim-Tipin in the reconstruction into account. Thus, only relatively long ssDNA (here 60 nt) is capable to span all across the complex and to form an adequate interface with Tim-Tipin. Based on the FA and EMSA analysis, the minimum requirement on ssDNA length for detectable Tim-Tipin interaction lies in the range of 15-30 nt. Increasing DNA length provides an additional contribution to Tim-Tipin-DNA complex formation, which is also reflected in the 6-fold higher affinity to 60 nt than to 30 nt ssDNA.

However, another explanation for the weak ssDNA binding affinity may be the acidic nature of the Tim-Tipin complex, which has a pI of 5.2. Further structural analysis is required to gain evidence about Tim-Tipin's length-dependent ssDNA binding.

3.7. Implications for Tim-Tipin-RPA and Tim-Tipin Function in the DNA Replication Checkpoint and DNA Repair

Collective observations show that the Tim-Tipin complex also interacts with components of the DNA replication checkpoint pathway [27, 61, 81, 83]. One key structure activating this pathway is accumulated ssDNA, which occurs when the DNA polymerase stalls due to DNA damage or reduced nucleotide levels causing replicative stress [143]. These impediments lead to the uncoupling of the polymerase from the helicase activity

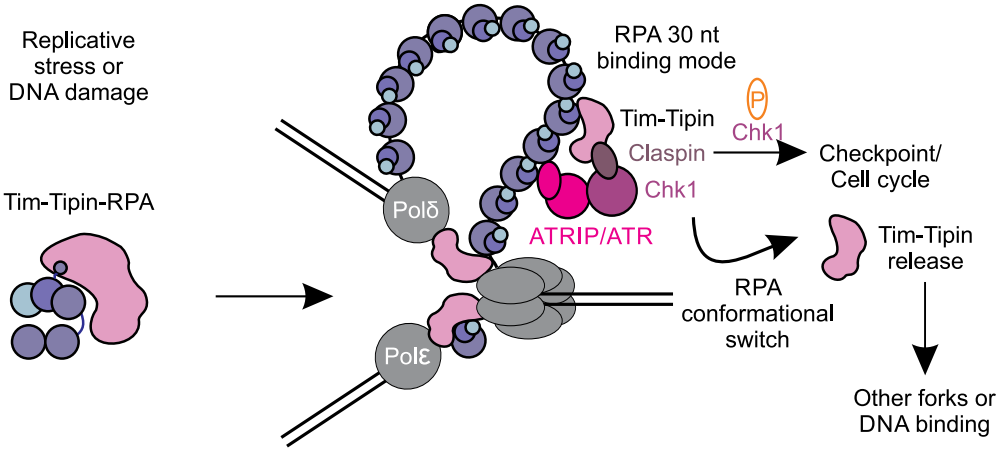


Figure 3.5.: Hypothetical role of Tim-Tipin-RPA at the replication fork under replicative stress conditions or after DNA damage. In response to replicative stress or DNA damage, accumulated ssDNA occurs, which is covered by RPA. Tim-Tipin could be recruited to these sides by RPA to mediate Chk1 phosphorylation in the DNA replication checkpoint protein complex [81]. RPA conformational switch induced by RPA-protein interactions or phosphorylation [139] could lead to Tim-Tipin release, which in turn pursues RPA-independent functions like DNA binding, helicase-polymerase coupling or association with other forks. Adopted from [128].

generating ssDNA stretches, which are immediately covered by RPA [143]. Although a continuous Tim-Tipin association with the replisome is suggested to stabilize the structure of stalled replication forks and to prevent the disassembly of the components until the damage is resolved [18], our study also suggests that Tim-Tipin may be recruited by RPA to the sides of accumulated ssDNA. Tim-Tipin associated with ssDNA-coated RPA could mediate the effective phosphorylation of Chk1 by ATR and thus contribute to the activation of the DNA damage response [61, 83]. Based on the SEC results, Tim-Tipin could be recruited and released from these sites in an RPA conformation-dependent manner. The conformational change could be induced by phosphorylation [139] or interactions with other proteins (discussed on next page). Thus, Tim-Tipin-RPA-ssDNA could serve as an adapter unit to physically link ATR and Chk1 and thus facilitate Chk1 phosphorylation. The release of Tim-Tipin from RPA could allow additional Chk1 phosphorylation by association with other forks (Figure 3.5).

The EMSA analysis in this thesis using branched DNA show that Tim-Tipin has a broad specificity to fork-like and splayed DNA junctions. This result together with the FA and SEC analysis using ssDNA hint at a RPA-independent role. Similar Tim-Tipin association with single-, double-stranded, and fork-like DNA is also reported *in vitro*

by Yilmaz *et al.* supporting the hypothesis that Tim-Tipin may use its affinity to recognize checkpoint inducing DNA substrates [144]. Nevertheless, further investigations are required to provide insights into the RPA-independent function of Tim-Tipin in checkpoint signaling.

Single-stranded DNA structures do not only occur as checkpoint inducing substrates after uncoupling of the polymerase and helicase due to replicative stress and DNA damage but also appear as intermediate structures in DNA repair mechanisms, namely nucleotide excision repair (NER), base excision repair (BER), and double-strand break (DSB) repair by homologous recombination. They usually emerge from oligomers being excised at the damaged DNA strand. RPA is suggested to stabilize these intermediate structures and to serve as an anchor point for the assembly of proteins involved in DNA repair [21]. The SEC and EMSA results on the ssDNA binding characteristics of the Tim-Tipin-RPA complex hint at a RPA-mediated recruitment of Tim-Tipin to the sides of the DNA damage. This hypothesis is supported by the observations that Tipin shares a similar binding interface on the RPA32WH domain like the DNA repair enzymes XPA and UNG2 and displays a sequence similarity to these enzymes in the region responsible for RPA binding [20, 34]. In addition, the MST measurement in this study determines a sub-micromolar binding affinity of Tim-Tipin to RPA (K_D = 0.18 µM), which is comparable to the binding affinity of XPA to RPA (K_D = 0.02 µM) [145]. It should be noted, that the XPA binding affinity highly depends on the length of the used protein constructs. Thus, binding studies using only XPA fragments, which have a sequence similarity to Tipin (e.g. human XPA, aa 1-98), showed a much lower affinity to RPA and were determined to be in the micro-molar range (K_D = 5 µM) [34], indicating that an additional contribution to RPA binding is made by other parts of XPA. This is consistent with the reported interaction of XPA's C-terminal region with RPA70 [145]. High K_D values were also reported for short UNG2 constructs (human UNG2, aa 73-88, $K_D \leq 10^{-6}$ M) [34].

Interestingly, several studies suggest that the association of XPA, UNG2, Rad52 [34], and the SV40 Tag helicase [116, 142] with the C-terminal region of RPA32 or with RPA70AB and RPA70N is attended by a conformational switch of RPA from the 30 nt to the 8 nt binding mode and by the dissociation of the RPA trimerization core, which is weakly bound to ssDNA (K_D = 2-10 µM) [97, 146]. It is speculated that the binding of these proteins places constraints on RPA's conformation and increases the distance between RPA70AB and the trimerization core. The resulting steric hindrances then prevent the re-association of the trimerization core with the ssDNA and thus may even

lead to a conformation induced displacement of RPA from DNA [97]. Unlike these suggestions of Brosey *et al.* and the sequence similarity of Tipin to XPA and UNG2, this thesis shows that the Tim-Tipin-RPA complex formation on ssDNA is only achieved with RPA in the 30 nt binding mode and is negatively affected by the 30 nt to the 8 nt conformational switch. The hypothesis that the RPA32WH domain acts as a general anchor and exchange point for RPA accessory proteins may be extended by this study [34]. The results lead to speculations that RPA's conformation also plays a critical role in the assembly of RPA accessory protein during DNA repair.

Further, the ubiquitous occurrence of ssDNA as an intermediate structure in normal and perturbed DNA replication, in DNA repair, and as a checkpoint inducing structure suggests that Tim-Tipin-RPA-ssDNA formation may also serve as a link between DNA replication, the S phase checkpoint pathway, and DNA repair.

3.8. Dynamic Interactions as Mechanistic Basis for the Replication Fork Organization

In summary, the Tim-Tipin complex is recruited to ssDNA by a specific interaction with RPA, which is formed by a similar RPA binding site as observed for key repair proteins. However, the formation and dissociation of Tim-Tipin-RPA-ssDNA is further based on RPA's conformation/binding mode and occurs in a highly dynamic manner. This result provides a mechanistic basis for the correct arrangement of the involved proteins at the replication fork. In general, dynamic interactions may support a high turnover of proteins and may be advantageous for an efficient and fast adaptation of the involved proteins at the replication fork to incidents during DNA replication. This study indicates that the recruitment of RPA binding proteins to RPA/RPA-ssDNA takes place in a sensitive equilibrium and hint at an elaborated organization of the replisome accessory components to ensure correct duplication of the genome.

3.9. Open Questions and Outlook

The Tim-Tipin complex fulfill a tethering and stabilizing function at the replication fork and is suggested to be recruited by a specific interaction with RPA. The EM 3D reconstruction of the Tim-Tipin-RPA complex shows that RPA adopts a U-shape conformation resembling the conformation when RPA is bound to ssDNA. Determination of the crystal structure of Tim-Tipin alone and in complex with RPA will explore what

is the driving force for the preference of this specific compact RPA architecture. In addition to the SEC experiments of this study, further binding studies are required to conclusively answer the question if a second interface between RPA70AB and Tim-Tipin/Tim is existing. To explicitly determine the K_D between Tim-Tipin and RPA, a label-free method (e.g. isothermal titration calorimetry, ITC) could be applied in addition to the MST measurements.

Moreover, the Tim-Tipin-RPA complex lacking the RPA70N and RPA32N regions could be used for *in vivo* experiments to characterize the role of these regions in recruitment and release of Tim-Tipin and other accessory proteins. In addition, interaction studies like MST or ITC using full-length, truncated or phospho-mimetic proteins could be performed to better quantify the impact of the RPA70N and RPA32N regions and to further characterize the influence of the phosphorylation on Tim-Tipin-RPA complex formation and ssDNA-binding activity *in vitro*.

Besides RPA, Tim-Tipin binds a number of replication factors like polymerases, the helicase and Claspin. Structural characterization of Tim-Tipin in complex with key or accessory replication protein is crucial to advance the understanding of the concerted assembly and function in the progressing replisome.

Further, the biochemical analysis in this thesis shows an intact Tim-Tipin-RPA complex formation with ssDNA under particular conditions. These results could serve as a lead to define the putative ssDNA path in Tim-Tipin-RPA and could help to conclusively answer the question, if Tim-Tipin contacts the ssDNA in the complex. It will be also interesting to further examine Tim-Tipin's length-dependent ssDNA binding by advanced binding experiments or structural characterization. These approaches could provide the answer to the question if Tim-Tipin harbors a long or multiple DNA binding sites for long DNA substrates spanning across the complex, which can form an adequate interface with Tim-Tipin. In addition, the results revealing a broad specificity of Tim-Tipin in DNA substrate recognition could be expand to characterize the putative RPA-independent roles of Tim-Tipin.

Moreover, Tim-Tipin is suggested to fulfill further functions in DNA processing. In addition to tethering and stabilizing function at the replisome, Tim-Tipin is involved in the DNA replication checkpoint promoting Chk1 activation. Tipin's sequence similarity to DNA repair enzymes and its similar binding interface on RPA32WH hints at a third function in DNA repair mechanisms. How these three pathway are interconnected and how Tim-Tipin-RPA or individual Tim-Tipin possibly link these three pathways by potentially favoring particular protein interactions of one pathway and disfavoring of the

other is poorly understood. The RPA conformation-induced association and dissociation of Tim-Tipin-RPA and the highly dynamic nature of the Tim-Tipin-RPA complex formation on ssDNA examined in this thesis provide a mechanistic basis for the recruitment and organization of other factors involved in the DNA processing pathways. Competitive pull-down experiments using purified proteins could help answering the question, if RPA/RPA-ssDNA/RPA-damaged DNA in complex with Tim-Tipin serves as an assembly platform for other proteins or if RPA recruits Tim-Tipin and other factors in a mutual exclusive manner? SEC analysis as performed in this study could investigate the highly dynamic complex assembly with additional factors. Finally the identification of new Tim-Tipin/Tim-Tipin-RPA interaction partners and further investigations of Tim-Tipin in the fully assembled DNA replisome could shed light on the elaborated organization of normal and perturbed DNA replication.

4. Materials and Methods

4.1. Materials

4.1.1. Consumables and Chemicals

All chemicals were purchased from Biomol (Hamburg, Germany), Merck (Darmstadt, Germany), Qiagen (Venlo, The Netherlands), Roche Diagnostics (Mannheim, Germany), Roth (Karlsruhe, Germany), Serva (Heidelberg, Germany), and Sigma-Aldrich (München, Germany). Transfection and ECL reagents were ordered from Invitrogen (Darmstadt, Germany) and GE Healthcare (München, Germany). Enzymes were purchased from Merck (Darmstadt, Germany), NEB (Schwalbach, Germany), Roche Diagnostics (Mannheim, Germany), and Thermo Scientific (Epson, UK). Antibodies were obtained from Santa Cruz Biotechnology (Santa Cruz, CA, USA). PVDF membranes were provided by Merck (Darmstadt, Germany). Template DNAs were purchased from Thermo Scientific (Epson, UK). Gel and plasmid extraction and fluorescent labeling kits were purchased from Qiagen (Venlo, The Netherlands) and Jena Bioscience (Jena, Germany). Buffer exchange columns, ultrafiltration devices and standard proteins were obtained from Bio-Rad Laboratories (Hercules, CA, USA), Sartorius Stedim Biotech, (Göttingen, Germany), and Thermo Scientific (Epson, UK). Chromatography material was purchased from GE Healthcare (München, Germany) and Qiagen (Venlo, The Netherlands). DNA oligonucleotides were synthesized by Eurofins Genomics (Ebersberg, Germany) and Purimex (Grebenstein, Germany). Bacterial and insect cell expression flasks were obtained from Corning B.V. Life Sciences (Amsterdam, The Netherlands), Sarstedt (Nümbrecht, Germany), and Sigma-Aldrich (München, Germany). Crystallization screens and cryo-loops were purchased from Hampton Research (Aliso Viejo, CA, USA) and Qiagen (Venlo, The Netherlands). Pre-cast gels and gel running buffers were obtained from Invitrogen (Darmstadt, Germany).

4.1.2. Equipment

Instrument	Supplier
Ultrasonic homogenizer	Bandelin electronic, Berlin, Germany
Emulsiflex C3 homogenizer	Avestin, Mannheim, Germany
Dounce homogenizer	Thermo Scientific, Epson, UK
ÄKTA purification systems	GE Healthcare, München, Germany
High pressure liquid chromatography	Waters, Milform, MA, USA
TDA302 detector array	Viscotek, Malvern, Herrenberg, Germany
NanoDrop spectrophotometer	Thermo Scientific, Epson, UK
ImageQuant LAS 4000 imager	GE Healthcare, München, Germany
Typhoon FL 7000 phosphoimager	GE Healthcare, München, Germany
Genios Pro fluorescence reader	Tecan, Männedorf, Switzerland
Monolith NT.115	NanoTemper,München, Germany
Gradient station	Biocomp, Fredericton, NB, Canada
Gradient fractionator	Biocomp, Fredericton, NB, Canada
SW40 rotor	Beckmann Coulter, Krefeld, Germany
Plasma cleaner chamber	Harrick Plasma, Ithaca, NY ,USA
Vitrobot cryo-station	FEI, Eindhoven, The Netherlands
CM200-FEG electron microscope	FEI, Eindhoven, The Netherlands
Tecnai F20 electron microscope	FEI, Eindhoven, The Netherlands
Phoenix nanodispenser robot	Art Robbins Instruments, Sunnyvale, CA, USA
PXII beam line	SLS, Villigen, Switzerland

4.1.3. Software

Software	Supplier/Developer
BSOFT	Heymann and Belnap 2007 [147]
CTFFIND3	Mindell and Grigorieff [148]
EMAN	Ludktke *et al.* 1999 [149]
Max Quant	Cox and Mann 2008 [150]
OmniSEC 4.5	Viscotek, Malvern, Herrenberg, Germany
Origin 8.1	Origin-Lab, Northampton, MA, USA
SPARX	Hohn *et al.* 2007 [151]
SPIDER	Frank *et al.* 1996 [152]
UCSF Chimera	Pettersen *et al.* 2004 [132]
VectorNTI	Life Technologies, Darmstadt, Germany
XDS	Kabsch 2010 [153]

4.1.4. Media and Buffers

Medium	Composition	Cell type
LB	1% (w/v) Bacto Trypton, 0.5% (w/v) Yeast extract, 0.5% (w/v) NaCl, pH 7.2	*E. coli*
TB	1.2% (w/v) Bacto Trypton, 2.4% (w/v) Yeast extract, 0.004% (v/v) glycerol, supplemented with phosphate buffer (10% (v/v), 0.17 M KH_2PO_4, 0.72 M K_2HPO_4 x 3 H_2O)	*E. coli*
SOC	2% (w/v) Bacto Trypton, 0.5% (w/v) Yeast extract, 10 mM NaCl, 10 mM $MgSO_4$, 2.5 mM KCl, 1 mM $MgCl_2$, 0.4% glucose, pH 7.2	*E. coli*
Sf-900 II SFM	Serum-free, protein-free insect cell culture medium	Insect cells

4.1.5. Antibiotic Solutions

Antibiotic	Stock solution	Final concentration
Ampicillin	100 mg/ml	100 µg/ml
Kanamycin	50 mg/ml	50 µg/ml
Streptomycin	50 mg/ml	50 µg/ml
Chloramphenicol	34 mg/ml	34 µg/ml

4.1.6. Bacterial Strains and Cell Lines

Table 4.1.: Bacterial stains

Strain	Species	Genotype
DH5α	*E. coli*	F$^-$ ϕ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96, *rel*A1 λ^-
BL21 DE3 gold	*E. coli*	F$^-$ *omp*T *hsd*S(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ(DE3) *end*A Hte
BL21 DE3 gold pLysS	*E. coli*	F$^-$ *omp*T *hsd*S(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ(DE3) *end*A Hte [pLysS Cam^r]
BL21 DE3 gold pRARE	*E. coli*	F$^-$ *omp*T *hsd*S(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ(DE3) *end*A Hte [pRARE Cam^r]
BL21 DE3 pRIL	*E. coli*	F$^-$ *omp*T *hsd*S(r_B^- m_B^-) *gal dcm* λ(DE3) [pRIL Cam^r]
DH10Bac	*E. coli*	F$^-$ *mcr*A, Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara*, *leu*)7697 *gal*U *galK*λ- *rps*L, *nup*G /pMON14272 /pMON7124

Table 4.2.: Cell lines

Cell line	Description
Sf21	Ovarian cells isolated from *Spodoptera frugiperda* (fall armyworm); used for transfection with bacmid DNA and propagation of recombinant baculovirus stocks
Sf9	originated from Sf21; used for transfection with bacmid DNA and propagation of recombinant baculovirus stocks
High Five	Ovarian cells derived from the parental *Trichopulsia ni* cell line (cabbage looper); used for the expression of recombinant proteins using the baculovirus expression system

4.1.7. Antibodies

Antibody (ab)	Specificity	Order number	Dilution
anti-Tipin	primary ab, polyclonal IgG from goat	sc-160865	1:4000
anti-RPA70	primary ab, monoclonal IgG_1 from mouse	sc-166023	1:5000
anti-goat	secondary ab, IgG-HRP from donkey	sc-2020	1:10000
anti-mouse	secondary ab, IgG-HRP from goat	BML-SA204-0100	1:10000

4.1.8. Plasmids

Table 4.3.: Vectors

Name	Application	Tag	Cleavable
pEC-A-HT-His	Bacterial expression of proteins	N-terminal 6xHis-Tag	TEV
pEC-K-3C-His	Bacterial expression of proteins	N-terminal 6xHis-Tag	3C
pEC-S-CDF-His	Bacterial expression of proteins	N-terminal 6xHis-Tag	Tev
pEC-K-3C-GST	Bacterial expression of proteins	N-terminal 6xHis-GST-Tag	3C
pCoofy27	Insect cell expression of proteins	N-terminal 6xHis-Tag	3C

Table 4.4.: Cloned Plasmids

Name	Vector	Primer	Insertion sites
Tim 992-1134	pEC-A-HT-His	17JW, 18JW	LIC TEV
Tim 1008-1134	pEC-K-3C-GST	25JW, 3JW	LIC 3C
Tim 1008-1125	pEC-K-3C-GST	25JW, 2JW	LIC 3C
Tim 872-949	pEC-K-3C-GST	14JW, 15JW	LIC 3C
Tim 872-1134	pEC-K-3C-GST	14JW, 3JW	LIC 3C
Tim 810-1134	pEC-K-3C-GST	12JW, 3JW	LIC 3C
Tim 818-1138	pEC-K-3C-GST	24JW, 23JW	LIC 3C
Tim 1-655	pEC-S-CDF-His	37JW, 41JW	LIC TEV
Tim 1-585	pEC-K-3C-GST	4JW, 7JW	LIC 3C
Tim 1-529	pEC-S-CDF-His	37JW, 48JW	LIC TEV
Tim 1-332	pEC-K-3C-GST	4JW, 6JW	LIC 3C
Tim 1-267	pEC-S-CDF-His	37JW, 38JW	LIC TEV
Tim 1-241	pEC-S-CDF-His	37JW, 49JW	LIC TEV
Tim 1-1134	pEC-S-CDF-His	37JW, 18JW	LIC TEV
Tim 1-1197	pCoofy27	9JW, 10JW	SLIC
Tipin 55-127	pEC-K-3C-GST	34JW, 64JW	LIC 3C
Tipin 55-136	pEC-K-3C-GST	34JW, 65JW	LIC 3C
Tipin 55-150	pEC-K-3C-GST	34JW, 35JW	LIC 3C
Tipin 55-169	pEC-K-3C-GST	34JW, 66JW	LIC 3C
Tipin 55-220	pEC-K-3C-GST	34JW, 36JW	LIC 3C
Tipin 55-278	pEC-K-3C-GST	34JW, 29JW	LIC 3C
Tipin 1-127	pEC-K-3C-GST	28JW, 64JW	LIC 3C
Tipin 1-136	pEC-K-3C-GST	28JW, 65JW	LIC 3C
Tipin 1-169	pEC-K-3C-GST	28JW, 66JW	LIC 3C
Tipin 1-278	pEC-K-3C-GST	28JW, 29JW	LIC 3C
RPA70 1-623	pEC-S-CDF-His	89JW, 57JW	LIC TEV
RPA70 190-623	pEC-S-CDF-His	78JW, 57JW	LIC TEV
RPA70 445-623	pEC-S-CDF-His	56JW, 57JW	LIC TEV
RPA70 190-431	pEC-K-3C-His	78JW, 79JW	LIC 3C
RPA32 1-270	pEC-A-HT-His	90JW, 59JW	LIC TEV
RPA32 43-270	pEC-A-HT-His	58JW, 59JW	LIC TEV
RPA32 173-270	pEC-K-3C-GST	42JW, 43JW	LIC 3C
RPA14 1-121	pEC-K-3C-His	53JW, 54JW	LIC 3C

4.1.9. Oligonucleotides for Cloning

Name	Sequence (5'→3')
2JW	cagaccgccaccgactgcttaTTCTGTCCGGTGCTCCTCAC
3JW	cagaccgccaccgactgcttaACGGGCTGACAGAAGGGC
4JW	ccaggggcccgactcgatgGACTTGTACATGATGAACTGTGAAC
6JW	cagaccgccaccgactgcttaGCGATGGACAGACAGCTCC
7JW	cagaccgccaccgactgcttaCTCTGAGGCCGCATCAAAGG
9JW	aagttctgttccaggggcccATGGACTTGTACATGATGAACTGTGAACTTCTAG
10JW	ccccagaacatcaggttaatggcgTCAGTCATCCTCATCCTCAATCTGAAACC
12JW	ccaggggcccgactcgAGCCACAGAGCTCCTCTGTG
14JW	ccagggagcagcctcgctggaagttctgttccaggggcccGAGTTCCAGAAGAGGAAAGGG
15JW	gcaaagcaccggcctcgttaAGAAGGCGCCAGCTTCTTTC
17JW	ccagggagcagcctcgctggaagttctgttccaggggcccGGTCAGGGCAGCTCATCTC
18JW	gcaaagcaccggcctcgttaACGGGCTGACAGAAGGGC
23JW	cagaccgccaccgactgcttaTGCTTTCCTCTTACGGGCTG
24JW	ccaggggcccgactcgatgCCTGAGGAAGAGGCCCAG
25JW	ccaggggcccgactcgatgCGTCAGGAAGGCCTCTCTG
28JW	ccaggggcccgactcgatgTTAGAGCAAGAAGAAAATGGCTTGTTTG
29JW	cagaccgccaccgactgcttaGTCTAAGTTAGTATGGTCCAGTTG
34JW	ccaggggcccgactcgatgAGAACAGTTAAAAGGAATCTACCTAAGC
35JW	cagaccgccaccgactgcttaATCTTCATGTACAATAGGCAGATCAAG
36JW	cagaccgccaccgactgcttaCTGGCTATTACTCAGTAGCTTTGC
37JW	ccagggagcagcctcgatgctggaagttctgttccaggggcccATGGACTTGTACATGATGAACTGTGAAC
38JW	gcaaagcaccggcctcgttaCCGAGCTCTCTTCTCCGC
41JW	gcaaagcaccggcctcgttaCTCCTGCTGCCGGGG
42JW	ccaggggcccgactcgatgAACAGCCAGGCCTCTGCAG
43JW	cagaccgccaccgactgcttaCTCTGCATCTGTAGACTTAAAGTG
48JW	gcaaagcaccggcctcgttaTCTTTTGTTCTGCACCATCAGGTTC
49JW	gcaaagcaccggcctcgttaGCGTCCCTGCCCTACTC
53JW	ccaggggcccgactcgatgGAGGACATAATGCAGCTCCC
54JW	cagaccgccaccgactgcttaTTCATGTTGTGGAAGCCCTACAG
56JW	ccagggagcagcctcgatgctggaagttctgttccaggggcccGGAGGGGGCAACACCAAC
57JW	gcaaagcaccggcctcgttaCATGTTCTTCCTGATGTTCGCG
58JW	ccagggagcagcctcgatgctggaagttctgttccaggggcccGCCCAGCATATTGTGCCCTG
59JW	gcaaagcaccggcctcgttaCTCTGCATCTGTAGACTTAAAGTG
64JW	cagaccgccaccgactgcttaTAGGTTTTCGACTCTGTCAATAAAATCTTC
65JW	cagaccgccaccgactgcttaACAGGTCTGAACCTCCTTTTTATTTC
66JW	cagaccgccaccgactgcttaAGTAGCAGATACATCTGGGCC
78JW	ccaggggcccgactcgatgCAGTCCAAAGTGGTGCCCATC
79JW	cagaccgccaccgactgcttaTAAGGCTTGTCCTTCTGAGTCAAAC
89JW	ccagggagcagcctcgatgGTGGGACACCTGAGCGAG
90JW	ccagggagcagcctcgatgTGGAATAGCGGATTCGAAAGCTTC

Small letters show LIC and SLIC overhangs.
Capital letters correspond to the sequence of the gene of interest.

4.1.10. Buffers for Protein Purification

Buffer	Composition	Protein	Application
PD 1	20 mM Tris pH 7.5, 150 mM NaCl, 2 mM imidazole, 10% glycerol, 0.1% NP-40	all tested constructs	Expression/ solubility test
Lys 1	50 mM Tris pH 7.4, 400 mM NaCl, 25 mM imidazole, 10% glycerol, 1 mM β-ME	Tim-Tipin, RPA70/32/14, RPA32/14, Tipin	Cell lysis
Lys 2	50 mM Tris pH 7.5, 400 mM NaCl, 25 mM imidazole, 10% glycerol, 1 mM β-ME	RPA70FL/32FL/14	Cell lysis
Lys 3	50 mM Tris pH 7.4, 300 mM NaCl, 10 mM imidazole, 10% glycerol, 1 mM β-ME	RPA70C/32/14	Cell lysis
Lys 4	50 mM Tris pH 7.5, 300 mM NaCl, 14 mM β-ME, 10% glycerol, 1 mM EDTA	RPA32WH, C-term. Tim	Cell lysis
Lys 5	50 mM Tris pH 7.5, 300 mM NaCl, 10 mM imidazole, 15% glycerol, 1 mM β-ME, 1 mM EDTA	RPA70AB	Cell lysis
RSB	10 mM Tris pH 7.5, 10 mM NaCl, 1 mM $MgCl_2$	Tim	Cell lysis
His 1	50 mM Tris pH 7.4, 200 mM NaCl, 25 mM imidazole, 5% glycerol, 1 mM β-ME	Tim-Tipin, RPA70/32/14, RPA32/14, Tipin	IMAC
His 2	50 mM Tris pH 7.4, 200 mM NaCl, 500 mM imidazole, 5% glycerol, 1 mM β-ME	Tim-Tipin, RPA70/32/14, RPA32/14, Tipin	IMAC
His 3	50 mM Tris pH 7.5, 200 mM NaCl, 25 mM imidazole, 5% glycerol, 1 mM β-ME	RPA70FL/32FL/14	IMAC
His 4	50 mM Tris pH 7.5, 200 mM NaCl, 500 mM imidazole, 5% glycerol, 1 mM β-ME	RPA70FL/32FL/14	IMAC
His 5	50 mM Tris pH 7.4, 300 mM NaCl, 10 mM imidazole, 5% glycerol, 1 mM β-ME	RPA70C/32/14, RPA70AB	IMAC
His 6	50 mM Tris pH 7.4, 300 mM NaCl, 500 mM imidazole, 5% glycerol, 1 mM β-ME	RPA70FC/32/14, RPA70AB	IMAC
His 7	50 mM Tris pH 7.4, 500 mM NaCl, 40 mM imidazole, 1 mM β-ME	Tim	IMAC
His 8	50 mM Tris pH 7.4, 500 mM NaCl, 500 mM imidazole, 1 mM β-ME	Tim	IMAC
HS 1	His-buffer 1 supplemented with 1 M NaCl	Tim-Tipin	Wash
HS 2	His-buffer 1 supplemented with 3 M NaCl	RPA70/32/14, RPA32/14	Wash
HS 3	His-buffer 1 supplemented with 2.5 M KCl	RPA70/32/14	Wash
HS 4	His-buffer 4 supplemented with 1 M NaCl	RPA70C/32/14	Wash
HS 5	His-buffer 5 supplemented with 1 M NaCl	RPA70AB	Wash
ATPW 1	His-buffer 1 supplemented with 2 mM ATP, 10 mM $MgSO_4$, 50 mM KCl	Tim-Tipin, RPA70/32/14, RPA32/14	Wash
ATPW 2	His-buffer 4 supplemented with 2 mM ATP, 10 mM $MgSO_4$, 50 mM KCl	RPA70C/32/14	Wash
Q 1	25 mM Bis-Tris pH 6.5, 150 mM NaCl, 14 mM β-ME	Tim-Tipin, Tipin	AEX
Q 2	25 mM Bis-Tris pH 6.5, 1000 mM NaCl, 14 mM β-ME	Tim-Tipin, Tipin	AEX
Q 3	25 mM Bis-Tris pH 7.0, 150 mM NaCl, 14 mM β-ME, 5% glycerol, 0.01 mM $Zn(OAc)_2$	RPA70/32/14, RPA32/14	AEX
Q 4	25 mM Bis-Tris pH 7.0, 1000 mM NaCl, 14 mM β-ME, 5% glycerol, 0.01 mM $Zn(OAc)_2$	RPA70/32/14, RPA32/14	AEX

Buffer	Composition	Protein	Application
Q 5	25 mM Bis-Tris pH 7.3, 100 mM NaCl, 14 mM β-ME, 5% glycerol, 0.01 mM $Zn(OAc)_2$	RPA70FL/32FL/14	AEX
Q 6	25 mM Bis-Tris pH 7.3, 1000 mM NaCl, 14 mM β-ME, 5% glycerol, 0.01 mM $Zn(OAc)_2$	RPA70FL/32FL/14	AEX
Q 7	20 mM Tris pH 7.6, 100 mM NaCl, 5 mM DTT, 5% glycerol, 0.01 mM $Zn(OAc)_2$	RPA70C/32/14	AEX
Q 8	20 mM Tris pH 7.6, 1000 mM NaCl, 5 mM DTT, 5% glycerol, 0.01 mM $Zn(OAc)_2$	RPA70C/32/14	AEX
Q 9	50 mM BTP pH 7.0, 50 mM NaCl, 10 mM β-ME	RPA32WH	AEX
Q 10	25 mM Bis-Tris pH 7.0, 150 mM NaCl, 5% glycerol, 5 mM β-ME	RPA70AB	AEX
Q 11	25 mM Bis-Tris pH 7.0, 1000 mM NaCl, 5% glycerol, 5 mM β-ME	RPA70AB	AEX
Q 12	50 mM Bis-Tris pH 7.0, 100 mM NaCl, 5% glycerol, 10 mM β-ME	Tim	AEX
Q 13	50 mM Bis-Tris pH 7.0, 1000 mM NaCl, 5% glycerol, 5 mM β-ME	Tim	AEX
Q 14	50 mM Bis-Tris pH 6.0, 50 mM NaCl, 14 mM β-ME	C-term Tim	AEX
Q 15	50 mM Bis-Tris pH 6.0, 1000 mM NaCl, 14 mM β-ME	C-term Tim	AEX
Dia 1	50 mM Tris pH 7.3, 150 mM NaCl, 14 mM β-ME	Tim-Tipin	Dialysis
SEC 1	20 mM Hepes pH 7.0, 125 mM NaCl, 2 mM DTT	Tim-Tipin	SEC
SEC 2	20 mM Hepes pH 7.5, 125 mM NaCl, 2 mM DTT, 0.01 mM $Zn(OAc)_2$	RPA70/32/14, RPA70FL/32FL/14	SEC
SEC 3	20 mM Hepes pH 7.5, 150 mM NaCl, 5 mM DTT, 2.5% glycerol, 0.01 mM $Zn(OAc)_2$	RPA70C/32/14, RPA32/14	SEC
SEC 4	25 mM Hepes pH 7.0, 150 mM NaCl, 2 mM DTT	RPA32WH, C-term. Tim	SEC
SEC 5	25 mM Hepes pH 7.0, 300 mM NaCl, 2 mM DTT	Tim	SEC

IMAC, immobilized metal ion affinity chromatography; β-ME, β-mercaptoethanol; ATP, adenosine triphosphate; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; AEX, anion-exchange chromatography; SEC, size exclusion chromatography

4.2. Methods

4.2.1. Cloning Procedures

Template DNA

Genes encoding full-length mouse RPA70 (BC019119), full-length mouse RPA32 (BC004578), full-length mouse RPA14 (BC028489) and full-length mouse Tipin (BC016211) were purchased from Thermo Scientific (Epson, UK). The full-length Tim gene (AB019001) was a generous gift from Dr. Achim Kramer, Charité Berlin, Germany.

Polymerase Chain Reaction

Genes were amplified from template DNA using *Pfu* DNA Polymerase (Thermo Scientific Fermentas, Epson, UK) and a standard PCR and amplification protocol (Table 4.4 and 4.5). After amplification the PCR product was analyzed by agarose gel elec-

trophoresis (1% w/v) in 1x TAE (40 mM Tris pH 8.0, 0.1% HOAc, 1 mM EDTA), cut from the gel and purified using a gel extraction kit (Qiagen, Venlo, The Netherlands).

Table 4.4.: PCR reaction

Component	Stock concentration	Final concentration	Amount
dNTPs	2 mM	0.2 mM	5 µL
Pfu buffer	10x	1x	5 µL
Template DNA	∼50-200 ng/µL	∼100 ng	1 µL
5' Primer	10 µM	0.4 µM	2 µL
3' Primer	10 µM	0.4 µM	2 µL
Pfu polymerase	2.5 U/µL	2.5 U	1 µL
ddH_2O		add to 50 µL	34 µL

Table 4.5.: PCR amplification protocol

Step	Temperature [°C]	Time	Cycles
1	95	180 sec	1
2	95	30 sec	
3	55	30 sec	
4	72	60 sec/500 bp	30x to step 2
5	72	10 min	1
6	4	∞	

Cloning in pEC Vectors

Amplified genes (Section 4.2.1) were cloned into pEC series of vectors generated by Jerôme Basquin and Florence Martin (MPIB), which were designed for Ligase Independent Cloning (LIC) [154]. The LIC principle utilizes the 3' to 5'-exonuclease activity of the T4 DNA polymerase (Novagen, Merck KGaA, Darmstadt, Germany) to create 12-15 base pairs (bp) long single-stranded overhangs on the linearized vector and the PCR product, which are complementary. By incubation of the processed vector and PCR product, the compatible overhangs anneal forming a circular product and can be directly transformed in *E. coli* without requiring DNA ligase.

For cloning, the LIC vector was linearized for 3 hours at 37 °C using restriction enzymes (NEB, Schwalbach, Germany) as described in Table 4.6. After restriction, the cleavage reaction was loaded onto an agarose gel and the DNA was isolated using a gel extraction kit (Qiagen). The genes of interest were amplified using oligonucleotides with specific LIC overhangs as described in Section 4.2.1. For LIC insert and LIC vector processing, the reaction was prepared as described in Table 4.7 and 4.8. After 30 min incubation at RT, the enzyme was inactivated at 75 °C for 20 min.

Note, 2 µL processed insert and 1 µL processed vector were mixed and incubated for 10 min at RT. After addition of 1 µL EDTA (6.25 mM final concentration), the mixture

Table 4.6.: Vector linearization using restriction enzymes

Component	Stock concentration	Amount
Vector		6 µg
10x NEBuffer 4	10x	10 µL
ZraI or SacII	10 U/µL or 20 U/µL	10-5 µL
ddH_2O		add to 100 µL

Table 4.7.: LIC insert processing using a 20 µL reaction mix

Component	Stock concentration	Amount
PCR product		600 ng (30 ng/µL)
T4 DNA polymerase buffer	10x	2 µL
dATP	25 mM	2 µL
DTT	100 mM	1 µL
T4 DNA Pol. LIC qualified	2.5 U/µL	0.4 µL
ddH_2O		add to 20 µL

Table 4.8.: LIC vector processing using a 30 µL reaction mix

Component	Stock concentration	Amount
Linearized pEC vector		450 ng (15 ng/µL)
T4 DNA polymerase buffer	10x	3 µL
dTTP	25 mM	3 µL
DTT	100 mM	1.5 µL
T4 DNA Pol. LIC qualified	2.5 U/µL	0.6 µL
ddH_2O		add to 30 µL

was incubated 10 min at RT and 2 µL of that mixture were transformed into competent *E. coli* cells (Section 4.2.1). Ten colonies were picked and tested for false positives via PCR, in which a single colony was used as template DNA. Two positive clones were amplified for plasmid purification using a plasmid DNA extraction kit (Qiagen, Venlo, The Netherlands) and the cloning success was additionally confirmed by DNA sequencing (Biochemistry Core Facility (MPIB) or Eurofins Genomics, Ebersberg, Germany).

Cloning in pCoofy Vectors

Amplified genes were cloned into pCoofy series of vectors generated by the Biochemistry Core Facility (MPIB) [155], which were designed for Sequence and Ligation-Independent Cloning (SLIC). SLIC cloning represents a parallel PCR-based method and uses the principle of homologous recombination and single-strand annealing, which is catalyzed by RecA [156].

For cloning, the pCoofy vector was linearized by PCR as described in Section 4.2.1 using oligonucleotides with specific SLIC overhangs. The gene of interest was also amplified via PCR using oligonucleotides with complementary overhangs to the linearized vector. After PCR, both products were purified via gel-extraction (Qiagen) and used

for the SLIC reaction as described in Table 4.9. The reaction mixture was incubated for 30 min at 37 °C and transformed into competent *E. coli* cells (Section 4.2.1). The clones were tested for false positives as described in Section 4.2.1.

Table 4.9.: SLIC reaction

Component	Stock concentration	Amount
Linearized pCoofy vector		100 ng
PCR product		300 ng
RecA buffer	10x	1 µL
RecA	2 µg/mL	1 µL
ddH_2O		add to 10 µL

Transformation of Competent Cells

Transformation-competent cells were prepared as described in the Qiagen *E. coli* Handbook and stored at -80 °C until further use. For transformation, cells were thawed on ice and ∼100 ng of plasmid were added to 50 µL *E. coli* cells. The mixture was incubated for 30 min on ice, transformed at 42 °C for 45 sec and chilled for 2 min on ice. After adding 500 µL of SOC medium, the cells were shaken at 37 °C for 80 min at 1000 rpm. The bacteria were plated on LB agar containing the appropriate antibiotic (Table 4.1.5) and incubated over night at 37 °C.

Baculovirus Stocks

All baculovirus constructs were cloned into pFastBac vectors (here the term pCoofy is used for pFastBac) expressing the gene of interest under the control of the *Autographa californica* multiple nuclear polyhedrosis virus polyhedrin promoter, which is highly induced in insect cells. DH10Bac *E. coli* cells were used as host for the pCoofy vector to generate recombinant bacmid DNA, which was then transfected into insect cells to produce the first recombinant baculovirus. After amplification of the virus, the stock was used to infect insect cells for large-scale expression of the recombinant protein.

Recombinant bacmids and baculovirus stocks were generated as described in the Bac-to-Bac manual (Invitrogen, Darmstadt, Germany). Briefly, after isolation of the recombinant bacmid DNA, Sf9 or Sf21 cells ware transfected with the bacmid DNA using Cellfectin II Reagent (Invitrogen). Baculoviral particles were released after transfection into the medium by cell lysis and were harvested from the supernatant. This baculovirus stock (P1) was used to amplify the viral stock and to obtain higher titer viruses (P2, P3) (Table 4.10).

To define the optimal amount of P3 virus for recombinant protein expression, 3 mL High Five cells (1x10^6 cells/mL) were infected with varying amounts of P3 at 26 °C and 130 rpm in 25 mL tubes (Sarstedt, Nümbrecht, Germany) as described in Table 4.11. After 72 h, 400 µL of the cultures were harvested by centrifugation, the pellets were dissolved in 125 µL 1 x SDS loading dye and analyzed on SDS-gels.

Table 4.10.: Generation of recombinant baculovirus

Final viral stock	Virus origin	Cells	Cell culture	Amplification procedure
P1	Transfected cells	Sf9 or Sf21	2mL/well, 0.4x10^6 cells/mL	Adherent cells, 72 h, 26 °C
P2	1mL P1	Sf9 or Sf21	10 mL, 1x10^6 cells/mL	Cell suspension, 96 h, 85 rpm, 26 °C
P3	0.5mL P2	Sf9 or Sf21	500 mL, 0.4x10^6 cells/mL	Cell suspension, 72 h, 85 rpm, 26 °C

Table 4.11.: Defining optimal virus amounts

Condition	P3 virus	Amount [µL]
1	1:100 dilution	15
2	1:100 dilution	150
3	non-diluted	15
4	non-diluted	75
5	non-diluted	150
6	non-diluted	300

4.2.2. Protein Expression

Expression in E. coli

Large scale protein expression was typically carried out in 2.5 L Tunair shake flasks (Sigma-Aldrich, Germany) containing 1 L TB medium supplemented with appropriate antibiotics (Table 4.1.5) and was inoculated using a small scale pre-culture (1:100 dilution). For the pre-culture, 100 mL LB medium with the appropriate antibiotics were inoculated with colonies of freshly transformed *E. coli* cells and grown for 3 h at 37 °C and 220 rpm.

Inoculated large scale bacterial cultures were cultivated at 37 °C and 220 rpm until they reached an optical density of 1.8-2.2. The temperature was then reduced to 18 °C and protein expression was induced with 0.1 mM Isopropyl-β-D-thiogalactopyranoside (IPTG) for 16-18 h. For RPA70 DBD-AB, the expression was induced with 1 mM IPTG for 4 h at 37 °C. The cells were harvested by centrifugation (8000 x g, 10 min) and kept at -80 °C until further use.

Small scale protein expression was performed to analyze expression and solubility

of the designed constructs (Section 4.2.3). Typically, the expression was carried as described for large scale expression with the exception, that 200 mL Erlenmeyer flasks containing 20 mL TB medium were used.

Table 4.12.: Conditions for large scale expression

Expressed construct	*E. coli* strain	Condition
Tim-Tipin	BL21 (DE3) pRARE	0.1 mM IPTG, 16-18 h, 18 °C
Tipin	BL21 (DE3) pRARE	0.1 mM IPTG, 16-18 h, 18 °C
C-terminal Tim	BL21 (DE3) pRIL	0.1 mM IPTG, 16-18 h, 18 °C
RPA70/32/14,	BL21 (DE3) gold	0.1 mM IPTG, 16-18 h, 18 °C
RPA70FL/32FL/14	BL21 (DE3) gold	0.1 mM IPTG, 16-18 h, 18 °C
RPA70C/32/14	BL21 (DE3) gold	0.1 mM IPTG, 16-18 h, 18 °C
RPA32/14	BL21 (DE3) gold	0.1 mM IPTG, 16-18 h, 18 °C
RPA32WH	BL21 (DE3) pLysS	0.1 mM IPTG, 16-18 h, 18 °C
RPA70AB	BL21 (DE3) pLysS	1 mM IPTG, 4 h, 37 °C

Expression in Insect Cells

Protein expression in insect cells was carried out in 3 L Fernbach flasks (Corning B.V. Life Sciences, Amsterdam, The Netherlands) using 500 mL High Five cell suspension with 1x10^6 cells/mL infected with ∼5 mL P3 virus. The cells were grown for 72 h at 26 °C under constant agitation (85 rpm) and were harvested by centrifugation (2000 x g, 15 min). The cell pellets were frozen in liquid nitrogen and kept at -80 °C until further use.

4.2.3. Protein Purification

All protein purification steps were carried out at 4 °C. The buffers used during purification are detailed in Table 4.1.10. Typically, loading of the lysate was performed using a peristaltic pump. All other purification steps were carried out on ÄKTA purification systems (GE Healthcare). The washing and elution steps were performed using buffer mixtures of buffer 1 and buffer 2 (e.g. His 1 and His 2, Q 1 and Q 2). The buffer mixture is noted in percentage.

Pull-down Experiments for Expression and Solubility Analysis

Pull-down experiments were performed for all *E. coli* expressed constructs in this thesis to analyze expression and solubility. Initially, the same amount of *E. coli* cells was harvested (corresponding to 20 mL of a culture with an optical density of 1), resuspended in pull-down buffer 1 (PD 1) and sonicated briefly (Bandelin Sonoplus, tip VS70T, pulse ON/OFF 0.2/0.8, 20% amplitude, Bandelin electronic, Berlin, Germany). The

cell lysate was clarified by centrifugation (21000 x g, 30 min) and the supernatant was incubated with 50 µL Glutathione sepharose 4B slurry (GE Healthcare, München, Germany) or Ni-NTA agarose (Qiagen) (50% v/v) for 1 h at 4 °C under constant agitation. After centrifugation (500 x g, 2 min), the supernatant was discarded and the beads were washed three times using 500 µL PD 1. The protein was eluted using 25 µL PD 1 supplemented with either 400 mM imidazole (Ni-beads) or 20 mM Glutathion (GSH beads) and shaking for 5 min at 30 °C. The collected eluate was concentrated in a vacuum centrifuge, mixed with 2 x SDS loading dye and loaded on SDS-gels.

Purification of Tim-Tipin

The pellet of recombinantly co-expressed His-GST-Tipin (aa 1-278) and His-Tim (aa 1-1134) was resuspended in lysis buffer 1 (Lys 1) supplemented with 1 mM phenylmethanesulfonylfluoride (PMSF) and 3 mg DNase I (Roche Diagnostics, Mannheim, Germany). The resuspended cells were lysed by sonication (Bandelin Sonoplus, tip MS72, pulse ON/OFF 0.5/0.5, 40% amplitude, 15 min) and the lysate was clarified by centrifugation (75000 x g, 50 min). The soluble fraction was loaded on Ni-NTA-sepharose (2 x 5 mL HisTrap FF column, GE Healthcare) and washed with His-buffer 1 (His 1) followed by a high-salt (HS 1) and ATP-wash (ATPW 1). The immobilized protein complex was further washed with 15% His-buffer 2 (His 2) (100 mM imidazole) and eluted by an imidazole gradient to 500 mM. The eluate was desalted using a Sephadex G-25 Fine desalting column (GE Healthcare) to Q-buffer 1 (Q 1), loaded on a Q-sepharose column (5 mL HiTrap Q HP, GE Healthcare), washed with 10% Q-buffer 2 (Q 2) (250 mM NaCl), and eluted using a gradient to 50% Q 2 (650 mM NaCl). His-tag and GST-tag were cleaved with 3C protease (final concentration 0.006 mg/ml) over night at 4 °C during dialysis against dialysis buffer 1 (Dia 1). The cleavage mixture was concentrated by ultra-filtration and subjected to SEC (Superdex 200 16/60 column, GE Healthcare) pre-equilibrated with SEC buffer 1 (SEC 1). For GST-pull-downs, the GST-tag was left uncleaved. Finally, the purified protein was concentrated by ultra-filtration, flash frozen in liquid nitrogen and stored at -80 °C.

Purification of Tipin

The purification protocol of His-GST-Tipin (aa 1-278) is similar to the protocol described for the Tim-Tipin complex (Section 4.2.3). Briefly, the protein was expressed in *E. coli*. After cell lysis in Lys 1 and centrifugation, the clarified lysate was loaded on Ni-NTA-Sepharose (2 x 5 mL HisTrap FF column, GE Healthcare), washed with 10%

His-buffer 2 (His 2) (75 mM imidazole), and eluted by an imidazole gradient to 500 mM. The eluate was dialysed over night against buffer Q 1 and the His-GST-tag was removed by adding 3C protease (final concentration 0.006 mg/ml). The cleavage mixture was loaded on a Q-sepharose column (5 mL HiTrap Q HP, GE Healthcare), washed with 5% Q 2 (200 mM NaCl), and eluted from the anion-exchanger using a gradient to 65% Q 2 (800 mM NaCl). The protein was further subjected to SEC (Superdex 200 16/60 column, GE Healthcare) pre-equilibrated with SEC buffer 1 (SEC 1). Finally, the purified protein was concentrated by ultra-filtration, flash frozen in liquid nitrogen and stored at -80 °C.

Purification of RPA70/32/14

The resuspended cell pellet of recombinantly co-expressed His-RPA70 (aa 190-623), His-RPA32 (aa 43-270) and His-RPA14 (aa 1-121) was lysed by sonication (Bandelin Sonoplus, tip MS72, pulse ON/OFF 0.5/0.5, 40% amplitude, 15 min) in Lys 1 supplemented with 1 mM PMSF, 3 mg DNaseI (Roche), 100 mM Urea and two tablets of EDTA-free protease inhibitor cocktail (Roche). After centrifugation (75000 x g, 50 min), the clarified fraction was loaded on Ni-NTA-Sepharose (2 x 5 mL HisTrap FF column, GE Healthcare) and the resin was first washed with His-buffer 1 (His 1), then with His 1 either supplemented with high salt (HS 2, 3.2 M NaCl; HS 3, 2.5 M KCl) or ATP (ATPW 1), and finally with 10% His 2 (75 mM imidazole) prior to the elution. The His-tagged complex was eluted by an imidazole gradient to 500 mM and the eluate was desalted using a Sephadex G-25 Fine desalting column (GE Healthcare) to Q-buffer 3 (Q 3). The protein was further applied on a Q sepharose column (HiTrap Q HP, GE Healthcare), washed with 4% Q-buffer 4 (Q 4) (190 mM NaCl), and eluted using a linear gradient to 20% Q 4 (350 mM NaCl). Finally, the complex was purified by SEC (Superdex 200 16/60 column, GE Healthcare) pre-equilibrated with SEC buffer 2 (SEC 2), concentrated by ultra-filtration, flash frozen, and stored at -80 °C.

Purification of RPA70FL/32FL/14

The purification protocol of co-expressed His-RPA70FL (aa 1-623), His-RPA32FL (aa 1-270) and His-RPA14 (aa 1-121) complex was the same as outline for RPA70/32/14 (Section 4.2.3) with the exception that buffers with higher pH values were used (Table 4.1.10), because of an increased isoelectric point (pI 6.2) compared to RPA70/32/14 (pI 5.8).

Purification of RPA70C/32/14

The co-expressed protein complex consisting of His-RPA70C (aa 445-623), His-RPA32 (aa 43-270) and His-RPA14 (aa 1-121) was purified using a similar protocol detailed for RPA70/32/14 (Section 4.2.3). Briefly, the resuspended cells were lysed by sonication in Lys 3 and the proteins were immobilized by IMAC followed by a high-salt and ATP-wash step (His 5, HS 4 and ATPW 2). The column was further washed with 10% His 6 (60 mM imidazole) prior to elution of the protein using a gradient to 500 mM imidazole (100% His 6). After desalting to low-salt buffer (Q 7), the protein was applied on a Q sepharose column (GE Healthcare) and eluted by a gradient to 20% Q 8 (300 mM NaCl). Finally, the protein was purified by SEC (Superdex 200 16/60 column, GE Healthcare) using buffer SEC 3, concentrated, flash frozen, and stored at -80 °C .

Purification of RPA32/14

For the purification of co-expressed His-RPA32 (aa 43-270) and His-RPA14 (aa 1-121), the resuspended cells were lysed in Lys 1 and the clarified cell lysate was loaded on Ni-NTA-sepharose followed by a high-salt and ATW-wash (HS 2, ATPW 1). After washing with 10% His 1 (75 mM imidazole) and elution by an imidazole gradient to 500 mM imidazole, the complex was desalted to low-salt buffer (Q 3). The protein was applied on Q sepharose column (HiTrap Q HP, GE Healthcare), washed with 5% Q 4 (200 mM NaCl), and eluted by a linear gradient to 100% Q 4 (1 M NaCl). The His-tag was removed by incubation with 3C protease (final concentration 0.006 mg/ml) over night at 4 °C and the protein complex was further purified by SEC (Superdex 200 16/60 column, GE Healthcare, buffer SEC 3), concentrated, flash frozen, and stored at -80 °C.

Purification of RPA32WH

For the purification of His-GST-RPA32WH (aa 173-270), resuspended cells were lysed by sonication in Lys 4 and bound to GSH sepharose (GE Healthcare). After washing with Lys 4 supplemented with 1 M NaCl and elution by Lys 4 supplemented with 20 mM Glutathion, the His-GST-tag was cleaved with 3C protease (final concentration 0.006 mg/ml) during over night dialysis against low salt buffer (Q 9). The protein was further passed through a Q-sepharose column. The flow through after Q-column containing RPA32WH was further purified by SEC (Superdex 200 10/300 GL column, GE Healthcare) using SEC 4, concentrated, flash frozen, and stored at -80 °C.

Purification of RPA70AB

Resuspended cells of over-expressed His-RPA70AB (aa 190-431) were lysed by sonication in Lys 5. The protein was immobilized on Ni-NTA-Sepharose (2 x 5 mL HisTrap FF column, GE Healthcare), washed with His 5, high-salt (HS 5), 5% His 6 (35 mM imidazole), and eluted by a gradient to 500 mM imidazole. After dialysis against low-salt buffer (Q 10) and His-tag cleavage using 3C protease (final concentration 0.006 mg/ml) over night, the protein was applied on a Q sepharose column (HiTrap Q HP, GE Healthcare), washed with 4% Q 11 (190 mM NaCl), and eluted by a linear gradient to 20% Q 11 (350 mM NaCl). Finally, the protein was subjected to SEC (Superdex 200 10/300 GL column, GE Healthcare) using SEC 4, concentrated, flash frozen, and stored at -80 °C.

Purification of Tim

His-Tim (aa 1-1197) was expressed in insect cell. The cells were osmomechanically lysed after resuspending the pellet in RSB buffer (7.5 mL/1 g cells) supplemented with 1 mM PMSF and three tablets of EDTA-free protease inhibitor cocktail (Roche). The crude lysate was homogenized using a Dounce homogenizer (Thermo Scientific, Epson, UK) by moving the pestle 20 times up and down. The lysate was adjusted to 500 mM NaCl and 40 mM imidazole and centrifuged (3220 x g, 15 min). The supernatant was decanted and subjected to a second centrifugation (75000 x g, 50 min). The lysate was applied on Ni-NTA-sepharose (2 x 1 mL HisTrap FF column, GE Healthcare), washed with His 7, His 7 supplemented with 1 M NaCl, and 7% His 8 (75 mM imidazole) prior to elution by a gradient to 100% His 8 (500 mM imidazole). The His-tag was cleaved using 3C protease (final concentration 0.006 mg/ml) during over night dialysis against Q 12 and was then applied on a Q sepharose column (HiTrap Q HP, GE Healthcare). Bound Tim was eluted using a gradient to 100% Q 13 (1 M NaCl) and subjected to SEC (Superdex 200 10/300 GL column, GE Healthcare) using SEC 5. Finally, the protein was concentrated, flash frozen, and stored at -80 °C.

Purification of C-terminal Tim Constructs

This purification protocol applies to all C-terminal Tim constructs. His-GST-tagged C-terminal Tim fragments were expressed in *E. coli*. The cell pellets were resuspended in Lys 4 supplemented with 3 mg DNase I (Roche) and passed twice through an Emulsiflex C3 homogenizer (Avestin, Mannheim, Germany) at a pressure of 15000-20000 psi. After

centrifugation (75000 x g, 50 min) the soluble fraction was applied on GSH sepharose (GE Healthcare) and the resin was washed with Lys 4 and Lys 4 supplemented with 1 M NaCl. Elution was carried out using Lys 4 supplemented with 20 mM Glutathion. The His-GST-tag was cleaved with 3C protease (final concentration 0.006 mg/ml), which was directly applied on the column. Both ends of the column were connected, the flow rate was reduced to 0.3 mL/min, and the on-column tag cleavage was performed over night. The flow through was desalted to Q 14 and applied on a Q sepharose column (HiTrap Q HP, GE Healthcare). The bound fraction was washed with Q 14 and eluted by a gradient to 100% Q 15 (1 M NaCl). Finally, SEC was performed using a Superdex 75 16/60 column (GE Healthcare) pre-equilibrated with SEC 4. The purified protein was concentrated, flash frozen, and stored at -80 °C.

4.2.4. Protein Biochemistry

Determination of Protein Concentration and Purity

Protein concentrations were determined by absorbance at 280 nm using the NanoDrop spectrophotometer (Thermo Scientific, Epson, UK). The extinction coefficients were calculated using VectorNTI (Life Technologies, Darmstadt, Germany). The purity of proteins was assessed by monitoring the absorbance at 280 nm and 260 nm during ion exchange chromatography and SEC, the 260/280 ratio, and SDS-PAGE.

Polyacrylamide Gel Electrophoresis

Proteins were analyzed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) using 12%-18% SDS-gels as described by Laemmli, 1970 [157] or using commercially available pre-cast SDS-gels (4-12% Bis-Tris in 1 x MES buffer, Invitrogen). SDS-gels were stained with Coomassie brilliant blue R-250.

SDS-gels loaded with Cy3-labeled RPA (Section 4.2.5) were first analyzed by fluorescence imaging (excitation 532 nm, filter 580 nm, Typhoon FL 7000 phosphoimager, GE Healthcare) before Coomassie staining.

For silver staining, the gels were soaked in fixing solution for 10 min and washed two times in ddH_2O for 5 min. For sensitization, the gels were soaked in 0.02% $Na_2S_2O_3$ for 1 min. After washing two times in ddH_20, the gels were soaked in fresh 0.1% $AgNO_3$ solution for 10 min in the dark, rinsed quickly with ddH_20, and with a small volume of developing solution. The gels were further soaked in fresh developing solution until bands appeared. The developing solution was discarded and stop solution was added to

block development. This protocol was provided by Dr. Christian Biertümpfel (MPIB).

Table 4.13.: Silver stain solutions

Solution	Component	Stock concentration	Final concentration
Fixing solution	Methanol	100%	40%
	Formalin	100%	13.5%
Developing solution	Na_2CO_3	powder	3%
	Formalin	100%	0.05%
	$Na_2S_2O_3$	powder	0.000016%
Stop solution	Ethanol	100%	50%
	HOAc	100%	15%

Native Gel Electrophoresis

To assess the homogeneity of individual proteins and protein complexes and for electromobility shift assays (EMSA), commercially available native 4-12% Tris-Glycine gels (Invitrogen, Darmstadt, Germany) were used. Note that 50% glycerol with bromphenol blue was used as 6 x native loading dye.

For EMSA, the native gels were first pre-equilibrated for 90 min in 1 x Tris-Glycine running buffer pH 8.3 (Invitrogen) at 125 V and 4 °C. After loading the samples, electrophoresis was performed at 4 °C and run at 90 V. Before Coomassie staining, the gels were analyzed by fluorescence imaging (Section 4.2.5).

For homogeneity analysis, the native gels were used without pre-equilibration. Electrophoresis was performed at 4 °C and run at 125 V for 2.5 h. Typically 3 µg were loaded per well. The native gels were stained with Coomassie. Native gels loaded with Cy3-labeled RPA (Section 4.2.5) were first analyzed by fluorescence imaging (excitation 532 nm, filter 580 nm, Typhoon FL 7000 phosphoimager, GE Healthcare) before Coomassie staining.

Immunoblot

For immunoblot analysis, samples were separated on native 4-12% Tris-Glycine gels as described in Section 4.2.4. After electrophoresis, the gels were incubated in 0.1% SDS for 15 min and the proteins were transferred onto 0.2 µm polyvinylidene fluoride (PVDF) membranes (Merck Millipore, Darmstadt, Germany) for 80 min at 200 mA using a transfer buffer consisting of 20 mM Tris, 150 mM glycine, 5% methanol, and 0.02% SDS. To fix the transferred proteins, the membranes were incubated in 10% HOAc for 15 min and air dried. After rehydration in 1 x Tris-buffered saline (TBS, 50 mM Tris pH 7.5, 150 mM NaCl) the membranes were blocked in 10% non-fat dry milk in 1 x TBS

for 1 h at RT, washed two times with TBS containing 0.05% Tween 20 (TBS-T), and once with TBS. The membranes were incubated with a primary antibody in 3% BSA in 1 x TBS for 1 h at RT. After two times washing with TBS-T and once with TBS, the membranes were further incubated with a horseradish peroxidase-conjugated secondary antibody for 1 h at RT, subsequently washed two times with TBS-T, and once with TBS. Bound antibodies were detected by chemiluminescence using ECL reagents (GE Healthcare) and an ImageQuant LAS 4000 imager (GE Healthcare).

Limited Proteolysis

For limited proteolysis, 10 µL protein was diluted to 0.6 mg/mL in SEC 2, mixed with 3 µL protease solution and incubated for 30 min on ice. Trypsin, Subtilisin, Elastase, GluC and Chymotrypsin were diluted from a 1 mg/ml stock using 20 mM Hepes pH 7.5, 50 mM NaCl and 10 mM $MgSO_4$ and were used at three different concentrations (0.1 mg/mL, 0.01 mg/mL, 0.001 mg/mL). The reaction was stopped using 1 µL 100 mM AEBSF (4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride). The degradation products were separated by SDS-PAGE and stained with Coomassie for visualization or peptide fingerprinting. The samples were also transferred to PVDF membranes for N-terminal sequencing. To determine the total mass of the fragments, the stopped reaction was analyzed by ESI-TOF mass spectrometry (MS). Peptide finger printing was performed by Dr. Cyril Boulègue or Nagarjuna Nagaraj, PhD (MPI Biochemistry Core Facility) (Section4.2.4). N-terminal Sequencing was carried out by Isabella Mathes, Dr. Reinhard Mentele or Dr. Josef Kellermann (MPI Biochemistry Core Facility) and ESI-TOF MS by Elisabeth Weyher-Stingl (MPI Biochemistry Core Facility).

Mass Spectrometry Analysis

For mass spectrometry analysis protein sample bands from SDS-gels were cut out. Alternatively, a liquid protein sample was analyzed. The samples were digested with trypsin [158], the peptides were analyzed by Orbitrap mass spectrometry [159], and identified using the Max Quant software [150]. The stoichiometry of the protein complexes was determined by dividing the sum of all peptide peak intensities by the number of theoretically observable tryptic peptides (intensity-based absolute quantification (iBAQ)) [130].

Static Light Scattering

To determine the molecular weight and the monodispersity of a protein sample, static light scattering (SLS) was performed. During a SLS experiment, the sample passes through a size exclusion chromatography column, which is directly coupled to a detector monitoring light scattering (LS), ultraviolet absorbance (UV) and refractive index (RI). Specifically, a high pressure liquid chromatography system from Waters (Milform, MA, USA) (pump, ultraviolet (UV)) coupled to a TDA302 detector array (Viscotek, Malvern, Herrenberg, Germany) was used. Proteins samples (10 µL, 2 mg/mL) were injected onto a Superdex 200 5/150 column coupled to the detector. Bovine serum albumin was used as a standard and the refractive index increment (dn/dc) was set to 0.180 ml/g for calculations. Data analysis was performed using OmniSEC 4.5 software (Viscotek). SLS measurements were conducted either by Dr. Claire Basquin, Dr. Jörg Tittor at MPIB or by myself.

Sucrose-Gradient Sedimentation

Sucrose-gradient sedimentation was used to determine the molecular weight of proteins. Sucrose gradients (5-20% w/v) were prepared in centrifuge tubes (Seton open-top polyclear centrifuge tubes, 14 x 95 mm) in 50 mM Hepes pH 7.5 and 150 mM NaCl using the Gradient Station (Biocomp, Fredericton, NB, Canada). The protein sample at 1 mg/ml or at 3.5 mg/ml (for standard proteins mixture) was applied on top of the gradient solution together with a 300 µL buffering cushion (50 mM Hepes pH 7.5, 150 mM NaCl, 2.5% sucrose). The samples were sedimented at 4 °C in a swing out rotor (Beckmann SW40 rotor, Beckmann Coulter, Krefeld, Germany) for 18 h at 35000 rpm (217290 x g). After harvesting using the Gradient Fractionator (Biocomp, Fredericton, NB, Canada), the fractionated samples were analyzed by SDS-PAGE (4-12% Bis-Tris in 1 x MES, Invitrogen). For gradient calibration, standard proteins (Bio-Rad Laboratories) were used: Myoglobin (17 kDa), Ovalbumin (44 kDa), γ-globulin (158 kDa for the entire molecule of two light and two heavy chains) and Thyroglobulin (670 kDa for the entire homo-dimeric molecule).

4.2.5. Binding Experiments

GST-Pull-Down Experiments

Prior to GST-pull-down experiments, Glutathione sepharose was blocked to reduce unspecific protein binding. Thus, 667 µL of 75% Glutathione sepharose 4B slurry (GE

Healthcare) were washed with 900 µL of 20 mM Hepes pH 7.5, 150 mM NaCl, and 0.1% NP-40 (blocking buffer) and resuspended in 500 µL of 20 mM Hepes pH 7.5, 550 mM NaCl, 0.1% NP-40, 0.08 mg/mL glycogen carrier, 0.08 mg/mL tRNA, and 0.8 mg/mL bovine serum albumin. After two hours incubation at 4 °C, the beads were washed three times with 900 µL blocking buffer and stored at 4 °C.

All GST pull-down experiments were performed at 4 °C. The GST-Tim-Tipin complex (6 µg, 4.2 µM final concentration) was mixed with 2-fold molar excess of RPA (input sample) in 30 µL, supplemented with 2 x buffer to a total volume of 60 µL containing the final concentrations of 20 mM Hepes pH 7.5, 125 mM NaCl, 12.5% glycerol, 0.1% NP-40, 1 mM DTT, and 0.01 mM $Zn(OAc)_2$. After incubation at 30 °C for 20 min, 12 µL blocked Glutathione sepharose (GE Healthcare, 50% v/v) and 200 µL buffer P1 (20 mM Hepes pH 7.5, 125 mM NaCl, 12.5% glycerol, 0.1% NP-40, 1 mM DTT, 0.01 mM $Zn(OAc)_2$) were added and the proteins were immobilized for 1 h under constant rotation at 4 °C. The beads were washed three times with buffer P1. Bound proteins were eluted using 20 µL buffer P2 (30 mM Tris pH 8.8, 150 mM NaCl, 14% glycerol, 0.1% NP-40, 1 mM DTT, 0.01 mM $Zn(OAc)_2$, 2 mM imidazole, 20 mM Glutathione) and shaking for 5 min at 30 °C. The eluate was collected, concentrated in a vacuum centrifuge, mixed with 2 x SDS loading dye, analyzed on 18% SDS-polyacrylamide gels, and stained with Coomassie blue.

Electrophoretic Mobility Shift Assay

For EMSA analysis, protein samples were incubated with DNA substrates and the protein-DNA complexes were separated by native PAGE (Section 4.2.4). DNA substrates were fluorescently labeled at the 5'-end with fluorescein (5-FAM) and purchased from Purimex (Grebenstein, Germany) or Eurofins Genomics (Ebersberg, Germany). Annealed DNA substrates (spDNA, RFDNA, dsDNA) and DNA substrates named 'CB' were a kind gift of Dr. Christian Biertümpfel (MPIB).

Typically, 16 µM of protein was incubated with the indicated amounts of ssDNA substrate (64, 32, 16 or 4 µM). For EMSA using spDNA, RFDNA and dsDNA, the DNA substrate concentration was kept constant at 80 nM and the total protein concentration was varied (0.08, 0.24, 0.8, 4.8 µM). To keep the final buffer concentrations at 20 mM Hepes pH 7.5, 125 mM NaCl, 2 mM DTT, and 0.01 mM $Zn(OAc)_2$ (SEC 2), 7.5 x buffer was added and the mixture was supplemented with 6 x native loading dye. The samples were incubated 15 min at RT and 15 min at 4 °C prior to separation by native PAGE (Section 4.2.4). Gels were analyzed by fluorescence imaging (excitation 473 nm, filter

520 nm, Typhoon FL 7000 phosphoimager, GE Healthcare) followed by Coomassie staining.

Table 4.14.: DNA substrates for EMSA

DNA substrate	Name	Sequence
8 nt ssDNA	8merFAM	5'-FAM-ATCCCTAA-3'
14 nt ssDNA	14merFAM	5'-FAM-GACGGCATCCCTAA-3'
30 nt ssDNA	CB209short	5'-FAM-ACGCTGCCGAATTCTACCAGTGCCTTGCTA-3'
60 nt ssDNA	CB209	5'-FAM-ACGCTGCCGAATTCTACCAGTGCCTTGCTA-GGACATCTTTGCCCACCTGCAGGTTCACCC-3'
spDNA	CB209	5'-FAM-ACGCTGCCGAATTCTACCAGTGCCTTGCTA-GGACATCTTTGCCCACCTGCAGGTTCACCC-3'
	CB212	5'-CGATAGTCGGATCCTCTAGACAGCTCCATGAGCA-AGGCACTGGTAGAATTCGGCAGCGT-3'
RFDNA	CB209	see spDNA
	CB212	see spDNA
	CB217	5'-CATGGAGCTGTCTAGAGGATCCGACTATCG-3'
	CB218	5'-GGGTGAACCTGCAGGTGGGCAAAGATGTCC-3'
dsDNA	CB209	see spDNA
	CB215	5'-GGGTGAACCTGCAGGTGGGCAAAGATGTCCTA-GCAAGGCACTGGTAGAATTCGGCAGCGT-3'

Fluorescence Anisotropy Measurements

For Fluorescence anisotropy measurements, 5'-FAM-labeled DNA substrates were used (Table 4.14). The DNA concentration was kept constant at 10 nM, while the protein concentrations varied between 1 nM and 10 µM. FA measurements were performed on a Genios Pro (Tecan, Männedorf, Switzerland). The protein-DNA mixture was incubated for 15 min at 30 °C in SEC 2 in a total volume of 50 µL before anisotropy reading. The excitation and emission wavelengths were 485 nm and 535 nm. Each titration point was carried out three times using 10 reads with an integration time of 40 µs. The K_D was calculated by directly fitting the curve using the Hill function (Equation 4.1) included in the program Origin 8.1 (Origin-Lab, Northampton, MA, USA),

$$A_{total} = A_0 + (A_{max} - A_0)\frac{x^n}{k^n + x^n} \tag{4.1}$$

where A_{total} is measured anisotropy, A_0 is the intrinsic anisotropy of the DNA substrate, A_{max} is the anisotropy of the saturated protein-DNA complex, n is the Hill coefficient and x is the concentration of the protein. To assess the quality of the regression, residual plot analysis was performed and the agreement between the observed and calculated binding isotherms was evaluated.

SEC-Based Protein-Protein Binding Experiments

For preparative formation of Tim-Tipin-RPA complexes, Tim-Tipin was mixed with RPA at a 1:1.5 (Tim-Tipin:RPA) molar ratio in a final volume of 200 µL. Typically ∼45 µM Tim-Tipin and ∼68 µM RPA (final concentrations) were mixed and supplemented with SEC 2 (Table 4.1.10) to the final volume. After 15 min incubation at RT and 15 min on ice, the mixture was injected on a Superdex 200 10/300 column equilibrated with SEC 2.

For analytical formation of Tim-Tipin-RPA complexes, Tim-Tipin was mixed with RPA at a 1:1 (16 µM:16 µM, ∼65 µg Tim-Tipin and ∼38.5 µg RPA) or 1:2 (16 µM:32 µM, ∼65 µg Tim-Tipin and ∼77 µg RPA) molar ratio in 25 µL final volume. The mixture was incubated as described above and loaded on a Superose 6 3.2/PC column on ÄKTAmicro (GE Healthcare) equilibrated with SEC 2. For the formation of Tim-Tipin-RPA32WH complexes, the molar amount of used protein mixture was increased to give final concentrations of 64 µM:64 µM and 64 µM:128 µM (Tim-Tipin:RPA32WH) because of low content of tyrosine or tryptophan in RPA32WH and resulting low absorbance at 280 nm.

SEC-Based Protein-ssDNA Binding Experiments

For analytical complex formation analysis of Tim-Tipin, RPA, and Tim-Tipin-RPA with ssDNA substrates, the protein sample was mixed with excess or sub-stoichiometirc amounts of ssDNA. The final concentrations were 16 µM:32 µM and 32 µM:8 µM (protein:ssDNA) for 60 nt ssDNA and 16 µM:32 µM and 16 µM:8 µM (protein:ssDNA) for 31 nt and 14 nt ssDNA. For SEC analysis non-labeled ssDNA substrates were used (Table 4.15). To estimate if ssDNA was bound to the protein, the ratio of the absorption at 260 nm and 280 nm was calculated and compared to the ratio of purified DNA-free proteins.

Table 4.15.: DNA substrates for SEC analysis

ssDNA substrate	Sequence
14 nt	5'-AGAAGAGCCAAAAC-3'
31 nt	5'-CGGGATCCCAGCCAGCGATGTCTCAAGCTGC-3'
60 nt	5'-CAGACCGCCACCGACTGCTTAGATATTTAA-GTTTTCTAATTTTTCATTGAAAGCATTAAG-3'

Calibration of SEC Columns

All SEC columns used in this thesis (Superdex 200 16/60, Superdex 75 16/60, Superdex 200 10/300 GL, Superose 6 3.2/PC, GE Healthcare) were calibrated using molec-

ular mass standard proteins (Bio-Rad Laboratories, Hercules, CA, USA): Vitamin B_{12} (1.35 kDa), Myoglobin (17 kDa), Ovalbumin (44 kDa), γ-globulin (158 kDa) and Thyroglobulin (670 kDa).

Fluorescent Labeling and Microscale Thermophoresis

To determine the binding affinity of Tim-Tipin to RPA, microscale thermophoresis was performed. MST measures the motion of molecules along an infrared (IR) laser induced temperature gradient. This thermophoretic movement depends on the hydration shell, the charge or size of the molecule and changes due to alterations in the structure upon binding of an interaction partner. The motion of the protein can be monitored by the fluorescence of an attached dye.

For MST measurements, RPA was fluorescently labeled using the Cy3 protein labeling kit according to manufacturer's protocol (Jena Bioscience, Jena, Germany). Briefly, 6 µL of purified RPA at 11 mg/mL (final concentration, 31 µM, 3 mg/mL) were mixed with 2 µL of 1 M Na_2CO_3 (100 mM final concentration) in a total volume of 20 µL. The pH was raised to an optimal value of 8.3-8.5 by the addition of Na_2CO_3. The supplied dye was diluted to the same concentration like RPA (3 mg/mL) and 2 µL of that dilution were added to the RPA solution. The mixture was incubated for 1 h in the dark and the conjugate was purified from unbound dye using buffer exchange spin columns (Micro Bio-Spin 6 columns, Bio-Rad) equilibrated with SEC 2.

The ideal degree of labeling (DOL) should correspond to approximately one fluorophore per 200 aa and was determined to be 4.4/858 aa ($\hat{=}$ 1/200 aa) for RPA. The average number of lysines labeled per RPA was calculated by Equation 4.2,

$$DOL = \frac{A_{max} * \epsilon_{280}(RPA)}{(A_{280} - A_{max} * CF) * \epsilon_{max}} \tag{4.2}$$

where A_{max} is the absorbance of the conjugate solution (Cy3-RPA) measured at 550 nm using NanoDrop (Thermo Scientific), ϵ_{280} is the extinction coefficient of RPA (ϵ_{280} = 69670 $cm^{-1}M^{-1}$), A_{280} is the absorbance of the conjugate solution measured at 280 nm, CF (correction factor), and ϵ_{max} (extinction coefficient of Cy3) are the intrinsic properties of the dye (CF = 0.08, ϵ_{max} = 150000 $cm^{-1}M^{-1}$).

For MST measurements, 100 nM Cy3-RPA were titrated with varying amounts of non-labeled Tim-Tipin (25-80550 nM) in SEC 2 and incubated for 10 min at RT. The thermophoresis measurements were performed at 50% LED and 65% IR-Laser power using a Monolith NT.115 (NanoTemper, München, Germany). Laser-On time was 40 s,

and Laser-Off time was 5 s. The experiments were repeated three times.

The results were analyzed and processed using Origin 8.1 (OriginLab). The baseline was subtracted from each individual experiment and each individual repetition was divided by its amplitude (the difference between the bound and unbound state). Three measurements were averaged and the standard deviation was determined. The K_D was calculated by directly fitting the curve using the Hill function (Equation 4.1) of Origin.

4.2.6. EM Structure Determination

Sample Preparation Using GraFix

Gradient Fixation (GraFix) is a method of mild intramolecular fixation and concurrent purification of a protein sample in a density gradient for EM analysis. This method is used to stabilize macromolecular complexes and to reduce heterogeneity of the sample due to particle dissociation during sample preparation on EM grids. GraFix was performed as described [129]. Briefly, sucrose gradients (5-30% (w/v) for Tim-Tipin-RPA, 5-20% (w/v) for Tim-Tipin and RPA) combined with a glutaraldehyde gradient (cross-linker, 0-0.2% (v/v)) were prepared in 50 mM Hepes pH 7.5, 150 mM NaCl using a standard gradient mixer (MPI workshop) and filled into centrifuge tubes (Beckmann, 50 Ultra-Clear Tubes, 14 x 95 mm). A cushion solution (300 µL) consisting of 50 mM Hepes pH 7.5, 150 mM NaCl, and 2.5% sucrose was applied on top of the gradient followed by the protein sample (200 µL, 1 mg/mL). Cross-linking was carried out while sedimenting the sample at 12 °C in a swing out rotor (Beckmann SW40 rotor, Beckmann Coulter, Krefeld, Germany) for 16 h at 35000 rpm (217290 x g). In parallel, the same sample was sedimented in a sucrose gradient without the cross-linker and was used as control. Samples were harvested from the bottom of the tube and glutaraldehyde was neutralized by adding 1 M Tris to a final concentration of 80 mM. The fractions were analyzed by SDS-PAGE (4-12% Bis-Tris in 1 x MES, Invitrogen). This sample was directly used for negative stain EM analysis. For cryo-EM observation, sucrose was removed using a Zeba Spin Desalting Column (Thermo Scientific) and samples were concentrated to 0.3 mg/ml using a Vivaspin 500 ultrafiltration device (Sartorius Stedim Biotech, Göttingen, Germany). Mass spectrometry analysis was performed to identify the presence of all components in the cross-linked samples (Section 4.2.4).

Negative Stain EM

Negative-stain EM was used as a first step for 3D structure determination and prior to grid preparation for cryo-EM. For negative staining, carbon-coated copper grids were glow-discharged in a plasma cleaner chamber (Harrick Plasma, Ithaca, NY ,USA) for 45 sec. The cross-linked sample from GraFix preparation (4 µL) (Section 4.2.6) was applied onto the grids for 30 s. Excess solution was blotted off with filter paper and the grid was washed four times with ddH_2O before applying negative stain solution (1% (w/v) uranyl acetate). After blotting the grid was air-dried. Data were collected on a CM200-FEG (FEI, Eindhoven, The Netherlands) operated at 160 kV and 50000 x magnification in a defocus range between 2 µm and 4 µm using an Eagle CCD camera (FEI) with a 2.16 Å pixel size.

Cryo-EM

For cryo-EM, 5 µL of the cross-linked sample (Section 4.2.6) were applied to glow-discharged Quantifoil holey carbon grids and vitrified in either liquid ethane or liquid ethane/propane mixture using a vitrobot cryo-station (FEI). Data of the vitrified specimens were collected on a Tecnai F20 electron microscope (FEI) at 200 kV and 50000 x magnification with a GATAN 626 cryo-holder. Data (673 images) were recorded using an Eagle CCD camera (FEI) with a pixel size of 2.21 Å at the specimen level.

Image Processing

Data processing was done using BSOFT [147], SPIDER [152], EMAN [149] and SPARX [151]. Micrographs were visualized and single particles were selected using BSHOW and BPICK of the BSOFT software package. For negative stained images, 13311 particles for Tim-Tipin-RPA, 10872 particles for Tim-Tipin, 1673 particles for RPA, and 1914 particles for the non-cross-linked Tim-Tipin-RPA sample were selected. For cryo-EM analysis, 39679 particles of cross-linked Tim-Tipin-RPA were chosen. The particles were boxed out with the size of 128 pixels. Contrast transfer function (CTF) was estimated and corrected by flipping phases for the images of vitrified specimen using the program CTFFIND3 [148] and BCTF of the BSOFT software package. For the initial assessment of the particles, selected Tim-Tipin-RPA, Tim-Tipin and RPA particles were subjected to reference-free classification using BSOFT or SPARX leading to 74 class averages. The alignment and classification was iteratively optimized.

3 D maps were displayed using UCSF Chimera [132]. Further, maps were segmented

using the segment map option of the Segger package to estimate volumes of the sub-complexes [160]. Crystal structures were docked into the 3D map by manual fitting of the models using Chimera.

Initial Model Generation Using Random Conical Tilt

The initial 3D model of Tim-Tipin-RPA was obtained by the Random Conical Tilt (RCT) method [161]. For RCT processing, image pairs of tilted (45°) and untilted (0°) micrographs of negative stained Tim-Tipin-RPA were collected. 857 particle tilt pairs were selected and 8 initial models were generated as described in Radermacher, 1988 [161]. Two most similar models with prominent features were merged and used as the initial model. First, the RCT model was refined against the class averages of the negative stained data by comparing the reprojections of the RCT reconstruction to the class averages for verification. The resulting 3D map was further refined against the raw data of the untilted 13311 particles by iterative reprojection matching using the merged RCT reconstruction as a reference.

Cryo-EM Reconstruction

Reference-based cryo-EM reconstructions were performed using SPARX. The reconstruction of the negative stain data (Section 4.2.6) was used as initial reference and was low-pass filtered to 50 Å. For the first cycles of iterative projection matching, the data was binned by two (4.42 Å/pixel) and 10 iterations per one cycle of iterative projection matching refinement were performed using angular increments from 5 to 2 degrees. For the last cycle of iterative refinement, non-binned original data with 2.21 Å/pixel size was used and the final 3D reconstruction was performed with an angular step of 2 degrees over the course of 10 iterations.

The resolution was estimated by Fourier shell correlation (FSC) curves between reconstructions from halves of the data set [162]. The data set was split into two and two independent reconstructions were iteratively calculated using the RCT reconstructions filtered to 50 Å as reference. The final resolution was estimated to be 17.3 Å with the FSC = 0.5 criteria. The amplitudes of the final density maps were corrected by using the amplitudes from an atomic model with a similar size/protein density (PDB code: 2BR2, exosome core) using the 'bampweigh' command in BSOFT and the structure was low-pass filtered to 17 Å. The angular distribution of the images used for the refinement was assessed using 'sxplot projs distrib.py' command in SPARX.

Antibody Labeling

For the mapping of RPA70, 0.23 µM of Tim-Tipin-RPA were mixed with 0.077 µM of an antibody against RPA70 (sc-166023, Santa Cruz Biotechnology, Santa Cruz, CA, USA, Table 4.1.7) and incubated at RT for 30 min. For the mapping of Tipin, 0.54 µM of the Tim-Tipin-RPA complex were mixed with 1.25 µM of an antibody against Tipin (sc-160865, Santa Cruz Biotechnology, Table 4.1.7), incubated on ice for 1 h and injected on a Superose 6 3.2/PC column. The protein-antibody mixture or the eluted immune complex were negatively stained with uranyl acetate (1% w/v, Section 4.2.6). Images were collected as described in Section 4.2.6. Sixty-six particles for RPA70 and 347 particles for Tipin localization were selected with the size of 128 pixels. Image processing was performed as described in Section 4.2.6.

Segmentation of the Tim-Tipin-RPA 3D Map

To assign the electron density corresponding only to Tim-Tipin in the Tim-Tipin-RPA reconstruction, 3D models missing the putative RPA density were generated (Tim-Tipin-ΔRPA models). The negative stain Tim-Tipin-RPA 3D map was sub-divided in eleven segments using the segment map option in Chimera [132, 160, 163]. Next, the volumes of the eleven segments were determined and segments enclosing the RPA volume were merged given two constraints: the final merged volume must enclose the calculated volume of ∼1.2 e^5 $Å^3$ and the merged segments must be connected. RPA's volume (∼1.2 e^5 $Å^3$) was calculated using its molecular weight (96000 Da) and the specific volume of 0.81 $Å^3$/Da. The merged volumes corresponding to the putative RPA density were deleted from the Tim-Tipin-RPA reconstruction resulting in Tim-Tipin-ΔRPA models. Finally, the alignment of Tim-Tipin 2D class averages to the reprojections of the Tim-Tipin-ΔRPA models was performed using the multi-reference alignment script in SPIDER. The total number of the averages with the best reference-fit was plotted as bar chart.

4.2.7. Crystallization Procedures

Initial Screening

Initial crystallization experiments were performed in 96 well sitting drop plates using commercially available or self-made screens. Typically, 100 nL protein solution was mixed with 100 nL of the reservoir solution using a Phoenix nanodispenser robot (Art Robbins Instruments, Sunnyvale, CA, USA). Two temperatures (20 °C and 4 °C) were

tested for screening. Initial hits were optimized in 96 well sitting or 24 well hanging drop plates. The plates were automatically imaged using the imaging system of the crystallization facility (MPIB). All crystallization experiments were set up by Dr. Karina Valer-Saldaña or Sabine Pleyer from the crystallization facility (MPIB).

Cryo-Protection

For cryo-protection, crystals were transferred from the original drop to a 5 µL drop containing the original drop solution supplemented with 5% of the cryo-protectant (ethylene glycol). In the following, the concentration of the cryo-protectant in the drop was gradually increased to reach the final concentration of 25% ethylene glycol by successively adding and removing of a series of cryo-protectant solutions with increasing amounts of ethylene glycol. Finally, the crystals were picked up from the cryo-protectant using suitably sized nylon cryo-loops (Hampton Research, Aliso Viejo, CA, USA) and flash cooled by plunging into liquid nitrogen.

Data Collection and Processing

Data from crystals mounted in cryo-loops was collected at 100 °K at the Swiss Light Source (SLS, Villigen, Switzerland) using the PXII beamline and a MarCCD detector. The data was indexed using XDS [153]. The Matthews probability analysis was performed using the website http://www.ruppweb.org/mattprob/default.html.

Acknowledgments

The present thesis would be nothing without the support of my supervisors, colleagues, family, and friends.

First of all, I would like to thank Elena Conti. Thank you a lot for giving me the opportunity to perform my research in your department at the Max Planck Institute of Biochemistry in Martinsried and for your infinite support during my thesis. I esteem the professional and very productive environment in your research group, which is not given to every PhD student.

I would like to express my deepest gratitude to Naoko Mizuno, who was a great supervisor. Thank you for helping me to develop new ideas on my project and to realize them immediately. With your great help and your motivating support, I was able to push the project in a consequent and productive way to a very successful end.

I am also very grateful to Eva Wolf for supervising this thesis. Thank you for giving me the opportunity to work on this challenging research topic and thank you a lot for your trust and freedom.

I have benefited a lot from the resources and the infrastructure at the Max Planck Institute of Biochemistry and from the support of the people in the department of Elena Conti. First, I thank the people in the different service groups and specially the Core Facility of the MPIB for the great methods portfolio and relying analyses. Moreover, I would like to acknowledge all people from the Conti group for helpful suggestions, discussions, and hands-on assistance. Besides, I would like to thank Judith Ebert for her trust, Claire Basquin and Jörg Tittor for their biophysical expertise, Peter Reichelt for help with columns and purification systems, Christian Benda and Walter Erhardt for IT- and technical support, Karina Valer-Saldaña, Sabine Pleyer, and Jerôme Basquin for setting up and running such a powerful crystallization facility, Felix Halbach and Sebastian Falk for their far-reaching expertise and critical discussions, and Michaela Rode for support in the insect cell culture lab. I thank Debora Makino for the inspiring discussions about functional assays for RPA, Ajla Hrle for always being willing to help with any kind of problem, Marc Baumgärtner for helping out with aliquots, Gretel

Buchwald and Christian Benda for introduction to GraFix, and Fabien Bonneau for providing protocols for EMSA and pull-down experiments. I am much obliged to Petra Lee and Ulrike Goldschmitt for their help with organizational issues.

I greatly acknowledge Christian Biertümpfel and Shun-Hsiao Lee for sharing their DNA substrates with me, Kristina Weber for introducing me to MST, and Julia Adam for help with the EM software.

I would like to thank the lab members of Naoko Mizuno for their support in EM, Esben Lorentzen, Christian Biertümpfel, and their lab members for kindly sharing the equipment with me.

Thanks a lot to Carina Vraschek, Eva Kowalinski, Julia Adam, Kristina Weber, Melanie Vetter, Stefanie Wachter, Michael Taschner, and Tim Still for proofreading of my thesis.

Furthermore, I would like to thank Nicole Kucera, Helena Riffkin, and Sven Hennig. I have appreciated a lot their warm welcome at the beginning of my PhD and thanks a lot for the great atmosphere in our group at the Max Planck Institute of Molecular Physiology in Dortmund.

I thank Anna Czarna, Eva Kowalinski, and Andrea Murachelli. I benefited a lot from their support in the department and I enjoyed their company here in Munich. The same holds for Carina Vraschek.

My sincere thanks to J. Rajan Prabu for his infinite help and support inside and outside MPI. Finally, I am believing that it is forming a complex.

Dear Mela, Krissy, and Marc, I can't image the second floor without you.

Dear Ira, you have been a good colleague and became a very good friend in the past years. Gemeinsam sind wir stark.

Dear Janina, it all started at the retreat. Thank you for your patience and support when I was working long in the lab.

Zudem möchte ich meiner Familie für die andauernde liebevolle Unterstützung, für das Verständnis und die vielen Möglichkeiten, die sie mir eröffnet hat, danken.

Schließlich möchte ich mich bei meinem besten Freund und Partner bedanken. Vielen Dank, daß Du mich die letzen Jahre begleitet, in vieler Hinsicht unterstützt und mich immer wieder zum Lachen gebracht hast.

Abbreviations

2D	Two-dimensional
3D	Three-dimensional
aa	Amino acid
AEBSF	4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride
AEX	Anion-exchange chromatography
Arm	Armadillo
ATP	Adenosine triphosphate
ATR	Ataxia telangiectasia mutated and Rad3-related protein
ATRIP	ATR interacting protein
β-ME	β-mercaptoethanol
BER	Base excision repair
Bis-Tris	Bis(2-hydroxyethyl)-amino-tris(hydroxymethyl)-methane
BSA	Bovine serum albumin
CTF	Contrast transfer function
d	Drosophila
DBD	DNA-binding domain
DDT	DNA-binding homeobox and different transcription factors
dNTP	Deoxynucleotide triphosphate
DSB	Double-strand break
ds	Double-stranded
DTT	Dithiothreitol
E	Glutamate
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
ESI-MS	Electrospray ionization mass spectrometry
FA	Fluorescence anisotropy
fl	Full-length
FPC	Fork protection complex

GraFix	Gradient fixation
HOAc	Acetic acid
HEAT	Huntingtin, Elongation factor 3, subunit of protein phosphatase 2A and yeast kinase TOR1
Hepes	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HR	Homologous recombination
HRP	Horseradish peroxidase
iBAQ	Intensity based absolute quantification
IMAC	Immobilized metal ion affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
IR	Infrared
KCl	Potassium chloride
K_D	Dissociation constant
m	Mammalian
MES	2-(N-morpholino)ethanesulfonic acid
$MgCl_2$	Magnesium chloride
MCM	Mini-chromosome maintenance
$MgSO_4$	Magnesium sulfate
MPIB	Max Planck Institute of biochemistry
MS	Mass spectrometry
MST	Microscale thermophoresis
NaCl	Sodium chloride
Na_2CO_3	Sodium carbonate
NCBI	National center for biotechnology information
NER	Nucleotide excision repair
Ni	Nickel
NLS	Nuclear localization signal
NP-40	Nonidet P-40 (Octylphenoxy)polyethoxyethanol
$Na_2S_2O_3$	Sodium thiosulfate
nt	Nucleotide
OB	Oligonucleotide/Oligosaccharide-binding
P	Proline
PAGE	Polyacrylamide gel electrophoresis
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction

PEG	Polyethylene glycol
pfam	Protein family
pI	isoelectric point
PMSF	Phenylmethanesulfonylfluoride
Pol	Polymerase
Pfu	*Pyrococcus furiosus*
psi	Pound-force per square inch
PSI-BLAST	Position-specific iterated basic local alignment search tool
PSIPRED	PSI-BLAST based secondary structure prediction
PVDF	Polyvinylidene fluoride
RF	Replication fork
RFC	Replication factor C
RPA	Replication protein A
rpm	Revolutions per minute
RT	Room temperature
SAXS	Small angle X-ray scattering
SCN	Suprachiasmatic nucleus
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SLIC	Sequence and ligation-independent cloning
sp	Splayed
SSB	Single strand-binding protein
S. pompe	*Schizosaccharomyces pombe*, fission yeast
S. cerevisiae	*Saccharomyces cerevisiae*, budding yeast
ss	Single-stranded
TAE	Tris-acetate-EDTA
Tim	Timeless
Tipin	Timeless-interacting protein
Tris	Tris(hydroxymethyl)aminomethane
UNG2	Uracil-DNA glycosylase 2
WH	Winged-helix
XPA	Xeroderma pigmentosum complementation group A protein
X-ray	Röntgen radiation
$Zn(OAc)_2$	Zinc acetate

A. Multiple Sequence Alignments

A.1. Tim Orthologs

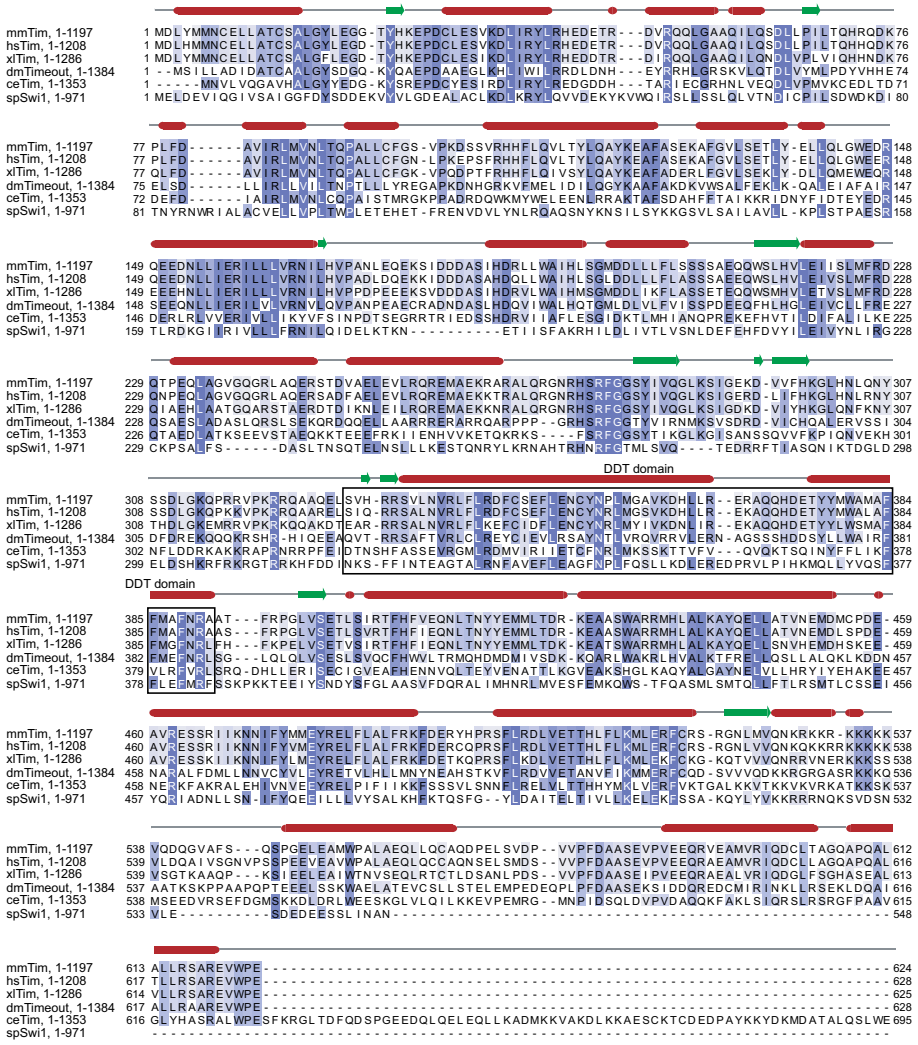


Figure A.1.: See after next page for figure legend.

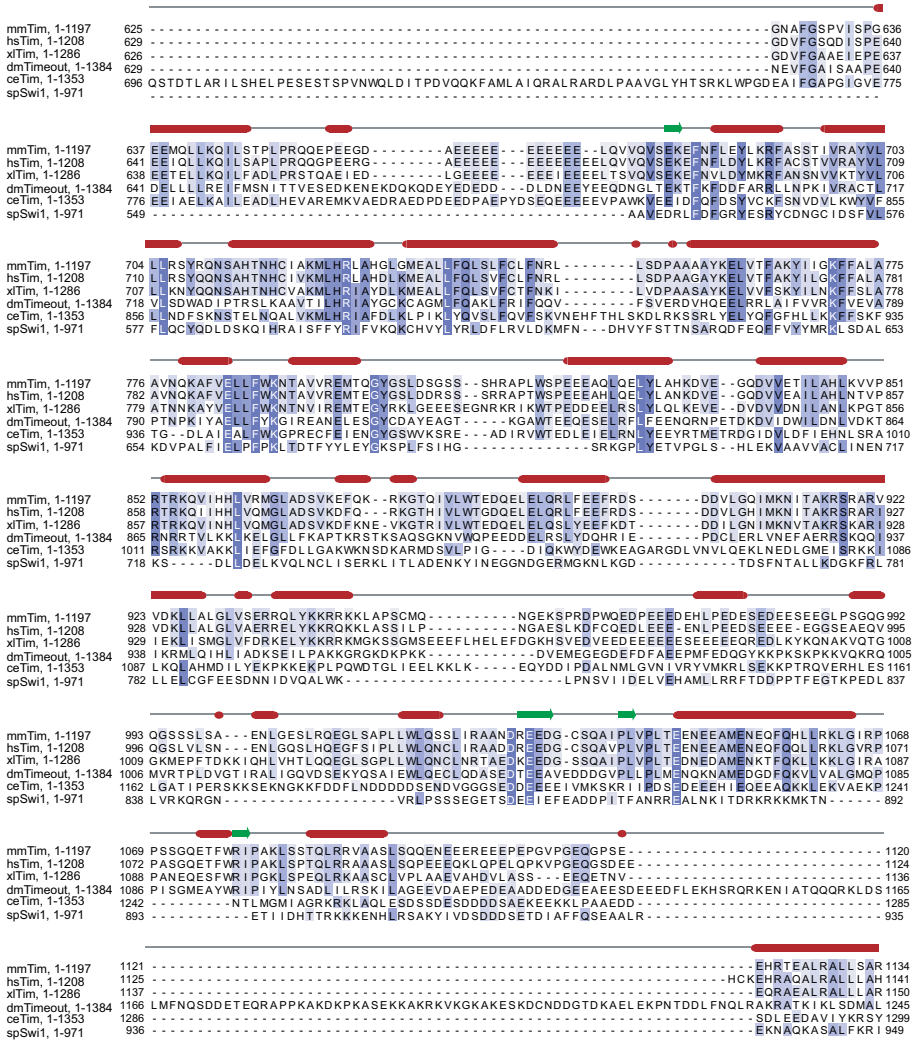


Figure A.1.: See next page for figure legend.

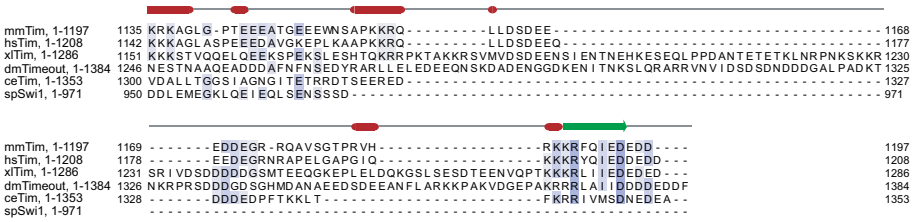


Figure A.1.: Multiple sequence alignment of Tim orthologs. Dark blue indicates strictly conserved residues; white variable residues. The black box indicates the DDT domain. The orthologs were identified using BLASTP [102, 164], aligned with CLUSTALW2 [165, 166], and displayed using Jalview [167]. The secondary structure elements are indicated and were calculated using Jpred [168]. The alignment shows the sequences from *Mus musculus (mm), Homo sapiens (hs), Xenopus laevis (xl), Drosophila melanogaster (dm), Caenorhabditis elegans (ce), and Schizosaccharomyces pombe (sp)*. UniProt and NCBI accession numbers [164, 169] from top to bottom: Q9R1X4, ACC80011.1, Q3LGB9, AAF73481.1, AAF13189.1, CAB44362.1.

A.2. Tipin Orthologs

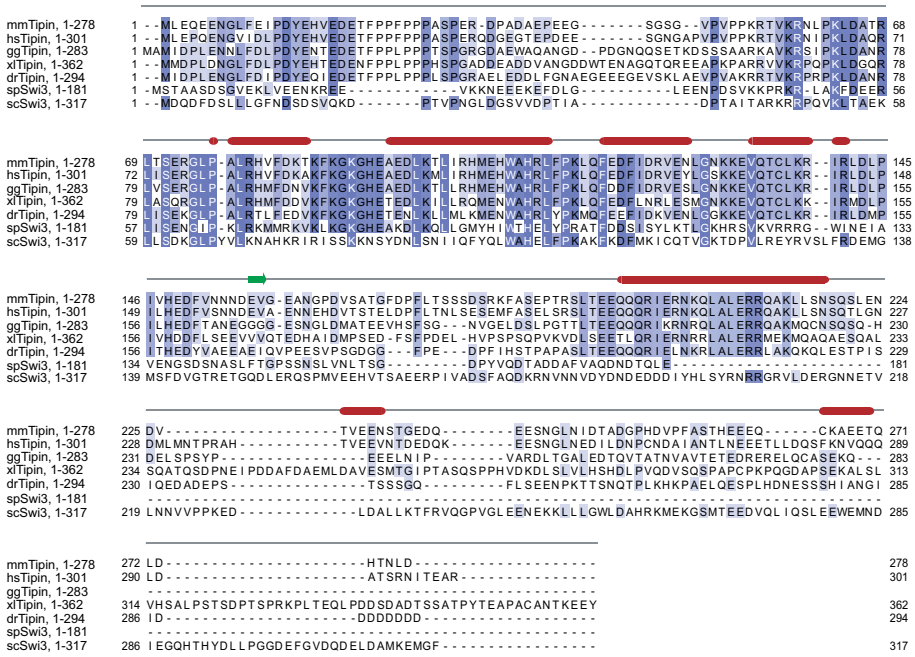


Figure A.2.: Multiple sequence alignment of Tipin orthologs. The orthologs were identified, aligned, and displayed as in Appendix Figure A.1. The alignment shows the sequences from *Mus musculus (mm), Homo sapiens (hs), Gallus gallus (gg), Xenopus laevis (xl), Danio rerio (dr), Schizosaccharomyces pombe (sp), and Saccharomyces cerevisiae (sc)*. UniProt and NCBI accession numbers [164, 169] from top to bottom: DAA01364.1, DAA01365.1, Q5F416, NP_001081090.1, AAI63996.1, AAS77252.1, DAA09947.1.

A.3. RPA14 Orthologs

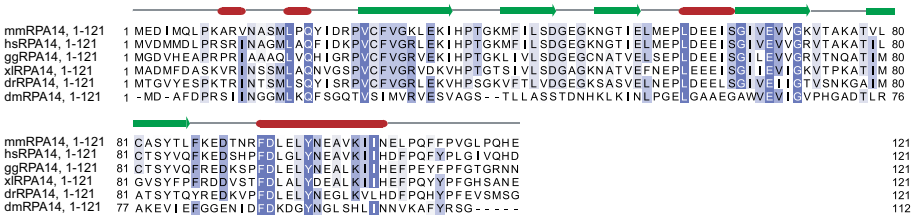


Figure A.3.: Multiple sequence alignment of RPA14 orthologs. The orthologs were identified, aligned, and displayed as in Appendix Figure A.1. The alignment shows the sequences from *Mus musculus (mm), Homo sapiens (hs), Gallus gallus (gg), Xenopus laevis (xl), Danio rerio (dr), and Drosophila melanogaster (dm)*. UniProt accession numbers [164] from top to bottom: Q9CQ71, A4D105, E1C4C7, Q6AZP5, A5PMY4, Q9VYW3.

A.4. RPA32 Orthologs

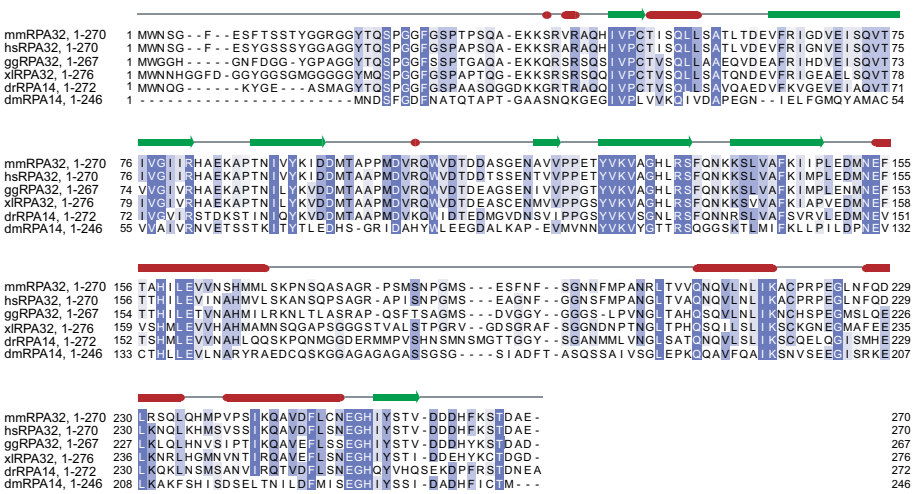


Figure A.4.: Multiple sequence alignment of RPA32 orthologs. The orthologs were identified, aligned, and displayed as in Appendix Figure A.1. The alignment shows the sequences from *Mus musculus (mm), Homo sapiens (hs), Gallus gallus (gg), Xenopus laevis (xl), Danio rerio (dr), and Drosophila melanogaster (dm)*. UniProt accession numbers [164] from top to bottom: Q62193, P15927, Q5ZLH1, Q6IP18, Q9DDG4, Q8SXG3.

A.5. RPA70 Orthologs

Figure A.5.: Multiple sequence alignment of RPA70 orthologs. The orthologs were identified, aligned, and displayed as in Appendix Figure A.1. The alignment shows the sequences from *Homo sapiens (hs), Mus musculus (mm), Gallus gallus (gg), Xenopus laevis (xl), Danio rerio (dr), and Drosophila melanogaster (dm).* UniProt and NCBI accession numbers [164, 169] from top to bottom: NP_002936.1, NP_080929, Q5ZJJ2, Q01588, Q6NY74.1, Q24492.

B. Difference Maps and 2D Class Averages

B.1. Tim-Tipin-ΔRPA Maps

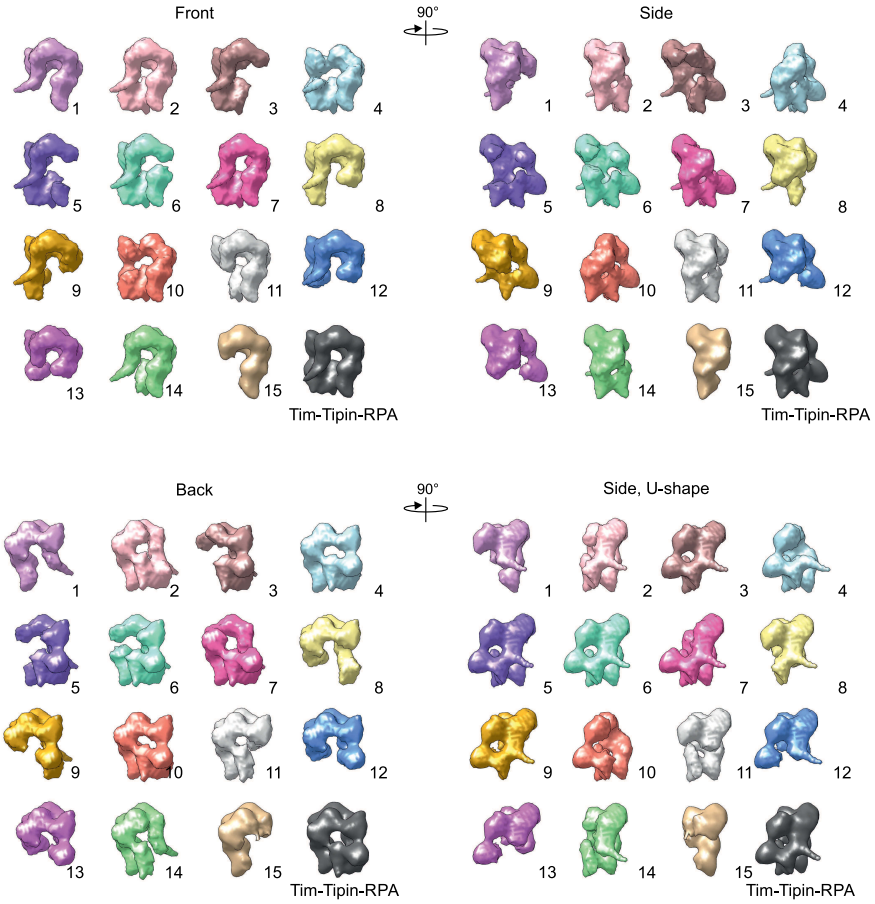


Figure B.1.: Difference maps in four different views (front, side, back, side, U-shape) related by rotation around the y-axis as indicated. The difference maps (Tim-Tipin-ΔRPA) were generated by subtraction of the putative volume enclosing RPA from the negative stain Tim-Tipin-RPA 3D reconstruction (black model).

B.2. Negative Stain Tim-Tipin 2D Class Averages

Figure B.2.: Negative stain Tim-Tipin 2D class averages used for difference mapping.

Bibliography

[1] Johnson, A. and O'Donnell, M. *Annual Review of Biochemistry* **74**(1), 283–315 (2005).

[2] Bell, S. P. and Stillman, B. *Nature* **357**(6374), 128–134 (1992).

[3] Maiorano, D., Moreau, J., and Mechali, M. *Nature* **404**(6778), 622–625 (2000).

[4] Liu, J., Smith, C. L., DeRyckere, D., DeAngelis, K., Martin, G., and Berger, J. M. *Molecular Cell* **6**(3), 637–648 (2000).

[5] Fernandez-Cid, A., Riera, A., Tognetti, S., Herrera, M., Samel, S., Evrin, C., Winkler, C., Gardenal, E., Uhle, S., and Speck, C. *Molecular Cell* **50**(4), 577–588 (2013).

[6] Costa, A., Ilves, I., Tamberg, N., Petojevic, T., Nogales, E., Botchan, M. R., and Berger, J. M. *Nature Structural & Molecular Biology* **18**(4), 471–477 (2011).

[7] Ilves, I., Petojevic, T., Pesavento, J. J., and Botchan, M. R. *Molecular Cell* **37**(2), 247–258 (2010).

[8] Walter, J. and Newport, J. *Molecular Cell* **5**(4), 617–627 (2000).

[9] Wold, M. S. *Annual Review of Biochemistry* **66**(1), 61–92 (1997).

[10] Dornreiter, I., Erdile, L. F., Gilbert, I. U., von Winkler, D., Kelly, T. J., and Fanning, E. *The EMBO Journal* **11**(2), 769–776 (1992).

[11] Waga, S. and Stillman, B. *Nature* **369**(6477), 207–212 (1994).

[12] Pursell, Z. F., Isoz, I., Lundström, E.-B., Johansson, E., and Kunkel, T. A. *Science* **317**(5834), 127–130 (2007).

[13] McElhinny, S. A. N., Gordenin, D. A., Stith, C. M., Burgers, P. M., and Kunkel, T. A. *Molecular Cell* **30**(2), 137–144 (2008).

[14] Prelich, G., Kostura, M., Marshak, D. R., Mathews, M. B., and Stillman, B. *Nature* **326**(6112), 471–475 (1987).

[15] Bauer, G. A. and Burgers, P. M. *Proceedings of the National Academy of Sciences* **85**(20), 7506–7510 (1988).

[16] Tsurimoto, T. and Stillman, B. *Journal of Biological Chemistry* **266**(3), 1950–1960 (1991).

[17] Bowman, G. D., O'Donnell, M., and Kuriyan, J. *Nature* **429**(6993), 724–730 (2004).

[18] Leman, A. R. and Noguchi, E. *Cell Cycle* **11**(21), 3945–3955 (2012).

[19] Gotter, A. L., Suppa, C., and Emanuel, B. S. *Journal of Molecular Biology* **366**(1), 36–52 (2007).

[20] Ali, S. I., Shin, J.-S., Bae, S.-H., Kim, B., and Choi, B.-S. *The International Journal of Biochemistry & Cell Biology* **42**(7), 1210–1215 (2010).

[21] Sancar, A., Lindsey-Boltz, L. A., Ünsal-Kacmaz, K., and Linn, S. *Annual Review of Biochemistry* **73**(1), 39–85 (2004).

[22] Feng, W., Collingwood, D., Boeck, M. E., Fox, L. A., Alvino, G. M., Fangman, W. L., Raghuraman, M. K., and Brewer, B. J. *Nature Cell Biology* **8**(2), 148–155 (2006).

[23] Sogo, J. M., Lopes, M., and Foiani, M. *Science* **297**(5581), 599–602 (2002).

[24] Namiki, Y. and Zou, L. *Proceedings of the National Academy of Sciences* **103**(3), 580–585 (2006).

[25] Zhao, H. and Piwnica-Worms, H. *Molecular and Cellular Biology* **21**(13), 4129–4139 (2001).

[26] Kumagai, A. and Dunphy, W. G. *Molecular Cell* **6**(4), 839–849 (2000).

[27] Serçin, O. and Kemp, M. G. *Cell Cycle* **10**(10), 1599–1606 (2005).

[28] Shechter, D., Costanzo, V., and Gautier, J. *Nature Cell Biology* **6**(7), 648–655 (2004).

[29] Smith, J., Tho, L. M., Xu, N., and Gillespie, D. A. *Advances in Cancer Research* **108**, 73–112 (2010).

[30] Clugston, C. K., McLaughlin, K., Kenny, M. K., and Brown, R. *Cancer Research* **52**(22), 6375–6379 (1992).

[31] Burns, J. L., Guzder, S. N., Sung, P., Prakash, S., and Prakash, L. *Journal of Biological Chemistry* **271**(20), 11607–11610 (1996).

[32] He, Z., Henricksen, L. A., Wold, M. S., and Ingles, C. J. *Nature* **374**(6522), 566–569 (1995).

[33] Wakasugi, M. and Sancar, A. *Journal of Biological Chemistry* **274**(26), 18759–18768 (1999).

[34] Mer, G., Bochkarev, A., Gupta, R., Bochkareva, E., Frappier, L., Ingles, C., Edwards, A. M., and Chazin, W. J. *Cell* **103**(3), 449–456 (2000).

[35] Nagelhus, T. A., Haug, T., Singh, K. K., Keshav, K. F., Skorpen, F., Otterlei, M., Bharati, S., Lindmo, T., Benichou, S., Benarous, R., and Krokan, H. E. *Journal of Biological Chemistry* **272**(10), 6561–6566 (1997).

[36] Golub, E. I., Gupta, R. C., Haaf, T., Wold, M. S., and Radding, C. M. *Nucleic Acids Research* **26**(23), 5388–5393 (1998).

[37] Sugiyama, T. and Kowalczykowski, S. C. *Journal of Biological Chemistry* **277**(35), 31663–31672 (2002).

[38] Stauffer, M. E. and Chazin, W. J. *Journal of Biological Chemistry* **279**(24), 25638–25645 (2004).

[39] Koike, N., Hida, A., Numano, R., Hirose, M., Sakaki, Y., and Tei, H. *FEBS Letters* **441**(3), 427–431 (1998).

[40] Sangoram, A. M., Saez, L., Antoch, M. P., Gekakis, N., Staknis, D., Whiteley, A., Fruechte, E. M., Vitaterna, M. H., Shimomura, K., King, D. P., Young, M. W., Weitz, C. J., and Takahashi, J. S. *Neuron* **21**(5), 1101–1113 (1998).

[41] Zylka, M. J., Shearman, L. P., Levine, J. D., Jin, X., Weaver, D. R., and Reppert, S. M. *Neuron* **21**(5), 1115–1122 (1998).

[42] Cermakian, N. and Sassone-Corsi, P. *Nature Reviews Molecular Cell Biology* **1**(1), 59–67 (2000).

[43] Hardin, P. E. *Current Biology* **15**(17), R714–R722 (2005).

[44] Saez, L. and Young, M. W. *Neuron* **17**(5), 911–920 (1996).

[45] Ceriani, M. F., Darlington, T. K., Staknis, D., Más, P., Petti, A. A., Weitz, C. J., and Kay, S. A. *Science* **285**(5427), 553–556 (1999).

[46] Yildiz, O., Doi, M., Yujnovsky, I., Cardone, L., Berndt, A., Hennig, S., Schulze, S., Urbanke, C., Sassone-Corsi, P., and Wolf, E. *Molecular Cell* **17**(1), 69–82 (2005).

[47] Hennig, S., Strauss, H. M., Vanselow, K., Yildiz, O., Schulze, S., Arens, J., Kramer, A., and Wolf, E. *PLoS Biol* **7**(4), e1000094 04 (2009).

[48] Kucera, N., Schmalen, I., Hennig, S., Öllinger, R., Strauss, H. M., Grudziecki, A., Wieczorek, C., Kramer, A., and Wolf, E. *Proceedings of the National Academy of Sciences* **109**(9), 3311–3316 (2012).

[49] Field, M. D., Maywood, E. S., O'Brien, J. A., Weaver, D. R., Reppert, S. M., and Hastings, M. H. *Neuron* **25**(2), 437–447 (2000).

[50] Kume, K., Zylka, M. J., Sriram, S., Shearman, L. P., Weaver, D. R., Jin, X., Maywood, E. S., Hastings, M. H., and Reppert, S. M. *Cell* **98**(2), 193–205 (1999).

[51] Shearman, L. P., Sriram, S., Weaver, D. R., Maywood, E. S., Chaves, I., Zheng, B., Kume, K., Lee, C. C., van der, G. T. J., Horst, Hastings, M. H., and Reppert, S. M. *Science* **288**(5468), 1013–1019 (2000).

[52] Barnes, J. W., Tischkau, S. A., Barnes, J. A., Mitchell, J. W., Burgoon, P. W.,

Hickok, J. R., and Gillette, M. U. *Science* **302**(5644), 439–442 (2003).

[53] Darlington, T. K., Wager-Smith, K., Ceriani, M. F., Staknis, D., Gekakis, N., Steeves, T. D. L., Weitz, C. J., Takahashi, J. S., and Kay, S. A. *Science* **280**(5369), 1599–1603 (1998).

[54] Jin, X., Shearman, L. P., Weaver, D. R., Zylka, M. J., Vries, G. J. D., and Reppert, S. M. *Cell* **96**(1), 57–68 (1999).

[55] Reppert, S. M. and Weaver, D. R. *Annual Review of Physiology* **63**(1), 647–676 (2001).

[56] Li, Z., Stuart, R. O., Qiao, J., Pavlova, A., Bush, K. T., Pohl, M., Sakurai, H., and Nigam, S. K. *Proceedings of the National Academy of Sciences* **97**(18), 10038–10043 (2000).

[57] Gotter, A. L., Manganaro, T., Weaver, D. R., Kolakowski, L. F., Possidente, B., Sriram, S., MacLaughlin, D. T., and Reppert, S. M. *Nature Neuroscience* **3**(8), 755–756 (2000).

[58] Benna, C., Scannapieco, P., Piccin, A., Sandrelli, F., Zordan, M., Rosato, E., Kyriacou, C., Valle, G., and Costa, R. *Current Biology* **10**(14), R512–R513 (2000).

[59] Benna, C., Bonaccorsi, S., Wülbeck, C., Helfrich-Förster, C., Gatti, M., Kyriacou, C. P., Costa, R., and Sandrelli, F. *Current Biology* **20**(4), 346–352 (2010).

[60] Gotter, A. L. *Journal of Molecular Biology* **331**(1), 167–176 (2003).

[61] Yoshizawa-Sugata, N. and Masai, H. *Journal of Biological Chemistry* **282**(4), 2729–2740 (2007).

[62] Chou, D. M. and Elledge, S. J. *Proceedings of the National Academy of Sciences* **103**(48), 18143–18147 (2006).

[63] Kemp, M. G., Akan, Z., Yilmaz, S., Grillo, M., Smith-Roe, S. L., Kang, T.-H., Cordeiro-Stone, M., Kaufmann, W. K., Abraham, R. T., Sancar, A., and Ünsal-Kaçmaz, K. *Journal of Biological Chemistry* **285**(22), 16562–16571 (2010).

[64] Aguilera, A. and Garcia-Muse, T. *Annual Review of Genetics* **47**(1), 19–50 (2013).

[65] Krings, G. and Bastia, D. *Proceedings of the National Academy of Sciences* **101**(39), 14085–14090 (2004).

[66] Noguchi, E., Noguchi, C., Du, L.-L., and Russell, P. *Molecular and Cellular Biology* **23**(21), 7861–7874 (2003).

[67] Foss, E. J. *Genetics* **157**(2), 567–577 (2001).

[68] Park, H. and Sternglanz, R. *Yeast* **15**(1), 35–41 (1999).

[69] Chan, R. C., Chan, A., Jeon, M., Wu, T. F., Pasqualone, D., Rougvie, A. E., and Meyer, B. J. *Nature* **423**(6943), 1002–1009 (2003).

[70] Numata, Y., Ishihara, S., Hasegawa, N., Nozaki, N., and Ishimi, Y. *Journal of Biochemistry* **147**(6), 917–927 (2010).

[71] Nedelcheva, M. N., Roguev, A., Dolapchiev, L. B., Shevchenko, A., Taskov, H. B., Shevchenko, A., Stewart, A. F., and Stoynov, S. S. *Journal of Molecular Biology* **347**(3), 509–521 (2005).

[72] Bando, M., Katou, Y., Komata, M., Tanaka, H., Itoh, T., Sutani, T., and Shirahige, K. *Journal of Biological Chemistry* **284**(49), 34355–34365 (2009).

[73] Gambus, A., Jones, R. C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R. D., and Labib, K. T. *Nature Cell Biology* **8**(4), 358–366 (2006).

[74] Cho, W.-H., Kang, Y.-H., An, Y.-Y., Tappin, I., Hurwitz, J., and Lee, J.-K. *Proceedings of the National Academy of Sciences* **110**(7), 2523–2527 (2013).

[75] Aria, V., De Felice, M., Di Perna, R., Uno, S., Masai, H., Syväoja, J. E., van Loon, B., Hübscher, U., and Pisani, F. M. *Journal of Biological Chemistry* **288**(18), 12742–12752 (2013).

[76] Sommariva, E., Pellny, T. K., Karahan, N., Kumar, S., Huberman, J. A., and Dalgaard, J. Z. *Molecular and Cellular Biology* **25**(7), 2770–2784 (2005).

[77] Noguchi, E., Noguchi, C., McDonald, W. H., Yates, J. R., and Russell, P. *Molecular and Cellular Biology* **24**(19), 8342–8355 (2004).

[78] Smith, K. D., Fu, M. A., and Brown, E. J. *The Journal of Cell Biology* **187**(1), 15–23 (2009).

[79] Errico, A., Costanzo, V., and Hunt, T. *Proceedings of the National Academy of Sciences* **104**(38), 14929–14934 (2007).

[80] Urtishak, K. A., Smith, K. D., Chanoux, R. A., Greenberg, R. A., Johnson, F. B., and Brown, E. J. *Journal of Biological Chemistry* **284**(13), 8777–8785 (2009).

[81] Ünsal-Kaçmaz, K., Mullen, T. E., Kaufmann, W. K., and Sancar, A. *Molecular and Cellular Biology* **25**(8), 3109–3116 (2005).

[82] Fanning, E., Klimovich, V., and Nager, A. R. *Nucleic Acids Research* **34**(15), 4126–4137 (2006).

[83] Ünsal-Kaçmaz, K., Chastain, P., Qu, P.-P., Minoo, P., Cordeiro-Stone, M., Sancar, A., and Kaufmann, W. *Molecular and Cellular Biology* **27**(8), 3131–3142 (2007).

[84] Mirkin, E. V. and Mirkin, S. M. *Microbiology and Molecular Biology Reviews* **71**(1), 13–35 (2007).

[85] Calzada, A., Hodgson, B., Kanemaki, M., Bueno, A., and Labib, K. *Genes & Development* **19**(16), 1905–1919 (2005).

[86] Wold, M. S. and Kelly, T. *Proceedings of the National Academy of Sciences* **85**(8), 2523–2527 (1988).

[87] Henricksen, L. A., Umbricht, C. B., and Wold, M. S. *Journal of Biological Chemistry* **269**(15), 11121–32 (1994).

[88] Habel, J. E., Ohren, J. F., and Borgstahl, G. E. O. *Acta Crystallographica Section D* **57**(2), 254–259 (2001).

[89] Brill, S. J. and Stillman, B. *Nature* **342**(6245), 92–95 (1989).

[90] Kim, C., Paulus, B. F., and Wold, M. S. *Biochemistry* **33**(47), 14197–14206 (1994).

[91] Kim, C. and Wold, M. S. *Biochemistry* **34**(6), 2058–2064 (1995).

[92] Bochkarev, A., Pfuetzner, R., Edwards, A., and Frappier, L. *Nature* **385**(6612),

176–181 (1997).

[93] Bochkarev, A., Bochkareva, E., Frappier, L., and Edwards, A. *The EMBO Journal* **18**(16), 4498–4504 (1999).

[94] Bochkareva, E., Belegu, V., Korolev, S., and Bochkarev, A. *The EMBO Journal* **20**(3), 612–618 (2001).

[95] Bochkareva, E., Korolev, S., Lees-Miller, S., and Bochkarev, A. *The EMBO Journal* **21**(7), 1855–1863 (2002).

[96] Fan, J. and Pavletich, N. P. *Genes & Development* **26**(20), 2337–2347 (2012).

[97] Brosey, C. A., Yan, C., Tsutakawa, S. E., Heller, W. T., Rambo, R. P., Tainer, J. A., Ivanov, I., and Chazin, W. J. *Nucleic Acids Research* **41**(4), 2313–2327 (2013).

[98] Adachi, Y. and Laemmli, U. K. *The Journal of Cell Biology* **119**(1), 1–15 (1992).

[99] Cardoso, M., Leonhardt, H., and Nadal-Ginard, B. *Cell* **74**(6), 979–992 (1993).

[100] Tanaka, T. and Nasmyth, K. *The EMBO Journal* **17**(17), 5182–5191 (1998).

[101] Nakaya, R., Takaya, J., Onuki, T., Moritani, M., Nozaki, N., and Ishimi, Y. *Journal of Biochemistry* **148**(5), 539–547 (2010).

[102] Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. *Nucleic Acids Research* **25**(17), 3389–3402 (1997).

[103] Jones, D. T. *Journal of Molecular Biology* **292**(2), 195–202 (1999).

[104] Andrade, M. A. and Bork, P. *Nature Genetics* **11**, 115–116 (1995).

[105] Vodovar, N., Clayton, J. D., Costa, R., Odell, M., and Kyriacou, C. P. *Current Biology* **12**(18), R610 – R611 (2002).

[106] Kippert, F. and Gerloff, D. L. *Current Biology* **14**(16), R650 – R651 (2004).

[107] Noguchi, C., Rapp, J. B., Skorobogatko, Y. V., Bailey, L. D., and Noguchi, E. *PLoS ONE* **7**(8), e43988 08 (2012).

[108] Doerks, T., Copley, R., and Bork, P. *Trends in Biochemical Sciences* **26**(3), 145–146 (2001).

[109] Engelen, E., Janssens, R. C., Yagita, K., Smits, V. A. J., van der Horst, G. T. J., and Tamanini, F. *PLoS ONE* **8**(2), e56623 02 (2013).

[110] Bochkarev, A. and Bochkareva, E. *Current Opinion in Structural Biology* **14**(1), 36–42 (2004).

[111] Theobald, D. L., Mitton-Fry, R. M., and Wuttke, D. S. *Annual Review of Biophysics and Biomolecular Structure* **32**(1), 115–133 (2003).

[112] Blackwell, L. J. and Borowiec, J. A. *Molecular and Cellular Biology* **14**(6), 3993–4001 (1994).

[113] Blackwell, L. J., Borowiec, J. A., and Mastrangelo, I. A. *Molecular and Cellular Biology* **16**(9), 4798–4807 (1996).

[114] Lavrik, O., Kolpashchikov, D., Nasheuer, H.-P., Weisshart, K., and Favre, A. *FEBS Letters* **441**(2), 186–190 (1998).

[115] Bastin-Shanower, S. A. and Brill, S. J. *Journal of Biological Chemistry* **276**(39), 36446–36453 (2001).

[116] Arunkumar, A. I., Stauffer, M. E., Bochkareva, E., Bochkarev, A., and Chazin, W. J. *Journal of Biological Chemistry* **278**(42), 41077–41082 (2003).

[117] Bochkareva, E., Korolev, S., and Bochkarev, A. *Journal of Biological Chemistry* **275**(35), 27332–27338 (2000).

[118] Bochkareva, E., Kaustov, L., Ayed, A., Yi, G.-S., Lu, Y., Pineda-Lucena, A., Liao, J. C. C., Okorokov, A. L., Milner, J., Arrowsmith, C. H., and Bochkarev, A. *Proceedings of the National Academy of Sciences* **102**(43), 15412–15417 (2005).

[119] Daughdrill, G. W., Ackerman, J., Isern, N. G., Botuyan, M. V., Arrowsmith, C., Wold, M. S., and Lowry, D. F. *Nucleic Acids Research* **29**(15), 3270–3276 (2001).

[120] Kim, D.-K., Stigger, E., and Lee, S.-H. *Journal of Biological Chemistry* **271**(25), 15124–15129 (1996).

[121] Din, S., Brill, S. J., Fairman, M. P., and Stillman, B. *Genes & Development* **4**(6), 968–977 (1990).

[122] Liu, V. F. and Weaver, D. T. *Molecular and Cellular Biology* **13**(12), 7222–7231 (1993).

[123] Brush, G. S., Anderson, C. W., and Kelly, T. J. *Proceedings of the National Academy of Sciences* **91**(26), 12520–12524 (1994).

[124] Brush, G., Morrow, D., Hieter, P., and Kelly, T. *Proceedings of the National Academy of Sciences* **93**(26), 15075–15080 (1996).

[125] Lee, S.-H., Kim, D.-K., and Drissi, R. *Journal of Biological Chemistry* **270**(37), 21800–21805 (1995).

[126] Matthews, B. *Journal of Molecular Biology* **33**(2), 491–497 (1968).

[127] Kantardjieff, K. A. and Rupp, B. *Protein Science* **12**(9), 1865–1871 (2003).

[128] Witosch, J., Wolf, E., and Mizuno, N. *Nucleic Acids Research* (2014). doi: 10.1093/nar/gku960.

[129] Stark, H. *Methods in Enzymology* **481**, 109 – 126 (2010).

[130] Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. *Nature* **473**(7347), 337–342 (2011).

[131] Radermacher, M., Wagenknecht, T., Verschoor, A., and Frank, J. *Journal of Microscopy* **146**(2), 113–136 (1987).

[132] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. *Journal of Computational Chemistry* **25**(13), 1605–1612 (2004).

[133] Melby, T. E., Ciampaglio, C. N., Briscoe, G., and Erickson, H. P. *The Journal of Cell Biology* **142**(6), 1595–1604 (1998).

[134] Simon, A. C., Zhou, J. C., Perera, R. L., van Deursen, F., Evrin, C., Ivanova, M. E., Kilkenny, M. L., Renault, L., Kjaer, S., Matak-Vinkovic, D., Labib, K.,

Costa, A., and Pellegrini, L. *Nature* **510**(7504), 293–297 (2014).

[135] Gomes, X. V., Henricksen, L. A., and Wold, M. S. *Biochemistry* **35**(17), 5586–5595 (1996).

[136] Brosey, C. A., Chagot, M.-E., Ehrhardt, M., Pretto, D. I., Weiner, B. E., and Chazin, W. J. *Journal of the American Chemical Society* **131**(18), 6346–6347 (2009).

[137] Pretto, D. I., Tsutakawa, S., Brosey, C. A., Castillo, A., Chagot, M.-E., Smith, J. A., Tainer, J. A., and Chazin, W. J. *Biochemistry* **49**(13), 2880–2889 (2010).

[138] Deng, X., Habel, J. E., Kabaleeswaran, V., Snell, E. H., Wold, M. S., and Borgstahl, G. E. *Journal of Molecular Biology* **374**(4), 865–876 (2007).

[139] Binz, S. K., Lao, Y., Lowry, D. F., and Wold, M. S. *Journal of Biological Chemistry* **278**(37), 35584–35591 (2003).

[140] Binz, S. K., Sheehan, A. M., and Wold, M. S. *DNA Repair* **3**, 1015–1024 (2004).

[141] Xu, X., Vaithiyalingam, S., Glick, G. G., Mordes, D. A., Chazin, W. J., and Cortez, D. *Molecular and Cellular Biology* **28**(24), 7345–7353 (2008).

[142] Jiang, X., Klimovich, V., Arunkumar, A. I., Hysinger, E. B., Wang, Y., Ott, R. D., Guler, G. D., Weiner, B., Chazin, W., and Fanning, E. *The EMBO Journal* **25**(23), 5516–5526 (2006).

[143] Byun, T. S., Pacek, M., Yee, M.-c., Walter, J. C., and Cimprich, K. A. *Genes & Development* **19**(9), 1040–1052 (2005).

[144] Yilmaz, S., Sancar, A., and Kemp, M. G. *PLoS ONE* **6**(7), e22986 07 (2011).

[145] Saijo, M., Kuraoka, I., Masutani, C., Hanaoka, F., and Tanaka, K. *Nucleic Acids Research* **24**(23), 4719–4724 (1996).

[146] Bochkareva, E., Frappier, L., Edwards, A. M., and Bochkarev, A. *Journal of Biological Chemistry* **273**(7), 3932–3936 (1998).

[147] Heymann, J. B. and Belnap, D. M. *Journal of Structural Biology* **157**(1), 3–18

(2007).

[148] Mindell, J. A. and Grigorieff, N. *Journal of Structural Biology* **142**(3), 334–347 (2003).

[149] Ludtke, S. J., Baldwin, P. R., and Chiu, W. *Journal of Structural Biology* **128**(1), 82–97 (1999).

[150] Cox, J. and Mann, M. *Nature Biotechnology* **26**(12), 1367–1372 (2008).

[151] Hohn, M., Tang, G., Goodyear, G., Baldwin, P., Huang, Z., Penczek, P. A., Yang, C., Glaeser, R. M., Adams, P. D., and Ludtke, S. J. *Journal of Structural Biology* **157**(1), 47–55 (2007).

[152] Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. *Journal of Structural Biology* **116**(1), 190–199 (1996).

[153] Kabsch, W. *Acta Crystallographica Section D* **66**(2), 125–132 (2010).

[154] Aslanidis, C. and de Jong, P. J. *Nucleic Acids Research* **18**(20), 6069–6074 (1990).

[155] Scholz, J., Besir, H., Strasser, C., and Suppmann, S. *BMC Biotechnology* **13**(12) (2013).

[156] Li, M. Z. and Elledge, S. J. *Nature Methods* **4**(3), 251–256 (2007).

[157] Laemmli UK. *Nature* **227**(5259), 680–685 (1970).

[158] Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. *Analytical Chemistry* **68**(5), 850–858 (1996).

[159] Olsen, J. V., de Godoy, L. M. F., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. *Molecular & Cellular Proteomics* **4**(12), 2010–2021 (2005).

[160] Pintilie, G. D., Zhang, J., Goddard, T. D., Chiu, W., and Gossard, D. C. *Journal of Structural Biology* **170**(3), 427–438 (2010).

[161] Radermacher, M. *J. Electron. Microsc Tech.* **9**(4), 359–394 (1988).

[162] Scheres, S. H. *Journal of Structural Biology* **180**(3), 519–530 (2012).

[163] Volkmann, N. *Journal of Structural Biology* **138**(1-2), 123–129 (2002).

[164] The UniProt Consortium. *Nucleic Acids Research* **42**(D1), D191–D198 (2014).

[165] Larkin, M., Blackshields, G., Brown, N., Chenna, R., McGettigan, P., McWilliam, H., Valentin, F., Wallace, I., Wilm, A., Lopez, R., Thompson, J., Gibson, T., and Higgins, D. *Bioinformatics* **23**(21), 2947–2948 (2007).

[166] McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P., and Lopez, R. *Nucleic Acids Research* **41**(W1), W597–W600 (2013).

[167] Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. *Bioinformatics* **25**(9), 1189–1191 (2009).

[168] Cole, C., Barber, J. D., and Barton, G. J. *Nucleic Acids Research* **36**(Web Server issue), W197–W201 (2008).

[169] Geer, L. Y., Marchler-Bauer, A., Geer, R. C., Han, L., He, J., He, S., Liu, C., Shi, W., and Bryant, S. H. *Nucleic Acids Research* **38**(Database issue), D492–D496 (2010).