The structure of Poz1-Tpz1 reveals a dimerization module in the fission yeast shelterin complex

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Cian Stutz

aus Basel, Schweiz

Basel 2014

Genehmigt von der Philosophisch-Naturwissenschaftlichen Falkultät auf Antrag von

Prof. Dr. Susan M. Gasser

Prof. Dr. Peter Baumann

Basel, den 26.02.2013

<div align="right">

Prof. Dr. Jörg Schibler

Dekan

</div>

Summary

Linear eukaryotic chromosomes are characterized by having distinct DNA ends. These ends, designated telomeres, are threatened by continued shortening in replication, nucleolytic degradation and unwanted fusion by the DNA-repair machinery. In order to counteract these processes, eukaryotic cells have evolved a protective cap, designated shelterin, which is assembled at telomeric repeats. In fission yeast, shelterin consists of the Taz1-Rap1 front-end that binds double strand telomeric repeats and the Pot1-Tpz1-Ccq1 assembly that is associated with the single strand 3' overhang. A small protein designated Poz1 bridges the two halves and has been shown to negatively regulate telomerase, the enzyme that synthesizes telomeric repeats. We set out to investigate the molecular details of the poorly characterized protein Poz1 and its interaction with Tpz1 in an effort to gain insight into the mechanisms of telomere length regulation.

Here we present the crystal structure of $spPoz1^{30\text{-}249}$ bound to $spTpz1^{475\text{-}508}$ at 2.4 Å resolution. Our structure remarkably resembles the structure of the TRF-homology domain (TRFH) of the human shelterin components TRF1 and TRF2. TRFH acts as a dimerization module in human shelterin. We speculated that also Poz1 functions as a dimerization module in fission yeast shelterin, a hypothesis that we subsequently validated by mutational analysis and size-exclusion chromatography. We revealed that Poz1 by itself is monomeric and upon binding of Tpz1 a Poz1-Tpz1 heterotetramer is formed. Furthermore, we showed that also *in vivo* heterotetramerization is essential for maintaining proper telomere length. Together with the previously reported dimeric existence of Taz1 and Rap1, as well as the recent discovery that Pot1 can dimerize upon telomere binding, these findings suggest an overall dimeric arrangement of the components in the fission yeast shelterin complex.

We replaced Poz1 by the structurally similar human TRF2H-dimerization domain and its binding partner hApollo. While *poz1Δ* strains show very long telomeres, the TRF2H-Apollo strain showed slow progressive telomere shortening. We suggest that shelterin function was mostly restored because TRF2H-Apollo imitates the primary function of Poz1, which is linking the double-strand and the single-strand binding halves of shelterin complex in a dimeric fashion. Likely, slow telomere shortening is observed because the interactions between Rap1-Poz1-Tpz1 are likely dynamic and regulated, but TRF2H and Apollo were fused to Rap1 and Tpz1 respectively, not allowing association and dissociation, leading to disrupted telomere regulation.

Furthermore, the structure revealed a zinc binding site at the binding interface of Poz1 and Tpz1. Disruption of the zinc site does not affect Poz1-Tpz1 interaction, but leads to decreased solubility of the complex *in vitro* and long telomeres *in vivo*. Consequently, binding of zinc is essential for the structural integrity of the complex.

Ultimately, by determining the crystal structure of $Poz1^{30-249}+Tpz1^{475-508}$, we uncovered new details of the architecture of the fission yeast shelterin complex and revealed striking similarities to structural elements found in human shelterin. Furthermore we examined the delicate interplay between the shelterin components and how essential its integrity is for proper telomere length homeostasis.

Table of contents

List of figures

List of tables

List of abbreviations

ATM	Ataxia telangiectasia mutated
ATR	ataxia telangiectasia and Rad3-related protein
Bp	Base pair
BRCA1	Breast cancer type 1 susceptibility protein
BRCT	BRCA1 C Terminus domain
DC	Dyskeratosis congenital
DNA	Desoxyribonucleic acid
DSB	Double-strand break
Est	Ever-shorter telomeres
EXAFs	Extended X-ray absorption fine structure
HEAT	Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A), and the yeast kinase TOR1
HMM	Hidden-Markov-Model
Hph	Hygromorphine
HPLC	High-performance liquid chromatography
HR	Homologous recombination
Kan	Kanamycin
Kbp	Kilo base pair
LC/MS	Liquid chromatography coupled to mass spectroscopy
LiOAc	Lithium acetate
MALDI-TOF	Matrix-assisted laser desorption/ionization time-of-flight
MES	2-Morpholino-ethane-sulfonic acid
MPD	(4s)-2-methyl-2,4-pentanediol
MRN	Mre11-Rad50-Nbs1
MRX	Mre11-Rad50-Xrs2
Myb	Myb proto-oncogene protein
Nat	Nourseothricin
NHEJ	Non-homologous-end-joining
NLS	Nuclear localization signal
NMR	Nuclear magnetic resonance

OB	Oligonucleotide/oligosaccharide binding domain
PBD	Poz1 binding domain
PCR	Polymerase chain reaction
PMSF	Phenylmethanesulfonylfluoride
RCT	Rap1 C-terminal domain
RNA	Ribonucleic acid
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC-MALS	Size exclusion chromatography with multi-angle light scattering
SLS	Swiss light source
ssDNA	Single-strand DNA
TCEP	Tris(2-carboxyethyl)phosphine
TEBP	Telomere end-binding protein
TERT	Telomerase reverse transcriptase
TEV	Tobacco etch virus
TRF	Telomere recognition factor
TRFH	TRF-homology
TRIS	Tris-(hydroxymethyl)-aminomethane
UTR	Untranslated region
YES	Yeast extract with supplements

Chapter 1

Introduction

1. Introduction

1.1 Telomeres, their function and difficulties they need to overcome

The genomes of prokaryotes and eukaryotic organelles are usually circular, as are most plasmids and viral genomes. Eukaryotic genomes on the other hand, are organized into multiple linear chromosomes. The ends of linear DNA create two major problems that the cells must counteract. First the DNA-replication machinery is unable to replicate linear DNA to completion, referred to as the end-replication problem. Without a mechanism in place to solve this problem, each round of DNA replication, would result in a loss of a short stretch of DNA, which eventually would lead to death by senescence. Secondly, without a protection mechanism, chromosome ends would be recognized as double strand breaks (DSBs) by the DNA-repair machinery and would be subjected to exonucleolytic degradation, recombination and end-to-end joining. The formation of dicentric chromosomes would have fatal consequences for the cell by causing severe problems during chromosome segregation.

Most organisms share a common solution for protecting their chromosomal ends from the threatening cellular environment. This is the telomerase-based solution. Here, the ends of chromosomal DNA are organized into repetitive protein-DNA complexes know as the telomeres. These allow the regulation of the reverse transcriptase telomerase, which can synthesize telomeric DNA thereby counteracting the loss of DNA caused by incomplete replication. Furthermore, the proteins that bind to the telomeric DNA have the ability to repress the DNA-repair machinery and in this manner, degradation, recombination and end-fusions are inhibited.

While the proteins that bind the telomeres and regulate the telomerase enzyme vary to some extent in composition and function from organism to organism, the principal of synthesizing telomeric repeats to the chromosomal ends by means of telomerase is conserved among many species ranging from yeast over ciliates and mice to humans. Although rare, some plants and animals have lost the telomerase enzyme during the course of evolution. These organisms maintain their chromosome ends by telomerase-independent mechanisms. Some plants such as the family of nightshades *(Solanaceae)* (Fajkus *et al.*, 1995; Peska *et al.*, 2008; Sykorova *et al.*, 2003; Watson & Riha, 2010) and onion *(Allium cepa)* (Pich *et al.*, 1996; Pich & Schubert, 1998) have been reported to lack telomerase. While for *Solanaceae*

the compensation mechanism for maintaining the telomeres remains unknown, *A. cepa* telomeres are thought to be maintained by transposable sequence-elements within the telomeres or homologous recombination between satellite telomere sequences. A far better characterized mechanism for telomerase-free telomere maintenance is present in the fruit fly *(Drosophila)*. *Drosophila* lacks telomerase (Sasaki *et al.*, 2000) and telomeres are elongated by occasional transposition of specialized retroelements (Mason *et al.*, 2008; Capkova *et al.*, 2008; Pardue *et al.*, 2008). Furthermore, the *Drosophila* telomeres are bound by a multi-protein complex called terminin in a sequence independent manner, which has been shown to be essential for protection against telomere fusions (Cenci *et al.*, 2005). The terminin components are not conserved outside of Drosophila. It has been suggested that during the course of evolution the telomerase-based telomere elongation mechanism was lost and *Drosophila* rapidly evolved the terminin complex to protect its telomeres (Raffa *et al.*, 2011).

Since this work centers around the telomeres of fission yeast *(Schizosaccharomyces pombe)*, the introduction of my thesis will focus on the telomerase-based maintenance system. I will describe the architecture of the telomeres with the proteins that bind them and what is known about the regulatory mechanisms that maintain the integrity of the chromosome ends. Human telomeres will also be highlighted as they share many structural and functional similarities with fission yeast telomeres. I will also give an overview of structural data that is available to date on the proteins associated with the telomeres.

1.1.1 Telomeric DNA sequences

In most eukaryotes, telomeric sequences are organized in a similar manner: short tandem repeats, which are G-rich in the strand containing the 3' end. This strand is referred to as the G-strand, while the complementary strand is called the C-strand. However the actual telomeric sequence varies among different species, as does the number of telomeric repeats. In cells that express telomerase, telomere length is kept within a narrow species- and cell-type specific range. The length of the double strand telomeric repeats ranges from <30 bp in some ciliates (Klobutcher *et al.*, 1981), ~300 bp in yeast (Shampay *et al.*, 1984), 5-15 kbp in humans (Moyzis *et al.*, 1988) and up to ~150kbp in mice (Kipling, 1990). Fission yeast telomeres were first cloned and sequenced in the 1980s by Sugawara and Szostak. They are

comprised of a degenerate telomeric repeat (G_{2-6}TTAC[A]) spanning the terminal ~300 bp of each chromosome. Irregular telomeric repeats have also been found in budding yeast (Szostak & Blackburn, 1982; Shampay *et al.,* 1984; Forstemann & Lingner, 2001). Mammals use the sequence TTAGGG for the telomeric repeats at their chromosome ends (Moyzis *et al.*, 1988; Meyne *et al.*, 1990). The length of the telomeric repeat tracts varies among different mammals, but for humans is generally in the range of 5-15 kb. The advantages of the extraordinarily long telomeres found in mice and rats are currently unknown (Kipling, 1990).

The actual terminus of chromosomal ends is not blunt but consists of an extension of the 3' end of the G-strand known as the 3' overhang or G-tail. This feature is conserved among most species in the eukaryotic kingdom. For *Drosophila* that lacks telomerase, a 3' overhang has also been suggested given the lower occupancy of some of the telomere binding factors at the extreme ends, but this hypothesis has not been confirmed (Gao *et al.*, 2010). The 3' overhang in yeast ranges from 12 to 100 bases in length (Wellinger *et al.*, 1992 and 1993), while in human cells the overhang can range between 50-500 nucleotides, which is considerably longer than in other eukaryotes (Makarov *et al.*, 1997; Stewart *et al.*, 2003). In budding yeast it has been shown that the 3' overhang length depends on the cell cycle stage with longer tails during S-phase (30-100 bases) and shorter tails found during the rest of the cell cycle (12-14 bases) (Wellinger *et al.*, 1993; Larrivee *at al.*, 2004). G-tails are suggested to be involved in the protection of chromosome ends either by T-loop formation to sequester the DNA ends and/or by binding of specific single-strand binding proteins (mechanisms described in **section 1.4**). The formation of the G-tails in *S. cerevisiae* (Dionne & Wellinger, 1996; Larrivee et al, 2004) and humans (Chai et al, 2006) is thought to be achieved by C-strand resection dependent on the Mre11-Rad50-Xrs2 (MRX) and Mre11-Rad50-Nbs1 (MRN) complexes respectively. In fission yeast, the Rad50-Rad32-Nbs1 (MRN) complex as well as the nuclease Dna2 have been implicated in the resection of the C-rich strand (Tomita *et al.*, 2004).

1.1.2 The end-replication problem

The end-replication problem was first described arising from the fact, that DNA-polymerases cannot replicate linear DNA to completion (**figure 1.1**). DNA-polymerases

synthesize DNA only in the 5'→ 3' direction, allowing normal replication of the leading strand. Lagging strand synthesis, on the other hand, cannot be performed in this manner and requires short RNA primers. These are used to prime lagging strand DNA synthesis, and are subsequently removed. The gaps are filled by DNA-polymerase δ. However, the most distal primer cannot be replaced with DNA after removal, leading to a loss of 8-12 nucleotides at the 5' end (**figure 1.1B**). Early on, this was thought to be the cause of progressive telomere shortening with each round of replication (Watson, 1972; Olovnikov, 1973; Cavalier-Smith, 1974; Bateman, 1975). Later on however, it was discovered that rather than being a lagging-strand synthesis issue, the progressive loss of DNA at the chromosome ends arises from the exonucleolytic degradation. Because leading-strand synthesis results in blunt-ended DNA, the 3' overhang must be generated. This is achieved by 5'→ 3' exonucleolytic degradation occurring on the telomere on which leading strand synthesis has taken place and results in a lack of template during the next leading strand synthesis (**figure 1.1C**) (Lingner *et al.*, 1995). With each round of replication, the DNA would get progressively shorter.

However, cells have a mechanism in place to counteract the progressive DNA shortening due to incomplete replication and resection with a key player being telomerase.
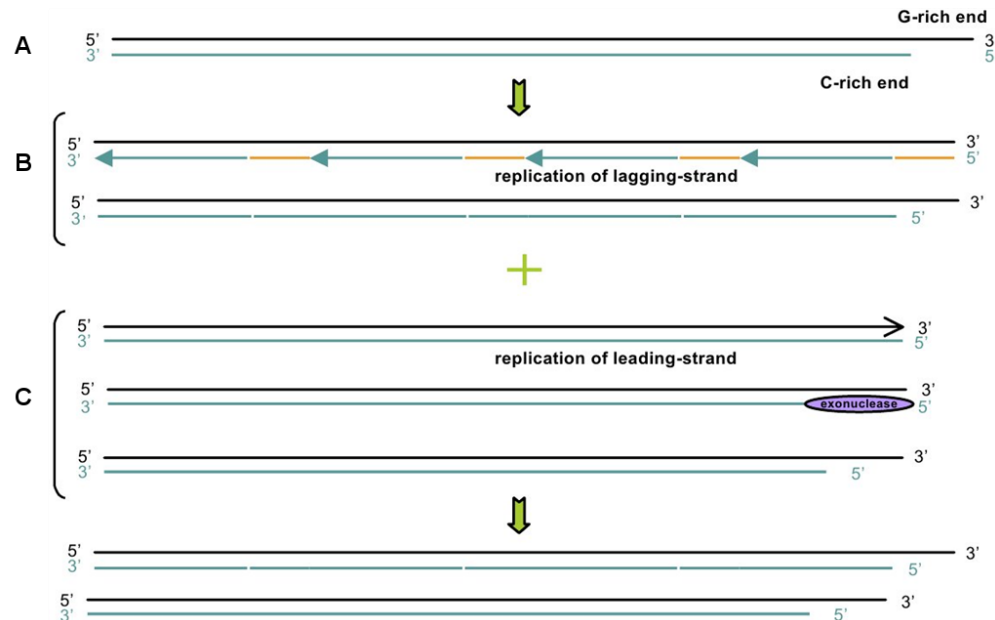
Figure 1.1. The DNA end-replication problem.

(A) The initial DNA double-strand with the 3' overhang. Due to the 5'-3' directionality of DNA polymerases, replication of the lagging strand is achieved by short RNA primers, which are subsequently removed and filled by the DNA polymerases (B). The very distal RNA primer however cannot be replaced. The problem unfolds upon replication of the leading-strand: exonucleolytic degradation on the telomere on which leading strand synthesis has taken place produces an overhang that results in a lack of template during the next leading strand synthesis (C). Without a mechanism to counteract this problem, at each round of replication the DNA would get progressively shorter. Adapted from Grandin and Charbonneau, 2008).

1.1.3 DNA-damage response in telomere regulation

Another threat to the liner DNA-ends is the DNA-repair machinery. DNA double-strand breaks (DSBs) must be repaired to maintain genomic integrity. This is achieved by joining of the DNA ends by homologous recombination (HR) and non-homologous-end-joining (NHEJ). Since telomeres are natural DNA termini, they could potentially be recognized as DNA double-strand breaks (DSBs) by the checkpoint machinery. This would lead to inadvertent joining of the chromosomal ends resulting in dicentric chromosomes and breakage-fusion-cycles during cell division (McClintock 1938 and 1941). However, this scenario is circumvented through telomeres. Specific proteins assemble at the telomeric repeats and protect them from the DNA repair machinery.

In general, in the presence of DNA damage, the activation of checkpoint proteins pauses the cell cycle, giving the cell sufficient time to repair the damage before starting DNA replication and mitosis (Sancar *et al.*, 2004; Humpal *et al.*, 2009). Mammalian cells are alerted to lesions in their genome by two phosphatidylinositol 3-kinase-related protein kinases called ATM and ATR (**figure 1.2**) (reviewed by Shiloh 2003). While the ATM pathway is thought to respond primarily to double-strand breaks (DSBs), ATR comes in response to the formation of ssDNA.
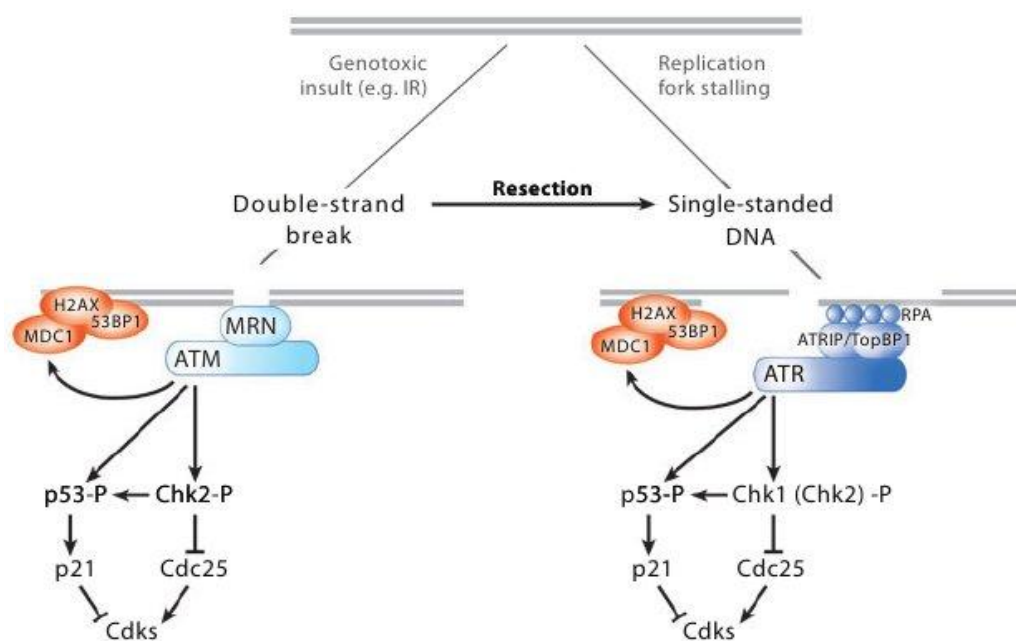
Figure 1.2. The DNA-damage response.
ATM and ATR pathways in humans. Figure adapted from Palm and de Lange, 2008.

These complexes promote the accumulation of other DNA damage response factors and they phosphorylate two nucleoplasmic effector kinases, Chk1 and Chk2, which can block cell cycle progression until damaged DNA has been fixed.

In fission yeast, these DNA damage sensor kinases are called Tel1[ATM] and Rad3[ATR] (where the superscript corresponds to the human orthologue). Paradoxically, while Tel1[ATM] and Rad3[ATR] activate the DNA repair machinery and thereby present a threat to linear DNA ends, it has been shown in fission yeast that they are also required to be present at the telomeres in order to phosphorylate the telomere protein Ccq1 (Moser *et al.*, 2011). Failure to phosphorylate Ccq1 results in the failure of the recruitment of telomerase to telomeres and leads to gradual shortening of telomeres (Yamazaki *et al.*, 2012) (more in **section 1.4.3**).

1.2 Telomerase

Cells require a special mechanism to counteract the progressive shortening of the ends of linear DNA (described in **section 1.1.2**). A key player in this mechanism is the ribonucleoprotein complex known as telomerase (**figure 1.3**), first discovered in 1985 in *Tetrahymena thermophila* by Carol Greider and Elizabeth Blackburn (Greider and Blackburn, 1985). Telomerase catalyzes the addition of telomeric repeats to the 3' overhang and consists of a reverse transcriptase protein and a large RNA component, which provides the template sequence for the telomeric repeat (Greider and Blackburn 1989, Lingner *et al.*, 1997).
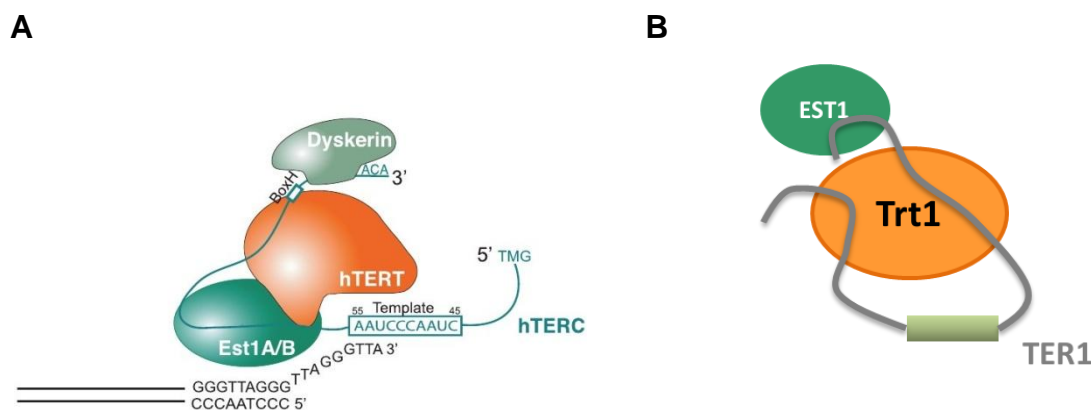
Figure 1.3. The telomerase reverse transcriptase complex.
Schematic view of **A)** the human and **B)** fission yeast telomerase complex. Figures adapted from Smogorzewska and Lange, 2004)

In humans, telomerase is composed of the reverse transcriptase TERT and the RNA moiety TERC, which contains the template for synthesis of the telomeric repeats (Feng, 1995; Lingner, 1997; Nakamura, 1997). While the TERT protein subunit is highly conserved among eukaryotes, the RNA component of telomerase diverged quite quickly in evolution and differs in length and sequence among different species. The telomerase RNAs do however share some common features among different species. This includes an open loop containing the template and the pseudoknot. In fission yeast the protein subunit is called Trt1 (Nakamura, 1997), while the RNA component had resisted discovery for a long time. Only recently the Baumann and Zakian groups (Leonardi, 2008; Webb, 2008) identified fission yeast TER1 as a 1213 nucleotide RNA, which is significantly longer than its human counterpart TERC (which comprises 451 nucleotides). Ter1 mediates interactions between Trt1 and Est1, Est1 being a telomerase accessory protein. Homologous Est telomerase accessory proteins are also found associated with telomerase in *S. cerevisiae* and humans. In *S. cerevisiae* it has been shown that Est proteins are not required for telomerase catalysis in vitro (Cohn, 1995; Lingner, 1997), but *in vivo* they have been shown to be essential and their deletion leads to progressive telomere shortening, known as the ever-shorter-telomeres (est) phenotype (Lundblad, 1989; Lendvay, 1996). In fission yeast, it has been shown that the interaction of Est1 with the telomere binding protein Ccq1 mediates telomerase recruitment to the telomeres (Webb and Zakian, 2012) (see also **section 1.4.3**). Moser and colleagues have shown that phosphorylation of Ccq1 by Tel1 and/or Rad3 is required for the Ccq1-Est1 interaction which goes in line with the failure to recruit telomerase in a Ccq1 mutant that cannot be phosphorylated (Moser *et al.*, 2011).

In 1997, Prescott and Blackburn found, that *S. cerevisiae* telomerase functions at least as a dimeric complex with two active sites and that it can bind multiple telomere substrates simultaneously. Similarly, Wang *et al.* (2002) showed also that telomerase in the ciliate *Euplotes crassus* has the ability to form dimers and furthermore human telomerase was shown to oligomerize (Beattie *et al.*, 2001). Two models are proposed for the coordinated action in a telomerase dimer where in one model the dimeric complex binds two substrates simultaneously with its two active sites and in a second model the two catalytic sites act sequentially on a single substrate in a template switching hand-off mechanism (Prescott and Blackburn 1997, Wenz *et al.*, 2001). Whether telomerase in fission yeast also exists as a multimeric complex remains to be determined, but it is certainly likely given the conservation of multimerization from ciliates over yeast to humans.

1.3 The shelterin complex

DNA ends of chromosomes are threatened by the end-replication problem and the DNA-damage repair machinery. In order to regulate the action of telomerase which deals with telomere shortening, and inhibit the DNA-repair machinery, a number of proteins need to assemble at the telomeric repeats. These proteins together form a large complex, found in many species, and is known as the shelterin complex (**figure 1.4**).
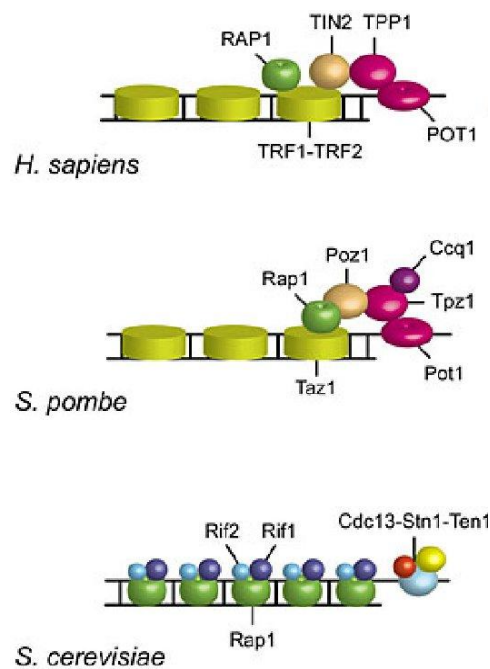
Figure 1.4. Shelterin complexes across different species.
While the complexes in humans and fission yeast *(S. pombe)* share several similarities, the *S. cerevisiae* complex has strongly diverged in composition. Figure adapted from Longhese *et al.*, 2012.

The shelterin complex is best described in *S. cerevisiae,* fission yeast (*S. pombe*) and humans. While the protein composition in *S. cerevisiae* has strongly diverged, the complexes in fission yeast and humans are more alike. Since my thesis centers on the fission yeast shelterin complex, the introduction will mainly focus on this complex and I will and the comparison to the human shelterin complex.

1.3.1 Fission yeast shelterin

In fission yeast, the shelterin complex is composed of Taz1-Rap1-Poz1-Tpz1-Ccq1-Pot1. Some of the components have only recently been discovered and are not well characterized as of yet. Taz1, a 663 amino acid protein, binds to telomeric repeats via its N-terminal Myb-related DNA binding domain, also known as the telobox (Bilaud *et al.*, 1996) which is also found in the homologous human telomeric proteins TRF1 and TRF2 (Cooper *et al.*, 1997) (**figure 1.5**). Taz1 has been shown, as TRF1 and TRF2, to form homodimers in order to enable binding double-strand telomeric repeats (Spink *et al.*, 2000). In TRF1 and TRF2, dimerization occurs via the TRF homology domain (TRFH) (Broccoli *et al.*, 1997), which can be observed in the X-ray structure (Fairall *et al.*, 2001). The TRFH domain has also been predicted in Taz1 and it is assumed that also here, this domain is responsible for dimerization (**figure 1.5**) (Li *et al.*, 2000). *Taz1Δ* fission yeast strains generate severely elongated telomeres suggesting a role for Taz1 in restricting telomere elongation (Cooper *et al.*, 1997). Taz1 has furthermore been implicated in efficient replication fork progression through the telomere, and loss of Taz1 leads to stalled replication forks (Miller *et al.*, 2006).
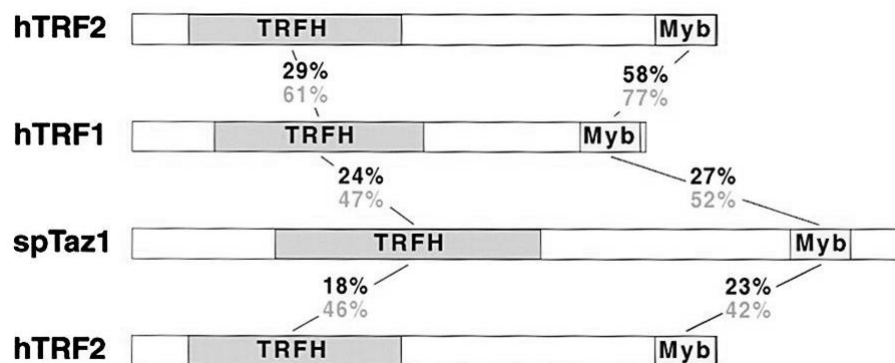
Figure 1.5. Sequence alignment of spTaz1, hTRF1 and hTRF2.

The TRF homology module and the Myb DNA-binding domain are found in each of the proteins. The sequence identity (black) and similarity (grey) are given in percent. Adapted from Li *et al.*, 2000.

Rap1, a 693 amino acid protein binds via its C-terminus to the central region of Taz1 (residues 365-396) and has been suggested to bind as a dimer to dimeric Taz1 (Chikashige and Hiraoka, 2001). Despite similarities between different species, Rap1 in *S. cerevisiae* functions differently in that it directly binds to the DNA, while in fission yeast and in humans, Rap1 does not have DNA binding ability and instead indirectly binds to the

telomeres (see **figure 1.4**). Fission yeast Rap1 contains an N-terminal BRCT domain followed by a Myb-domain, both also present in *S. cerevisiae* and human RAP1 (**figure 1.6**). BRCT domains (first described in the breast cancer protein BRCA1, Koonin et al, 1996) are conserved domains frequently found in DNA damage-responsive cell cycle checkpoint proteins (Bork *et al.*, 1997). BRCT domains bind phosphoproteins and phosphopeptides containing the pS-X-X-F recognition motif (Glover *et al.*, 2004) but have also been reported to allow protein-protein recognition via direct BRCT-BRCT contacts (Zhang *et al.*, 1998). Furthermore, the Crb2 fission yeast DNA-damage response protein has been shown to homodimerize via its BRCT domain (Willson *et al.*, 1997). The hypothesis that spRap1 dimerizes via its BRCT domain remains to be proven.

Rap1Δ strains show long telomeres similar to *Taz1Δ* (Chikashige *et al.*, 2001). Furthermore, Rap1 has been shown to collectively with Taz1 regulate 3' overhang formation as well as prevent telomere fusions (Miller *et al.*, 2005). Rap1 also has functions outside telomere length regulation: The meiotic proteins Bqt1 and Bqt2 together form a link between Rap1 and spindlepole body protein Sad1, and these interactions have been shown to be essential for bouquet formation (bundle arrangement of chromosomes) which plays an important role in homologous chromosome pairing and progression of meiosis (Chikashige *et al.*, 2006).
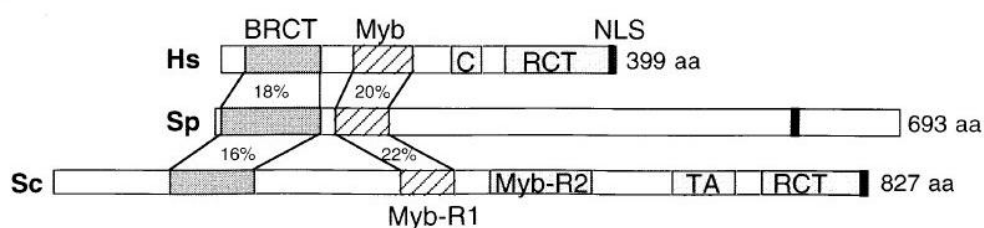
Figure 1.6. Sequence alignment of human, fission yeast and *S. cerevisiae* Rap1.

The BRCT domain and the Myb-domain are shared features. Adapted from Chikashige and Hiraoka, 2001.

In 2011, Chen *et al.* published the NMR solution structure of the C-terminal domain of Rap1 in complex with a peptide of Taz1 (**figure 1.7**). It revealed remarkable similarities of the architecture and interaction of Rap1-Taz1 compared to the human homologues RAP1-TRF2, whose crystal structure was solved in the same study.
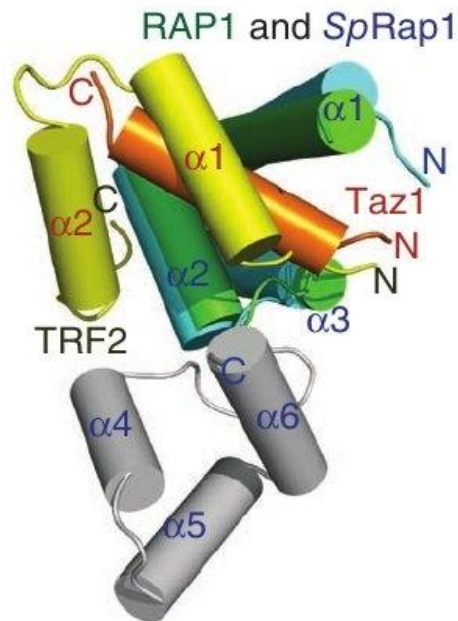
Figure 1.7. Superposition of spRap1$^{639-693}$-Taz1$^{365-396}$ and hRAP1$^{303-399}$-hTRF2$^{275-316}$.
SpRap1 is colored in blue, hRAP1 in green and grey, spTaz1 in orange and hTRF2 in yellow. (Chen *et al.*, 2011).

Poz1, a short 249 amino acid protein binds to Rap1 (Miyoshi *et al.*, 2008). The protein sequence of Poz1, allowed no predictions about sequence motifs or structural domains. *Poz1Δ* cells exhibit very long telomeres (up to 2000bp). Thus Poz1 has been proposed to be implicated in telomere shortening by negatively regulating telomerase. This is furthermore underlined by *Poz1Δ*/*TrtΔ* (telomerase) double mutants that do not show the long telomere effect (Miyoshi *et al.*, 2008). Poz1 binds to a 508 amino acid protein called Tpz1.

Tpz1 contains a predicted N-terminal OB-fold (oligonucleotide/oligosaccharide binding domain, Murzin *et al.*, 1993) of which sequence and secondary structure elements align well with the structure of the OB-fold of the human Tpz1-homologue TPP1 (**figure 1.8**) (Wang *et al.*, 2007). Tpz1, in turn, makes interactions with Poz1 via its C-terminal residues (379-508) and the interaction with Pot1 (described later) via the N-terminal OB-fold (Miyoshi *et al.*, 2008). *Tpz1Δ* cells do not grow or form only very small colonies. Surviving cells exhibit a complete loss of telomeric DNA as well as self-circularization, similar to *Pot1Δ* cells (discussed below), indicating that Tpz1 is required for the protection of DNA ends against unwanted fusions (Miyoshi *et al.*, 2008). Furthermore, Tpz1 is found at the core of the shelterin complex and it is not surprising that its removal causes major telomere dysfunction.
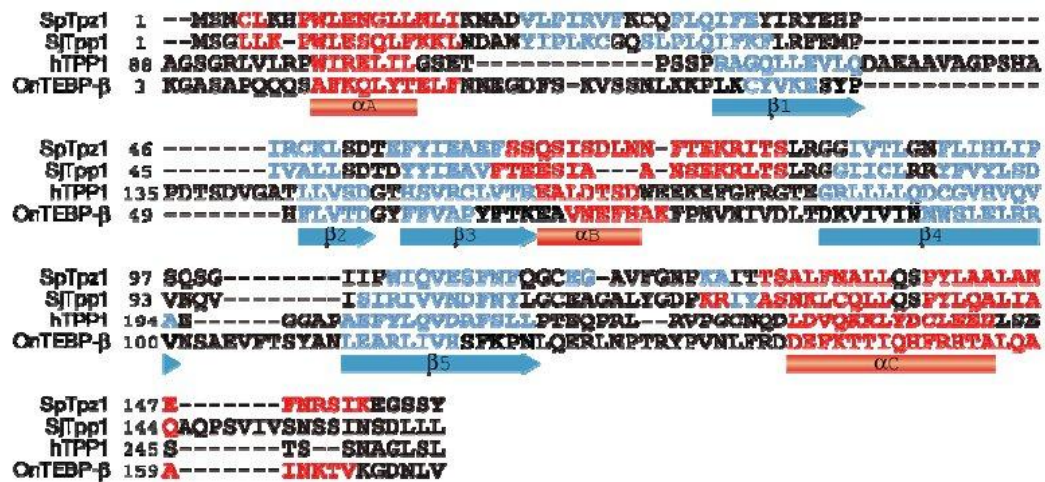
Figure 1.8. Sequence alignment of spTpz1 and its homologues.

SjTpp1 is from *Saccharomyces japonicus*, a closely related fission yeast species. OnTEBP-β is the telomere end-binding protein from *Oxytricha nova,* a ciliated protozoan (Horvath *et al.*, 1998, Gray *et al.*, 1991). Red (bar) and blue (arrow) letters represent α-helices and β-strands, respectively. SpTpz1 and sjTpp1 1 share the α−β−β−β−α−β−β−α structure that is commonly observed in OB folds of telomere-related proteins, including hTPP1 and OnTEBP-β.

The binding site for Ccq1 is also found within the C-terminal region (residues 379-508) of Tpz1. Ccq1 is a 735 amino acid protein. Residues 1-436 are predicted to contain a HEAT repeat domain and its C-terminal residues 500-720 are predicted to form a coiled-coil domain (Flory *et al.*, 2004). Ccq1 has been implicated not only in telomere regulation, but also plays a critical role during meiosis (Flory *et al.*, 2004). To date, no proteins homologous to Ccq1 have been found in *S. cerevisiae* or humans. *Ccq1Δ* cells show shortened telomeres (by ~200bp) which implicated Ccq1 as a positive regulator of telomerase and therefore responsible for telomere elongation (Miyoshi *et al.*, 2008). Furthermore, Ccq1 has been shown to be required for proper telomerase recruitment to the telomeres (Webb and Zakian, 2012) (see also **section 1.4.3**).

The sixth component of the fission yeast shelterin complex is a protein called Pot1 (Baumann and Cech, 2001). Pot1 is 555 amino acids long and binds via its C-terminal region to the   N-terminal OB-fold in Tpz1 (Miyoshi *et al.*, 2008). Pot1 binds to the single-stranded G-rich 3' overhang of the telomeric repeats via its N-terminus and the crystal structure thereof was solved by Lei *et al.* in 2003 (**figure 1.9**).
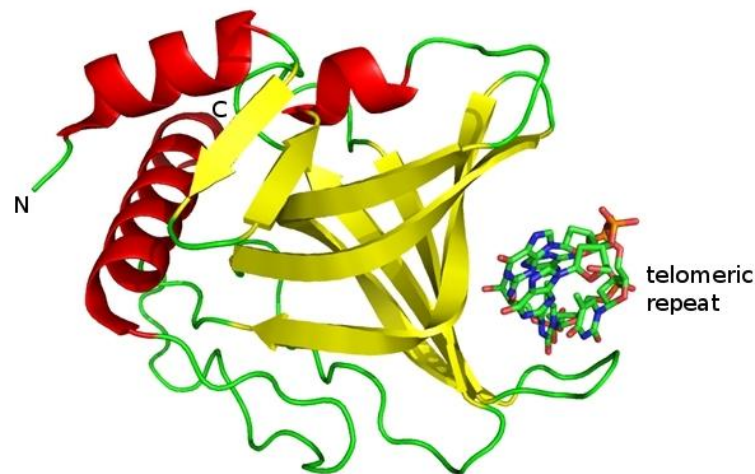
Figure 1.9. The structure of spPot1[1-185] (N-terminal OB-fold) in complex with a ssDNA telomeric repeat (GGTTAC). α-Helices of Pot1 are colored in red, β-sheets in yellow. DNA is shown as a stick model with carbon in green, nitrogen in blue, oxygen in red and phosphorus in orange (Lei *et al.*, 2003).

While the structure of Pot1 solved only shows the N-terminal OB-fold bound to a telomeric repeat, it was later shown that fulllength Pot1 has two N-terminal OB-folds and binds to two telomeric repeats via OB1 and OB2, thereby increasing its binding affinity (Trujillo *et al.*, 2005). Binding via two OB-folds has been suggested to accommodate telomeric substrates with different spacer lengths: As opposed to mammalian telomeres, fission yeast telomeres contain spacers of different lengths in between the telomeric repeats. (reviewed in Zakian, 1995). The structure of human POT1 revealed how two N-terminal OB-folds bind ssDNA (Lei *et al.*, 2004). Recently it was shown, that besides binding two telomeric repeats via OB1-OB2, spPot1 can dimerize upon binding to the DNA, thereby binding two telomeric repeats via the OB1 of each molecule (Nandakumar and Cech, 2012). The Cech group reported, that when the 3' overhang is very short (too short to be accommodated by OB1 and OB2), then two spPot1 molecules bind two telomeric repeats in a dimeric fashion via their OB1-folds. One explanation states that this strategy could have evolved to maintain end-protection of short overhangs that cannot be bound by a Pot1 monomer due to the incomplete DNA-binding template. Furthermore, they suggested that a Pot1-dimer could serve as a platform to recruit two Tpz1 and two Ccq1 molecules that help recruit dimeric telomerase (should it be dimeric in fission yeast, see **section 1.2**), which in turn lengthens the telomeres with short overhangs. Dimerization has also been observed for another telomeric ssDNA-binding protein, scCdc13, found in *S. cerevisiae*, which in turn also

possesses a heterogeneous telomere sequence (Sun *et al.*, 2011). *Pot1Δ* strains show a complete loss of telomeric DNA and these cells have been shown to survive by chromosome circularization (Baumann and Cech, 2001). Thus Pot1 is essential for the protection of telomere ends from unwanted fusion events and degradation.

Taken together, the fission yeast shelterin complex can be divided into the double-strand binding subcomplex Taz1-Rap, and the single-strand 3' overhang binding subcomplex Pot1-Tpz1-Ccq1, with Poz1 bridging the two subcomplexes.

1.3.2 The human shelterin complex

While the human shelterin complex (Liu *et al.* 2004b; O'Connor *et al.* 2004; Ye *et al.* 2004a) has a composition that resembles the fission yeast complex in many ways, some components have diverged. The major difference being that the human double-strand telomeric repeats are bound by two dimeric proteins instead of one (Taz1 in fission yeast), and referred to as TRF1 and TRF2 (Zhong *et al.*, 1992; Chong *et al.* 1995; Ludérus *et al.*, 1996; Bilaud *et al.*, 1996). These proteins share a common domain structure consisting of the TRF homology domain (TRFH) and a C-terminal Myb DNA-binding domain that have also been predicted to exist in spTaz1 (**figure 1.5**). The structure of the Myb DNA-binding domain of TRF1 was solved in 2001 (Nishikawa *et al.*, 2001) (**figure 1.10**).
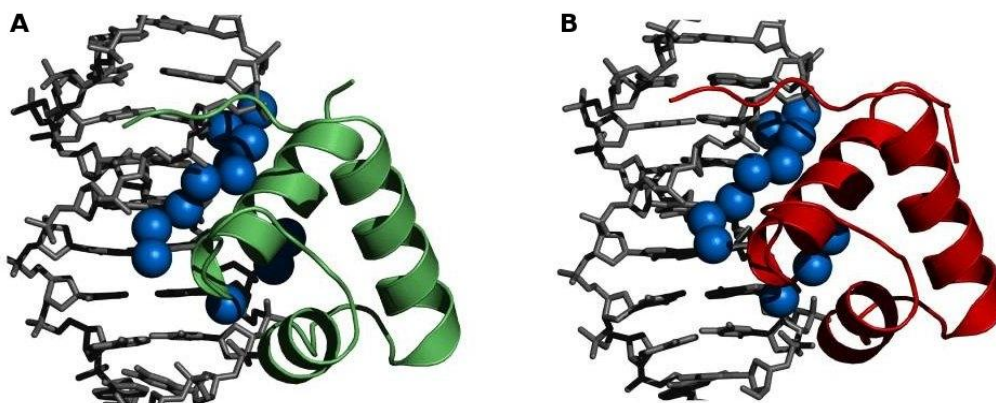
Figure 1.10. The structures of the Myb DNA-binding domains of A) hTRF1 and B) hTRF2 in complex with DNA.

While the sequence identity between TRF1 and TRF2 is 56%, the structures and contacts are almost identical. TRF1 is in green, TRF2 in red. DNA is presented in grey. The blue spheres represent conserved water molecules that mediate contacts between the protein and the telomeric repeat sequence and increase the specificity and affinity (Court *et al.*, 2005).

TRF1 and TRF2 have the ability to dimerize via the TRFH domain (Broccoli *et al.*, 1997) and the structures of the dimerization modules have been solved (Fairall et al 2001). Recently, the structure of TRF1 was solved in complex with a peptide of its binding partner TIN2 and TRF2 with a peptide from the 5' exonuclease Apollo (**figure 1.11**) (Chen *et al.*, 2008; Demuth *et al.*, 2004). Apollo is a shelterin associated factor and is essential for proper replication of telomeres as well as end protection (Touzote *et al.*, 1999). TRF1 and TRF2 are subject to extensive post-translational control. For TRF1, nine different sites have been reported to be phosphorylated by six different kinases which modulates its binding to telomeric DNA and its stability (Walker and Zhu, 2012). The majority of the sites localize to a flexible linker region between the DNA-binding Myb-domain and the TRFH domain. Control of TRF2 by phosphorylation is less well understood. Phosphorylation of T188 in TRF2 by ATM has been shown to be implicated in the DNA-damage response (Huda *et al.*, 2009). Furthermore TRF1 and TRF2 have been shown to be ubiquitylated which targets them for protease mediated degradation (Chang *et al.*, 2003; Fujita *et al.*, 2010). Many more post-translational modifications have been reported that together weave an extensive regulation network for TRF1 and TRF2 (Walker and Zhu, 2012).
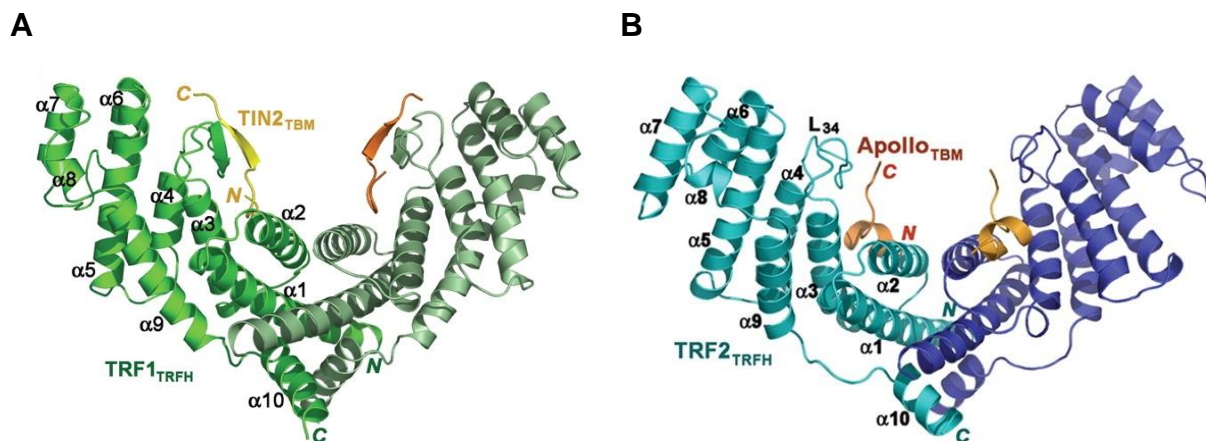
Figure 1.11. Crystal structures of the TRFH-domains of hTRF1 and hTRF2.

The structures of **A)** hTRF1[62-268] in complex with hTIN2[256-276] and **B)** hTRF2[86-287] in complex with hApollo[496-532] . Chen *et al.*, 2008.

TIN2 (Kim *et al.*, 1999) is a small protein that simultaneously binds and bridges TRF1 and TRF2 and is viewed as the functional homologue of Poz1, despite any obvious sequence or predicted structural similarity between them. Given the sequence and structural similarity between TRF1 and TRF2, it was assumed that TIN2 would bind the TRFH domain

in both proteins. While TIN2 binds in the TRFH domain in TRF1, the TRFH domain in TRF2, however, is not required for binding to TIN2. Instead, TRF2[352-365], a short region that lies C-terminally of TRF2H is required for TIN2 interaction. Thus, the same molecular surface that in TRF1 binds TIN2, binds the 5' exonuclease Apollo in TRF2, while TRF2H cannot bind TIN2. The Lei group also showed that TRFH in both, TRF1 and TRF2, attracts a number of other shelterin-associated proteins containing a F/Y-X-L-X-P sequence motif (also present in TIN2 and Apollo) but with distinct specificities (Chen *et al.*, 2008). These results suggest that the TRFH domains of TRF1 and TRF2 function as protein docking sites that recruit different telomeric factors with distinct functions.

Besides binding TIN2, TRF2 also binds RAP1 (Li *et al.*, 2000). Human RAP1, as its homologues in *S. cerevisiae* and fission yeast, contains an N-terminal BRCT domain followed by a Myb-domain (**figure 1.6**). Like fission yeast Rap1, human RAP1 does not bind DNA directly as Rap1 in *S. cerevisiae* does, but rather binds the telomeres indirectly through interaction with TRF2. The interaction is mediated by the very C-terminal residues 303-399 of RAP1 and a central region of TRF2 which corresponds to residues 275-316. The X-ray structure of this interaction (described in **section 1.3.1** (**figure 1.7**)) strongly resembles the interaction of the homologues in fission yeast, Rap1-Taz1. RAP1 has been implicated in inhibiting chromosome fusions by non-homologous end joining and is recruited to the telomere by TRF2 (Sarthy *et al.*, 2009).

TPP1 is the human homologue of fission yeast Tpz1 (Liu *et al.*, 2004, Ye *et al.*, 2004) and binds to TIN2 via its C-terminal region (334-544). An OB-fold had been predicted in its N-terminus and was later verified by a crystal structure (**figure 1.12a**) (Wang *et al.*, 2007). A remarkable similarity of TPP1 to *Oxytricha nova* telomere end-binding protein TEBP-β was observed despite the low sequence identity of only 11% (**figure 1.12b**). This includes also three peripheral α-helices outside the OB-fold, suggesting that TPP1 is not merely another OB-fold containing protein, but shares common ancestry with onTEBP-β. Given the remarkable conservation from human to *Oxytricha nova* and the prediction of the OB-fold in the orthologous spTpz1 it can be assumed that Tpz1 might have a similar architecture, although no structure thereof has been solved to date. TPP1 together with POT1 have been shown to increase the processivity of telomere extension by telomerase (Wang *et al.*, 2007). Furthermore, it seems that a subdomain of human telomerase (TERT) called DAT

(dissociates activities of telomerase), interacts with TPP1 thereby recruiting telomerase to the telomere (Armbruster *et al.*, 2004; Zaug *et al.*, 2010).
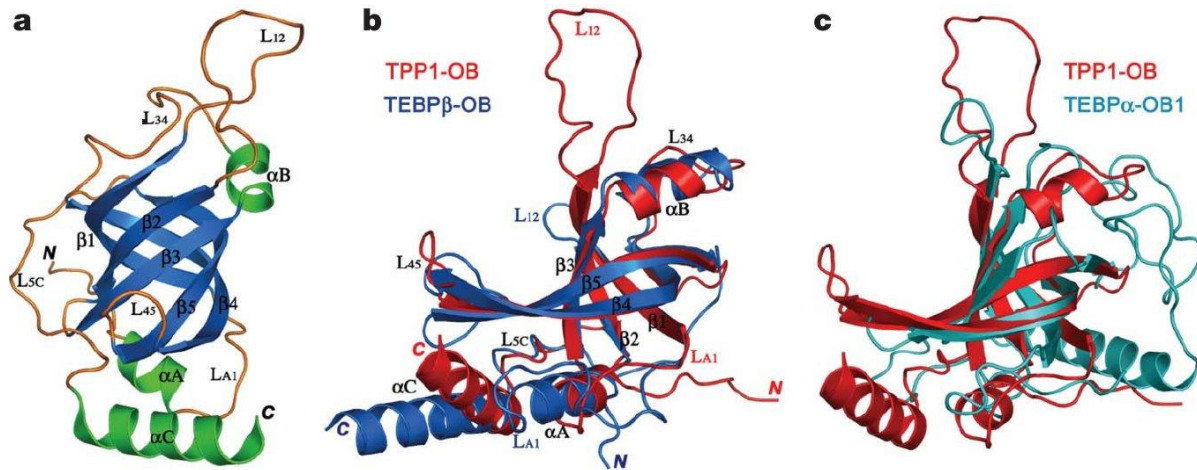
Figure 1.12. Comparison of hTPP1 to its homologues.

a) The crystal structure of hTPP1-OB (87-344). **b)** Superposition of hTPP1-OB onto the crystal structure of the OB-fold of onTEBP-β. **c)** the more distantly related second *Oxytricha nova* telomere end-binding protein TEBP-α. (Wang *et al.*, 2007).

The final component of the human shelterin complex is POT1 (Baumann and Cech, 2001). Human POT1, similar to fission yeast Pot1, binds the telomeric single-strand G-rich 3' overhang. The structure of POT1 (**figure 1.13**) revealed two N-terminal OB-folds that both bind to the ssDNA (Lei *et al.*, 2004), a binding mode that has also been suggested for fission yeast Pot1 (Croy *et al.*, 2006). POT1 has been predicted to contain a third OB-fold in its C-terminus which constitutes an arrangement similar to the three OB-folds found in the *Oxytricha nova* homolog TEBP-α, suggesting common ancestry. However, the architecture of the chromosome ends in humans and *Oxytricha nova* have diverged. While humans have a long 3' G-rich single-strand overhang, the DNA-ends in *Oxytricha nova* merely consist of a 16-nucleotide, 3'-terminal, single strand $T_4G_4T_4G_4$ extension. This results in a different arrangement of the single-strand binding complexes: TEBP-α together with TEBP-β sits as a single complex at the very end of the on the 3' overhang and buries the 3' DNA-end within the complex preventing any access by telomerase (**figure 1.14**) (Horvath *et al.*, 1998). In humans, the mechanism has become more complex and hPOT1 coats the entire single-strand overhang (Loayza *et al.*, 2004) and cooperatively with the other shelterin components, the 3' end is sequestered in a loop structure termed the T-loop (Griffith *et al.*, 1999) (see **section 1.4**).
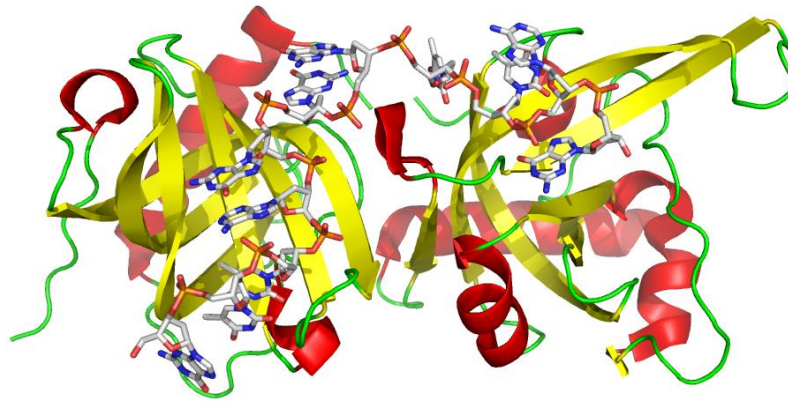
Figure 1.13. The crystal structure of hPOT1[5-299].

Telomeric ssDNA is bound by OB1 and OB2. A similar architecture has been predicted for fission yeast Pot1, though only the structure of spPot-OB1 has been solved. α-Helices of POT1 are colored in red, β-sheets in yellow. DNA is shown as a stick model with carbon in white, nitrogen in blue, oxygen in red and phosphorus in orange. (PDB=1XJV, Lei *et al.*, 2004. Figure generated with PyMOL, deLano, 2002).
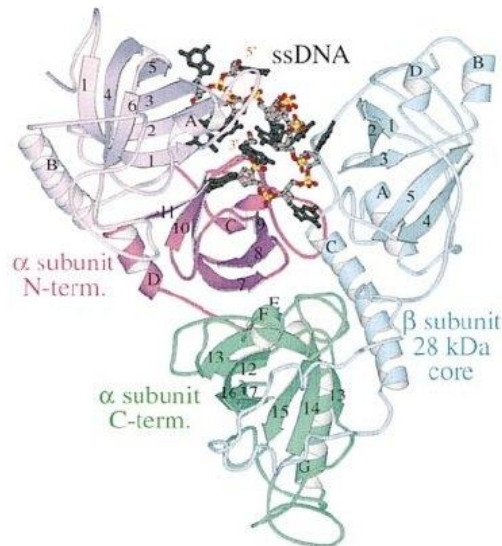
Figure 1.14. Crystal structure of the onTEBP-α/β telomere end-binding complex.

The very 3' end of the telomeric DNA is buried completely within in the complex. Unlike hPOT1 and spPot1 that coat the 3' overhang, this complex actually sits at the very end of the DNA (Horvath *et al.*, 1998).

1.4 Current models of telomere length regulation by shelterin

In yeast and mammals, the length of telomeres is maintained within a set size-range by a negative feedback loop that blocks the action of telomerase at individual chromosome ends. When a given telomere is very long, the telomerase pathway is restrained allowing the telomere to shorten. If the telomeric region is very short, the control is relaxed so that

telomerase can restore its length (Marcand *et al.*, 1997). Besides the telomerase enzyme, the shelterin complex is a key component in this regulation mechanism.

While much work has been carried out on uncovering details of the mechanisms of telomere length regulation in humans, less is known about the situation in fission yeast. In the following the current models in humans and fission yeast are discussed. Some aspects remain speculative and more evidence is required to prove that these mechanisms indeed function in the proposed manner.

1.4.1 Model for telomere length regulation in humans

The common model of telomere length regulation is a stochastic process and is also known as the "protein counting model" (Marcand *et al.*, 1997) because it is thought that cells determine the length of their telomeres by the number of shelterin complexes bound to their repeats. The longer the telomeres are, the more shelterin complexes are bound.

More shelterin complexes present at long telomeres (**figure 1.15a**) are thought to increase the loading of the single-strand binding component POT1 onto the telomeric overhang (de Lange, 2005). According to the model, the binding of shelterin complexes at long telomeres is thought to induce a conformational change in the DNA architecture where the chromosomal end is folded back forming the so called T-loop, a phenomenon that has been observed *in vitro* by electron microscopy (Griffith *et al.*, 1999). It was shown that the shelterin component TRF2 is able to induce T-loop formation of telomeric DNA in vitro (Stansel *et al.*, 2001). The 3' single-stranded overhang is thought to invade the duplex telomeric repeat array with a displacement D-loop at the invasion site (**figure 1.15b**). The T-loop is suggested to represent a protective cap at the chromosome end and in this conformation it is proposed that the DNA end is hidden inside the structure preventing access and its extension by the telomerase enzyme. In the case of short telomeres, fewer shelterin complexes are bound which decreases the loading of POT1 onto the single-strand overhang which in turn does not allow T-loop formation, leaving the DNA-ends in an open conformation, where they are readily accessible and extendable by telomerase.

A

B
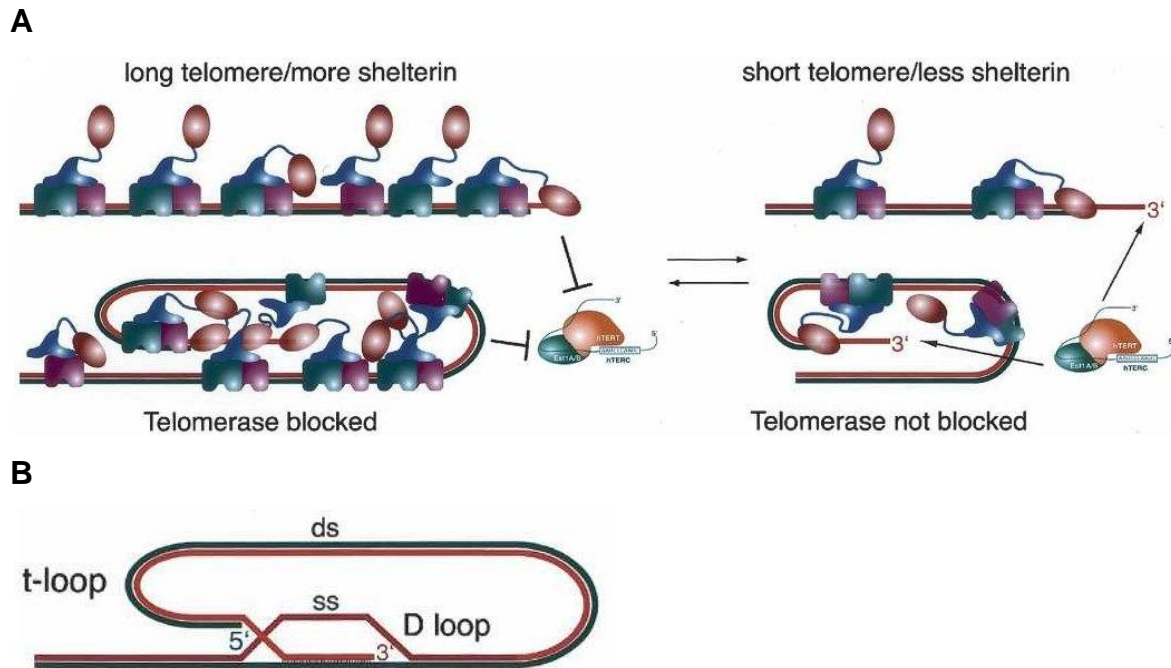
Figure 1.15. The "protein counting" and T-loop model for human telomere length-regulation.

A) Situation at the chromosome ends in the case of long and short telomeres. **B)** Schematic representation of the invasion of the duplex DNA by the 3' single-strand overhang. Adapted from de Lange, 2005.

1.4.2 Model for telomere length regulation in fission yeast

The mechanisms of telomere length regulation in fission yeast are less well studied. The Griffith group showed in 2004, that like TRF2, Taz1 is able to induce T-loop formation *in vitro*. However, the existence of T-loops *in vivo* has not yet been determined for fission yeast due to technical difficulties in isolating sufficient telomeric fragments for analysis (Tomaska *et al.*, 2004).

The common model states (**figure 1.16**) (Miyoshi *et al.*, 2008, Bianchi and Shore, 2008) that in the case of long telomeres, there are many double-strand telomeric repeats and therefore many Taz1-Rap complexes are bound. The telomeric repeats are counted allowing the cell to determine whether the telomeres are long or short. Given the high abundance of Taz1-Rap1 molecules bound on long telomeres, all Poz1-Tpz1-Ccq1-Pot1 complexes are captured through Rap1-Poz1 interaction and it has been suggested that in this conformation Ccq1 cannot carry out its recruiting/activating function on telomerase (see **section 1.4.3**), thereby inhibiting telomere extension and causing them to shorten. Hypothetically, the inhibition of Ccq1 function on telomerase could be achieved by a conformational change

transduced upon bridging of the two shelterin subcomplexes, the double-strand and single-strand binding halves, by Poz1. Additionally, should fission yeast telomeres also form T-loop structures, the DNA ends might fold back in a similar manner observed at human telomeres, physically sequestering the DNA ends from telomerase (Miyoshi *et al.*, 2008, Bianchi and Shore, 2008). A folding back of the budding yeast telomere had previously been reported by Strahl-Bolsinger *et al.* in 2007.
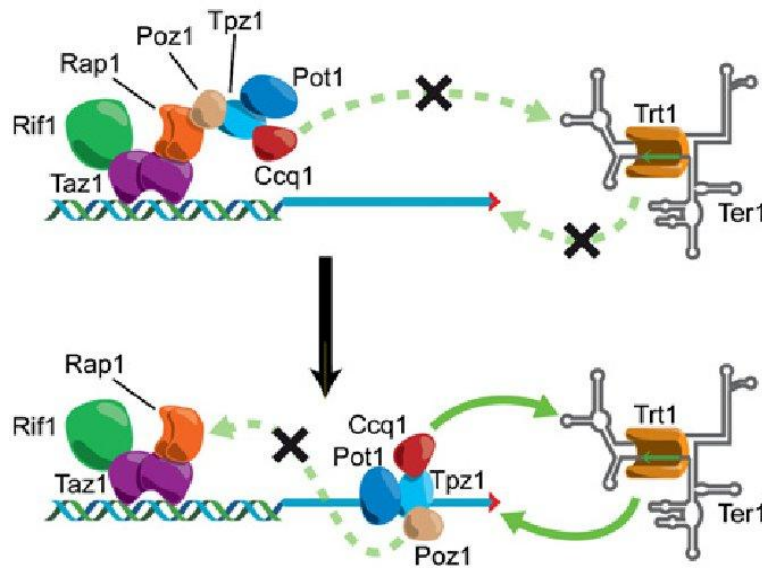
Figure 1.16. A model for telomere length regulation in fission yeast.

The upper figure represents the situation in the case of long telomeres, the lower the case of short telomeres. At long telomeres, many Taz1-Rap1 complexes are bound to the abundant number of telomeric repeats. Due to the resulting high concentration of Taz1-Rap1, the Poz1-Tpz1-Ccq1-Pot1 complexes are bound and sequestered. Given this conformation, the Ccq1 activating/recruiting function on telomerase is thought to be inhibited. At short telomeres, fewer Taz1-Rap1 complexes are present, resulting in free Poz1-Tpz1-Ccq1-Pot1 complexes bound to the 3' overhang and in this conformation Ccq1 can activate/recruit telomerase which in turn lengthens the telomere. Adapted from Bianchi and Shore, 2008.

In the case of short telomeres, fewer Taz1-Rap1 molecules are bound, not allowing binding and sequestering all available Poz1-Tpz1-Ccq1-Pot1 complexes. In this conformation, Ccq1 can carry out its recruiting/activating function on telomerase, thereby promoting the extension of the short telomere. Furthermore, should there be T-loop formation in fission yeast, the structure would be opened thereby presenting a free DNA-end readily accessible by telomerase (Miyoshi *et al.*, 2008, Bianchi and Shore, 2008).

1.4.3 Ccq1 and telomerase

The Est proteins (ever-shorter telomeres) are telomerase interacting accessory proteins and have been shown to be essential for telomere-lengthening in fission yeast as well as in *S. cerevisiae* and humans. Webb and Zakian showed (2012) that in fission yeast Est1 and Trt1 association with telomeres is dependent on the shelterin component Ccq1. They propose that telomerase recruitment to the telomeres is mediated by the interaction of Est1 with Ccq1 and that this interaction is required for the subsequent association of Trt1 with the telomere. Consistent with this model, another group reported that Ccq1 T93 phosphorylation by Tel1 (ATM) and/or Rad3 (ATR) is a critical step in telomerase-telomere association, as this phosphorylation is required for Ccq1-Est1-interaction (Moser *et al.*, 2011). Phosphorylation of T93 in Ccq1 for telomerase recruitment was reported shortly after also by Yamazaki *et al.* (2012). After the telomerase holoenzyme is recruited to the telomere, the interaction of Est1 with the telomerase RNA TER1 is thought to tether the complex to the telomere. Webb *et al.* suggest that after the recruitment of the telomerase holoenzyme, a conformational change takes place in the shelterin-complex that makes the end of the chromosome accessible to telomeric repeat addition by Trt1. Whether the conformational change occurs after recruitment of telomerase, or whether the conformational change (from the closed to the open state) needs to take place before Ccq1 becomes active and can recruit telomerase (as described in the model in **1.4.2**), requires further investigation.

1.5 The CST complex

In humans and fission yeast, shelterin is the primary telomere protein complex. In recent years an additional telomere-associated complex has been discovered. This complex is known as the CST complex and in humans consists of CTC1, STN1 and TEN1 with the later two being orthologous to Stn1 and Ten1 from *S. cerevisiae* (**figure 1.17**) (Surovtseva *et al.*, 2009; Miyake *et al.*, 2009; Casteel *et al.*, 2009). In *S. cerevisiae*, Stn1 and Ten1 assemble together with Cdc13 and bind the 3' overhang while the Rap1-Rif2-Rif1 complex binds to the double-strand telomeric repeats (Grandin *et al.*, 2001). Cooperatively, both complexes are essential for telomere integrity in *S. cerevisiae* (Bianchi and Shore, 2008; Giraud-Panis *et al.*, 2010; Anbalagan *et al.*, 2011). Deletion of any of the CST components leads to C-strand

degradation, DNA damage response and cell cycle arrest (Garvik *et al.*, 1995; Grandin *et al.*, 1997; Grandin *et al.*, 2001).
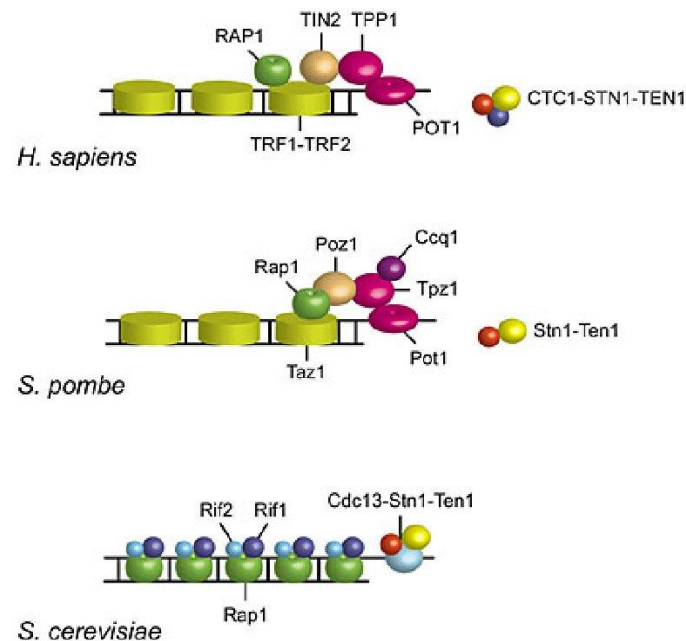
Figure 1.17. Protein-telomere complexes in different species.

Additionally to them being found at the *S. cerevisiae* telomeres, Stn1-Ten1 homologues have also been discovered in humans and fission yeast. Figure adapted from Longhese *et al.*, 2012.

For a long time it was thought that the CST complex was absent in humans because the homologues were difficult to detect in the human genome. It was assumed that in humans, the CST complex was replaced by POT1 which occupies the single-stranded 3' overhang. Following the discovery of the CST complex in humans (Surovtseva *et al.*, 2009; Miyake *et al.*, 2009; Casteel *et al.*, 2009) it has been suggested that multiple protein-telomere complexes work in concert in maintaining the chromosome ends (Price *et al.*, 2010). Knockdown of CST allows excessive telomerase activity, promoting telomere elongation. This suggests a role in restriction of telomerase. In light of the cooperativity of shelterin and CST, it has been shown that STN1 interacts with TPP1 (Wan *et al.*, 2009). Given the fact that TPP1 interacts with and stimulates telomerase, it has been proposed that STN1 is also involved in telomerase regulation, likely in negative regulation (Wang *et al.*, 2007; Xin *et al.*, 2007; Chen *et al.*, 2012).

In 2007, a CST-like complex was also discovered at the fission yeast telomeres (Martin *et al.*, 2007). The proteins spStn1 and spTen1 interact with each other and localize to

the telomeric 3' overhang. However, no interactions between fission yeast shelterin and the Stn1-Ten1 complex have been detected to date. It is likely that further interaction partners that link the two complexes remain to be discovered. Deletion of either Stn1 or Ten1 leads to a phenotype similar to that of *potΔ* strains: Loss of telomeric DNA is observed with survival achieved by chromosome circularization. The crystal structures of Stn1-Ten1 from fission yeast, *S. cerevisiae* and *C. tropicalis* revealed remarkable conservation amongst each other and with the Rpa2-Rpa3 complex (which bind single-strand DNA by means of its OB-folds) (Sun *et al.*, 2009; Bochkarev *et al.*, 1999). Thus, the CST complex has been suggested to function as a telomere-specific RPA complex (Gao *et al.*, 2007).

In summary, the discovery of the CST complex in organisms besides *S. cerevisiae*, reveals an aspect of telomere regulation that is not restricted to organisms that lack Pot1 afterall, but that is a more conserved mechanism than previously thought.

1.6 Telomere biology and disease

In recent years emerging evidence has been found linking age-related diseases to mechanisms of telomere maintenance. The first described telomere-mediated syndrome was dyskeratosis congenita (DC), a rare syndrome of premature aging (see Walne *et al.*, 2008 for a historical review). DC is thought to be caused by mutations in the gene encoding dyskerin which lead to compromised telomerase activity and short telomeres (Mitchell *et al.*, 1999). Since then, many more diseases have been attributed to the short telomere defect (Armanios and Blackburn, 2012) (**table 1.1**).

Gene	First diagnosis	Mechanism of telomere shortening	Reference
TERT; TR	Familial IPF Sporadic IPF Aplastic anaemia Autosomal-dominant dyskeratosis congenita Familial MDS-AML	• Partial loss-of-function • Haploinsufficiency	Hao *et al.*, 2005; Vulliamy *et al.*, 2001; Armanios *et al.*, 2005; Yamaguchi *et al.*, 2005; Alder *et al.*, 2008; Armanios *et al.*, 2007; Tsakiri *et al.*, 2007; Yamaguchi *et al.*, 2003; Calado *et al.*, 2009; Kirwan *et al.*, 2009; Vulliamy *et al.*, 2002; Fogarty *et al.*, 2003
DKC1	De novo dyskeratosis congenita X-linked recessive dyskeratosis congenita Hoyeraal-Hreiderasson syndrome	• Partial loss-of-function • Decreased TR stability and biogenesis	Heiss *et al.*, 1998; Mitchell *et al.*, 1999; Knight *et al.*, 1999; Yaghmai *et al.*, 2000
TINF2	De novo dyskeratosis congenita Autosomal-dominant dyskeratosis congenita Hoyeraal-Hreiderasson syndrome Revesz syndrome	• Not completely understood • Probably dominant-negative mutations	Walne *et al.*, 2007; Walne *et al.*, 2008; Chiang *et al.*, 2004
NOP10	Autosomal-recessive dyskeratosis congenita	Presumed loss of telomerase function	Walne *et al.*, 2007
NHP2	Autosomal-recessive dyskeratosis congenita	Presumed loss of telomerase function	Vulliamy *et al.*, 2008
TCAB1	Autosomal-recessive dyskeratosis congenita	Impaired TR trafficking; loss-of-function	Zhong *et al.*, 2011
CTC1	Coats plus syndrome Autosomal-recessive dyskeratosis congenita	Loss-of-function	Gu *et al.*, 2012; Keller *et al.*, 2012; Anderson *et al.*, 2012; Polvi *et al.*, 2012

Table 1.1. Disease spectrum and mechanism of telomere shortening in telomere syndromes

AML, acute myeloid leukaemia; CTC1, conserved telomere protection component 1; DKC1, dyskeratosis congenita 1; IPF, idiopathic pulmonary fibrosis; MDS, myelodysplastic syndrome; TCAB1, telomerase Cajal body protein 1; TINF2, TRF1-interacting nuclear factor 2. Table adapted from Armanios and Blackburn, 2012.

Furthermore, research in the field of telomere dysfunction has revealed the role of telomere regulation in the context of cancer. It has been shown that reactivation of telomerase in cells harboring critically short telomeres leads to the acquisition of unlimited proliferative capacity as is found in cancer cells (Ding *et al.* 2012). The current opinions on cancer in telomere biology are discussed in the following.

1.6.1 A "molecular clock"

While for germ line cells a high proliferative capacity is desired, somatic cells need to be limited in their replicative potential. It has been proposed, that the loss of telomeric repeats, when reaching a critical telomeric length (~ 3kb), induces a DNA-damage signal that results in the exit of the cell cycle and replicative senescence (Campisi, 1997). This lays the foundation for a "molecular clock" that determines the replicative life of a cell (Vaziri *et al.*, 1998). Consequently, cells can only replicate a certain number of times before they settle in a state of senescence and seize to proliferate. The fact that telomeres progressively shorten in somatic cells comes from the repression of telomerase in this cell type (Allsopp *et al.*, 1995). On the other hand telomerase is active in cells of the germinal line which have long telomeres (~10kb) ensuring continuous proliferative capacity (Hastie *et al.*, 1990). The importance of telomerase expression was most definitively shown by Bodnar *et al.* who reintroduced telomerase into primary human fibroblasts and endothelial cells that lack telomerase activity. They found elongated telomeric tracts resulting in a significant increase in their replicative life (Bodnar *et al.*, 1998).

1.6.2 Telomeres in cell immortalization and cancer

In normal somatic cells telomerase is repressed and telomeres progressively shorten, until they reach a critical length where after they enter a state of senescence. The mechanism of cell cycle exit in these cells is maintained by signals that activate the pathway of the tumor suppressor genes p53 and p16/Rb (Shay *et al.*, 1991). It is thought that inactivation of these pathways allows bypassing of replicative senescence and further telomere attrition. This process is characterized by a two-stage model of cell growth (see Shay *et al.*, 2011 for a recent review): The stage when all checkpoints are intact and regular replicative senescence occurs, is referred to as the mortality stage 1 (M1). Upon bypassing M1 by disruption of the p53 and/or p16/Rb pathways (caused by endogenous mutations or exogenous mutagens, Hollstein *et al.*, 1991, Counter *et al.*, 1992; Rogan *et al.*,1995), cells proliferate further until they reach the crisis stage (M2) (Girardi *et al.*, 1965, Wright *et al.*, 1989) (See also Noble *et al.*, 2004). Intriguingly, the crisis stage can act as a line of defense against cancer development because genomic instability that arises during this stage and DNA-damage

signaling kills off the vast majority of cells. However, rare immortalized clones can arise from crisis stage by telomere lengthening through reactivation of telomerase or in the absence of telomerase through ALT ("alternative lengthening of telomeres" without the necessity of telomerase via inter-telomeric recombination) (Murnane *et al.*, 1994; Dunham *et al.*, 2000). Reactivation of telomerase is thought to occur by mutation of additional oncogenes and tumor suppressor genes. This is supported by a study that showed that telomerase is active in up to 90% of human tumors with no telomerase activity found in any of the somatic tissue analyzed (Dhaene *et al.*, 2000).

While the mechanisms linking telomere dysfunction and cancer are complex and require more investigations, the progressive understanding of the interplay of telomere associated proteins lays the foundation for research on the manipulation of telomerase activity as a potential therapeutic target against cancer. Currently, therapeutics directed against telomerase have yet to show success in the clinic, thus further underlining the importance of gaining as much understanding as we can on the mechanisms involved (see Tárkány *et al.*, 2008 for a review).

1.7 Aim of this work

The telomerase enzyme adds the telomeric repeats to the ends of linear chromosomal DNA, and these in turn recruit and bind a large number of factors that protect telomeres from the DNA repair machinery, and maintain telomeric length. Some of these telomeric factors form the telomere binding complex known as shelterin. The proteins forming the shelterin complex have been shown to be essential for proper telomere maintenance. Here we are particularly interested in the molecular details of shelterin architecture and how they function in regulating telomeric length. The current models that explain telomere length regulation in fission yeast and humans still remain partially speculative and more work needs to be done in order to verify their correctness. Our goal is to improve our molecular understanding of shelterin components and their interactions through a combination of structural biology and *in vivo* manipulation of fission yeast.

To date, limited structural information is available on the shelterin proteins. Structures determined include the OB1 domain of spPot1 (Lei *et al.*, 2003) and the hPOT1 DNA-

binding domain (Lei *et al.*, 2004), the structures of the hTRF1 and hTRF2 DNA-binding domain (Court *et al.*, 2005) as well as their dimerization domain (TRF-homology domain) in complex with a peptide of hTIN2 and hApollo respectively (Chen *et al.*, 2008). Furthermore, the structure of the OB domain of hTPP1 has been determined (Wang *et al.*, 2007) as well as a fusion construct of the C-terminus of spRap1 bound to a peptide of spTaz1 (Chen *et al.*, 2011). While for some of the remaining components and domains structure predictions can be used, others remain enigmatic. One such poorly characterized protein is Poz1, a fission yeast shelterin component that has only recently been identified and supposedly is the human functional homolog to TIN2 (Miyoshi *et al.*, 2008). It is the structure and function of Poz1 around which my PhD-thesis is centered.

The current model for fission yeast length regulation proposes that the shelterin complex exists in two different conformations: on short telomeres, an open conformation where Taz1-Rap complexes are bound to the double-strand telomeric repeats and Poz1-Tpz1-Ccq1-Pot1 complexes are bound to the 3' single-strand overhang. On long telomeres, due to the high concentration of Taz1-Rap1, the Poz1-Tpz1-Ccq-Pot1 complexes are bound to Rap1 thereby forming the closed complex. In the open conformation, Ccq1 has been proposed to carry out its recruiting/activating function on telomerase whereas in the closed conformation this activity is inhibited. Additionally it is thought that in the closed conformation, the DNA-end folds back which physically sequesters the 3'ends and prevents them from being extended by telomerase. In this model, the Poz1-protein plays a crucial role, as it is the bridge between the double-strand and the single-strand binding part of shelterin and based on the model this is where the complex opens and closes.

By solving the structure of spPoz1 in complex with its binding partner spTpz1, we anticipated to uncover the molecular architecture of this complex, the function of which we aimed to address in vivo through mutation. Through this approach we set out to improve current models of telomere homeostasis.

Chapter 2

Material and methods

2. Material and methods

2.1 Boundary optimization by limited proteolysis and subsequent identification by LC/MS

2.1.1 Limited proteolysis of spPoz1$^{30-249}$+Tpz1$^{360-508}$ and "in-gel" protein identification

The protease trypsin was added to spPoz1$^{30-249}$+Tpz1$^{360-508}$ at 10 mg/ml in steps of increasing concentration from 0.003 to 3 % in a volume of 20 µl. The reactions were kept on ice for 10 minutes, after which the reaction was stopped by adding PMSF (phenylmethanesulfonylfluoride) (Sigma) to a final concentration of 5 mM. The samples were then immediately boiled at 110 °C for 5 minutes and quickly loaded on SDS-PAGE.

For "in-gel" protein identification, bands were excised from the SDS-PAGE gel and subjected to mass-spectroscopy. For generation of peptide fragments, the excised gel pieces were digested with trypsin (Promega) at 19 ng/µl overnight at 37°C. The samples were analyzed on a MALDI-TOF (Bruker, Ultraflex2) by Daniel Hess and Ragna Sack from the FMI protein analysis facility.

2.1.2 Identification of co-migrating fragments by size exclusion chromatography

SpPoz1$^{30-249}$+Tpz1$^{360-508}$ is generally stored in 50 mM Tris-HCl, 500 mM NaCl and 0.25 mM TCEP (tris(2-carboxyethyl)phosphine) unless otherwise stated. A volume of 500 µl of spPoz1$^{30-249}$+Tpz1$^{360-508}$ at 10mg/ml was incubated with 0.1 % trypsin for 10 minutes on ice. The reaction was stopped by adding PMSF to a final concentration of 5 mM where after the sample was immediately injected on a SD200 10/300 gelfiltration column (GE Healthcare).

2.1.3 Boundary identification of Tpz1-fragments by LC/MS

For precise boundary identification in solution we performed liquid chromatography coupled to mass spectroscopy (LC/MS) in collaboration with Daniel Hess and Dominique Klein from the FMI protein analysis facility. For separation of the fragments in solution, the proteolysed sample containing the fragments of interest was injected on a PLRP L50mm ID1mm 300A 5 um column (Polymer Laboratories) and run in an acetonitrile gradient on a HPLC system (Agilent 1100) coupled to a 4000 Q TRAP LC/MS/MS system (AB Sciex).

2.2 Cloning

Plasmids containing cDNA for full-length fission yeast *(S. pombe)* Tpz1 and Poz1 were provided by Peter Baumann (Stowers Institute, Kansas City). Poz1$^{30-249}$ was amplified from this plasmid by PCR using KOD polymerase (Novagen) and the following primers, and the resulting product was inserted into the pNT80e vector (Thomä Lab, His-tag vector derived from pColADuet-1, Novagen) by restriction ligation using the *Not1* and *Kpn1* sites.

Forward primer:
AAGAATGCGGCCGCGAATCTTCCATTGTGAACGC

Reverse primer:
CGGGGTACCTCACTAATTAATGTTTGAGGTAA

Tpz1$^{475-508}$ was PCR-amplified using KOD polymerase and the following primers and inserted into pNT10e (Thomä Lab, GST-tag vector derived from pGEX, GE) by restriction ligation using the *Not1* and *Kpn1* sites.

Forward primer:
AAGAATGCGGCCGCAACTCTGAGGCCTGTGAAAT

Reverse primer:
CGGGGTACCTCATTAGCTTTTGTTTCGAAACT

2.3 Mutagenesis

Dimerization and zinc-site mutants of Poz1-Tpz1 where generated by PCR amplifying the previously prepared plasmids (see **section 2.2**) with Phusion High-Fidelity polymerase (Finnzymes) and primers containing the appropriate mutations. The dimerization mutant Poz1$^{30-249}$_V34E_C37R was generated with the following primers using spPoz1$^{30-249}$ as template:

Forward primer:
TCTTCCATTGAAAACGCCAGGTTACGATAT

Reverse primer:
AGAAGGTAACTTTTGCGGTCCAATGCTATA

The zinc-site mutant was generated in two steps. First residues C479A and C482A were mutated by PCR amplifying Tpz1$^{475-508}$ using the following primers:

Forward primer:
TCTGAGGCCGCTGAAATGGCTCGGCTTGGG

Reverse primer:
AGACTCCGGCGACTTTACCGAGCCGAACCC

In a second step the mutation H488A was added using the following primers:

Forward primer:
CGGCTTGGGCTACCTGCTGGATCATTCTTTGAG

Reverse primer:
GCCGAACCCGATGGACGACCTAGTAAGAAACTC

The template was subsequently digested by Dpn1 (New England Biolabs) for 1 hour at 37°C. The mixture was then transformed in DH5α and clones were selected for antibiotic resistance on LB-plates.

2.4 Protein production

2.4.1 Native protein

Overexpression of spPoz1$^{30\text{-}249}$+Tpz1$^{475\text{-}508}$ was carried out in the *E. coli* strain BL21(DE3) Rosetta (Novagen). Cultures were grown at 37°C up to an OD$_{600}$ of 0.6 in Luria Bertani medium (LB, Sigma) supplemented with 25 µg/ml kanamycin, 34 µg/ml chloramphenicol and 50 µg/ml ampicilin. Expression was induced by addition of isopropyl-ß-D-thiogalactopyranoside (IPTG) to a final concentration of 0.25 mM. Cultures were grown at 20°C for an additional 18 hours. Cells were harvested by centrifugation at 2800 g at 4°C for 15 minutes. Bacterial pellet was resuspended in 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.25 mM TCEP (tris(2-carboxyethyl)phosphine), 1 mM PMSF, containing a protease inhibitor cocktail (Roche). Lysis was induced by sonication (Branson sonifier) for 4 minutes. The homogenate was clarified by ultra-centrifugation at 185000 g at 4°C for 30 minutes. The supernatant was applied to a 25 ml Ni^{2+}-chelating gravity flow column (Sigma) equilibrated in lysis buffer. Elution was performed using 120 mM imidazole. The eluent was incubated overnight at 4°C in the presence of 2 % TEV protease which resulted in efficient removal of the His6-tag and GST-tag from Poz1 and Tpz1 respectively. The sample was then diluted to 50 mM MES pH 6.5, 300 mM NaCl, 0.25 mM TCEP and loaded onto a 20ml Source S ionic exchange column (GE Healthcare) equilibrated in 50 mM MES pH 6.5 and 0.25 mM TCEP. The complex eluted at a NaCl concentration of 400 mM. The peak fractions were pooled and concentrated where after in a final step the sample was applied to a SD200 26/60 gelfiltration column (GE Healthcare). The protein purity was assessed by SDS-PAGE.

2.4.2 Selenomethionyl derivative protein

Production of selenomethionyl derivatized spPoz1$^{30\text{-}249}$+Tpz1$^{475\text{-}508}$ was performed as follows: Protein expression was carried out using the same expression plasmid that was used for native protein production. 6 L of 1x minimal medium (M9) were prepared with 1 L containing 1.6 g KH$_2$PO$_4$, 5.6 g K$_2$HPO$_4$, 0.8 g (NH$_4$)$_2$SO$_4$, 0.2 g citrate and 0.4 g MgSO$_4$. The amino acids Ala, Arg, Asp, Cys, Glu, Gly, His, Pro, Ser, Trp, Tyr, Gln and Asn were added to 40 mg/L. Ile, Leu, Lys, Phe, Thr and Val were added at 100 mg per 1 L media. 40 mg of thiamine and thymine each were added to each liter of M9 media. An overnight culture

at 37°C in regular LB medium was prepared. Once the M9 media cooled down after autoclaving, to each liter of 1x M9 medium glucose was added to a final concentration of 32 mM. 25 µg/ml kanamycin, 34 µg/ml chloramphenicol and 50 µg/ml ampicilin were added. 0.05 g of L-selenomethionine was dissolved in 3 ml of sterile water and added to 1 L of M9 medium. Each liter of M9 medium was inoculated with overnight culture at a dilution of 1:50 and grown at 37°C to an OD_{600} of 0.6. Protein production was induced with 0.25 mM IPTG and growth continued for another 18 hours at 20°C. After this point, harvesting of the cells, sonication and protein purification was carried out following the same procedure as for the native protein (see **2.3.1**). Due to the lower yield and reduced expression volume the gelfiltration step (**figure 2.2**) was carried out on the smaller SD200 10/300 25 ml gelfiltration column (GE). Subsequently the protein purity was assessed by SDS-PAGE.

2.5 Crystallization of spPoz1$^{30-249}$+Tpz1$^{475-508}$

Crystallization conditions for native spPoz1$^{30-249}$+Tpz1$^{475-508}$ were sought using commercial crystallization screens. Initial crystals were obtained from the Core3 crystal screen (Qiagen) in a 96-well sitting drop plate set up by a Phenix crystallization robot (Art Robbins). Crystallization conditions were refined in 24-well hanging drop plates (Hampton) with 500 µl crystallization solution in the well and 1 µl of protein mixed with 1 µl of mother liquor in the hanging drop.

Crystallization conditions for selenomethionyl derivatized were found by trying the same condition that worked best for the native protein. Initial crystals were obtained in 24-well hanging drop plates and further refined.

Best native and selenomethionyl crystals were briefly soaked in a drop containing mother liquor as well as 25 % ethylene glycol for cryoprotection and then flash frozen in liquid nitrogen in Hampton cryo loops.

2.6 Size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS)

For estimation of the molecular mass of spPoz1$^{30-249}$ and spPoz1$^{30-249}$+Tpz1$^{475-508}$ in order to determine their oligomeric state, proteins were injected on a SD200 10/300 gelfiltration column (GE) coupled to a miniDAWN TREOS multiangle-light-scattering detector (Wyatt) and an Optilab T-rEX refractive index detector (Wyatt) (experiment performed by Manjappa Lingaragu, FMI Basel). A total amount of 200 µg of each sample was injected in 50mM HEPES, 200 mM NaCl and 1 mM DTT.

2.7 Analytical size exclusion chromatography

For analytical size exclusion chromatography (SEC), a 25 ml SD200 10/300 gelfiltration column (GE) was equilibrated in 50 mM Tris-HCl pH 8.0, 500 mM NaCl and 0.25 mM TCEP. 1 mg of each sample, spPoz1$^{30-249}$, spPoz1$^{30-249}$+Tpz1$^{475-508}$ and spPoz1$^{30-249}$-V34E_C37R+Tpz1$^{475-508}$ was injected onto the column in a total volume of 100 µl and run at a flow rate of 0.5 ml/min. All runs were performed consecutively in a single day with a blank run of equilibration buffer in between each run.

2.8 Generating Poz1-Tpz1 fission yeast (*S. pombe)* mutant strains

2.8.1 Poz1-Tpz1 dimerization mutants

Gene targeting was carried out according to Bähler *et al*. (1998). Two different Poz1-Tpz1 dimerization mutants were generated. The first being fulllength Poz1$^{1-249}$ and the second being the N-terminally truncated Poz1$^{30-249}$. For each construct, we additionally designed a version tagged with the V5-epitope tag (GKPIPNPLLGLDST) for subsequent analysis by western blot. The tag was placed C-terminally linked by 6xGly residues. The following constructs were designed:

1. spPoz1-WT-V5

2. spPoz1-V34E-C37R

3. spPoz1-V34E-C37R-V5

4. spPoz1$^{30-249}$

5. spPoz1$^{30-249}$-V5

6. spPoz1$^{30-249}$-V34E-C37R

7. spPoz1$^{30-249}$-V34E-C37R-V5

In a first step, the Poz1 gene was deleted in *Schizosaccharomyces pombe* strain SPB72:h+leu1-32ade6-M216ura4-D18his3-D1 (provided Yukiko Shimada, Bühler Lab, FMI Basel). To achieve this, the kanamycin resistance cassette in pFA6a (Bühler Lab, FMI) was amplified with oligonucleotides containing 80 base pairs homology to the 5' and 3' UTR of spPoz1 respectively:

Forward primer:

ATCTCGTAGACCAACTTACATTGACTTTACCAACTTTTTATACTTTTATCCTTCGTGTATAGATTAG TTTTTTCCATAAACGGATCCCCGGGTTAATTAA

Reverse primer:

TAGTTTTAGACTTTTGACCCCTCCAGGAATTTAAAGATACCAAAAATTTATTAATATAAAGAGCTTA TCGTTATTCGATAGAATTCGAGCTCGTTTAAAC

The amplified product was then transformed in SPB72 according to the LiOAc protocol (Okazaki *et al.*, 1990, Kanter-Smoler *et al.*, 1994) and plated on plates containing YES-medium and kanamycin antibiotics. Colonies grew after three days. Positive clones were determined by PCR and restreaked in such a way that single colonies could be selected and restreaked once more on fresh plates. These cells were used for the following steps. From these isolated colonies, genomic DNA was extracted by boiling and vortexing a small amount of cells picked from the plate in 10 µl 20 mM NaOH.

In a second step, constructs 1-7 were re-inserted in the genome in the position of the kanamycin resistance cassette that had been inserted into the genome at the place of *poz1* in the first step.

Spliced Poz1 and Poz1_V34E_C37R where amplified with primers containing PacI and AscI restriction sites and subsequently were cloned into pMB22 (provided by Claudia Keller and Marc Bühler, FMI Basel) by restriction ligation with PacI and AscI (NEB). This vector has a terminator sequence and a nourseothricin (NAT) resistance cassette following the AscI restriction site. The primer combinations used for the various constructs are listed in **table 2.1**. Constructs that included a C-terminal V5-tag were amplified with primers containing the tag sequence.

Construct	Template	Forward primer	Reverse primer
1	spPoz1$^{1-249}$	CS1	CS2
2	spPoz1$^{1-249}$_V34E_C37R	CS1	CS3
3	spPoz1$^{1-249}$_V34E_C37R	CS1	CS2
4	No ligation required, spPoz1$^{1-249}$_pMB22 was used to amplify spPoz1$^{30-249}$ for homologous recombination		
5	No ligation required, spPoz1$^{1-249}$-V5_pMB22 was used to amplify spPoz1$^{30-249}$-V5 for homologous recombination		
6	No ligation required, spPoz1$^{1-249}$_V34E_C37R_pMB22 was used to amplify spPoz1$^{30-249}$_V34E_C37R for homologous recombination		
7	No ligation required, spPoz1$^{1-249}$_V34E_C37R-V5_pMB22 was used to amplify spPoz1$^{30-249}$_V34E_C37R-V5 for homologous recombination		

Table 2.1. Templates and primers used for cloning into the antibiotic resistance cassette vector.

CS1:

CCTTAATTAAATGAATGAGAAGATTCGTTC

CS2:

GGCGCGCCTTAGGTGCTATCCAGGCCCAGCAGCGGGTTCGGAATCGGTTTGCCGCCGCCGCC GCCGCCATTAATGTTTGAGGTAAGCA

CS3:

GGCGCGCCCTAATTAATGTTTGAGGTAA

Subsequently, the gene, terminator and NAT resistance were amplified from the previously generated plasmids using the primers listed in **table 2.2** containing the appropriate 5'- and 3' UTR homology sequences.

Construct	Template	Forward primer	Reverse primer
1	spPoz1$^{1-249}$-V5_pMB22	CS4	CS5
2	spPoz1$^{1-249}$_V34E_C37R_pMB22	CS4	CS5
3	spPoz1$^{1-249}$_V34E_C37R-V5_pMB22	CS4	CS5
4	spPoz1$^{1-249}$_pMB22	CS6	CS5
5	spPoz1$^{1-249}$-V5_pMB22	CS6	CS5
6	spPoz1$^{1-249}$_V34E_C37R_pMB22	CS7	CS5
7	spPoz1$^{1-249}$_V34E_C37R-V5_pMB22	CS7	CS5

Table 2.2. Templates and primers used to amplify the gene and resistance cassette to be integrated into the *S. pombe* genome.

CS4:

ATCTCGTAGACCAACTTACATTGACTTTACCAACTTTTTATACTTTTATCCTTCGTGTATAGATTAG TTTTTTCCATAAAATGAATGAGA AGATTCGTTC

CS5:

TAGTTTTAGACTTTTGACCCCTCCAGGAATTTAAAGATACCAAAAATTTATTAATATAAAGAGCTTA TCGTTATTCGATAGAATTCGAGCTCGTTTAAAC

CS6:

ATCTCGTAGACCAACTTACATTGACTTTACCAACTTTTTATACTTTTATCCTTCGTGTATAGATTAG TTTTTTCCATAAAATGGAATCTTCCATTGTGAACGC

CS7:

ATCTCGTAGACCAACTTACATTGACTTTACCAACTTTTTATACTTTTATCCTTCGTGTATAGATTAG TTTTTTCCATAAAATGGAATCTTCCATTGAAAACGC

The resulting products were transformed in the *poz1Δ* strain generated in the previous step. The transformation was plated on YES-medium plates containing NAT antibiotics. First colonies appeared after 3-4 days and were restreaked and isolated as described before. From isolated single colonies, genomic DNA was extracted by boiling and vortexing a small amount of cells picked from the plate in 10 µl 20 mM NaOH. The isolated DNA was used to amplify PCR products covering the entire gene, as well as at least 70 base pairs upstream of the 5' homolgous recombination site and covering the 3' end up to the resistance cassette. These amplicons were then sent for DNA sequencing (Novaseq, Novartis Basel sequencing service).

2.8.2 Zinc binding site mutant strains

Because the *tpz1Δ* fission yeast strain has been shown to be lethal or grow very slow (Miyoshi *et al.*, 2008), we decided to directly replace the C-terminal region of spTpz1 containing the sites of mutation in the genome, rather than generating an intermediate mutant as we did with the Poz1 dimerization mutants (see **2.8.1**). Also here we added V5-epitope tags to the C-terminus of the protein by including the sequence in the appropriate primers. The following spTpz1 constructs were designed:

1. spTpz1$^{475-508}$-V5
2. spTpz1$^{475-508}$-C482A-H488A
3. spTpz1$^{475-508}$-C482A-H488A-V5

Tpz1$^{475-508}$ and Tpz1$^{475-508}$-C482A-H488A were amplified from previously generated plasmids using the primers listed in **table 2.3** containing the PacI and AscI restrictions site respectively and restriction ligated into pMB22.

Construct	Template	Forward primer	Reverse primer
1	spTpz1$^{475-508}$	CS8	CS10
2	spTpz1$^{475-508}$_C482A-H488A	CS9	CS11
3	spTpz1$^{475-508}$_C482A-H488A-V5	CS9	CS10

Table 2.3. Templates and primers used for cloning into the antibiotic resistance cassette vector.

CS8:

CCTTAATTAAAACTCTGAGGCCTGTGAAATG

CS9:

CCTTAATTAAAACTCTGAGGCCGCTGAAATG

CS10:

*GGCGCGCCTTAGGTGCTATCCAGGCCCAGCAGCGGGTTCGGAATCGGTTTGCCGCCGCCGCC
GCCGCCGCTTTTGTTTCGAAACTCCT*

CS11:

GGCGCGCCTTAGCTTTTGTTTCGAAACTCCT

Gene, terminator and resistance cassette were subsequently amplified using the primers listed in **table 2.4**. The resulting products were transformed and colonies were isolated and their DNA sequenced as described in section **2.8.1**.

Construct	Template	Forward primer	Reverse primer
1	**spTpz1$^{475\text{-}508}$_pMB22**	**CS12**	**CS13**
2	**spTpz1$^{475\text{-}508}$_C482A-H488A_pMB22**	**CS12**	**CS13**
3	**spTpz1$^{475\text{-}508}$_C482A-H488A-V5_pMB22**	**CS12**	**CS13**

Table 2.4. Templates and primers used to amplify the gene and resistance cassette to be integrated into the *S. pombe* genome.

CS12:

CGCTGCAAGAGTTGTATGTTGAGCATCAGAGCAAGAAACGGCGTCTCGAACTATTTCAATTAACA AATAATCATCAAAAGAACTCTGAGGCCTGTGAAAT

CS13:

TCACTGTATGTCTGTAACAGTTAACTTCCGTACTTAGTAAAATGTTAGTAAAAAAGGAAGATATGT GATACAGCAATTGAGAATTCGAGCTCGTTTAAAC

2.8.3 SpTpz1-hApollo and spRap1-hTRF2H fusions

In order to replace Poz1 in the *S. pombe* genome by the structurally similar human shelterin components, TRF2H and Apollo, Tpz1$^{475\text{-}508}$ was amplified with by PCR where the reverse primer contained the DNA sequence for Apollo$^{495\text{-}530}$ followed by the V5-tag. The following primers were used:

Forward primer:
CCTTAATTAAACTGAATTCAGGGGTCTAGCA

Reverse primer:
GGCGCGCCTTAGGTGCTATCCAGGCCCAGCAGCGGGTTCGGAATCGGTTTGCCCGAGCCTTTA TGGTATTTTTCCACTT

This fusion construct was then restriction ligated with PacI and AscI into pMB22. The construct was then amplified from this plasmid with primers containing the appropriate 5'

and 3' 80bp sequence homology such that it would be inserted into the *Poz1Δ* strain used in **2.12.1** at the Tpz1 C-terminus with six glycine residues as a linker.

Forward primer:

GTCGGCTTGGGCTACCTCATGGATCATTCTTTGAGCTATTGCGAGATTGGAAAAAAATAGAGGAG
TTTCGAAACAAAAGCACTGAATTCAGGGGTCTAGC

Reverse primer:

TCACTGTATGTCTGTAACAGTTAACTTCCGTACTTAGTAAAATGTTAGTAAAAAAGGAAGATATGT
GATACAGCAATTGAGAATTCGAGCTCGTTTAAAC

The resulting strain was then used to insert the Rap1-TRF2H fusion. For this, TRF2[40-245] was amplified by PCR with a reverse primer containing the V5-tag and the product was inserted by restriction ligation with PacI and AscI into pMB35 (a vector similar to pMB22 but with a hygromorphine resistance cassette, provided by the Bühler lab).

Forward primer:

CCTTAATTAAGAGCGCGGCGCGGGGGAGGC

Reverse primer:

GGCGCGCCTTAGGTGCTATCCAGGCCCAGCAGCGGGTTCGGAATCGGTTTGCCCGAGCCTTTC
AAAGCCTTTTTGGCCA

The construct was then amplified from this plasmid with primers containing the appropriate 5' and 3' 80bp sequence homology such that it would be inserted into the Poz1Δ-Tpz1-Apollo strain at the Rap1 C-terminus with six glycine residues as a linker.

Forward primer:

ACAATTCCTTGAAGCTATGGAATCTACGGGTGGAAGAGTACGAATTGCCATTGCTAAACTACTTT
CAAAACAAACTTCTGAGCGCGGCGCGGGGGAGGC

Reverse primer:

TATGCATAAAAAGATTCGTAATATTGTACAAGTTTAGGTCTCTTTAGAGAAATAGAATTTGGGCAG
AGATGCTCGGCAATGAATTCGAGCTCGTTTAAAC

The resulting strain was then selected for antibiotic resistance against Kan-Nat-Hph. Only very few colonies grew on the YES-plates. Nevertheless, a single clone could be isolated harboring the correct sequences.

2.9 Telomere length assay

In order to determine the effects of the Poz1 and Tpz1 mutants on telomere regulation, the telomere length was determined as described in the following. These experiments were performed by Lili Pan in Peter Baumann's lab at the Stowers Institute, Kansas City, USA.

Fission yeast genomic DNA was prepared as follows (Bunch *et al.*, 2005): Frozen fission yeast cell pellets at around 5×10^8 cells were thawed, washed in 10ml Z buffer (50 mM sodium citrate, 50 mM sodium monohydrogen phosphate, 40 mM ethylenediaminetetraacetic acid) at pH 7.8 and resuspended in 2 ml Z buffer supplemented with 0.5 mg/ml Zymolyase-100T and 2 mM DTT. The cell suspension was incubated at 37°C for 1 hour and subsequently sodium dodecyl sulfate was added to 4 % (w/v). Samples were incubated at 65°C for 10 minutes. The volume was increased to 10ml with 5x TE (50 mM Tris-HCl, pH 8.0, 5 mM ethylenediaminetetraacetic acid) and proteinase K (Sigma) was added to 50 µg/ml. After 1 hour of incubation at 50°C, 3 ml of 5M potassium acetate solution was added and samples were incubated on ice for 30 minutes. After removal of the precipitate by centrifugation, the clarified supernatant was mixed with 1 volume of isopropanol to precipitate nucleic acids. Samples were kept for 20 minutes and ice and subsequently subjected to centrifugation at 10,000 g for 5 minutes. After briefly drying the DNA pellets they were resuspended in 0.5 ml 5x TE containing DNase free RNase A (50 µg/ml). After incubation at 37°C for 1 hour, organic extraction and ethanol precipitation, genomic DNA was resolubilized in 1x TE. The genomic DNA was then analyzed by Southern blotting.

Southern blotting was performed as follows (Baumann and Cech, 2000): 15 – 20 µg of genomic DNA was digested for 6 to 8 hours with EcoRI in the appropriate buffers supplied by the manufacturer (NEB). Restriction fragments were subsequently loaded onto 1.1 % agarose gels and run in 0.5x TBE at 2 V/cm for 16 hours. To confirm equal loading, gels

were stained for 30 minutes in ethidium bromide a 1 µg/ml, where after the DNA was visualized under UV. DNA was denatured by treatment with sodium hydroxide and then transferred onto a nylon membrane (Hybond NX, Amersham) in 10x SSC (1.5 M NaCl, 0.15 M sodium citrate) by capillary action. Hybridizing probes specific for the telomeric and telomere associated sequences (TAS) were created by the same method using gel-purified fragments of pNSU70 (Sugawara, 1988). Hybridizations were carried out in Church-Gilbert buffer at 65°C (Church and Gilbert, 1984).

2.10 Western blot

In order to verify that the Poz1-Tpz1 mutant proteins were expressed in fission yeast, western blots with anti-V5 antibody were performed as follows by Lili Pan (Bunch *et al.*, 2005).

$1x10^8$ fission yeast cells were lysed by vortexing with glass beads (0.5 mm in diameter) and in the presence of 10% trichloroacetic acid for 8 minutes at 4°C. Beads were washed in 10% trichloroacetic acid, and precipitated proteins were pelleted by centrifugation for 2 minutes at 16,000 g. Following an acetone wash, proteins were solubilized in 1 x protein sample buffer (1x NuPAGE LDS sample buffer containing 50 mM dithiothreitol and 2% sodium dodecyl sulfate). Samples were incubated at 75°C for 5 min and subsequently centrifuged at 16,000 g for 2 min. Soluble fractions were loaded onto a NuPAGE 4 to 12% Bis-Tris gel with 1x MOPS buffer and the gel was run according to the manufacturers recommendations. Proteins were then transferred to Protran nitrocellulose membrane in a Bio-Rad Mini Trans-Blot cell at 100 V for 1 h. Blocking of the membrane was carried out in 1x TTBS (20 mM Tris-HCl, pH 7.5, 0.1 % Tween, 137 mM NaCl) plus 5 % nonfat milk. Given the V5-tag we added to the Poz1-Tpz1 mutants, the monoclonal anti-V5 (Invitrogen, 460705) was used at a 1:5000 dilution as primary antibody. Horseradish peroxidase-conjugated goat anti-mouse (Thermo Scientific, pn31430) was used at 1:5,000 dilution as secondary antibody. Bands were detected by ECL2 substrate (Pierce).

Chapter 3

Results - Part I

3. Results – Part I: Fission yeast Poz1-Tpz1 structure solution

Crystallization of large protein complexes in their fulllength forms is often challenging. It is not uncommon to observe low expression levels, aggregation, insufficient solubility or poor stability during protein expression. All these factors reduce the likelihood of successful crystallization.

In order to overcome this impediment, we optimized the protein boundaries of the shelterin components by secondary structure prediction, as well as limited proteolysis. The underlying principle of limited proteolysis is the removal of unstructured regions, such as flexible loops, by addition of small amounts of protease to the protein. The resisting protein should in principle be of a more compact nature, and therefore more likely to result in successful crystal formation.

Ultimately, by applying this approach on the different fission yeast shelterin components in various combinations, we finally succeeded with Poz1 and Tpz1. The optimization procedures are described in the following.

3.1 Boundary optimization by limited proteolysis and subsequent identification by LC/MS

Poz1 is a 249 amino acid protein with a molecular weight of 30 kDa. The attempt to express and purify fulllength Poz1 failed due to the insoluble nature of this protein (**figure 3.1A**). Likewise, co-expression together with its binding partners, Rap1 or Tpz1, did not alleviate this problem. Following this, we designed a series of constructs of Poz1 based on secondary structure prediction (**figure 3.2**). The most promising of these, was Poz1$^{30-249}$ that showed good expression levels and solubility in *E.coli* (**figure 3.1B**).

A   **B**
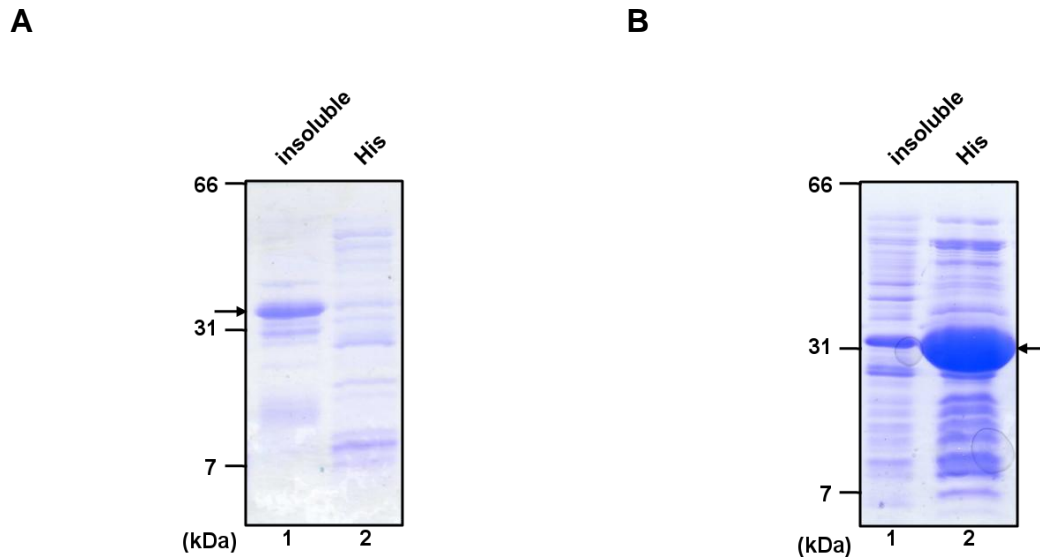
Figure 3.1. *E. coli* **expression test of spPoz1.**

SDS-PAGE of Ni^{2+}-affinity pulldowns of **A)** spPoz1$^{1-249}$ (30 kDa) and **B)** spPoz1$^{30-249}$ (26 kDa). Lanes 1 show the insoluble fraction and lanes 2 the affinity pulldown with Ni^{2+}-resin.
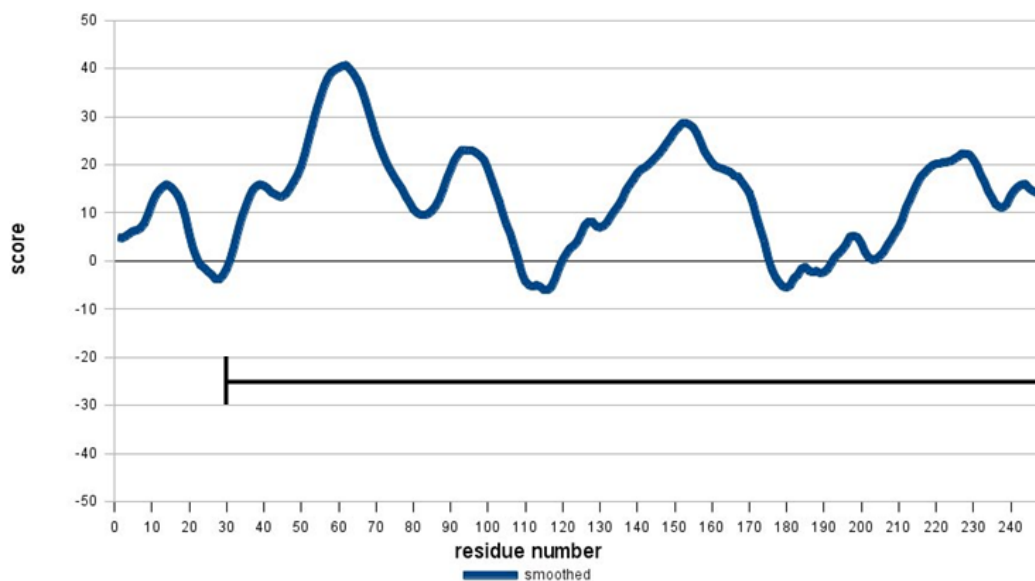
Figure 3.2. Structure propensity plot for spPoz1$^{1-249}$.

The structure propensity secondary structure prediction approach relies on the probability of a given amino acid to form secondary structures (α-helix, β-strand and coil). The probability parameters are derived from already determined protein structures. The secondary structure states of the neighboring amino acids are also taken into account. Ultimately, a prediction is made, on whether a given amino acid sequence is in a folded or unfolded state, allowing the identification of hypothetical protein domains (Chou and Fasman, 1974; Garnier *et al.*, 1996). In the structure propensity plot, values >0 predict ordered regions of a protein, while values <0 predict disordered regions. The boundary of the most promising construct of Poz1 with residues 30-249 is indicated.

Tpz1 is a 508 amino acid protein with a molecular weight of 58 kDa. While it can be expressed in its fulllength form in Hi-5 insect cells (*Trichopulsia ni*, cabbage looper ovary), it aggregates when purified by itself as well as in any combination with other shelterin

components. Tpz1 is predicted to contain a large unstructured segment in the central part of the protein which likely allows Tpz1 to exist in many different conformations due to the flexible nature of this region. This in turn is an unwanted characteristic for crystallization. As a consequence, we designed a series of constructs of Tpz1 based on secondary structure predictions (**figure 3.3**).
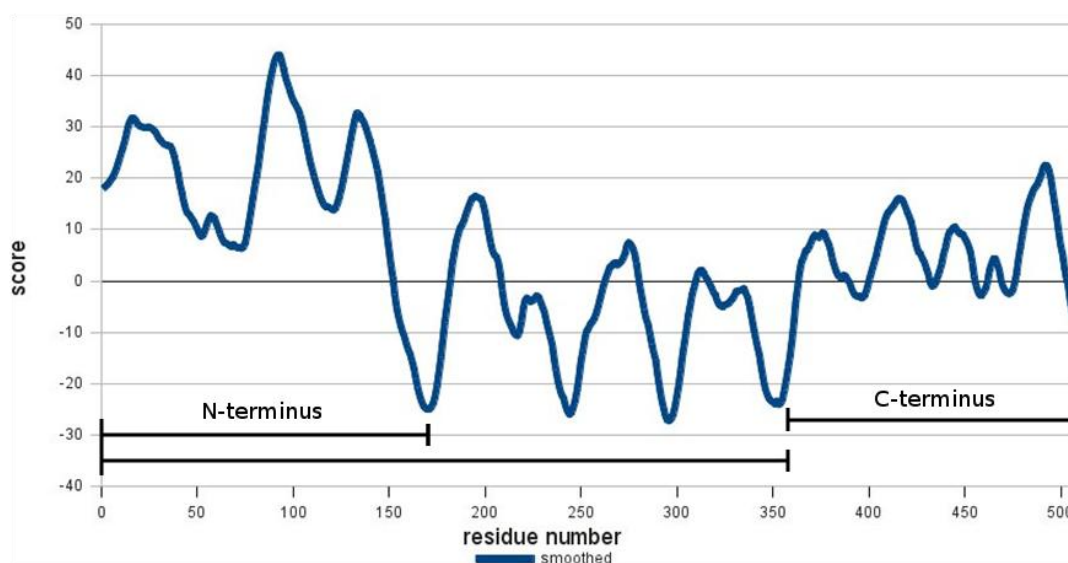
Figure 3.3. Structure propensity plot of spTpz1$^{1-508}$.

Some of the constructs that were designed are highlighted. The C-terminus was of special interest as it contains the Poz1-binding domain.

Our focus was on a C-terminal construct Tpz1$^{360-508}$, which we showed by affinity pulldowns, still harbors the Poz1-binding domain (PBD). Nevertheless, this construct remained aggregated/multimerized in complex with Poz1$^{30-249}$. Following this, we decided to pursue the limited proteolysis approach to remove unstructured regions that may be causing the aggregation.

3.1.1 Limited proteolysis of Poz1$^{30-249}$ +Tpz1$^{360-508}$

The protease trypsin was added to Poz1$^{30-249}$+Tpz1$^{360-508}$ in steps of increasing concentrations from 0.003 to 3 %. Emerging stable fragments were identified by SDS-PAGE (**figure 3.4**).
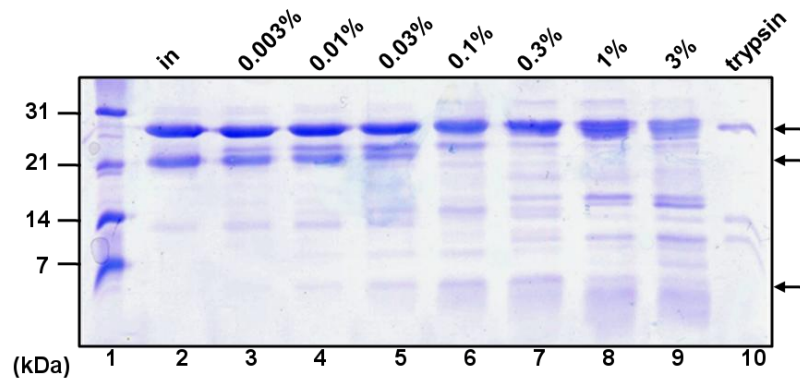
Figure 3.4. Limited proteolysis experiment with trypsin and Poz1$^{30-249}$ +Tpz1$^{360-508}$.
Bands of interest are indicated: Upper arrow: Poz1$^{30-249}$, middle arrow: Tpz1$^{360-508}$, lower arrow: Putative Tpz1-fragments.

Poz1$^{30-249}$ remained resistant to a protease concentration of up to 3 % trypsin, underlining the compact and well structured nature of this construct. Contrariwise, Tpz1$^{360-508}$ showed severe signs of degradation already at low protease concentrations, emphasizing the unstructured nature of this protein. At a protease concentration of 0.3 %, the majority of Tpz1$^{360-508}$ had been degraded to smaller fragments. In order to answer the question whether the low molecular fragments that appeared were originating from Tpz1 degradation, we identified the bands from the SDS-PAGE gel by mass-spectroscopy. The results confirmed that indeed these fragments are degradation products originating from Tpz1.

3.1.2 Identification of co-migrating fragments by size exclusion chromatography

To determine whether any of the low molecular weight fragments that appeared in limited proteolysis still bind to Poz1, we repeated the proteolysis experiment in a larger scale with 0.1 % trypsin (this being the concentration at which Tpz1 was degraded to smaller fragments and Poz1 remained to resist proteolysis (**Figure 3.4, lane 6**). The reaction was stopped by addition of PMSF, followed by immediate analysis of the sample by size-exclusion chromatography (**figure 3.5**).

A
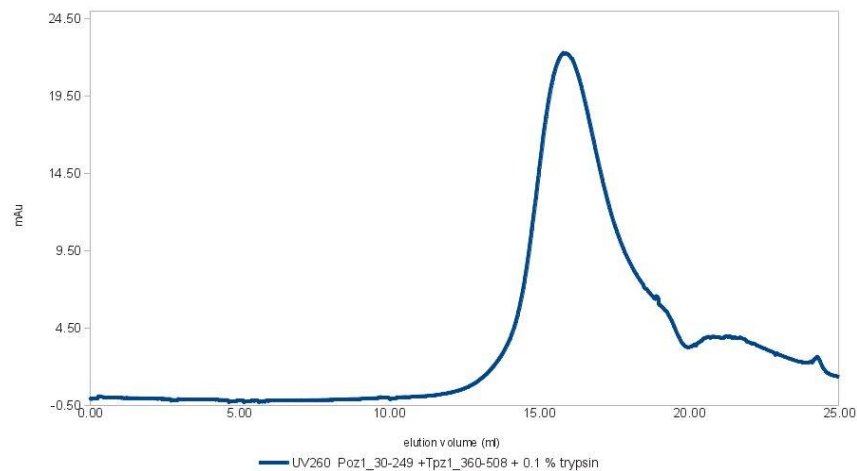
B
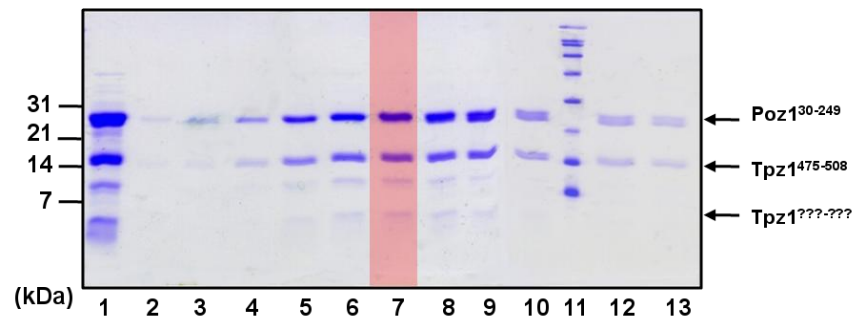
Figure 3.5. Size-exclusion chromatography of the large-scale limited proteolysis experiment: Poz1$^{30-249}$+Tpz1$^{475-508}$ + 0.1 % trypsin.

A) Gelfiltration chromatogram. **B)** SDS-PAGE: Poz1 and Tpz1 are indicated. The input is in lane (1). Fractions were collected in steps of 0.5 ml. Lanes (2) to (13) correspond to elution volumes ranging from 12 to 17 ml. Co-migrating fragments are observed at low molecular weight. The Tpz1 fragment of interest at approximately 5 kDa is indicated. The fraction in lane (7) was used for boundary identification by LC/MS.

Indeed we found a low molecular weight fragment co-migrating with Poz1$^{30-249}$. The band was excised and identified by mass spectroscopy as originating from Tpz1.

3.1.3 Boundary identification of Tpz1-fragments by LC/MS

In order to identify the exact boundaries of the co-migrating Tpz1-fragments, we applied liquid chromatography coupled to mass spectroscopy (LC/MS). The fragments contained in fraction 7 of the gelfiltration (**figure 3.5B**) could be separated by HPLC and its boundaries determined. Ultimately, the precise boundaries of the fragment of interest were assigned as Tpz1 residues 475-508.

3.1.4 Validation of the boundary optimized complex by recombinant expression

The behavior of the Poz1-Tpz1 complex with the optimized protein boundaries obtained by LC/MS was tested. A small scale expression test was performed in *E. coli* and the protein complex was purified by pulling on the tag of Tpz1 and subsequently on the tag of Poz1 (**figure 3.6)**
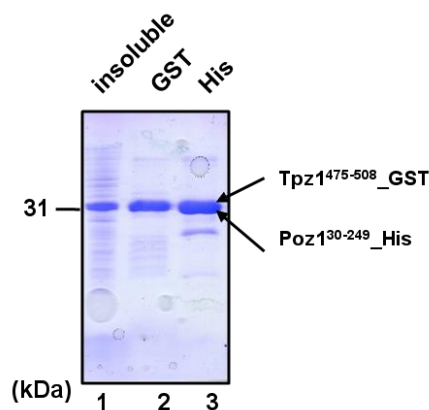
Figure 3.6. Expression test of spPoz1$^{30-249}$+Tpz1$^{475-508}$.

Lane (1) shows the insoluble fraction. Lane (2) the affinity pulldown with Glutathion-S-Transferase resin (GST) and lane (3) the subsequent pulldown with Ni^{2+} resin. Because of the similar sizes of 30 kDa for Tpz1 with the GST-tag and 26 kDa for His-tagged Poz1, the bands are not distinguishable on the scanned gel.

3.1.5 Protein purification of spPoz1$^{30-249}$+Tpz1$^{475-508}$

In order to obtain pure and sufficient amounts of protein for crystallization trials, Poz1$^{30-249}$+Tpz1$^{475-508}$ was expressed in a larger scale (12 L) in *E. coli* and purified in three chromatographic steps: N^{2+}-affinity chromatography, ion exchange chromatography and size-exclusion chromatography (**figure 3.7A**). The protein purity was assessed by SDS-PAGE (**figure 3.7B**). A total of 2 mg of pure protein was obtained per 1 L of *E. coli* culture.

A
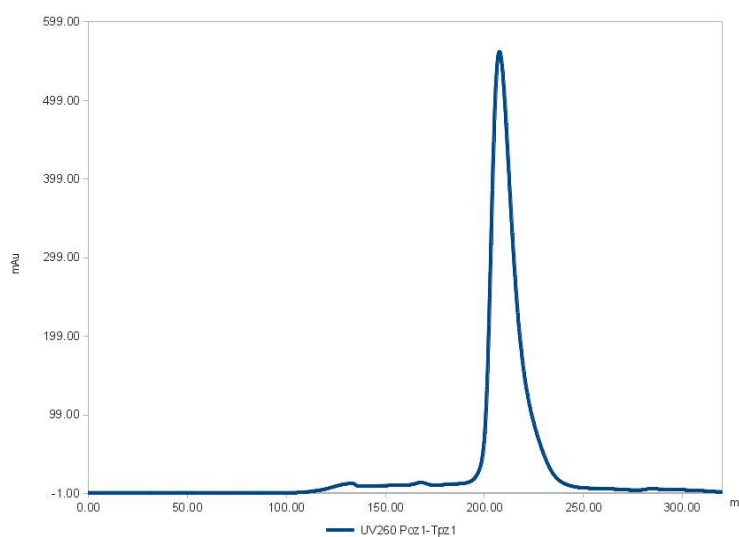
B
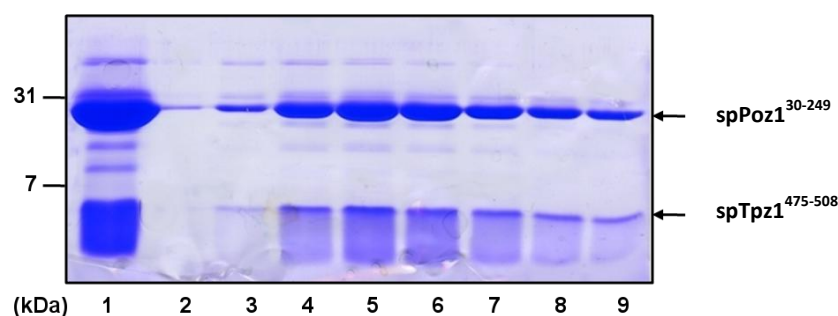
Figure 3.7. Protein purification of native spPoz1$^{30-249}$+Tpz1$^{475-508}$.

A) Chromatogram and **B)** SDS-PAGE of the gelfiltration step. The input is in lane (1). Fractions were collected in steps of 4 ml. Lanes (2) to (9) correspond to elution volumes ranging from 192 to 220 ml.

3.1.6 Protein purification of selenomethionyl derivatized spPoz1$^{30-249}$+Tpz1$^{475-508}$

When solving a structure with X-ray crystallography, one must deal with the "crystallographic phase problem". Information about the phase of a reflection cannot be determined from the diffraction image, which means the phases need to be determined by other means. The phase, however, is required for the calculation of the electron density (see Taylor, 2003, for an overview of the phase problem). The simplest method for determining the phases is "molecular replacement" (Rossman and Blow, 1962). The principle here is that the phases of structure factors from a known protein structure, similar to the one in question, are taken and used as initial estimates. The structure factor is a mathematical function describing the amplitude and phase of a wave diffracted from crystal. The known protein is termed the "phasing model". It has been shown, that the intensities of the unknown structure

contain enough information to reveal the differences between the phasing model and the unknown structure, when using the phases of the known model as initial estimates. The phasing model must exert a certain level of similarity. A very low correlation between the two proteins may result in very poor solutions produced by molecular replacement algorithms. As there were no solved structures in the protein data bank with any apparent similarity to Poz1-Tpz1 that could have been used for molecular replacement, it was necessary to utilize a different method for determining the phases, namely the experimental phasing method of multiwavelength anomalous dispersion (MAD) (Hendrickson, 1991).

MAD depends on the characteristic of an atom to absorb X-rays of a specific wavelength. The absorption of an element drops sharply just below their characteristic emission wavelength K_β. This sudden drop is called an elements absorption edge. Atoms show anomalous scattering if the energy of the X-rays is near the elements absorption edge. Elements absorb as well as emit X-rays. When the atom absorbs the energy, it is the re-emitted, but out of phase with the incident X-ray beam. The energy is also different, because some has been used for the transition. This effect is termed anomalous scattering. Data collection at and around the absorption edge and comparison of the resulting diffraction patterns, allows establishment of the phases of the reflections.

The most commonly used anomalous scatterer for phase determination by MAD is selenium. For the incorporation into protein, the sulfur atom in methionine residues is replaced by a selenium atom (Hendrickson *et al.*, 1990). Thus protein with selenomethionine residues instead of native methionine residues must be produced, purified and crystallized.

The same purification procedure was applied for selenomethionyl derivatized Poz1-Tpz1 (aside from the expression that was carried out in minimal media containing selenomethionine, see **section 2.4.2**). The complex behaved identical to the native protein, except for the lower, but still reasonable expression yield 1 mg of protein per liter of *E. coli* culture (**figure 3.8**).

A
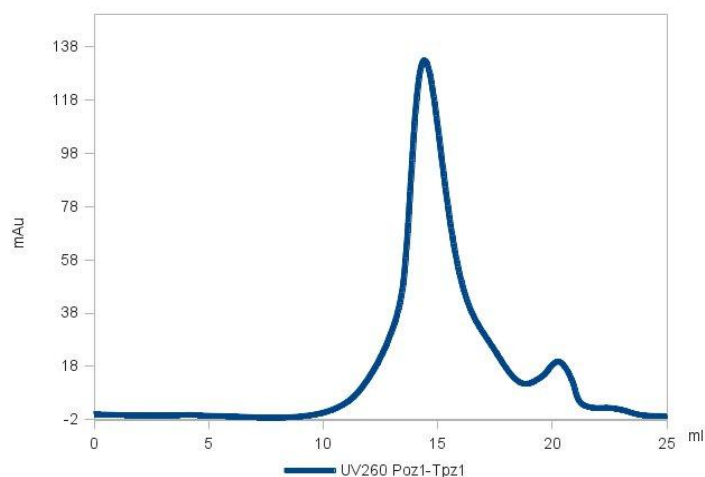
B
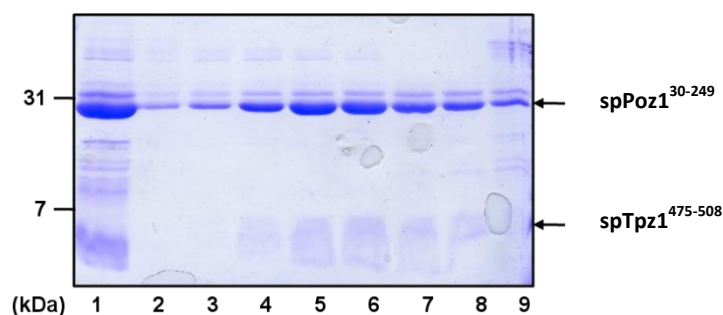
Figure 3.8. Protein purification of selenomethionyl derivatized spPoz1$^{30-249}$+Tpz1$^{475-508}$.

A) Chromatogram and **B)** SDS-PAGE of the gelfiltration step. The input is in lane (1). Fractions were collected in steps of 0.5 ml. Lanes (2) to (9) correspond to elution volumes ranging from 12.5 to 16 ml.

3.2 Crystallization

For protein structure determination by X-ray crystallography, crystals of Poz1-Tpz1 were grown and optimized. The successful procedure we applied is described in the following section.

3.2.1 Crystallization of native spPoz1$^{30-249}$+Tpz1$^{475-508}$

Initial crystals of native Poz1$^{30-249}$+Tpz1$^{475-508}$ grew from solutions containing 30 % 1,2-propanediol and 0.1 M MES pH 6.5. Rod-shaped crystals grew with various dimensions within 2-3 days after mixing. Refinement of these crystals required varying the 1,2-propanediol concentration, as well as screening pH and protein concentration. This resulted in

single crystals of larger size suitable for X-ray diffraction experiments. Best crystals were grown from 14 % 1,2-propanediol, 0.1 M MES pH 6.0 and of a protein concentration of 4.5 mg/ml. These crystals were flash frozen in liquid nitrogen and mother liquor supplemented with 25 % ethylene glycol for cryoprotection and subjected to diffraction tests at the synchrotron X-ray beamline (**figure 3.9**).
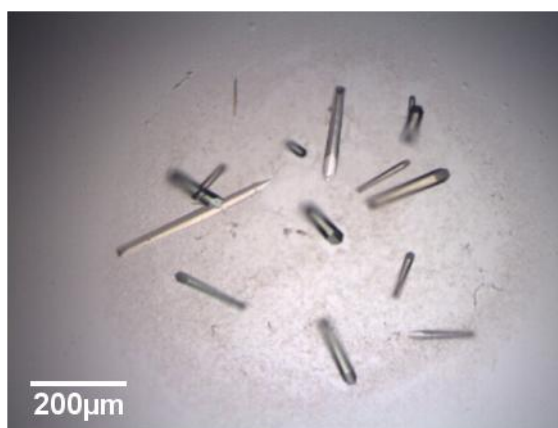
Figure 3.9. Optimized spPoz1$^{30-249}$+Tpz1$^{475-508}$ crystals used for native data collection.

The crystals show a hexagonal base and hexagonal elongated tip. Lower protein concentration (4.5 mg/ml) yielded fewer but larger crystals than higher (14 mg/ml) protein concentration.

3.2.2 Selenomethionyl crystals

The first step was to try and obtain crystals of selenomethionyl derivative Poz1$^{30-249}$+Tpz1$^{475-508}$ in the same crystallization conditions used for the native protein. Small and thin needle shaped crystals were obtained in the native crystallization condition containing 14 % 1,2-propanediol, 0.1 M MES pH 6.0 and a protein concentration of 4.5 mg/ml. Lowering the pH to 5.6 yielded fewer but bigger crystals, however the best diffracting selenomethionyl crystals were obtained by substituting 1,2-propanediol by the similar compound MPD ((4s)-2-methyl-2,4-pentanediol). Well diffracting crystals grew in 6 % MPD, 0.1 M MES pH 5.6 and protein at 5 mg/ml (**figure 3.10**).
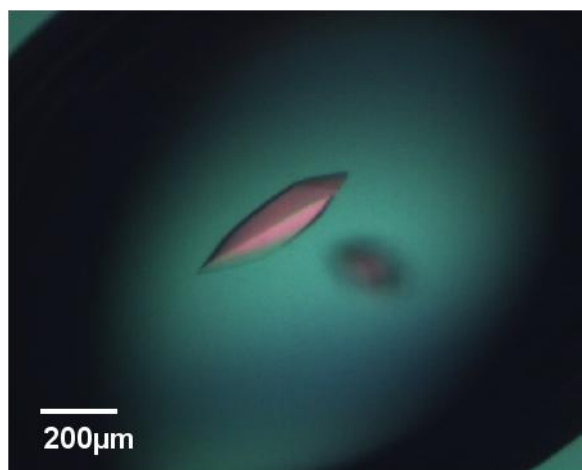
Figure 3.10. Selenomethionyl derivatized spPoz1$^{30\text{-}249}$+Tpz1$^{475\text{-}508}$ crystal.

Crystal grew in 6 % MPD, 0.1 M MES pH 5.6 and at protein concentration of 5mg/ml within 3 days.

3.3 Collection of native diffraction data

Two native data sets were collected. The first on the synchrotron beam line X06SA at the SLS (Villigen, Switzerland), which is fitted with a marCCD-225 detector. Data were collected for a total rotation range of 180° with an oscillation angle of φ=1.0° per frame and an exposure time of 0.5 s (**table 3.1**). A second data set was collected at the synchrotron beam line X10SA at the SLS, which is fitted with a Pilatus 6M detector. 1000 images were collected with an oscillation angle of φ=0.25° per frame and an exposure time of 0.25 s. Higher resolution was achieved for the second data set (**table 3.1**).

Native data sets		
	Data set 1	**Data set 2**
X-ray source	X06SA	X10SA
Detector	marCCD-225	Pilatus 6M
Wavelength	1.0 Å	1.0 Å
Distance	200.0 mm	380.0 mm
Exposure time	0.5 s	0.25 s
Oscillation	1.0°	0.25°
Number of images	200	1000

Table 3.1. Strategy for the collection of native diffraction data

The data images (**figure 3.11**) were processed using the XDS package (Kabsch, 2010). Crystals belonged to the space group $P3_121$ of the primitive hexagonal lattice family with unit cell dimensions a=66.39 Å, b=66.39 Å, c=126.61 Å, α=β=90° and γ=120°. Data processing statistics are given in **table 3.2**.
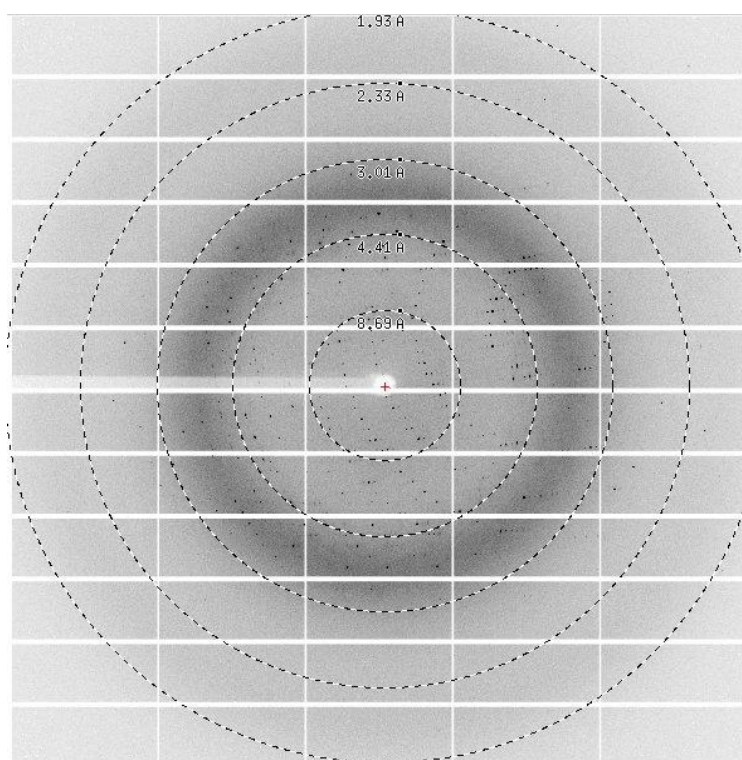
Figure 3.11. Diffraction pattern of spPoz1[30-249]+Tpz1[475-508] crystals.

A single X-ray diffraction pattern obtained from native Poz1-Tpz1 crystals on the X10SA synchrotron beamline at the SLS. Image generated with Adxv (http://www.scripps.edu/~arvai/adxv.html).

Native data processing statistics		
	Data set 1	**Data set 2**
Space group	$P3_121$	$P3_121$
Unit cell	a=b=66.10 Å, c=126.30 Å, α=β=90°, γ=120°	a=b=66.39 Å, c=126.61 Å, α=β=90°, γ=120°
Resolution (Å)	11.63-2.40 (2.67-2.60)	10.73-2.40 (2.46-2.40)
Unique reflections	10311 (739)	13038 (954)
R-factor (%)	7.9 (69.1)	5.9 (83)
Multiplicity	11.76 (12.00)	9.00 (9.63)
Completeness (%)	99.2 (97.9)	98.9 (97.9)
I/σ (I)	25.50 (4.79)	23.34 (4.28)

Table 3.2. Statistics for the processing of the high resolution data sets with XDS.

Values in parenthesis refer to reflections of the outermost resolution shell.

Calculation of the Matthew's coefficient (**table 3.3**) led to the assumption that there is one molecule in the asymmetric unit of this crystal form with a solvent content of 53.85 %.

Table 3.3. Calculation of the Matthews coefficient

Number of molecules per asymmetric unit	Matthews coefficient V_m [Da^3/Da]	Solvent content [%]	P(tot)
1	2.66	53.85	1.00
2	1.33	7.70	0.00

Space group and unit cell dimensions: P3$_1$21: a=66.39.4 Å, b=66.39 Å, c= 126.61 Å, α=β=90°, γ=120°
The Matthews coefficient VM is calculated with VM=V/MW*Z*N
V: Unit cell volume (483285.688 Å^3)
MW: Molecular weight of Poz1-spTpz1 protomer (30241.8 Da)
N: Number of molecules per asymmetric unit cell
Z: asymmetric units per unit cell

3.4 MAD data collection

The structure of spPoz1$^{30-249}$-Tpz1$^{475-508}$ was resolved using multiwavelength anomalous dispersion (MAD) on selenomethionyl derivatives. MAD diffraction data (**figure 3.12**) were collected at the SLS beamline X06SA at three wavelengths, at and around the absorption edge of selenium. In order to choose wavelengths that give optimal anomalous signals, a fluorescence scan (EXAFs=extended X-ray absorption fine structure) was performed on the crystal before data collection (**figure 3.13**). Based on the EXAFs scan the following energies were used: E=12.6631 eV (λ_{peak}=0.9791 Å) for the peak corresponding to the maximal f''. An energy of E=12.6566 eV ($\lambda_{inflexion}$=0.9796 Å) was chosen for the edge point representing the minimal f', and E=12.7635eV (λ_{remote}=0.9714 Å) for a high energy remote point. Data for a rotation range of 360° were collected for the peak and inflection. For the remote wavelength merely 180° were collected as signs of radiation damage were starting to appear. The input parameters for the MAD data collection are listed in **table 3.4**.
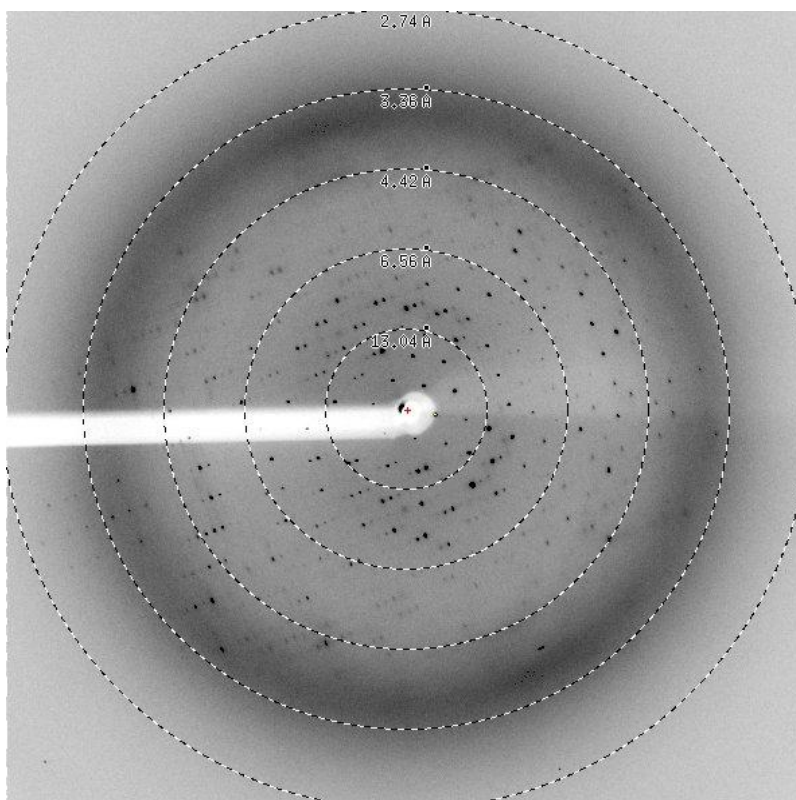
Figure 3.12. Diffraction image of a selenomethionyl derivatized spPoz1[30-249]+Tpz1[475-508] crystal.

A single X-ray diffraction image recorded during MAD data collection on the peak wavelength. Image generated with Adxv (http://www.scripps.edu/~arvai/adxv.html).
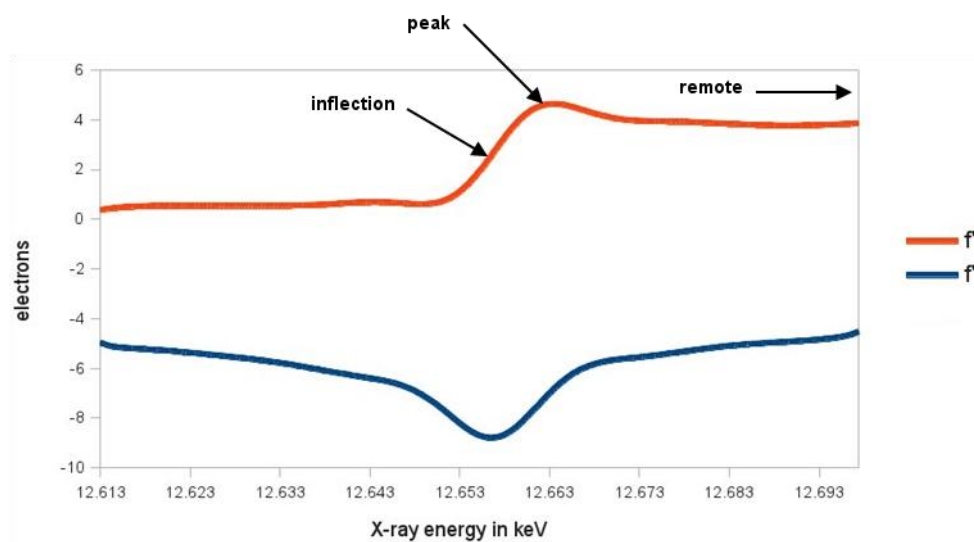
Figure 3.13. Absorption edge plot for selenium.

EXAFs scan on the selenomethionyl derivatized spPoz1[30-249]+Tpz1[475-508] crystal used for MAD data collection. f' is the real and f'' is the imaginary component of the anomalous signal of selenium. f' and f'' are plotted versus the X-ray energy. The individual X-ray energies used for data collection are indicated.

	Peak	**Inflection**	**Remote**
X-ray source	X06DA	X06DA	X06DA
Detector	marCCD225	marCCD225	marCCD225
Detector distance	300	300	300
Exposure time	1 s	1 s	1 s
Energy (ε)	12663.1 eV	12656.6 eV	12763.5 eV
$\Delta\varphi$	1.0	1.0	1.0
φ	360°	360°	180°

Table 3.4. Strategy for the collection of MAD data sets

3.5 MAD data processing

Data images from selenomethionyl derivatized protein crystals were processed using the XDS package (Kabsch, 2010). Crystals belong to the same space group as native crystals and data processing statistics are given in **table 3.5**. Statistics for the anomalous signal are given in **table 3.6**.

X-ray data statistics of MAD data			
	Peak	**Reflection**	**Remote**
Space group	P3$_1$21		
Unit cell	a=b=67.03, Å c=124.4 Å α=β=90, γ=120		
Beamline	X06DA		
Wavelength (Å)	0.9791	0.9796	0.9714
Overall resolution (Å)	15.65-3.5	16.10-3.6	16.10-3.6
Unique Bijvoet's	7784	7159	7138
R$_{sym}$(I)	11.4	11.2	18.0
Multiplicity	11.72	11.72	5.84
Completeness (%)	98.3	98.2	97.9
I/σ (I)	20.15	22.87	15.16
Outer resolution shell (Å)	3.59-3.50	3.69-3.60	3.69-3.60
Unique Bijvoet's	563	549	543
R$_{sym}$(I)	48.7	30.7	49.7
Multiplicity	11.8	11.72	5.76
Completeness (%)	98.8	98.9	97.8
I/σ (I)	4.03	6.35	4.53

Table 3.5. Statistics after processing of MAD data with XDS

Resolution [Å]		S_norm / S_ano	
	Peak	Inflection	Remote
9.3	5.61	2.35	3.18
6.6	4.44	2.00	2.22
4.5	1.43	0.90	0.97
3.6	0.91	0.82	0.82

Table 3.6. Estimation of the anomalous signal with XDS.

S_norm: mean value of Sigma(I) for acentric reflections assuming Friedel's law is valid.
S_ano: mean value of Sigma(I) for acentric reflections assuming Friedel's law is violated.
Anomalous scattering contributions to the intensities are indicated by S_norm/S_ano >1

3.6 Determination of the sub-structure of anomalous scatterers

The positions of the anomalous scatterers in the unit cell were determined with the software SOLVE (Terwilliger, T.C. and J. Berendzen, 1999) within the software package Phenix (Adams *et al.*, 2010) by inputting the three MAD wavelengths containing the raw, unmerged intensities. SOLVE searches for heavy atoms using difference Patterson methods and scores potential solutions on the basis of agreement with the Patterson map, the difference Fourier map, the presence of solvent and protein regions in a native electron density map, and the figure of merit of phasing. Since the calculated Matthew's coefficient predicted a single complex in the asymmetric unit (**table 3.3**), and spPoz1 contains 8, and spTpz1 1 methionine residues, a search for 9 heavy atom positions was carried out using data ranging from 20 to 3.6 Å resolution. 7 out of a total of 9 scatterers in the asymmetric unit were found.

Given the ambiguity in space group symmetry remaining at that point ($P3_121$ and $P3_221$ enantiomeric lattices), both direct and inverted positions of heavy atom clusters were assayed in phasing. Visual inspection of the resulting density maps finally revealed this crystal form to belong to the space group $P3_121$.

A first electron density map obtained in SOLVE was subsequently subjected to maximum likelihood density modification in RESOLVE (Terwilliger, 2000, 2003, 2004) in order to improve the quality of the map (**table 3.7**). We could readily identify several α-helices and the resulting map allowed model building (**figure 3.14A**).

	Resolution	FOM-prior	FOM-total
SOLVE/RESOLVE	All	0.46	0.71
SHARP/SOLOMON	All	0.42	0.91

Table 3.7. Phasing statistics after density modification.

FOM = Figure of merit. Values are shown for acentric reflections.

As a second approach, phases were independently calculated and refined in SHARP (Vonrhein *et al.*, 2007) using native data up to 3.5 Å resolution and derivative data to 3.5 Å resolution, and by supplying the heavy atom coordinates previously obtained in SOLVE. The resulting electron density maps calculated by SHARP had clearly identifiable solvent-protein regions and partially interpretable features. Subsequent solvent flattening using the SOLOMON procedure (Abrahams and Leslie, 1996) implemented in SHARP assuming a solvent content of 53.85 % and phase extension to 2.6 Å furthermore improved the quality of the secondary structure features in map (**table 3.7**) (**figure 3.14B**).

A   **B**
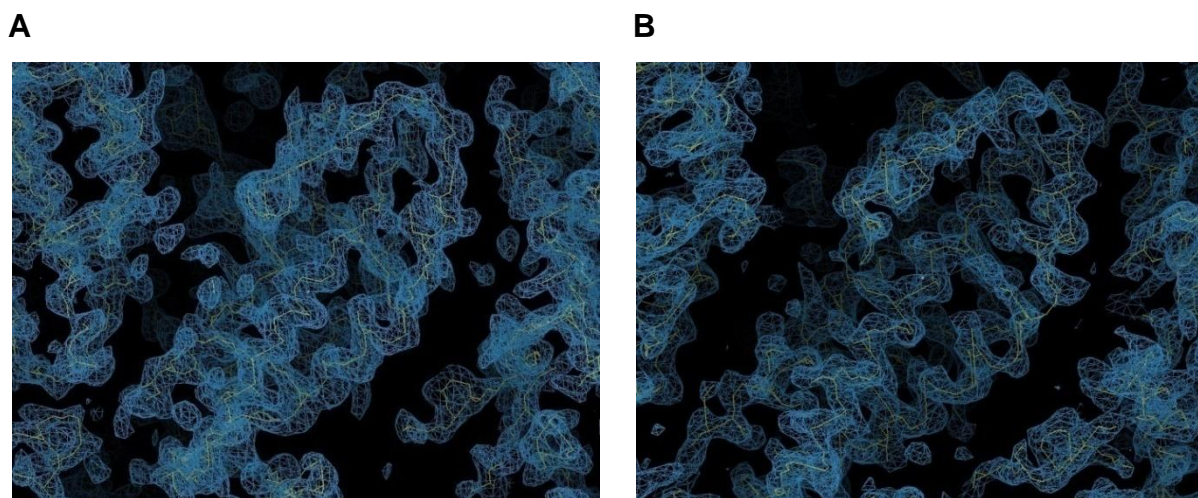
Figure 3.14. Initial electron density map obtained from A) SOLVE/RESOLVE at 3.6 Å maximum resolution and B) SHARP at 2.6 Å maximum resolution.
$2F_{obs}$-F_{calc} map is displayed at a contour level of 1.42 σ. A crude skeleton has been generated based on the map in COOT to visualize the readily identifiable α-helices.

Both electron density maps obtained from SHARP and RESOLVE were of reasonable quality to start model building. Combining the two maps for model building yielded optimal results. Some features where better identifiable in the map originating from SHARP, especially side-chains. On the other hand, flexible regions were better defined in the map obtained from RESOLVE.

3.7 Model building and refinement

As a first step towards model building, poly-alanine α-helices were manually placed in regions of the electron density maps that clearly displayed the typical α-helical appearance (COOT, Emsley and Cowtan, 2004; Emsley *et al.*, 2010). These were then progressively edited in iterative cycles of model building and refinement. The structure was refined against native data ranging from 20 Å to 2.6 Å resolution in REFMAC5 (Murshudov *et al.*, 1997). Due to the relatively high resolution of the native data, automatic model building using various programs such as RESOLVE, Buccaneer (Cowtan, 2006) and ARP (Lamzin and Wilson, 1997, Morris *et al.*, 2003) was attempted. These, however, did not yield any models that were more complete than those where α-helices were manually placed. After each cycle of model building and refinement, gradually more features appeared in the electron density, which allowed the addition of more residues, also in non-secondary structure regions. As more details appeared in the maps, some side chains could be built and in combination with the knowledge of the positions of the selenomethionine residues, the amino acid sequence of secondary structure elements could be determined, as well as the connectivity. Numerous cycles of model building and refinement resulted in a near complete model of the Poz1-Tpz1 complex. It was at this point that a second native data set was obtained with higher resolution, such that refinement could be carried out to 2.4 Å. This in turn allowed building some features that could not be identified in the previous electron density map. Finally, most of the model could be built with the exception of the residues that were not identifiable in the electron density, as they are most likely situated in unstructured loops. Solvent atoms were built with ARP/wARP (Lamzin and Wilson, 1997). The waters were subsequently visually inspected in COOT. The final model contained 178 amino acids in spPoz1 and 29 residues in spTpz1 and 22 water molecules. Residues 70-85, 118-121 and 236-249 of spPoz1 could not be built. Furthermore, three residues in the N-terminus of spTpz1, 475-477, and two residues in the C-terminus, 507-508 are missing in the final model (**table 3.8**) (see **section 4.1** for the final structure). The geometry of the structure was analyzed with the program RAMPAGE (Lovell *et al.*, 2003) (**figure 3.15**).

Number of residues in favored region (~98.0% expected): 201 (97.1%)

Number of residues in allowed region (~2.0% expected): 6 (2.9%)

Number of residues in disallowed region: 0 (0.0%)
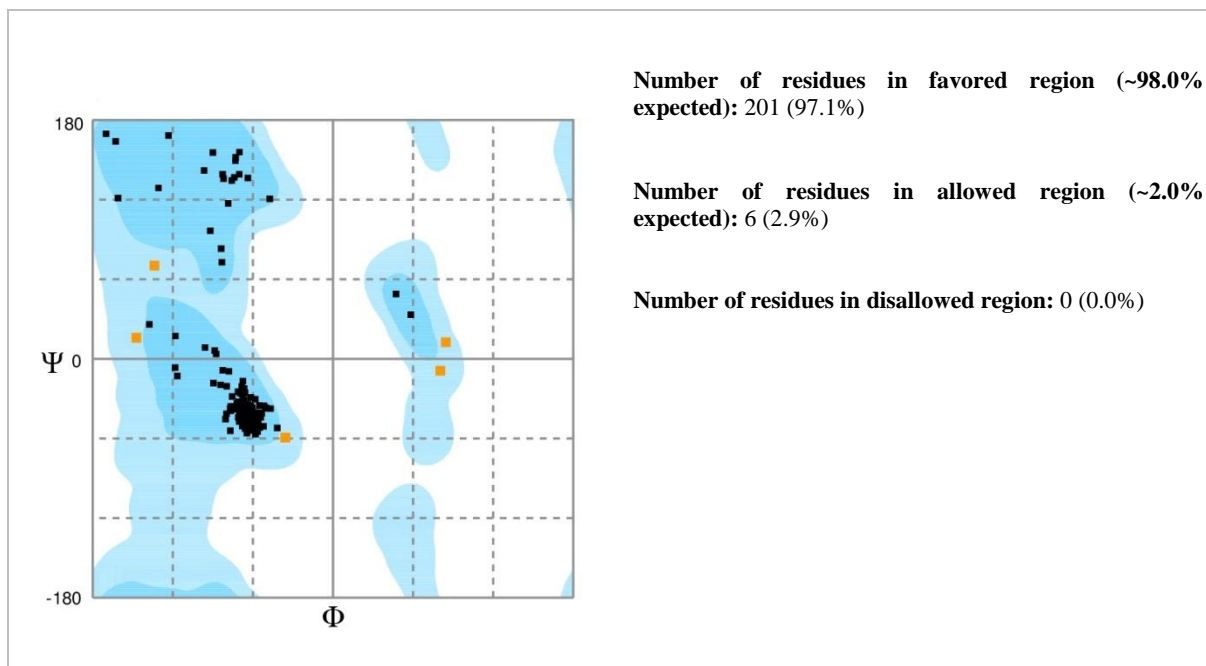
Figure 3.15. Analysis of the geometry of the final spPoz1$^{30-249}$+Tpz1$^{475-508}$ model.

Ramachandran plot (Ramachandran, 1963) generated with the program RAMPAGE.

Number of reflections in working / free set	13038 / 2473
Number of protein residues / solvent molecules	207 / 22
R-factor / R-free (%)	21.47 / 24.47
Rmsd bond length (Å) / bond angle (°)	0.010 / 1.10

Table 3.8. Final refinement statistics of spPoz1$^{30-249}$+Tpz1$^{475-508}$.

Chapter 4

Results - Part II

4. Results – Part II: Structure and function of fission yeast Poz1-Tpz1

4.1 The crystal structure of spPoz1$^{30\text{-}249}$ + Tpz1$^{475\text{-}508}$

The crystal structure of fission yeast *(Schizosaccharomyces pombe)* Poz1$^{30\text{-}249}$ in complex with Tpz1$^{475\text{-}508}$ has been elucidated at 2.4 Å resolution (**figure 4.1**) using multiwavelength anomalous dispersion (MAD) phases obtained from a selenomethionyl variant (see results **part I**).
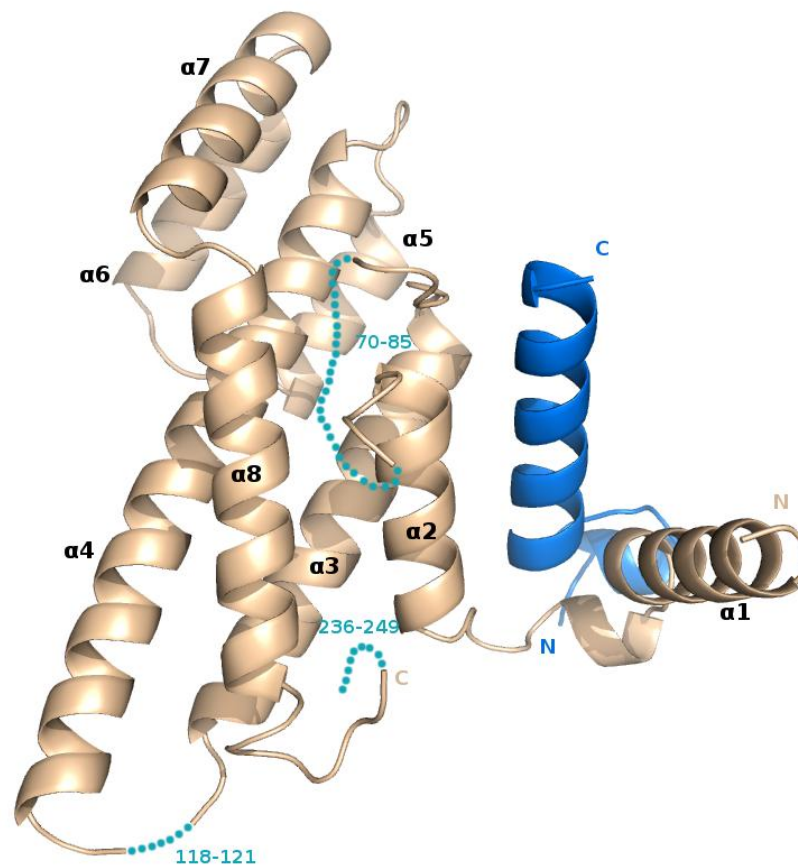
Figure 4.1: The crystal structure of spPoz1$^{30\text{-}249}$+Tpz1$^{475\text{-}508}$.

Poz1 is colored in wheat, Tpz1 in blue. Secondary structure elements are labeled in Poz1. The regions that were not built due to missing electron density are sketched in turquoise.

The complex shows a compact globular, exclusively α-helical structure. Poz1$^{30\text{-}249}$ is composed of 8 α-helices. The loops containing residues 70-85 connecting α2 and α3 and the loop containing residues 118-121 connecting α3 and α4 were not built due to the absence of electron density for these regions in the map. Additionally, the final Poz1 residues 236-249

were not visible in the electron density, most likely due to the flexible nature of the C-terminus. Helices α3, α4, α5 and α8 form a four helix bundle. Tpz1 is found locked in between α1 and α2 of Poz1. For Tpz1 three residues in the N-terminus, 475-477, and two residues in the C-terminus, 507-508 were not visible in the electron density map.
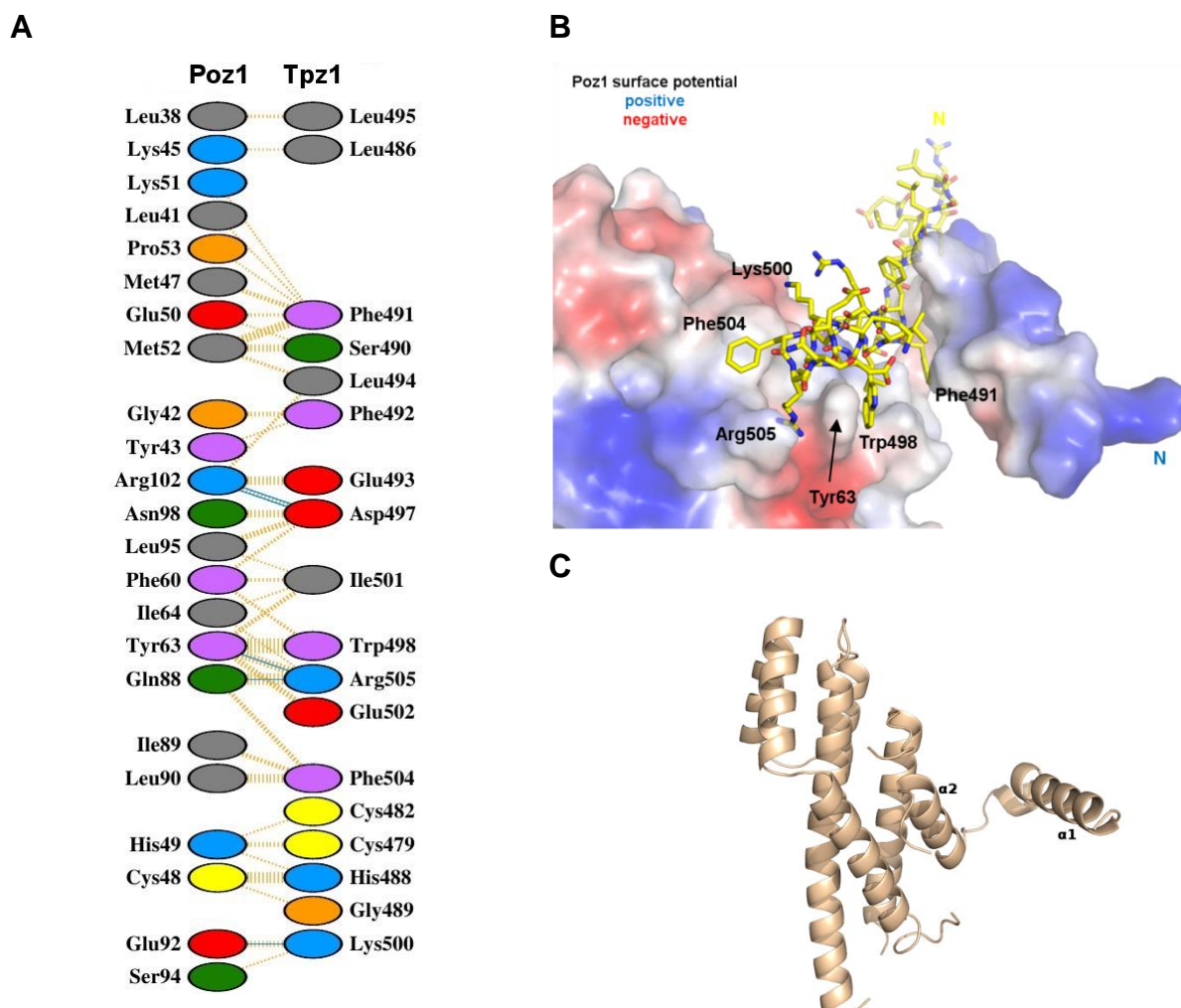
Figure 4.2. Interactions formed between Poz1[30-249] and Tpz1[475-508].

A) List of residues involved in intermolecular interactions. Hydrogen-bonds are shown by blue lines while non-bonded contacts are marked by orange dashed lines. Residues are color coded as follows: Positively charged in light blue, negatively charged in red, neutral in green, aliphatic in grey, aromatic in purple, Pro&Gly in orange and cysteine in yellow. Figure was generated by PDBsum (Laskowski 2001, http://www.ebi.ac.uk/pdbsum/). **B)** View of major contacts formed between Poz1 and Tpz1. Poz1 is shown as a surface model with the surface potential displayed. Stick model of Tpz1 is colored in yellow. **C)** Model of the Poz1 structure with Tpz1 removed.

Numerous interactions are formed at the interface of Poz1 and Tpz1, suggesting a tight complex (**figure 4.2A and B**). The interactions include a total of 5 hydrogen bonds and 122 non-bonded contacts including hydrophobic and ionic interactions as well as base-stacking.

Major contacts include Phe491 of Tpz1 that inserts in between helix α1 and α2 of Poz1 stabilizing α1. Asp497 of Tpz1 forms two hydrogen bonds with Arg102 of Poz1 (not visible in **figure 4.2B**). Lys500 of Tpz1 contacts Glu92 on the surface of Poz1 and Phe504 of Tpz1 makes several non-bonded contacts to the hydrophobic surface of Poz1. Trp498 of Tpz1 is involved in stacking interactions with Tyr63 of Poz1. Furthermore, Arg505 of Tpz1 points into a negatively charged pocket of Poz1.

Helix α1 of Poz1 is connected to α2 by an unstructured linker and is held in position by the interaction with Tpz1. In the absence of Tpz1-binding (**figure 4.2C**), we would expect α1 to be flexible allowing it to adopt numerous conformations, which poses a likely explanation why attempts to crystallize Poz1$^{30-249}$ by itself did not result in crystals.

4.2 Structural similarity to human TRF1 and TRF2

Preceding this work, little information was available on Poz1. BLAST (Altschul *et al.*, 1990) searches as well as Hidden-Markov-Model-(HMM)-based structure prediction (Biegert *et al*, 2006) did not yield any homologous proteins, nor did they reveal any known protein domains or sequence motifs. Most likely TIN2 acts as the functional homolog to Poz1 in the human shelterin complex. This assumption, however, was made based on the positional similarity in the shelterin complex and not on sequence similarity.

Solving the structure of Poz1-Tpz1 allowed us to perform a structure similarity search in DALI (Holm and Rosenström, 2010). The DALI server compares a query structure to the structures in the protein data bank (Berman, 2000). DALI found two relevant candidate structures, which were the "telomere repeat binding factors" TRF1 and TRF2 found in the human shelterin complex. The structures of TRF1 and TRF2 that were solved (Fairall *et al.*, 2001 and Chen *et al.*, 2008) correspond to the TRF-homology domains (TRFH). The TRFH domains mediate homodimerization and are structurally almost identical to TRF1 and TRF2. We subsequently superimposed our Poz1-Tpz1 structure onto TRF1 and TRF2 (**figure 4.3, table 4.1** and **figure 4.4** for the sequence alignment with secondary structures).
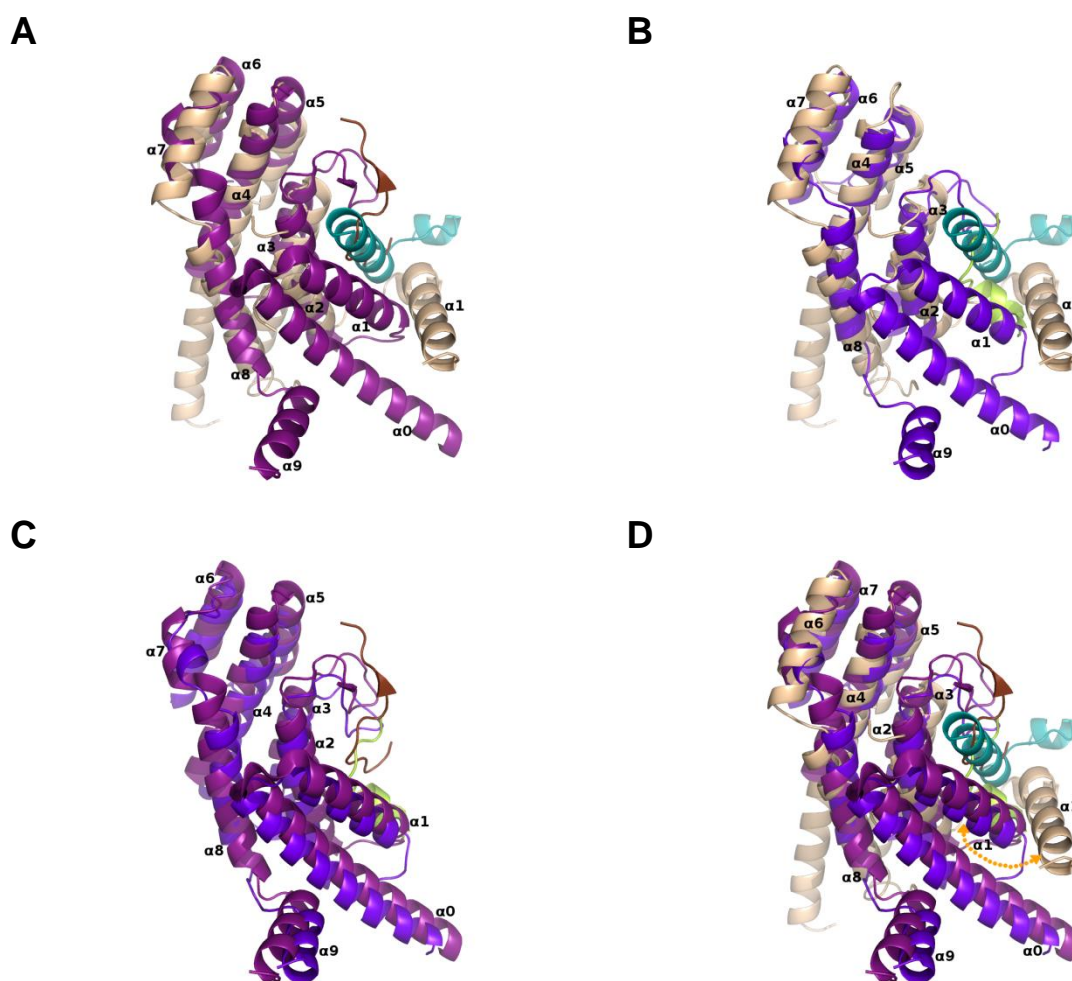
Figure 4.3. Superpositions of spPoz1$^{30-249}$-Tpz1$^{475-508}$ onto hTRF1$^{62-268}$-TIN2$^{256-268}$ and hTRF2$^{86-287}$-Apollo$^{497-510}$.

Poz1 is colored in wheat. Tpz1 in blue, TRF1 in magenta, TIN2 in brown, TRF2 in purple and Apollo in green. **A)** Poz1-Tpz1 aligned with TRF1, **B)** Poz1-Tpz1 and TRF2, **C)** TRF1 and TRF2 and **D)** Poz1-Tpz1, TRF1 and TRF2. An orange arrow indicates the difference in conformation of α1 in Poz1 compared to α1 in TRF1 and TRF2.

	spPoz1$^{30-235\ (54-235)}$	hTRF1$^{62-268\ (115-268)}$	hTRF2$^{86-287\ (143-273)}$
spPoz1$^{30-235\ (54-235)}$	-	3.171 (3.014)	2.768 (2.757)
hTRF1$^{62-268\ (115-268)}$	-	-	2.057 (2.054)
hTRF2$^{86-287\ (143-273)}$	-	-	-

Table 4.1. Rmsd (root-mean-square deviation) values [Å] for the superpositions of spPoz1, hTRF1 and hTRF2.
Residues and rmsd values in brackets are for superpositions using boundaries excluding outlier regions that are not commonly shared among these structures. Rmsd values were calculated in COOT (Emsley *et al.*, 2010).

The overall agreement of the C$_\alpha$-backbones is remarkable. Almost all secondary structure elements are present in each structure and only minor positional shifts are observed

(**figure 4.4**). This observation was surprising, as no apparent sequence similarity of Poz1 to TRF1 or TRF2 was found.

In the superpositions of Poz1-Tpz1 with TRF1-TIN2 and TRF2-Apollo (**figure 4.3**), helices α1 to α8 of Poz1 align well with the corresponding helices in the human proteins, albeit with some differences in length. Helices α3 and α4 are roughly double the length in Poz1 compared to their counterparts in the TRF proteins. The most significant difference between the two structures is α1 of Poz1. This helix is found disconnected from the core of the protein, while in the TRF proteins this helix is packed closely to the helix bundle (**figure 4.3D**). The region analogous to helix α0 of the TRF proteins is not present in the structure of Poz1-Tpz1, as it was necessary to truncate the first 29 amino acids of Poz1 due to the insolubility of the fulllength protein. Based on secondary structure prediction this region might also form an α-helix. Either the truncation of this section is the reason for the different conformation of α1 in Poz1 compared to α1 in TRF1 and TRF2, or this helix is naturally positioned differently than in the TRF proteins. The region that corresponds to the C-terminal helix α9 of TRF1 and TRF2 is not visible in the Poz1 structure, though secondary structure prediction suggests that it also forms a helix. **Figure 4.3C** shows how the structure of TRF1 and TRF2 superpose, despite the low sequence conservation.
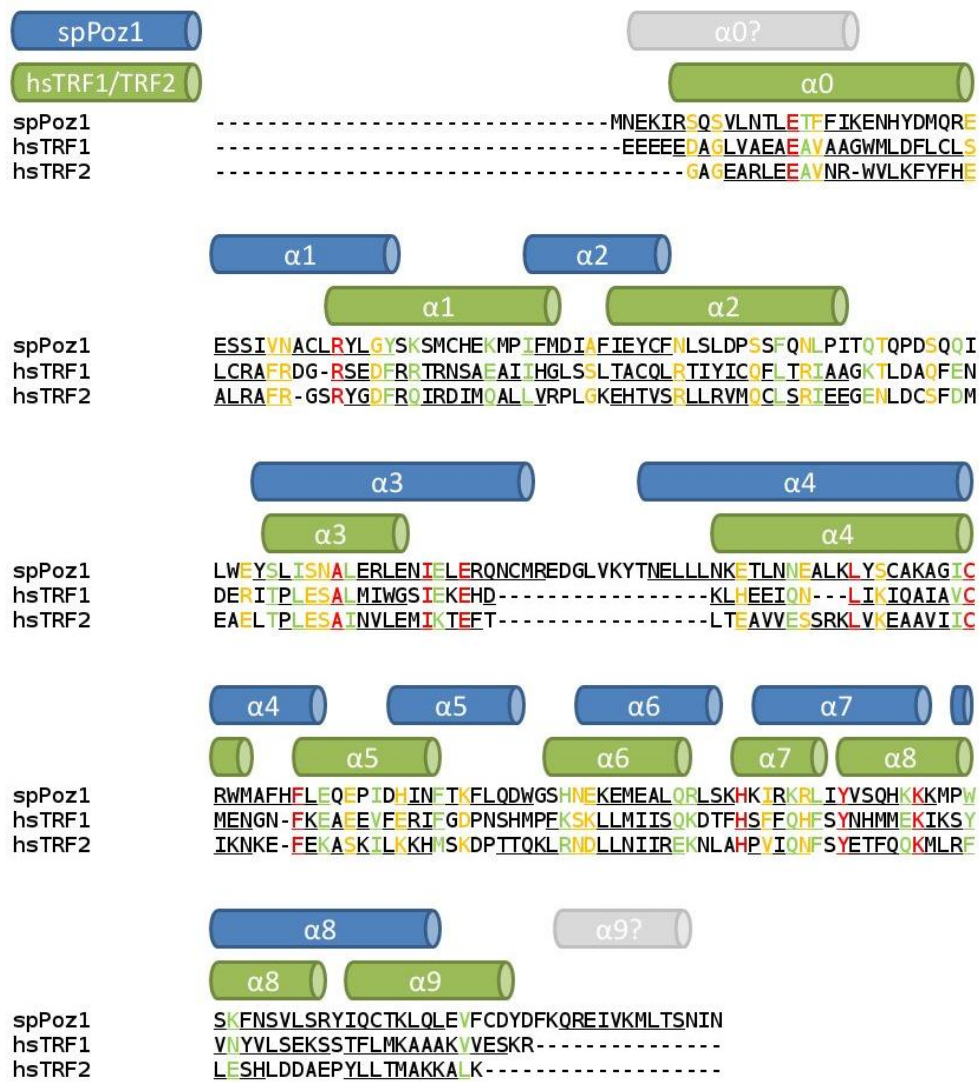
Figure 4.4. Sequence alignment with highlighted secondary structure elements.

The sequence of Poz1[1-249] was aligned with the sequence of TRF1[58-268] and TRF2[84-287] in ClustalW (Larkin *et al.*, 2007). Blue cylinders represent α-helices in Poz1, green cylinders α-helices in TRF1 (not shown for TRF2 for simplicity. They are almost identical to those in TRF1). Grey cylinders represent regions that are not present in the structure of Poz1 but are predicted by PSI-pred (Jones, 1999; Buchan *et al.*, 2010) as α-helices. Red letters show strictly conserved residues, green letters show residues that are conserved in an amino acid group and yellow letters represent residues that are conserved across groups.

4.3 A conserved binding pocket among Poz1, TRF1 and TRF2

Our structure further revealed the location of the Tpz1 binding region in Poz1. From the superpositions with the human TRF complexes, we observe that the binding pockets of TIN2 in TRF1 and Apollo in TRF2 co-localize to the binding pocket of Tpz1 in Poz1 (**figure 4.5**).
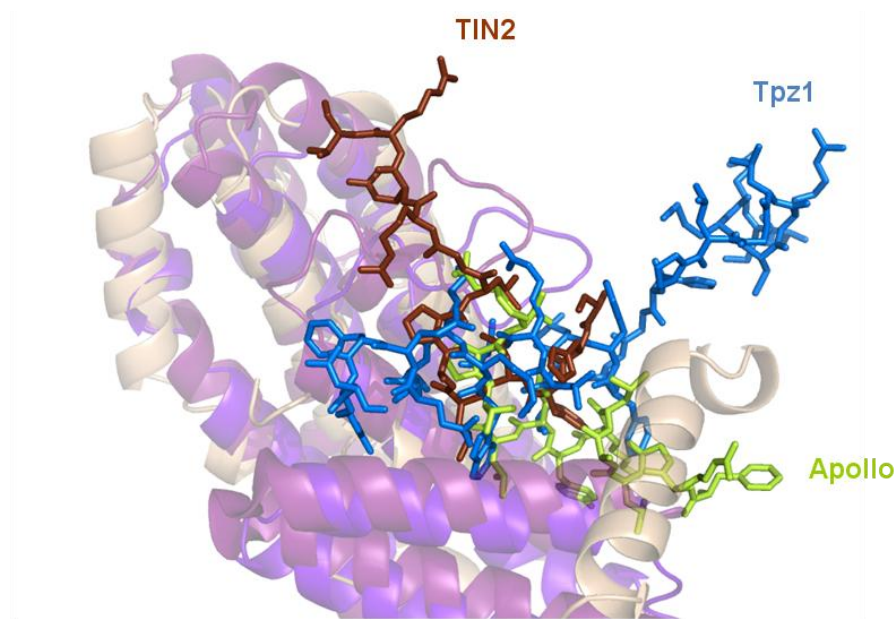
Figure 4.5. The binding pockets in spPoz1, hTRF1 and hTRF2.

Superposition reveals a common binding region found in all three proteins. Poz1 is colored in wheat, TRF1 in magenta and TRF2 in purple.

TRF1 and TRF2 use their binding pockets to bind a number of different telomere associated proteins which all share the common binding motif F/Y-X-L-X-P (where X may be any amino acid) (Chen *et al.*, 2008). The common binding region suggests that the binding pocket of Poz1 may also represent a module for binding different telomeric proteins, although to date only Tpz1 is known to bind Poz1 in this location. Assuming that the fission yeast binding motif contains similar residues to those found in the human TRFH binding motif, we searched for an analogous sequence in the Tpz1 residues bound to Poz1. We found that in the binding region, Tpz1 contains residues F-X-L while the remaining residues of the human motif are not present. One possibility is that F-X-L is indeed part of a recognition motif and the remaining residues are different compared to the human ones. A motif search in the

fission yeast genome (Hertz-Fowler *et al.*, 2004. http://old.genedb.org/) for F-X-L expectedly returned far too many results, preventing the validation of our hypothesis.

Upon close inspection of the hypothetical F-X-L motif in our structure (**figure 4.6**), we see that F492 of Tpz1 is positioned differently than their hypothetical counterparts F258 and Y504 of TIN2 and Apollo, respectively. This may be due to the different conformation of helix α1 in Poz1 with which it interacts. Furthermore, given the low sequence similarity of Poz1 and the TRF proteins, it is not surprising to find differences in the amino acids involved in the binding pocket. L494 superimposes with the position of F/Y. Residues W498 and K500 of Tpz1 superimpose with the conserved motif residues L260^{TRF2}/L506^{Apollo} and P262^{TRF2}/P508^{Apollo}.
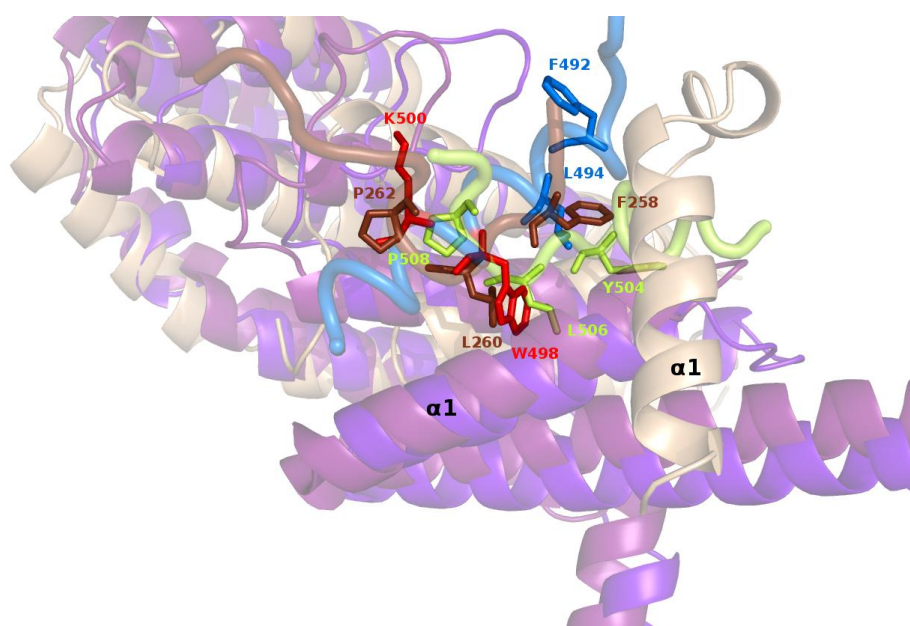
Figure 4.6: Molecular details of the substrates in the binding clefts.

Tpz1 is in blue, TIN2 in brown, Apollo in green. The residues that form the F/Y-X-L-X-P recognition motif in TIN2 and Apollo are shown as sticks. Residues that could be part of a motif and share the sequence with the human motif in Tpz1 and are shown in blue, residues in Tpz1 that align with the human motif but differ in sequence are in red.

From the superposition a hypothetical motif F-X-L-X-X-X-W-X-K was derived. A search for this motif in the yeast genome returned two results, midasin and Cwf5. Given the fact that F492 of the hypothetical motif is positioned at a larger distance from the binding pocket compared to F258 in TIN2 and Y504 in Apollo, we speculated that this residue might not be part of the motif in fission yeast, and performed another motif search excluding this

region (L-X-X-X-W-X-K). This search returned a total number of 127 proteins with a variety of functions. The most interesting finding was spTaf1, also known as Taz1 interacting factor 1 (Ueno et al, 2001). Taf1 as well as other candidates need to be further examined to determine whether they play a role at the telomeres.

4.4 Oligomeric state of Poz1-Tpz1

The structure of Poz1-Tpz1 revealed a remarkable similarity to the human shelterin components TRF1 and TRF2. The structures of TRF1 and TRF2 that had been solved, are the TRF-homology domains (TRFH), which in TRF1 and TRF2 are responsible for dimerization. TRF1 with TIN2 and TRF2 with Apollo form heterotetrameric arrangements. This finding lead to the hypothesis, that Poz1 like TRFH, also functions as a module for dimerization in the fission yeast shelterin complex.

4.4.1 Arrangement of Poz1-Tpz1 in the crystal lattice – The heterotetramer hypothesis

As an initial step, we analyzed the arrangement of the Poz1-Tpz1 molecules in the crystal lattice in an attempt to identify a hypothetical heterotetrameric composition. Should the complex indeed form heterotetramers, we would expect to observe such an arrangement in the crystal.

Visual inspection showed that throughout the $P3_121$ crystal lattice there are two Poz1-Tpz1 monomers that come together in a head-to-head manner in the region of the Poz1 N-terminus and the bound Tpz1 fragment. This could indeed represent a hypothetical heterotetrameric arrangement (**figure 4.7).**
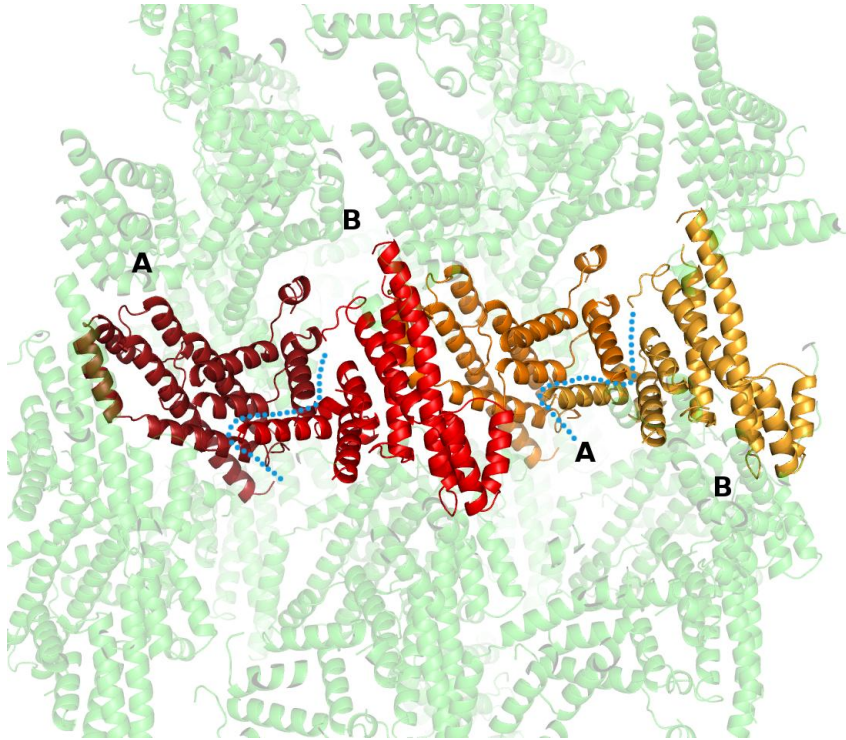
Figure 4.7. The P3₁21 crystal lattice of the spPoz1-Tpz1 crystals used for structure solution.

Two hypothetical heterotetramers are highlighted in the crystal lattice. One heterotetramer is composed of a Poz1-Tpz1 protomer *A* colored in dark red and a protomer *B* colored in light red. The protomers of a second heterotetramer are colored in dark yellow and light yellow. The heterotetramerization interfaces are indicated by blue dashed lines.

In a next step we submitted the structure of Poz1-Tpz1 to the PDBePISA server (Krissinel and Henrick, 2007) which analyzes macromolecular interfaces and performs prediction of probable quaternary structures. PISA returned a hypothetical stable heterotetrameric arrangement composed of two Poz1 and two Tpz1 molecules (**figure 4.8**), which corresponded to the heterotetramer we had suggested from visual inspection of the crystal lattice.
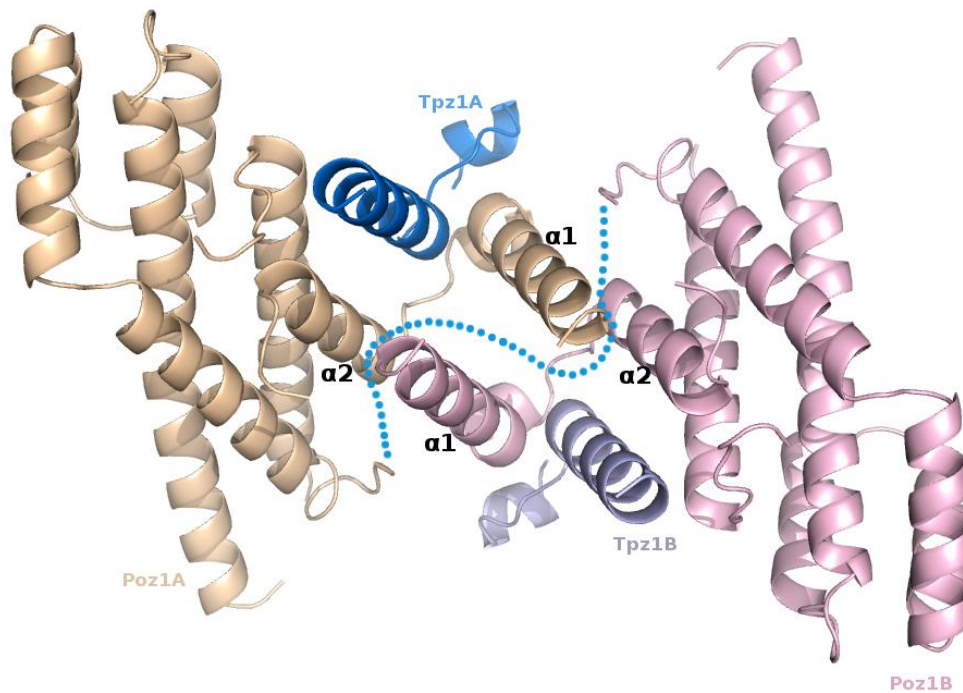
Figure 4.8. Stable spPoz1-Tpz1 heterotetramer as predicted by PISA and as found in the crystal lattice.
Helices forming major contacts to the opposite protomer are labeled. The interface boundary is indicated by a blue dashed line.

In this hypothetical heterotetramer, the interface is formed by α1 and α2 of Poz1 in each protomer. Following this, we set out to investigate whether there are any similarities of this architecture to the heterotetrameric arrangements found in TRF1 and TRF2. To examine this, the predicted heterotetrameric model of Poz1-Tpz1 was superimposed onto the TRF1 and TRF2 heterotetramers (**figure 4.9**) (for simplicity only the superposition with TRF1-TIN2 is shown as the superposition with TRF2-Apollo returned an identical result).

A

B
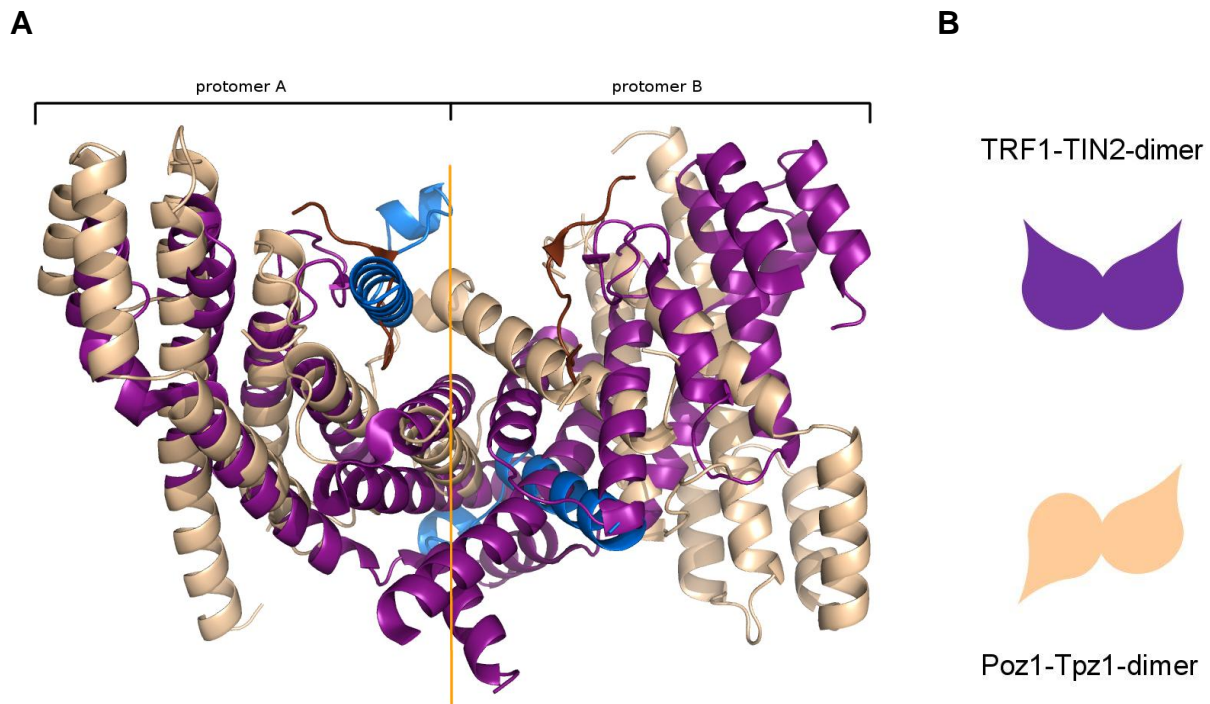
Figure 4.9. Comparison of the spPoz1-Tpz1 heterotetramer with the hTRF1-TIN2 heterotetramer.
Poz1 and Tpz1 are colored in wheat and blue, TRF1 and TIN2 in magenta and brown respectively. **A)** Superposition of Poz1-Tpz1 and TRF1-TIN2. The orange line represents a common divider between the protomers in each complex. **B)** Schematic representation of the protomer arrangement in the heterotetramer.

Of interest, The C_α-backbones of protomers *A* superimposed well as expected, while the second protomers *B* did not seem to structurally align. After close inspection of protomers *B*, it became clear that the disagreement arises from the fact that the Poz1-Tpz1 subunit *B* is rotated by 180° with respect to protomer B of TRF1. In the Poz1-Tpz1 heterotetramers, the protomers are not mirrored as in TRF1 and TRF2, however, are rather related by a 180° rotation around the heterotetramer interface (**figure 4.9B**). Besides this 180° rotation, we found that the heterotetrameric arrangements are similar and a common boundary between the protomers of Poz1-Tpz1 and TRF1-TIN2 can be observed (**figure 4.9A**).

Effectively, α-helix α1 in Poz1-Tpz1, which makes the major contributions to the heterotetramerization interface, also plays a role in forming the interface in the TRF complexes (**figure 4.10**). Helix α0 in the TRF proteins is also part of the multimerization interface. This helix is absent in our Poz1-Tpz1 structure, because in order to obtain a soluble protein, residues 1-29 in Poz1 corresponding to α0 in TRF were removed. Based on structure prediction, α0 is also present in Poz1. Thus, as in TRF, it is likely to be involved at the heterotetramerization interface.

In TRF, a short C-terminal helix α9 is also found at the heterotetramerization interface. In Poz1, this region has been predicted to form a short helix, however, we were not able to find any electron density for this part of the protein. This could be explained by the fact that the N-terminal helix of Poz1 was removed: In the TRF-heterotetramer, the N-terminal helix α0 interacts with the C-terminal helix α9 of the other monomer, keeping it packed close to the core of the protein (**figure 4.10B**). Possibly, this is also the case in Poz1, however, upon removal of the N-terminal helix, the C-terminal helix is no longer bound to the core and therefore inherits flexibility because it is merely connected to α8 via a flexible linker (**figure 4.10A**). Thus, due to the high degree of flexibility we fail to identify this region in the electron density.

Additionally, helix α2 in Poz1 is also part of the interface, whereas in TRF this helix is not involved. Overall, the observed differences between the heterotetramerization interfaces may be due to the N-terminal truncation of Poz1. On the other hand, it is also possible that the interfaces are intrinsically different. In any case, these results strengthen the likelihood of heterotetramerization in Poz1-Tpz1.

A
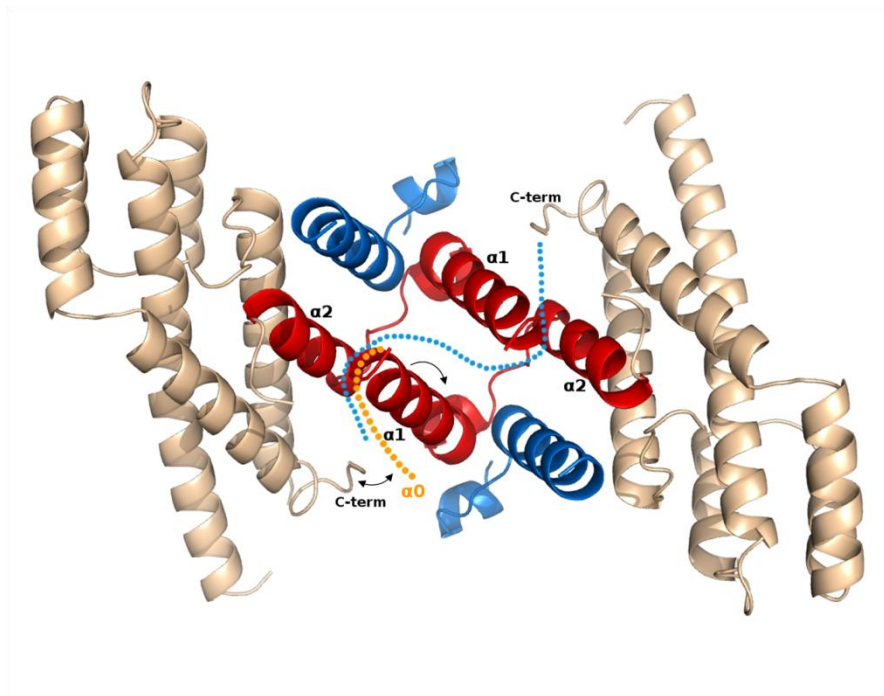
B
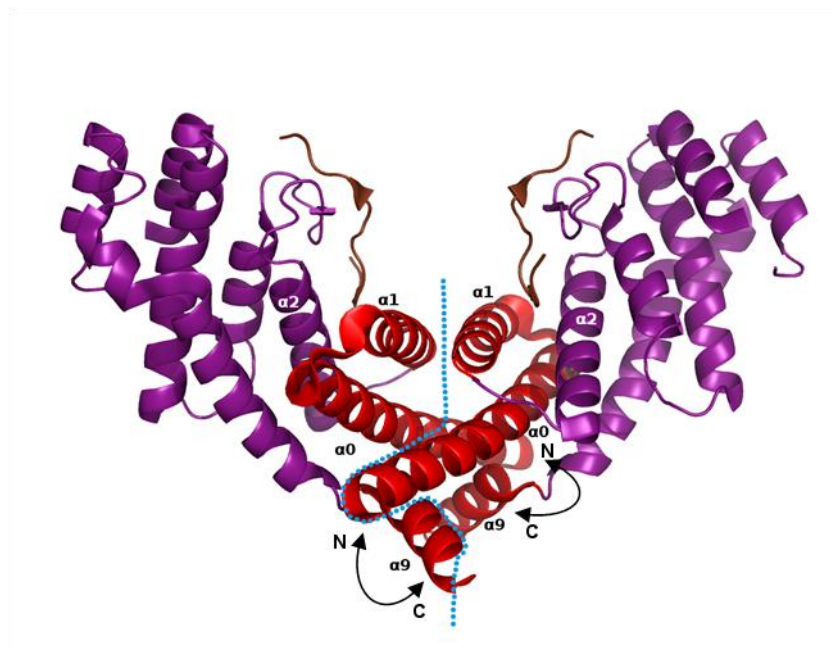
C
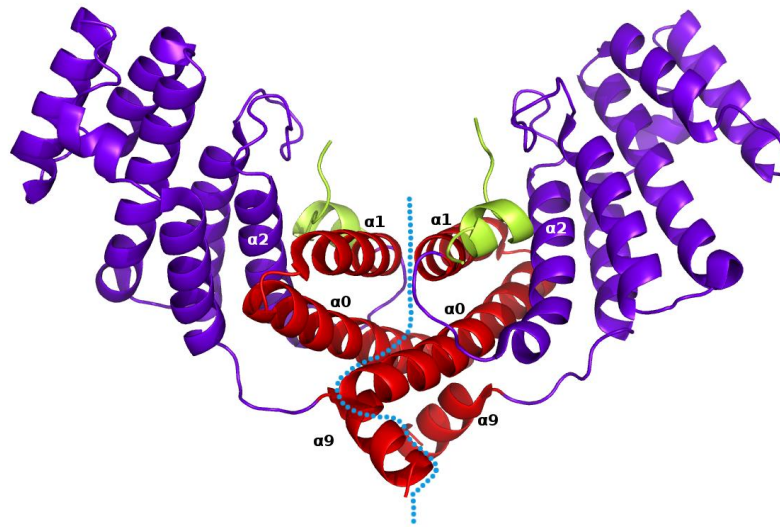
Figure 4.10. Overview of secondary structure elements involved at the heterotetramerization interface.

A) Poz1-Tpz1 in wheat and blue. A hypothetical position of α0 in the Poz1-Tpz1 heterotetramer is indicated in orange. In this conformation, α0 and α9 from another protomer could interact as α0 and α9 in TRF1 and TRF2 **B)** TRF1-TIN2 in magenta and brown **C)** TRF2-Apollo in purple and green. α-helices colored in red represent those that form interactions with the opposing molecule. A dashed blue line represents the boundary between protomers.

4.4.2 Oligomeric state of Poz1-Tpz1 in solution

In order to validate our heterotetramerization hypothesis, we designed a mutant complex in which the hypothetical multimerization interface observed in the crystal lattice was disrupted. Subsequently, the mutant complex was compare to the wildtype protein in size-exclusion chromatography (SEC). In SEC, proteins are separated based on their molecular weight. Larger proteins elute fast while smaller proteins get delayed in the resin matrix and elute later. Thus, the transition from a multimeric to a monomeric complex can be observed by a shift of the UV-peak to a lower elution volume.

4.4.2.1 The Poz1-Tpz1 heterotetramerization mutant

From analysis of the heterotetramerization interface in the Poz1-Tpz1 lattice (**figure 4.8**), we observed that most interactions are formed between helices α1 from the two protomers. Following this, mutations in helix α1 were introduced, such that two residues with the same charge are positioned closely opposite each other, which would lead to a charge mediated repulsion of the helices and in turn disruption of the heterotetramerization interface. In order to achieve this, residues Val34 and Cys37 in Poz1 were mutated to Glu34 and Arg37, respectively. This generated two negative and two positive charges facing each other in close proximity in the potential heterotetramer (**figure 4.11**). These mutants were subsequently subjected to size- exclusion chromatography and light scattering for analysis of their oligomeric state.
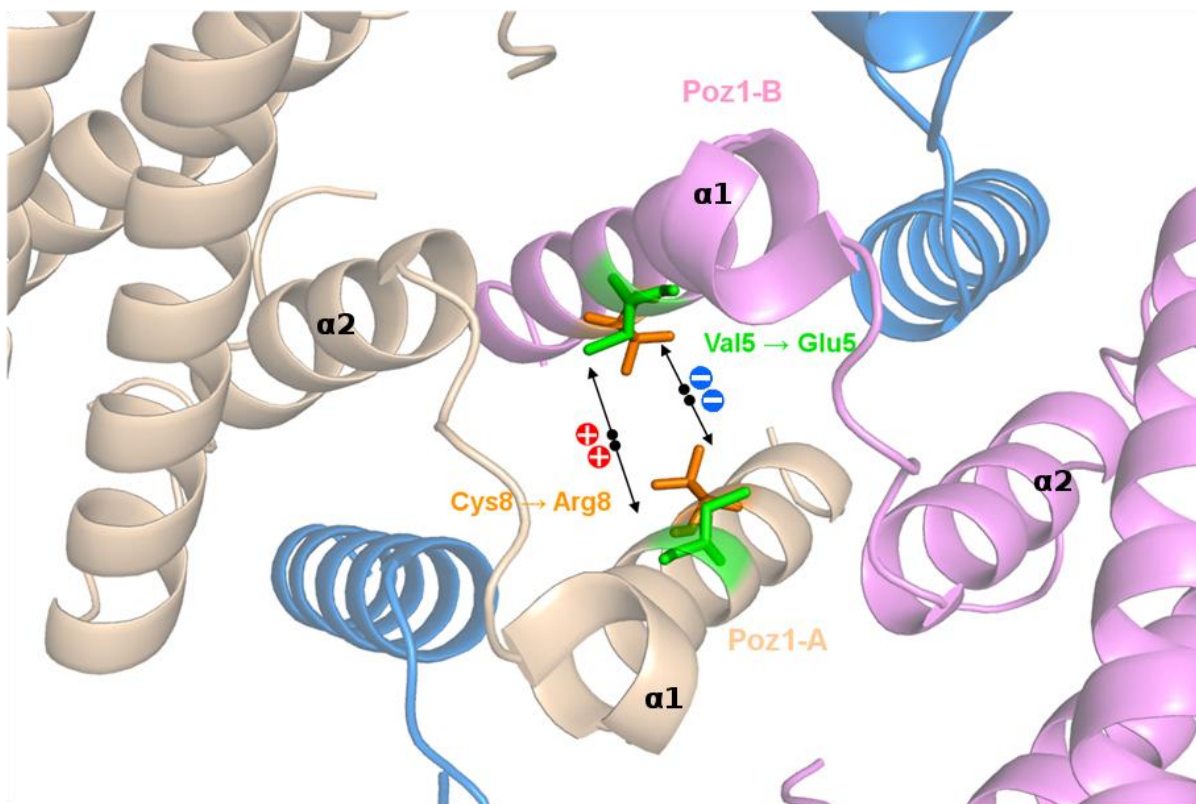
Figure 4.11. Close-up view of the Poz1-Tpz1 heterotetramerization interface.

Helices α1 form the majority of interactions at the heterotetramerization interface. The residues in α1 that are expected to disrupt heterotetramerization upon mutation are represented as sticks in the model.

4.4.2.2 Analytical size-exclusion chromatography

By analytical size-exclusion chromatography (SEC), we anticipated to observe shifts between the elution volumes between Poz1$^{30-249}$, Poz1$^{30-249}$+Tpz1$^{475-508}$ and the heterotetramerization mutant Poz1$^{30-249}$_V34E_C37R+Tpz1$^{475-508}$.

SEC of Poz1$^{30-249}$ reported an elution volume of 15.1ml on a SD200 gelfiltration column (**figure 4.12**). Based on the calibration run on this gelfiltration column, this elution volume corresponds to the molecular weight of Poz1$^{30-249}$ as a monomer.
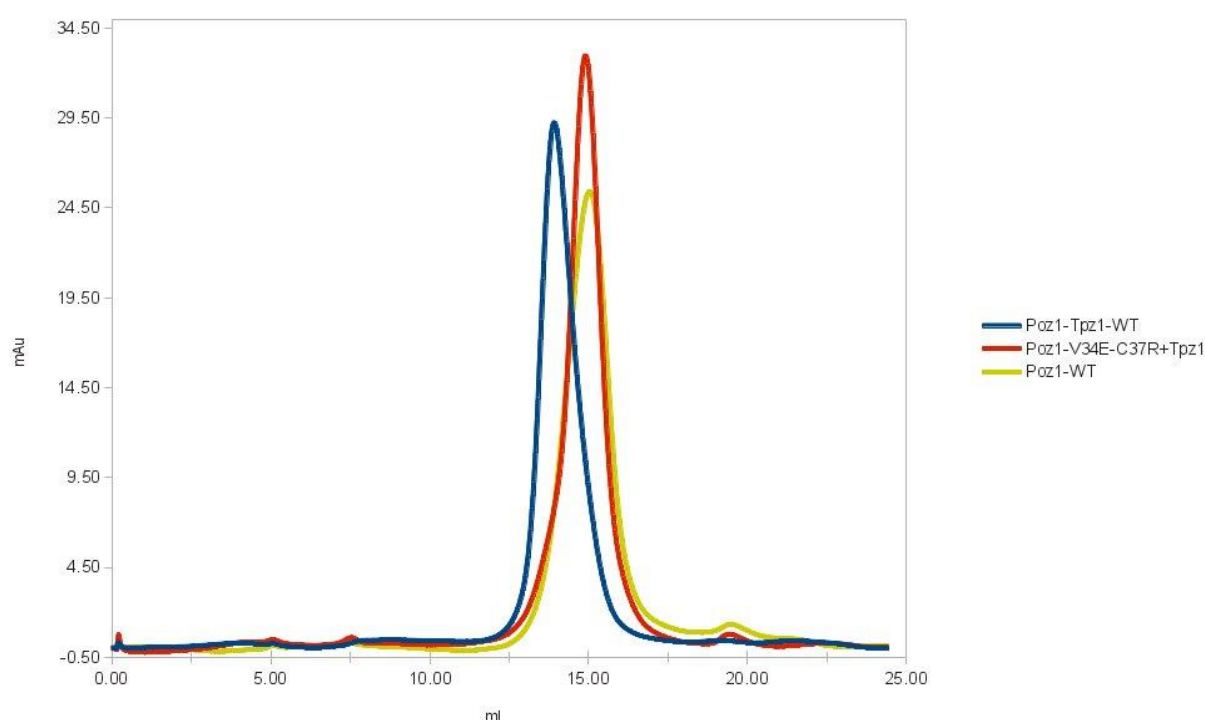
Figure 4.12. Size-exclusion chromatography for proving the heterotetramerization hypothesis.

The chromatograms of the three performed chromatography runs were overlaid to visualize the shifts in elution volume. Monomeric Poz1$^{30-249}$ *wildtype* (WT) elutes at similar volume as Poz1$^{30-249}$_V34E-C37E+Tpz1$^{475-508}$, while *wildtype* Poz1$^{30-249}$+Tpz1$^{475-508}$ elutes at significantly higher molecular weight, indicating dimer formation.

Subsequent injection of Poz1$^{30-249}$+Tpz1$^{475-508}$ returned an elution volume of 14ml, which was a significant shift towards higher molecular weight and higher than for a Poz1-Tpz1 monomer. Based on the elution volume, the estimated molecular mass corresponded to a Poz1-Tpz1 heterotetramer. In order to confidently assign the shift to heterotetramerization and not solely to the addition of Tpz1$^{475-508}$, a third SEC run with the heterotetramerization

mutant was performed. In gelfiltration, the heterotetramerization mutant $Poz1^{30-249}\_V34E\_C37R+Tpz1^{475-508}$ reported an elution volume of 14.9ml which was minimally lower than the elution volume of $Poz1^{30-249}$. This agreed with monomeric Poz1 now bound to a 34 amino acid Tpz1-fragment. However, the peak for the mutant was shifted to significantly higher elution volume compared to the *wildtype* Poz1-Tpz1 complex, implying that the *wildtype* complex is indeed a heterotetramer and that the mutant is not due to disrupted heterotetramerization.

These results demonstrate that in the absence of $Tpz1^{475-508}$, $Poz1^{30-249}$ is monomeric and upon binding to Tpz1 the complex becomes heterotetrameric. Subsequently, when introducing mutations into the heterotetramerization interface, the complex becomes heterodimeric (meaning one Poz1 bound to one Tpz1). These results confirm the heterotetramerization hypothesis in solution. To address the question whether heterotetramerization of Poz1-Tpz1 is relevant *in vivo*, we designed a number of mutant fission yeast strains and monitored telomere length regulation (see **section 4.4.3**).

4.4.2.3 SEC-MALS with Poz1-Tpz1 heterotetramerization mutant

In order to reinforce the results obtained by SEC, we performed size-exclusion chromatography coupled to multi-angle-light-scattering (SEC-MALS) to determine more precise masses for the complexes. Running $Poz1^{30-249}$ by itself on SEC-MALS gave an approximate molecular weight of 26 kDa which was in good agreement with the calculated theoretical mass of 26 kDa, further confirming its monomeric state (**figure 4.13**). In a second run, $Poz1^{30-249}+Tpz1^{475-508}$ was injected and returned a mass of 47 kDa which falls short of the calculated heterotetrameric mass of 60 kDa, but at the same time was a higher mass than for monomeric Poz1-Tpz1 (30 kDa). A possible explanation is that $Poz1^{30-249}+Tpz1^{475-508}$ occurs in equilibrium of the heterotetrameric and the monomeric form. The peaks for the monomer at 30 kDa and the heterotetramer at 60 kDa are overlapping, making it difficult to distinguish them and leading to an average peak at 47 kDa.

A

B
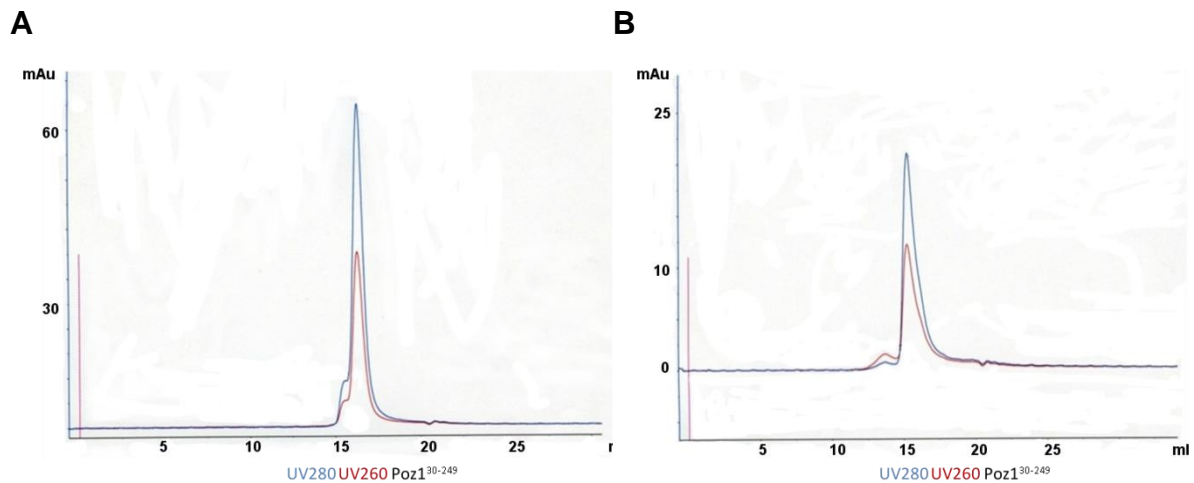
Figure 4.13. Chromatograms of the SEC-MALS experiment.

For **A)** Poz1$^{30-249}$ a molecular mass of 26kDa was reported and for **B)** Poz1$^{30-249}$_V34E_C37R+Tpz1$^{475-508}$ a molecular mass of 47kDa.

We suggest that due to the 29 amino acid N-terminal truncation in Poz1 we may have destabilized the heterotetramerization interface compared to the fulllength protein. After introduction of the heterotetramerization disrupting mutations, the heterotetrameric form is then finally completely eliminated. This goes in line with the *in vivo* experiments in fission yeast where the *Poz1Δ1-29* strain showed slight telomere deregulation, possibly due to partial, but not complete disruption of heterotetramerization (see **section 4.4.3**). We hypothesize that for fulllength Poz1 we would observe 100% of the heterotetrameric form in SEC-MALS, an experiment that has yet to be performed but has proven difficult because of the poor behavior of fulllength Poz1 in protein purification (**see section 3.1**).

Ultimately, the combination of SEC and SEC-MALS has revealed that in solution Poz1 by itself occurs monomeric and that binding of Tpz1 induces heterotetramerization of the complex.

4.4.3 Poz1-Tpz1 heterotetramerization *in vivo*

To determine whether heterotetramerization of Poz1-Tpz1 exists *in vivo*, the length of the telomeric repeats of the heterotetramerization disruption mutant was compared to that of *wildtype* fission yeast cells. As previously described by Miyoshi *et al.* (2008), *Poz1Δ* strains show strongly extended telomeric length compared to wild type cells. When introducing the charge repulsion mutations that disrupted heterotetramerization *in vitro* into the fission yeast genome, we observed severe telomere elongation, similar to what is observed in *Poz1Δ* cells (**figure 4.14**).
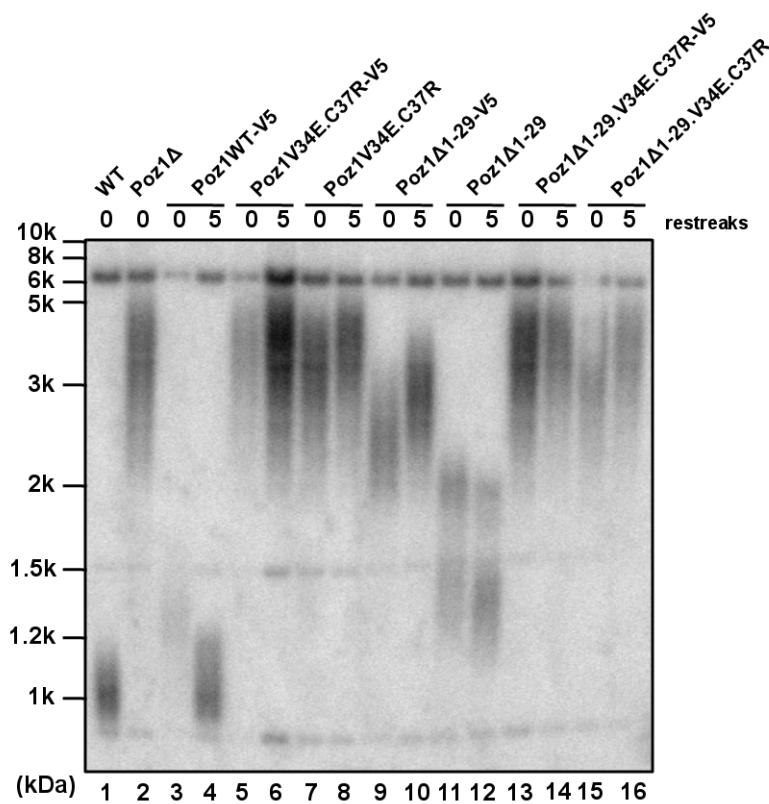
Figure 4.14. Telomere length assay for Poz1-Tpz1 heterotetramerization mutants.
Telomeres were analyzed at two time points: After 0 and after 5 restreaks in order to allow the telomere length to equilibrate. The *Poz1Δ* strain shows long telomeres (lane 2) compared to the *wildtype* strain. The Poz1 heterotetramerization mutant shows similarly long telomeres (lanes 7-8). Poz1 lacking the 29 N-terminal amino acids shows mildly elongated telomeres (lanes 11-12). After introduction of the heterotetramerization disrupting mutations, *Poz1Δ1-29* also shows severe telomere elongation (lanes 15-16). V5-tagged versions were designed for protein expression analysis by western blot. The V5-tag seems to have no effect on telomere length except in the case of the *Poz1Δ1-29* strain: Telomeres are elongated more strongly, almost as severely as in the *Poz1Δ* strain. This could be due to further destabilization of heterotetramerization because of the close proximity of the tag to the heterotetramerization interface.

This indicated that indeed Poz1-Tpz1 exists as heterotetramers *in vivo* and that upon disruption of heterotetramerization, shelterin function is disturbed. The expression of the mutant proteins was verified by western blot (**figure 4.15**).
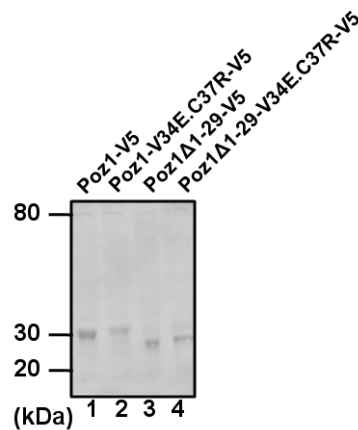
Figure 4.15. Western-blot for the Poz1-Tpz1 heterotetramerization mutants.

V5-tagged Poz1 is visible on the western blot. All bands are found at the expected sizes: (1) and (2) 31 kDa, (3) and (4) 27 kDa. The expression levels are not affected by the mutations (1, 4) and the N-terminal Poz1 truncation (3, 4) compared to the *wildtype* protein (1).

In order to assess whether the N-terminally truncated Poz1$^{30-249}$ can retain its function, we determined the telomeric length and found slightly elongated telomeres compared to *wildtype* cells (**figure 4.14**). When comparing the heterotetramerization interfaces of the TRF proteins and Poz1-Tpz1, we found the truncated N-terminal region of Poz1 (residues 1-29), in TRF corresponds to a helix that makes major contacts at the heterotetramerization interface. In Poz1, that sequence region has also been predicted to form a helix and given the highly conserved structural architecture of Poz1 and TRF1 and TRF2, it is highly probable that this region is also involved at the heterotetramerization interface of Poz1-Tpz1. Due to the missing helix in *Poz1Δ1-29*, there possibly is a slight destabilization of the heterotetramerization compared to the *wildtype* and this could explain why we observe slightly longer telomeres. Once the heterotetramerization disrupting mutations are introduced in *Poz1Δ1-29*, yet again we observed very long telomeres as in *Poz1Δ*. From this we conclude that even with its missing N-terminal region, Poz1Δ1-29 allows heterotetramerization *in vivo*. However, possibly a less tight complex is formed leading to the impaired, but not the fully deficient telomere length regulation observed when the heterotetramerization disrupting mutations are introduced.

Concluding, we discovered that Poz1-Tpz1 forms a heterotetrameric assembly *in vitro* as well as *in vivo*. Furthermore, these results show that multimerization of Poz1-Tpz1 is an essential function of this complex, as disruption of the protein-protein interactions leads to severe telomere regulation defects.

4.5 Zinc binding by Poz1-Tpz1

4.5.1 Modeling the zinc ion in the Poz1-Tpz1 structure

During the refinement of the Poz1-Tpz1 structure, we repeatedly encountered a large patch of positive electron density in the Fo-Fc difference map at a contour level of 1.5σ in an area at the interface between Poz1 and Tpz1 (**figure 4.16**). Despite testing many different conformations of the protein backbone and side chains, we could not satisfy the positive electron density.
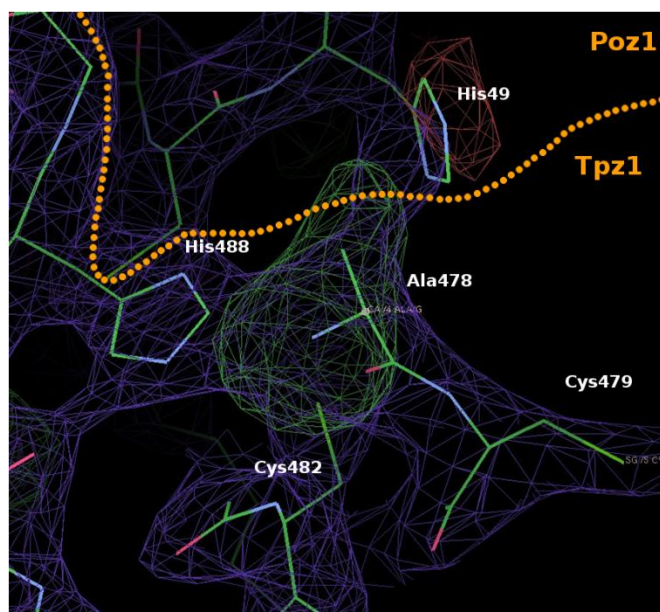
Figure 4.16. View of the electron density at the interface of Poz1 and Tpz1.

The 2Fo-Fc electron density is shown as a purple mesh. Red mesh indicates negative electron density in the Fo-Fc difference map, while green mesh indicates positive electron density. (Figure was generated in COOT, Emsley *et al.*, 2010. Contour level is at 1.5σ).

After closer inspection of the amino acid composition surrounding the positive electron density, the close proximity of two histidine and two cysteine residues became apparent. These residues are found in many proteins coordinating metal ions. Following this, we attempted to model a zinc ion into the positive electron density. After several steps of model building and refinement, the zinc atom could be correctly placed and coordinated by two histidine residues and two cysteine residues with ionic bond lengths that agree well with the ionic bonds in zinc binding sites of other proteins (**figure 4.17**).

A

B

C
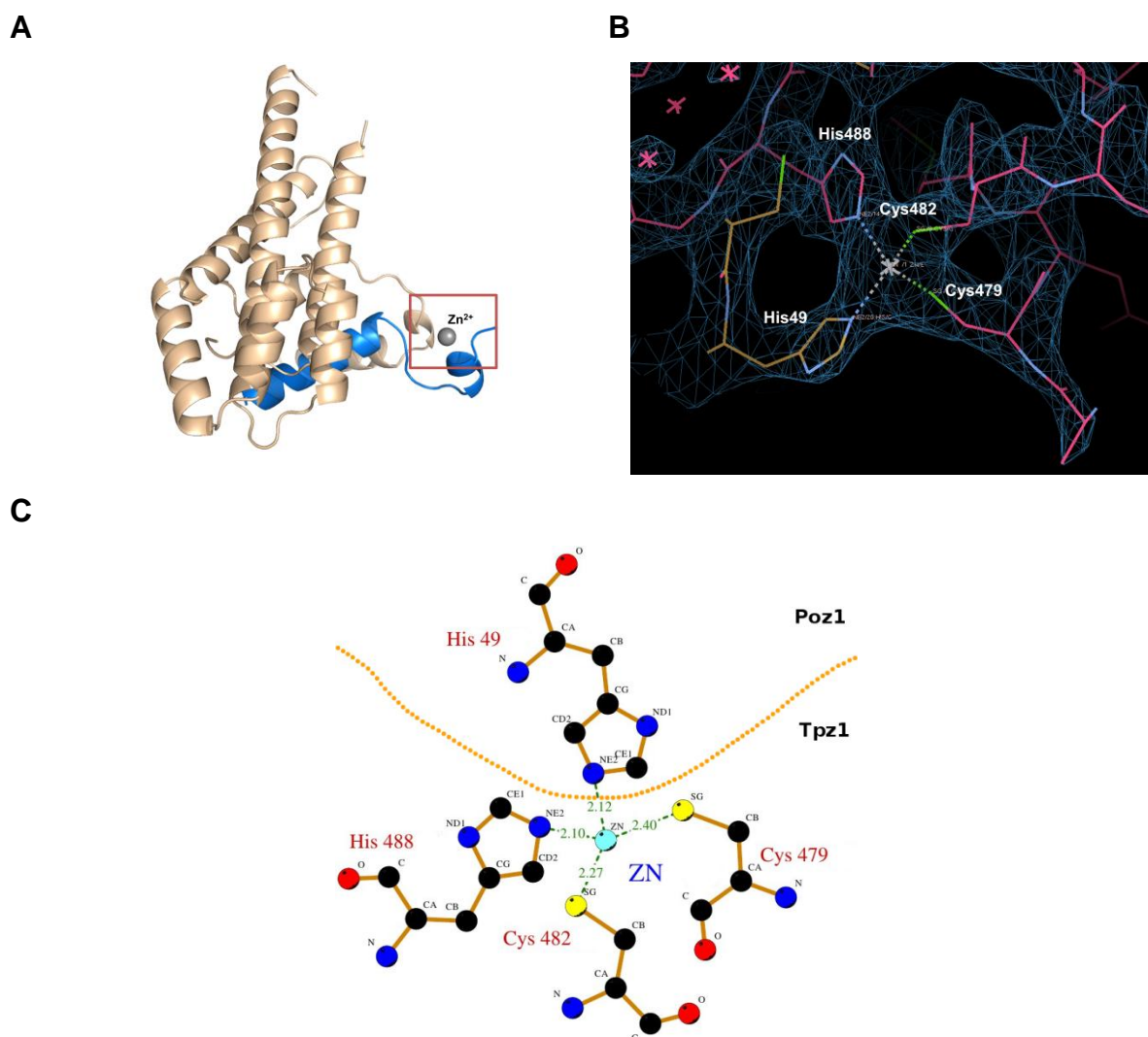
Figure 4.17. Zn^{2+}-ion coordinated at the interface of Poz1 and Tpz1.

A) Position of the zinc ion in the overall structure. Poz1 is colored in wheat and Tpz1 in blue. **B)** Close-up view of the refined electron density at the zinc-coordination site. **C)** Schematic representation of the Zn^{2+}-ion coordination (Figure generated by PDBsum, Laskowski, 2001).

In order to verify that a zinc ion rather than any other metal ion is bound, we carried out a fluorescence scan (EXAFs) across the absorption edge of zinc at the synchrotron beamline (**figure 4.18**). Based on the obtained absorption edge plot for zinc, we collected diffraction data at the peak wavelength (**table 4.2**). The anomalous scattering contribution was subsequently determined after processing with XDS. We observed a significant anomalous signal (**table 4.3**) which together with the expected absorption edge plot was evidence that a zinc ion indeed is bound to the complex.

Zn-site data set	
	Data set 2
X-ray source	X10SA
Detector	Pilatus 6M
Wavelength	1.2825Å
Distance	400.0mm
Exposure time	0.25s
Oscillation	0.25°
Number of images	1600

Table 4.2. Strategy for the collection of the Zn-site data set
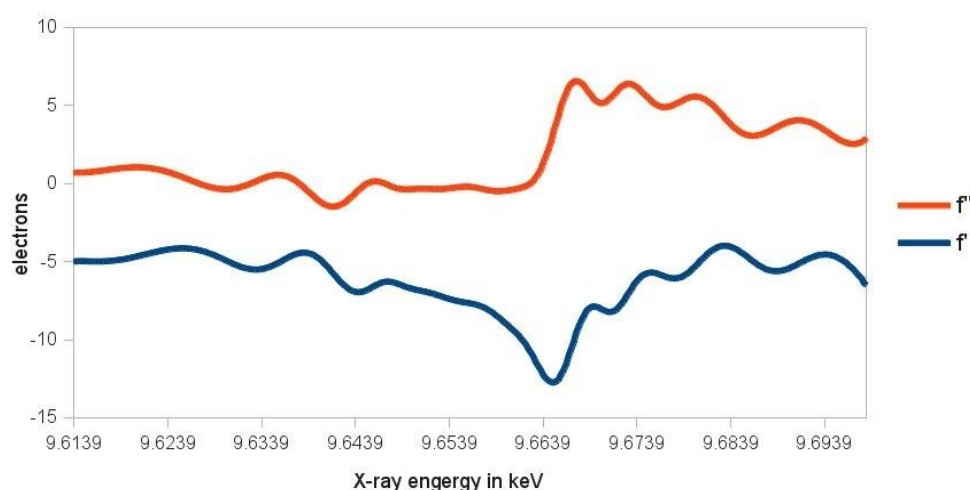
Figure 4.18. EXAFs scan over the absorption edge of zinc performed on a Poz1[30-249]+Tpz1[475-508] crystal.

f' is the real and f'' is the imaginary component of the anomalous signal of zinc. f' and f'' are plotted versus the X-ray energy.

Resolution [Å]	S_norm/S_ano (peak)
8.18	2.468
5.84	3.248
4.78	2.235
4.15	1.536
3.71	1.11

Table 4.3. Estimation of anomalous signal with XDS.

Interestingly, the zinc ion is bound at the interface between Poz1 and Tpz1. The metal ion is coordinated by residue His49 provided by Poz1 and residues Cys479, Cys482 and His488 provided by Tpz1 (**figure 4.17C**).

Taken together, we have shown that Poz1-Tpz1 binds a zinc ion immediately at the protein-protein interaction interface. However, the function of the coordinating Zn^{2+} ion remains unclear. It may play a structural role for complex integrity and assembly or it may have a regulatory role.

4.5.2 Functional role of the bound zinc ion

There is no zinc present in the crystallization conditions. It therefore most likely originates from the expression host. To address whether zinc binding is essential for protein expression, folding and complex formation, we designed a mutant of the $Poz1^{30-249}+Tpz1^{475-508}$ complex where metal ion coordination is disrupted. We mutated residues Cys479, Cys482 and His488 of Tpz1 to alanine residues, which would confidently disrupt the binding of zinc. Subsequently, we expressed the mutant variant of the Poz1-Tpz1 complex in *E. coli*, and pulled down the protein by Ni^{2+}- and GST-affinity and assessed the protein on SDS-PAGE (**figure 4.19**).

A

B
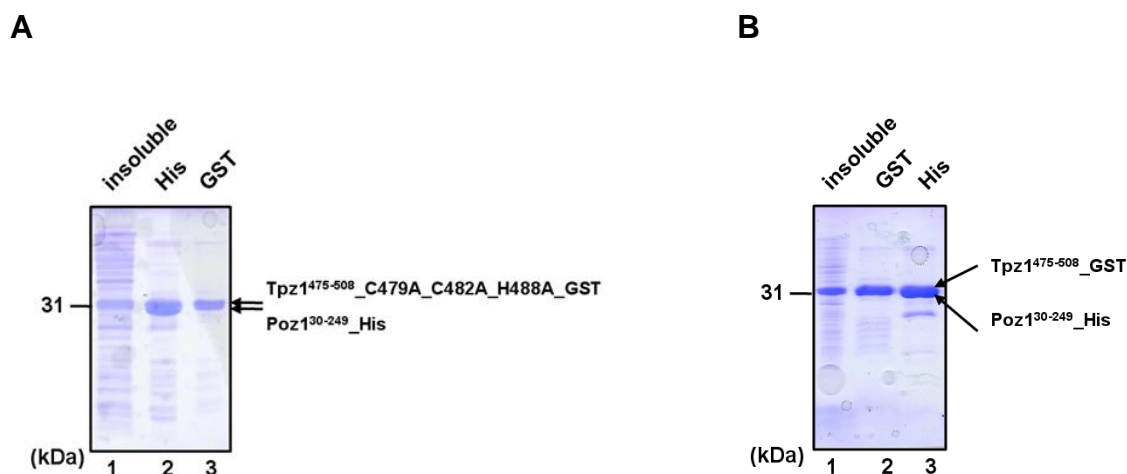
Figure 4.19. *E. coli* **expression test and affinity pulldown of the zinc-site mutant.**

Poz1 is N-terminally His-tagged while Tpz1 is GST-tagged. On SDS-PAGE, **A)** the mutant complex (2, 3) shows no difference compared to the **B)** *wildtype* complex (2, 3).

The Tpz1 zinc site mutant formed a complex with Poz1 that behaved identically as the *wildtype* proteins. Therefore we conclude that the binding of zinc is not required for expression and complex formation in solution. The zinc site is located at close proximity to

the Poz1-Tpz1 heterotetramerization interface. Possibly, failure to bind zinc interferes with the formation of the heterotetramer. Conformational changes could occur in the secondary structure elements involved in the interface no longer allowing heterotetramerization.

In order to test this hypothesis in gelfiltration, I expressed and purified the zinc site mutant Poz1-Tpz1 complex. While the expression level was good and I could readily pull down the complex in Ni^{2+}-affinity chromatography, the protein subsequently precipitated, leaving no material for size-exclusion chromatography.

Given the above findings, I propose that the failure to bind zinc interferes with the structural integrity of the complex, thereby reducing its solubility. Furthermore, heterotetramerization could be impaired, but this remains to be validated. These findings do not exclude that the binding of zinc has additional roles, for example as part of a regulation mechanism. To address the question whether the zinc site has any relevance in vivo, we determined the effects of zinc site disruption on telomere length regulation in fission yeast cells.

4.5.3 Zinc atom coordination *in vivo*

To investigate the relevance of zinc coordination by Poz1-Tpz1 *in vivo*, the zinc site disrupting mutations, C482A and H488A, were introduced into a fission yeast strain and the telomeric length compared to the wild type strain was evaluated (**figure 4.20**) (Due to technical reasons a double mutant was chosen as opposed to the triple mutant described in **4.5.2**). The mutant strain showed elongated telomeres, similar to the telomere length of the *Poz1Δ* strain. This result showed that the coordination of the zinc ion is essential *in vivo*. However, whether the 2-fold reduced expression of Tpz1_C482A_H488A observed by western blot has an effect on telomere length needs to be further investigated (**figure 4.21**). Firstly the experiment needs to be repeated to determine whether the difference is real, secondly mutations of Cys479, Cys482 and His488 to different residues (for example Arg) could be tried that disrupt zinc binding but do not change the expression level.
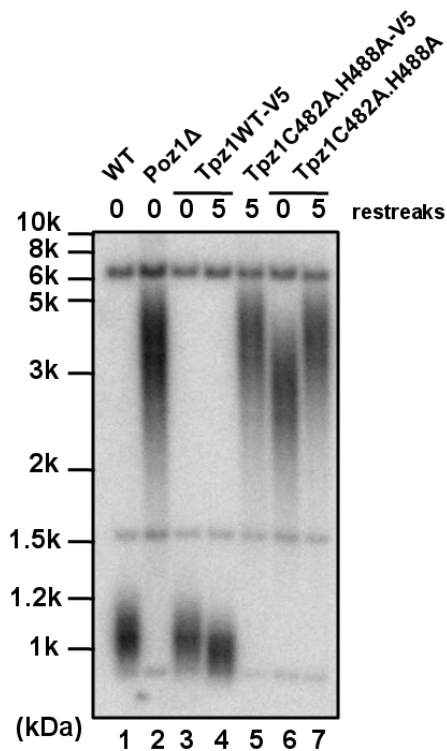
Figure 4.20. Telomere length assay for Tpz1 zinc-site mutants.

Telomeres were analyzed at two time points: After 0 and after 5 restreaks. Lane (1) shows telomere length for the *wildtype* strain, lane (2) shows telomere length for the *Poz1Δ* strain. The strain expressing Tpz1 with the disrupted zinc site shows telomere elongation (6-7). Strains with V5-tag were designed for subsequent expression level analysis by western blot.

In summary, as in *in vitro*, failure to bind zinc impairs the integrity of the Poz1-Tpz1 complex *in vivo*. Given the observed telomere elongation, it is possible that conformational changes in the zinc site mutant lead to impaired multimerization, which in turn leads to the same phenotype we saw for the heterotetramerization mutant.
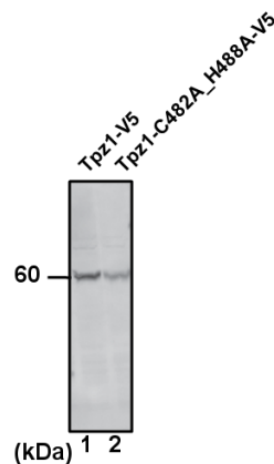
Figure 4.21. Western-blot for the Tpz1 zinc site mutant.

V5-tagged Tpz1 is visible on the western blot. The bands are found at the expected sizes: 58 kDa for (1) and (2). The expression levels of the mutant protein (2) are 2-fold reduced compare to *wildtype* (1).

4.6. The function of Poz1 as a molecular bridge

Poz1 bridges the double-strand and the single-strand binding halves of the shelterin complex. Based on the data presented in this work, it does so by linking a Rap1 dimer to two Tpz1 molecules in a dimeric manner. We were wondering whether Poz1 has any other functions besides acting as a "molecular-bridge" between Rap1-Taz1 and Tpz1-Ccq1-Pot1.

To examine this, we set out to construct a mutant fission yeast strain where the Poz1 protein is replaced by the structurally similar hTRF2H domain as well as its binding partner hApollo. In a first step, the *poz1* gene was deleted. In a subsequent step, TRF2$^{40-245}$ was fused to the C-terminus of Rap1 with a 6x glycine linker in between allowing a certain degree of flexibility so that it can find and form the interaction with Apollo. Finally, a minimal peptide of Apollo (495-530) that is known to interact with TRF2H (Chen *et al.*, 2008), was fused to the C-terminus of Tpz1. Again a 6x glycine linker was introduced in between Tpz1 and Apollo allowing the short binding domain of Apollo to protrude from the Tpz1 C-terminus und reach its binding pocket in TRF2H. In theory, Rap1-TRF2H should interact with Apollo-Tpz1 thereby reconstituting the bridge between the double- and single-strand binding complexes in a similar way in which Poz1 forms a bridge in *wildtype* cells (**figure 4.22**). Importantly, TRF2H-Apollo should merely restore the link between the subcomplexes, but they also represent a structurally highly similar substitute that should ensure the integrity of the dimeric arrangement throughout the shelterin complex.
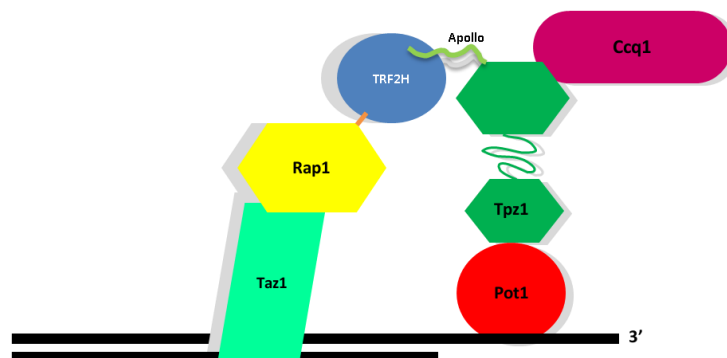
Figure 4.22. Schematic view of the shelterin complex of the generated spRap1-hTRF2H-hApollo-spTpz1 fission yeast strain.
The TRF2H-domain is fused to the C-terminus of Rap1 and the Apollo peptide is fused to the C-terminus of Tpz1.

The telomeric length of the resulting chimeric TRF2H-Apollo strain was determined and compared to that in *wildtype* cells (**figure 4.23**).
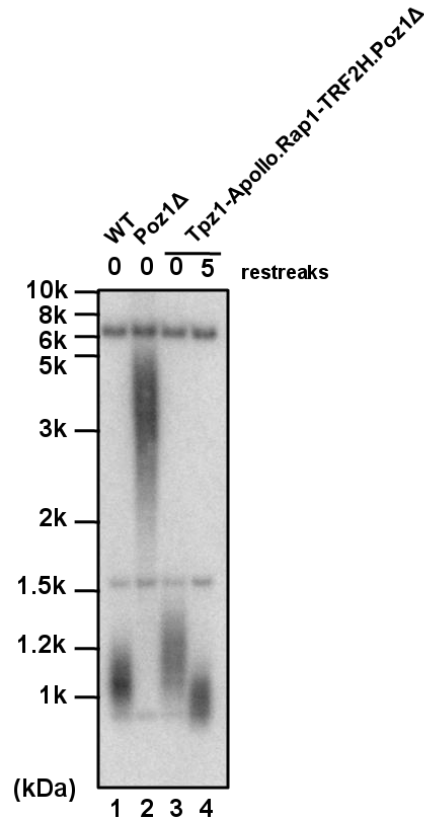
Figure 4.23. Telomere length assay for the Rap1-TRF2H-Tpz1-Apollo fusion strain.

Telomeres were analyzed at two time points: After 0 and after 5 restreaks. The strain containing the engineered TRF2-Apollo bridge shows slightly shorter telomeres (lane 4) than the *wildtype* strain after 5 restreaks (lane 1). Although the poz1 gene was deleted, No telomere elongation was observed in the *Rap1-TRF2H-Tpz1-Apollo* fusion strain (4) in contrast to the *Poz1Δ* strain (lane 2). Expression was verified by western blot.

At first, after 0 restreaks, the fusion strain showed slightly longer telomeres than *wildtype* cells which is likely due to the way the strain was generated, meaning the deletion of *poz1* in the first step lead to telomere elongation, and only later the fusions were inserted in the next two steps. After 5 restreaks we observed mild telomere shortening. Remarkably, the experiment reveals that the engineered TRF2H-Apollo-bridge can rescue the "very long telomeres" phenotype of *Poz1Δ* strains. Given these findings, the primary role of Poz1 must be the linkage of Rap1 and Tpz1. Furthermore, the experiment is so efficient, because the Poz1-bridge is of a dimeric nature and we replaced this architecture with a structurally very similar assembly that maintains the tetrameric arrangement. These results further support the hypothesis of Poz1-Tpz1 heterotetramer *in vivo*.

Moreover, it was shown that the telomeres of the TRF2H-Apollo fusion strain experience slow telomere shortening (after 5 restreaks), an observation that I attribute to the fact that the engineered bridge is a fixed construction that cannot be opened or closed. The association/dissociation between Poz1 and Rap1/Tpz1, however, must be regulated. The inability to support this, results in telomere dysfunction. To verify that the fusion construct was indeed produced in the cells, we verified the expression by western blot (**figure 4.24**). Tpz1-Apollo is expressed at a lower level than Rap1-TRF2H, but according to the results of the telomere length assay, sufficient protein is available to form the molecular bridge.
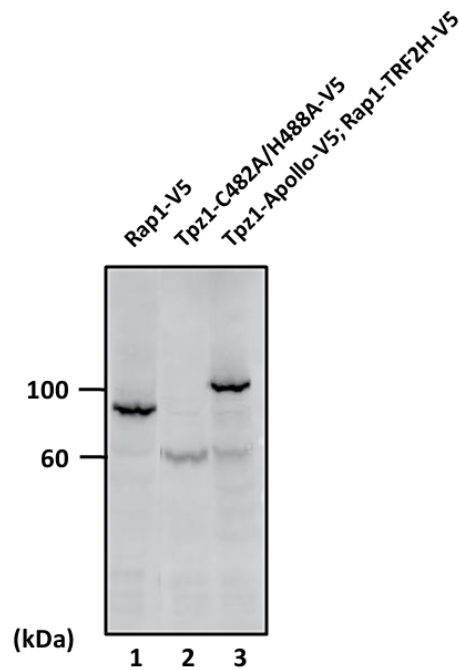
Figure 4.24. Western-blot for the TRF2H-Apollo fusion strain.

In lane (1) and (2) there is Rap1 (79 kDa) and Tpz1 (58 kDa) respectively as a control. V5-tagged Tpz1-Apollo (64 kDa) and V5-tagged Rap1-TRF2H (105 kDa) are in lane (3).

Chapter 5

Discussion and outlook

5. Discussion and outlook

Telomeres, the DNA-protein structures at the ends of linear chromosomal DNA have been shown to be of utmost importance for maintaining genomic integrity. As a consequence of the vast regulation mechanisms at the DNA ends, degradation and chromosome fusions are avoided. In recent years emerging evidence has revealed, that maintaining and protecting the telomeres is far more complex than the mere addition of telomeric repeats by telomerase. A myriad of telomere associated factors has been identified with likely more still waiting to be discovered. Given the fundamental importance of passing on intact genetic information upon cell division, our understanding of the underlying molecular mechanisms is required for the development of preventative measures for interrelated diseases. Importantly, the advances in the field of telomere biology research have uncovered the crucial role of telomere maintenance in development of cancer.

Here we focused on the telomere-protection complex known as shelterin, which is a large multi-subunit protein complex present in many organisms ranging from yeast to humans. Its components have been shown to promote essential functions in telomere length control. By gaining insight into the molecular structures of some of these components, we have identified indispensible features that ensure the proper function of shelterin complexes in telomere maintenance.

While work on the human shelterin complex presented itself difficult, ultimately investigations on the architecturally similar fission yeast complex led to successful structure solution of subcomplexes. The ease of *in vivo* genetic manipulation in yeast allowed us thereafter to follow up on hypothesizes based on the complex structures. An isolated shelterin subcomplex composed of spPoz1-Tpz1 with boundaries obtained by limited proteolysis was successfully used in structure determination by X-ray crystallography.

The Poz1-Tpz1 shelterin subcomplex

Preceding this work little was known about the structure and function of the fission yeast shelterin component Poz1 (Miyoshi *et al.*, 2008). Based on the protein sequence, no

motifs or sequence similarity to any other proteins could be identified. Through the results presented here, we have uncovered several new features of the Poz1-Tpz1 shelterin subcomplex. The structure allowed us to improve our knowledge of the overall shelterin architecture and revealed some striking similarities to the human shelterin proteins.

Similarity of Poz1-Tpz1 to the human shelterin components TRF1 and TRF2

Despite the absence of any sequence similarity to the human shelterin components TRF1 and TRF2, the crystal structure of Poz1$^{30\text{-}249}$-Tpz1$^{475\text{-}508}$ revealed a remarkable structural similarity to the TRF-homology domain (TRFH). The TRFH-domain is a conserved dimerization module found in both TRF1 and TRF2. Interestingly, Poz1-Tpz1 is not the functional homologues to TRF1 or TRF2 but, according to widely accepted models, are the homologues of hsTIN-TPP1 (**figure 5.1**).

Fission yeast
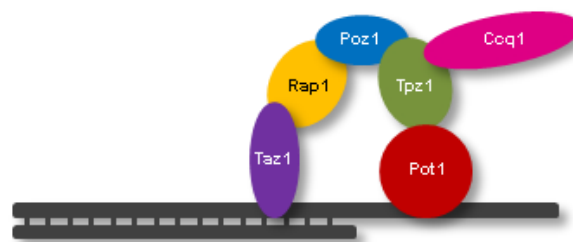
Human
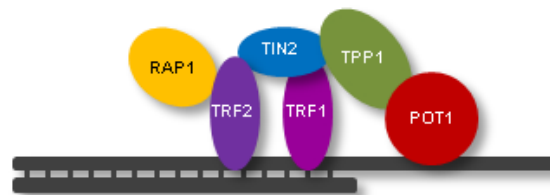
Figure 5.1. Comparison of the human and the fission yeast shelterin complexes.

Nevertheless, virtually every secondary structure element found in Poz1, aligns with those found in the human TRF proteins, albeit with some differences in length. If Poz1-Tpz1 is structurally similar to TRF1 and TRF2, it might also be similar in structure to the presumed homologues TIN2-TPP1 whose structure has not been determined. Perhaps the TRFH-domain is a conserved structural element found throughout the shelterin complexes among many different species.

A conserved binding pocket for Poz1-Tpz1, TRF1-TIN2 and TRF2-Apollo

Alignment of our Poz1-Tpz1 structure with those of TRF1-TIN2 and TRF2-Apollo revealed that the binding pocket for Tpz1 in Poz1 superimposes with little deviation onto the binding pockets of TIN2 in TRF1 and Apollo in TRF2 (**section 4.3**). Back in 2008, Chen *et al.* showed that TRF1 and TRF2 bind their interaction partners, TIN2 and Apollo, respectively, using a shared docking motif Y/F-X-L-X-P. This motif was also present in another TRF1 interacting protein called hPinX1. Given these findings it was proposed that the TRFH domains of TRF1 and TRF2 function as telomeric protein docking sites that recruit different shelterin-associated factors with distinct functions using a common recognition motif. The striking similarity of the Poz1-Tpz1 subcomplex to the TRF1-TIN2 and TRF2-Apollo structures, raises the question whether a similar binding mode exists in Poz1-Tpz1. In the search for a common binding motif, we compared residues involved in the binding pocket in the human complex with those in the fission yeast complex. Given that there is no sequence similarity between the TRF proteins and Poz1-Tpz1, it is not surprising that we did not observe identical residues in the binding site. By structural superposition of the human and fission yeast binding pockets we attempted to derive a hypothetical binding motif in Tpz1. This consisted of the sequence (F-X)-L-X-X-X-W-X-K. A motif search in the fission yeast genome for F-X-L-X-X-X-W-X-K returned two proteins: the first being midasin, a giant AAA family ATPase whose homologue Rea1 in *S. cerevisiae* drives removal of biogenesis factors during multiple stages of 60S ribosome assembly (Bassler et al 2010). The second hit was an RNA-binding protein designated spCwf5 which has been proposed to act as a part of a pre-spliceosome complex (Ohi *et al.*, 2002). Further work is required to determine whether these proteins are indeed involved with the shelterin complex. When searching for the less restrictive motif L-X-X-X-W-X-K, many more hits were found. The protein of interest among these was the protein spTaf1 which is known to interact with Taz1 (Ueno *et al*, 2001). This finding is particularly interesting because Taz1 has also been predicted to contain a TRFH-like domain (Li *et al.*, 2000). We propose that Taz1 also binds proteins in its TRFH-domain and recognizes them via a specific motif. Given the conservation of the TRFH-domain it is possible that the recognition motif in proteins binding to spTaz1 is similar to that in proteins that bind to Poz1. Since Taf1 binds Taz1, we do not expect it to bind Poz1 as well. Rather we suggest that Taf1 turned up in the motif search because of a similar but not identical recognition motif. Possibly, we have identified a part of a motif that is similar among Taz1- and Poz1-binding proteins, and the remaining sequence

specific for Poz1 remains unknown. Further investigations will help to define the motif more specifically. Moreover, examination of the proteins retrieved in the motif search is required to determine their possible involvement at the shelterin complex. The next step would be to determine whether Taf1 indeed binds to the Taz1 TRFH domain. Also, using the Poz1 TRFH domain as bait, we could try to fish for further hypothetical binding partners in yeast lysate and compare the results with the proteins retrieved in the motif search.

Poz1-Tpz1 heterotetramerization

The TRFH domain functions as a dimerization module in TRF1 and TRF2 (Spink *et al.*, 2000). Consequently, TRF1 and TIN2 as well as TRF2 and Apollo form heterotetramers. Given the high degree of structural similarity of Poz1-Tpz1 to the TRFH domain, we speculated that Poz1-Tpz1 might also be involved in heterotetramerization. In a first step we identified a hypothetical heterotetramer in the crystal lattice. Subsequent mutational analysis of the hypothetical heterotetramerization interface by analytical size exclusion chromatography and light scattering, revealed that indeed Poz1-Tpz1 forms heterotetramers in solution. However, we found that Poz1 by itself is monomeric and only upon binding to Tpz1 a heterotetrameric complex forms.

To address the question whether heterotetramerization is merely an artifact in solution, or whether Poz1-Tpz1 exists as heterotetramers *in vivo*, we examined telomere length regulation in a heterotetramerization mutant in fission yeast cells. Cells expressing the Poz1-Tpz1 heterotetramerization mutant, showed dramatically elongated telomeres compared to *wildtype* cells, demonstrating that heterotetramerization is essential for proper telomere regulation. Strains containing the N-terminally truncated Poz1$^{30-249}$ used for crystallization showed mildly longer telomeres compared to *wildtype* cells. It seems that here, the heterotetrameric complex loses some stability due to the missing N-terminal helix that, based on secondary structure predictions, is located at the heterotetramerization interface. This could explain the possible monomer-dimer equilibrium observed in the light scattering experiment. A less stable heterotetramer could lead to the mild telomere elongation phenotype that did not show complete telomere length deregulation.

If Poz1-Tpz1 is found as a heterotetramer in the shelterin complex, how does disrupting the heterotetramerization interface disturb the shelterin regulation system? It is well known that Taz1 binds the double-strand telomeric repeats as dimers, and that two spRap1 molecules bind the Taz1 dimer. In a recent study by Nandakumar and Cech (2012), it was reported that depending on the single-stranded telomeric substrate, Pot1 can dimerize upon DNA binding. It has been suggested that Pot1 may recruit two Tpz1 molecules and these in turn could bind two Ccq1 molecules assisting the recruitment of dimeric telomerase (Assuming that fission yeast telomerase is dimeric, as it has been shown for telomerase of many other organisms from ciliates to humans, **section 1.2**). We find that the heterotetramerization characteristics of Poz1-Tpz1 fits well with an overall dimeric shelterin complex: Our updated model suggests that a Taz1 dimer binds to the telomeric repeat and is bound by two Rap1 molecules. Then monomeric Poz1 binds to Tpz1 which induces heterotetramerization and Poz1 subsequently acts as a dimeric adaptor that can bind the two Rap1 molecules. Then two Tpz1 molecules interact with two Ccq1 molecules and two Pot1 molecules. Depending on the single-stranded telomeric substrate, only one of the Pot1 molecules can bind to the DNA or, given the right substrate, two Pot1 molecules, now in close proximity, could dimerize on the DNA upon binding. For Tpz1 the ability to self dimerize has not been shown. It is possible that the two Tpz1 molecules that are bound to the Poz1-dimer function as flexible arms. This would allow the Pot1 molecules to bind different telomeric repeats in the case where Pot1 does not dimerize (**figure 5.2C**).
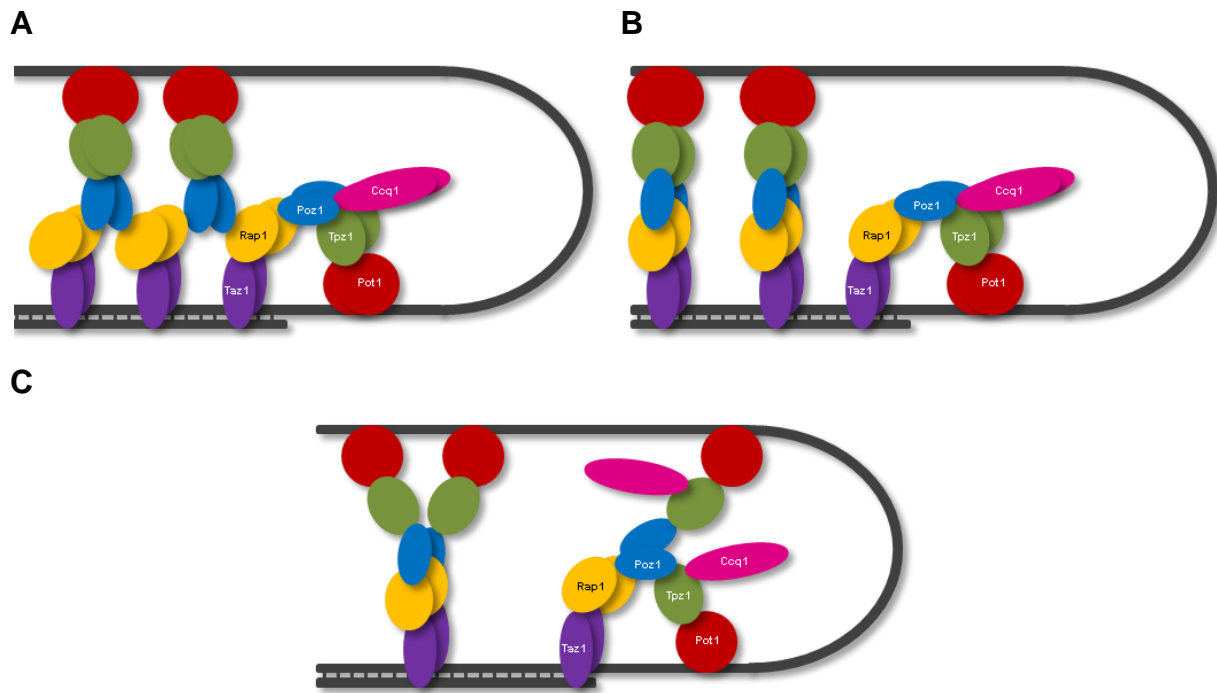
Figure 5.2. View of how the shelterin complexes could be arranged at the telomeric end.

A) Poz1-dimers links arrays of Taz1-Rap1 complexes. **B)** Poz1-Tpz1-heterotetramers bind Taz1-Rap1-heterotetramers and each shelterin complex stays separate. **C)** Two Tpz1 molecules are bound but are not dimeric and have the Pot1 molecules sitting on flexible arms allowing them to bind different telomeric repeats.

An alternative explanation could be that the Poz1-dimer acts as a bridge linking two Taz1-Rap1-heterotetramers in a similar way that TIN2 links the two human double-strand binding complexes, TRF1 and TRF2 (**figure 5.2A**). This would lead to arrays of Taz1-Rap1 molecules: A Rap1 molecule would be connected to a Rap1 molecule of a neighboring Taz1-Rap-heterotetramer. The formation of arrays is a possible explanation for the tendency of these proteins to aggregate during protein purification. Tpz1-Pot1 molecules could then be docked on top of the array and could bind the folded back 3' telomeric overhang to form the T-loop structure. Should it not be the case that Poz1-dimers link arrays of Taz1-Rap1 complexes-a Poz1-dimer might bind one Rap1-dimer and this would still allow binding of the folded back telomeric overhang (**figure 5.2B**).

Currently all of the scenarios described above are possible and further experiments should be performed in order to clarify which represents reality. Analysis of different combinations of shelterin components by electron microscopy could shed light on some higher order arrangements.

An essential zinc-binding site at the interface of Poz1 and Tpz1

Furthermore, our structure revealed a unique feature of Poz1-Tpz1 that is not present in the structurally similar TRF1 and TRF2, which is the binding site for a zinc ion. The metal ion is coordinated by three residues provided by Tpz1 and one residue by Poz1. An initial hypothesis, that the binding of zinc might be essential for Poz1-Tpz1 complex formation, was disproved by an affinity pulldown experiment with a mutant containing a disrupted zinc site. In vitro, the mutant showed no difference in expression level and interaction compared to the *wildtype* proteins implying that the complex was still readily formed. Surprisingly, the metal site was not essential for complex formation. Then what is its function? To follow up on this question we further determined the telomeric length of cells containing the zinc site disruption mutations in their genome. These cells exhibited dramatically elongated telomeres compared to *wildtype* cells, similar to what was observed in *Poz1Δ* strains. What does the inhibition of zinc binding disturb in the telomere length regulation system? A possible explanation for the observed effects is that failure to bind zinc might cause destabilization of the Poz1-Tpz1 heterotetramer. This assumption is supported by the fact that the zinc binding site lies in close proximity to the helix α1 of Poz1 that forms major contacts in the heterotetrameric assembly. The inability to bind zinc might cause a conformational change of the secondary structure elements involved at the heterotetramerization interface, no longer allowing multimerization. Using size-exclusion chromatography we attempted to test whether the zinc site mutant remains to support the formation of heterotetramers. Due to precipitation of the protein after affinity chromatography the experiment was unsuccessful. Possibly, the inability to retain the zinc ion disrupts the integrity of the complex (or at least parts of it since Poz1 stills binds Tpz1) and decreases protein solubility. We propose that failure to bind zinc also results in protein destabilization *in vivo* including disruption of the heterotetramer, leading to the long telomere phenotype. It should be noted that the protein destabilization was observed *in vitro* for the construct boundaries Poz1$^{30-249}$+Tpz1$^{475-508}$ and whether structural destabilization also occurs in the fulllength proteins with a disrupted zinc site remains currently unknown.

An alternative, more speculative explanation, is that instead of structural destabilization or disruption of the heterotetramerization interface, failure to bind the zinc ion, might result in a longer range conformational change in Tpz1. So far we have solved the structure a Tpz1 C-terminal fragment. The architecture of the remaining larger part of the

protein remains unknown. One possibility is that the zinc ion must be bound at all times to ensure the integrity of the protein structure. Another possibility is that association and dissociation of the zinc ion and conformational changes resulting thereof are part a shelterin regulation mechanism. Zinc transfer could be mediated by interaction with proteins such as metallotheioneins and theioneins that have been implicated in metal transfer from and to zinc binding proteins (Margoshes and Vallee 1995, Jacob *et al.*, 1998).

Poz1 as a bridge linking the double- and single-strand binding complexes

Poz1 bridges the double-strand and single-strand binding parts of the shelterin complex by linking Rap1 and Tpz1. Is this the only function of Poz1 and is the bridge a static construction, or can the bridge be opened and closed as part of a regulation mechanism? To address this question, we attempted to replace the Poz1-bridge by a structurally similar protein, namely human TRF2. By fusing the TRF2H dimerization domain to the C-terminus of Rap1, as well as a peptide of the TRF2H interacting protein Apollo to the C-terminus of Tpz1, we reconstructed the link that Poz1 forms in *wildtype* fission yeast (**figure 5.3**).
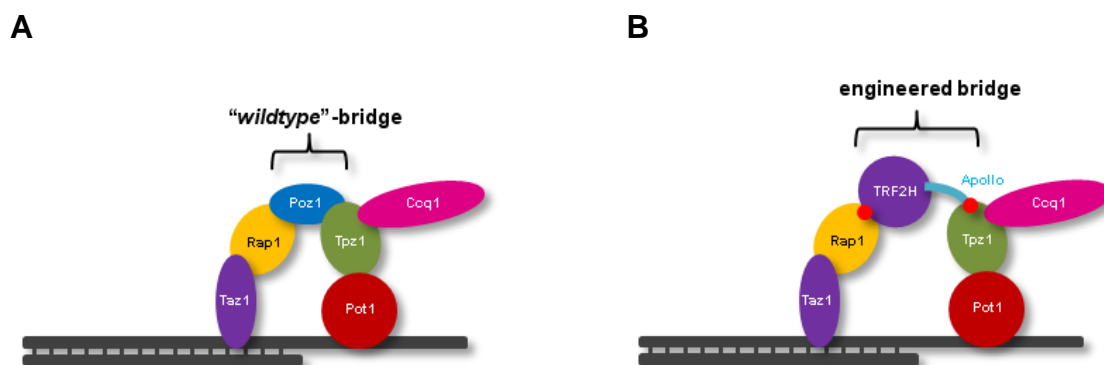
Figure 5.3. Replacing the "Poz1-bridge" by an engineered TRF2H-Apollo bridge.

A) In *wildtype* fission yeast, Poz1 forms a bridge between the double-strand binding shelterin complex Taz1-Rap1 and the single-strand binding complex Pot1-Tpz1. **B)** We removed Poz1 and replaced it by an engineered bridge composed of the human shelterin proteins TRF2H and hsApollo. Small dots indicate a covalent fusion.

Indeed fission yeast cells that harbored the engineered bridge did not show the "very long telomere" phenotype that we would observe in *Poz1Δ* cells. Instead, these cells showed slow telomere shortening after several restreaks. Possibly, the introduction of the structurally

similar TRF2H domain was able to compensate for most of the function of Poz1. But why can a foreign species protein, which is not a functional homologue, substitute for most of the function of Poz1? Poz1 links the double-strand and the single-strand halves of the shelterin complex. This is the role that the TRF2H dimerization module fused to Rap1 and the TRF2H binding peptide of Apollo fused to Tpz1 take over. Besides restoration of the link between Rap1 and Tpz1, the high degree of structural similarity of TRF2H-Apollo and Poz1-Tpz1, ensures littlest possible structural distortion in the shelterin complex, as the overall dimensions and even the Cα-backbone are very similar. Importantly, we propose that TRF2H-Apollo imitates the function of Poz1 so well, since it establishes the dimeric link between the Rap1-dimer and the two Tpz1 molecules. Thus, I propose the primary function of Poz1 to be the bridging of the double-strand and the single-strand halves of the shelterin complex in a dimeric manner.

However, the restoration of Poz1 function was not completely successful, as slow telomere shortening was observed in the TRF2H-Apollo strain. A likely explanation is that in *wildtype* cells, the interaction of Poz1 with Rap1 and Tpz1 is of a dynamic nature and the proteins associate and dissociate in a regulated manner. In the engineered fusion strain, the bridge is constitutive and the complex is therefore forced to remain in the closed conformation, which is a likely cause for the defects in telomere length homeostasis. This explanation is in line with the model of fission yeast telomere length regulation where in the "closed state" (meaning the Poz1-bridge is closed) there is no recruitment/stimulation of telomerase (see introduction **1.4**). Since the engineered TRF2H-Apollo-bridge represents the closed form, telomerase might not be recruited/stimulated as suggested in the model. This would lead to progressive telomere shortening. On the other hand, when the Poz1-bridge is opened, telomerase can be recruited/stimulated. The *Poz1Δ* strain showed strong telomere elongation, indicating that Poz1 is not required for telomere elongation. However, when telomeres are long, Poz1 is essential for closing the bridge and thereby inhibiting elongation by telomerase.
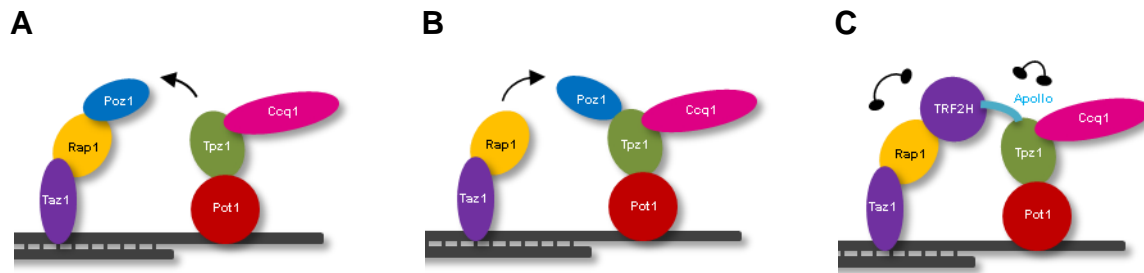
Figure 5.4. Bridging the double-strand and the single-strand shelterin complex.

A) and **B**) Opening of the shelterin complex by dissociation of Poz1 from either Rap1 or Tpz1. **C**) The shelterin complex is forced into the closed state as TRF2H is fused to Rap1 and Apollo is fused to Tpz1. Since TRF2H and Apollo are human proteins their interaction is not regulated, not allowing the opening of the bridge.

In the TRF2H-Apollo fusion strain we replaced Poz1 by an artificial bridge. Perhaps it is the missing Poz1 that is causing length deregulation and not the fact that the complex is kept in a closed conformation. As a future experiment we need to analyze Poz1 in the context of a forced "closed-state": If we fuse Poz1 to Rap1 and Tpz1 and we see similar effects than what was observed in the TRF2H-Apollo fusion strain this would indicate that the mere presence of Poz1 as a bridge is not sufficient to maintain proper shelterin function. Consequently, this would show that the association and dissociation of the bridge in a regulated manner is of crucial importance.

A future quest will be to determine additional domains of shelterin components besides Poz1 that are essential for shelterin function. One possible approach is to assemble a minimal shelterin complex: All structural elements that are required for assembly would be present and other regions would be systematically excluded. In such a manner we could identify currently unknown domains that are part or targets of regulation mechanisms.

Final remarks

By solving the structure of the Poz1-Tpz1 shelterin subcomplex and *in vivo* experiments based there on in fission yeast, we have gained insight into telomere length regulation and have highlighted how delicate the system is and how small changes can lead to severe regulation defects. Although the human shelterin complex differs from the fission yeast complex, many similarities can be observed and a better understanding of the system in yeast can be translated to the human system. Given the fact that telomere dysfunction,

especially length regulation, has been implicated in many human diseases including cancer, gaining as much knowledge as possible on the molecular workings of these mechanisms is paramount.

Acknowledgements

Research described in this thesis was carried out at the Friedrich Miescher Institute for Biomedical Research in Basel. I want to thank in particular Dr. Nicolas Thomä for giving me the opportunity to work in your group and for all your help and patience during my PhD. Thank you to Prof. Susan Gasser and Prof. Marc Bühler from the FMI and Prof. Peter Baumann from the Stowers Institute in Kansas City for accepting the role as thesis committee members and giving me ideas for my project through insightful discussions. Thank you to Peter Baumann and Lili Pan for performing the telomere length assays on my fission yeast mutants and for taking the time to discuss the meaning of the results. Thank you to former lab member Dr. Andrea Scrima for your help with software for structure solution. My sincere thanks go to Dr. Daniel Hess, Dominique Klein and Ragna Sack from the FMI massspec facility for your help with protein identification experiments. I am also very grateful to Dr. Heinz Gut from the FMI protein structure facility for the many trips to the SLS and for tips and tricks with protein crystals, data collection and structure solution. Thank you to Dr. Jeremy Keusch for help with the crystallization robot and the crystallization tray storage machine. I want to thank Claudia Keller and Yukiko Shimada from the Bühler lab at the FMI for the time you took to introduce me to the techniques of yeast genome manipulation and for supplying me with your reagents. Thank you to Frank Schmitges and Mahamadou Faty for introducing me to some techniques in the wet lab at the beginning of my PhD. Many thanks to my friends Markus, Philipp, Nicolas, Andrea, Frank, Renata, Jonathan, Mahamadou, Tianlai, Kerstin, Simone, Wassim, Raju, Richard, Eric and Julius from the Group Thomä for good times inside and outside the FMI as well as insightful discussions on problems and techniques in the lab. Thank you Elena for all your understanding, patience and love. You were always there despite our spatial distance. Last but not least I would like to dedicate this thesis to my parents who always believed in me and supported me wherever they could.

References

References

Abrahams, J. P. and A. G. Leslie (1996). "Methods used in the structure determination of bovine mitochondrial F1 ATPase." Acta Crystallogr D Biol Crystallogr 52(Pt 1): 30-42.

Adams, P. D., P. V. Afonine, et al. (2010). "PHENIX: a comprehensive Python-based system for macromolecular structure solution." Acta Crystallogr D Biol Crystallogr 66(Pt 2): 213-21.

Alder, J. K., J. J. Chen, et al. (2008). "Short telomeres are a risk factor for idiopathic pulmonary fibrosis." Proc Natl Acad Sci U S A 105(35): 13051-6.

Allsopp, R. C. and C. B. Harley (1995). "Evidence for a critical telomere length in senescent human fibroblasts." Exp Cell Res 219(1): 130-6.

Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." J Mol Biol 215(3): 403-10.

Anbalagan, S., D. Bonetti, et al. (2011). "Rif1 supports the function of the CST complex in yeast telomere capping." PLoS Genet 7(3): e1002024.

Anderson, B. H., P. R. Kasher, et al. (2012). "Mutations in CTC1, encoding conserved telomere maintenance component 1, cause Coats plus." Nat Genet 44(3): 338-42.

Armanios, M. and E. H. Blackburn (2012). "The telomere syndromes." Nat Rev Genet 13(10): 693-704.

Armanios, M., J. L. Chen, et al. (2005). "Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita." Proc Natl Acad Sci U S A 102(44): 15960-4.

Armanios, M. Y., J. J. Chen, et al. (2007). "Telomerase mutations in families with idiopathic pulmonary fibrosis." N Engl J Med 356(13): 1317-26.

Armbruster, B. N., C. M. Linardic, et al. (2004). "Rescue of an hTERT mutant defective in telomere elongation by fusion with hPot1." Mol Cell Biol 24(8): 3552-61.

Bähler, J., J. Q. Wu, et al. (1998). " Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe." Yeast 14(10):943-51.

Bassler, J., M. Kallas, et al. (2010). "The AAA-ATPase Rea1 drives removal of biogenesis factors during multiple stages of 60S ribosome assembly." Mol Cell 38(5): 712-21.

Bateman, A. J. (1975). "Letter: Simplification of palindromic telomere theory." Nature 253(5490): 379-80.

Baumann, P. and T. R. Cech (2000). "Protection of telomeres by the Ku protein in fission yeast." Mol Biol Cell 11(10): 3265-75.

Baumann, P. and T. R. Cech (2001). "Pot1, the putative telomere end-binding protein in fission yeast and humans." Science 292(5519): 1171-5.

Beattie, T. L., W. Zhou, et al. (2001). "Functional multimerization of the human telomerase reverse transcriptase." Mol Cell Biol 21(18): 6151-60.

Berman, H. M., J. Westbrook, et al. (2000). "The Protein Data Bank." Nucleic Acids Res 28(1): 235-42.

Bianchi, A. and D. Shore (2008). "How telomerase reaches its end: mechanism of telomerase regulation by the telomeric complex." Mol Cell 31(2): 153-65.

Bianchi, A. and D. Shore (2008). "Molecular biology. Refined view of the ends." Science 320(5881): 1301-2.

Biegert, A., C. Mayer, et al. (2006). "The MPI Bioinformatics Toolkit for protein sequence analysis." Nucleic Acids Res 34(Web Server issue): W335-9.

Bilaud, T., C. E. Koering, et al. (1996). "The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human." Nucleic Acids Res 24(7): 1294-303.

Bochkarev, A., E. Bochkareva, et al. (1999). "The crystal structure of the complex of replication protein A subunits RPA32 and RPA14 reveals a mechanism for single-stranded DNA binding." Embo J 18(16): 4498-504.

Bodnar, A. G., M. Ouellette, et al. (1998). "Extension of life-span by introduction of telomerase into normal human cells." Science 279(5349): 349-52.

Bork, P., K. Hofmann, et al. (1997). "A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins." Faseb J 11(1): 68-76.

Broccoli, D., L. Chong, et al. (1997). "Comparison of the human and mouse genes encoding the telomeric protein, TRF1: chromosomal localization, expression and conserved protein domains." Hum Mol Genet 6(1): 69-76.

Buchan, D. W., S. M. Ward, et al. (2010). "Protein annotation and modelling servers at University College London." Nucleic Acids Res 38(Web Server issue): W563-8.

Bunch, J. T., N. S. Bae, et al. (2005). "Distinct requirements for Pot1 in limiting telomere length and maintaining chromosome stability." Mol Cell Biol 25(13): 5567-78.

Bunch, J. T., N. S. Bae, et al. (2005). "Distinct requirements for Pot1 in limiting telomere length and maintaining chromosome stability." Mol Cell Biol 25(13): 5567-78.

Calado, R. T., J. A. Regal, et al. (2009). "A spectrum of severe familial liver disorders associate with telomerase mutations." PLoS One 4(11): e7926.

Campisi, J. (1997). "Aging and cancer: the double-edged sword of replicative senescence." J Am Geriatr Soc 45(4): 482-8.

Campisi, J. (1997). "The biology of replicative senescence." Eur J Cancer 33(5): 703-9.

Capkova Frydrychova, R., H. Biessmann, et al. (2008). "Regulation of telomere length in Drosophila." Cytogenet Genome Res 122(3-4): 356-64.

Casteel, D. E., S. Zhuang, et al. (2009). "A DNA polymerase-{alpha}{middle dot}primase cofactor with homology to replication protein A-32 regulates DNA replication in mammalian cells." J Biol Chem 284(9): 5807-18.

Cavalier-Smith, T. (1974). "Palindromic base sequences and replication of eukaryote chromosome ends." Nature 250(5466): 467-70.

Cenci, G., L. Ciapponi, et al. (2005). "The mechanism of telomere protection: a comparison between Drosophila and humans." Chromosoma 114(3): 135-45.

Chai, W., A. J. Sfeir, et al. (2006). "The involvement of the Mre11/Rad50/Nbs1 complex in the generation of G-overhangs at human telomeres." EMBO Rep 7(2): 225-30.

Chang, W., J. N. Dynek, et al. (2003). "TRF1 is degraded by ubiquitin-mediated proteolysis after release from telomeres." Genes Dev 17(11): 1328-33.

Chen, L. Y., S. Redon, et al. (2012). "The human CST complex is a terminator of telomerase activity." Nature 488(7412): 540-4.

Chen, Y., R. Rai, et al. (2011). "A conserved motif within RAP1 has diversified roles in telomere protection and regulation in different organisms." Nat Struct Mol Biol 18(2): 213-21.

Chen, Y., Y. Yang, et al. (2008). "A shared docking motif in TRF1 and TRF2 used for differential recruitment of telomeric proteins." Science 319(5866): 1092-6.

Chiang, Y. J., S. H. Kim, et al. (2004). "Telomere-associated protein TIN2 is essential for early embryonic development through a telomerase-independent pathway." Mol Cell Biol 24(15): 6631-4.

Chikashige, Y. and Y. Hiraoka (2001). "Telomere binding of the Rap1 protein is required for meiosis in fission yeast." Curr Biol 11(20): 1618-23.

Chikashige, Y., C. Tsutsumi, et al. (2006). "Meiotic proteins bqt1 and bqt2 tether telomeres to form the bouquet arrangement of chromosomes." Cell 125(1): 59-69.

Chong, L., B. van Steensel, et al. (1995). "A human telomeric protein." Science 270(5242): 1663-7.

Chou, P. Y. and G. D. Fasman (1974). "Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins." Biochemistry 13(2): 211-22.

Chou, P. Y. and G. D. Fasman (1974). "Prediction of protein conformation." Biochemistry 13(2): 222-45.

Church, G. M. and W. Gilbert (1984). "Genomic sequencing." Proc Natl Acad Sci U S A 81(7): 1991-5.

Cohn, M. and E. H. Blackburn (1995). "Telomerase in yeast." Science 269(5222): 396-400.

Cooper, J. P., E. R. Nimmo, et al. (1997). "Regulation of telomere length and function by a Myb-domain protein in fission yeast." Nature 385(6618): 744-7.

Counter, C. M., A. A. Avilion, et al. (1992). "Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity." Embo J 11(5): 1921-9.

Court, R., L. Chapman, et al. (2005). "How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA: a view from high-resolution crystal structures." EMBO Rep 6(1): 39-45.

Cowtan, K. (2006). "The Buccaneer software for automated model building. 1. Tracing protein chains." Acta Crystallogr D Biol Crystallogr 62(Pt 9): 1002-11.

Croy, J. E., E. R. Podell, et al. (2006). "A new model for Schizosaccharomyces pombe telomere recognition: the telomeric single-stranded DNA-binding activity of Pot11-389." J Mol Biol 361(1): 80-93.

de Lange, T. (2005). "Shelterin: the protein complex that shapes and safeguards human telomeres." Genes Dev 19(18): 2100-10.

DeLano W "The PyMOL Molecular Graphics System." DeLano Scientific LLC, San Carlos, CA, USA. http://www.pymol.org

Dehe, P. M. and J. P. Cooper (2010). "Fission yeast telomeres forecast the end of the crisis." FEBS Lett 584(17): 3725-33.

Demuth, I., M. Digweed, et al. (2004). "Human SNM1B is required for normal cellular response to both DNA interstrand crosslink-inducing agents and ionizing radiation." Oncogene 23(53): 8611-8.

Dhaene, K., E. Van Marck, et al. (2000). "Telomeres, telomerase and cancer: an up-date." Virchows Arch 437(1): 1-16.

Ding, Z., C. J. Wu, et al. (2012). "Telomerase reactivation following telomere dysfunction yields murine prostate tumors with bone metastases." Cell 148(5): 896-907.

Dionne, I. and R. J. Wellinger (1996). "Cell cycle-regulated generation of single-stranded G-rich DNA in the absence of telomerase." Proc Natl Acad Sci U S A 93(24): 13902-7.

Dunham, M. A., A. A. Neumann, et al. (2000). "Telomere maintenance by recombination in human cells." Nat Genet 26(4): 447-50.

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr 60(Pt 12 Pt 1): 2126-32.

Emsley, P., B. Lohkamp, et al. (2010). "Features and development of Coot." Acta Crystallogr D Biol Crystallogr 66(Pt 4): 486-501.

Fairall, L., L. Chapman, et al. (2001). "Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2." Mol Cell 8(2): 351-61.

Fajkus, J., A. Kovarik, et al. (1995). "Organization of telomeric and subtelomeric chromatin in the higher plant Nicotiana tabacum." Mol Gen Genet 247(5): 633-8.

Feng, J., W. D. Funk, et al. (1995). "The RNA component of human telomerase." Science 269(5228): 1236-41.

Flory, M. R., A. R. Carson, et al. (2004). "An SMC-domain protein in fission yeast links telomeres to the meiotic centrosome." Mol Cell 16(4): 619-30.

Fogarty, P. F., H. Yamaguchi, et al. (2003). "Late presentation of dyskeratosis congenita as apparently acquired aplastic anaemia due to mutations in telomerase RNA." Lancet 362(9396): 1628-30.

Forstemann, K. and J. Lingner (2001). "Molecular basis for telomere repeat divergence in budding yeast." Mol Cell Biol 21(21): 7277-86.

Fujita, K., I. Horikawa, et al. (2010). "Positive feedback between p53 and TRF2 during telomere-damage signalling and cellular senescence." Nat Cell Biol 12(12): 1205-12.

Gao, G., J. C. Walser, et al. (2010). "HipHop interacts with HOAP and HP1 to protect Drosophila telomeres in a sequence-independent manner." Embo J 29(4): 819-29.

Garnier, J., J. F. Gibrat, et al. (1996). "GOR method for predicting protein secondary structure from amino acid sequence." Methods Enzymol 266: 540-53.

Garvik, B., M. Carson, et al. (1995). "Single-stranded DNA arising at telomeres in cdc13 mutants may constitute a specific signal for the RAD9 checkpoint." Mol Cell Biol 15(11): 6128-38.

Girardi, A. J., F. C. Jensen, et al. (1965). "Sv40-Induced Tranformation of Human Diploid Cells: Crisis and Recovery." J Cell Physiol 65: 69-83.

Giraud-Panis, M. J., M. T. Teixeira, et al. (2010). "CST meets shelterin to keep telomeres in check." Mol Cell 39(5): 665-76.

Glover, J. N., R. S. Williams, et al. (2004). "Interactions between BRCT repeats and phosphoproteins: tangled up in two." Trends Biochem Sci 29(11): 579-85.

Grandin, N. and M. Charbonneau (2008). "Protection against chromosome degradation at the telomeres." Biochimie 90(1): 41-59.

Grandin, N., C. Damon, et al. (2001). "Ten1 functions in telomere end protection and length regulation in association with Stn1 and Cdc13." Embo J 20(5): 1173-83.

Grandin, N., S. I. Reed, et al. (1997). "Stn1, a new Saccharomyces cerevisiae protein, is implicated in telomere size regulation in association with Cdc13." Genes Dev 11(4): 512-27.

Gray, J. T., D. W. Celander, et al. (1991). "Cloning and expression of genes for the Oxytricha telomere-binding protein: specific subunit interactions in the telomeric complex." Cell 67(4): 807-14.

Greider, C. W. and E. H. Blackburn (1985). "Identification of a specific telomere terminal transferase activity in Tetrahymena extracts." Cell 43(2 Pt 1): 405-13.

Greider, C. W. and E. H. Blackburn (1989). "A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis." Nature 337(6205): 331-7.

Griffith, J., S. Michalowski, et al. (1999). "Electron microscopy of DNA-protein complexes and chromatin." Methods Enzymol 304: 214-30.

Gu, P., J. N. Min, et al. "CTC1 deletion results in defective telomere replication, leading to catastrophic telomere loss and stem cell exhaustion." Embo J 31(10): 2309-21.

Hao, L. Y., M. Armanios, et al. (2005). "Short telomeres, even in the presence of telomerase, limit tissue renewal capacity." Cell 123(6): 1121-31.

Hastie, N. D., M. Dempster, et al. (1990). "Telomere reduction in human colorectal carcinoma and with ageing." Nature 346(6287): 866-8.

Heiss, N. S., S. W. Knight, et al. (1998). "X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions." Nat Genet 19(1): 32-8.

Hendrickson, W. A. (1991). "Determination of macromolecular structures from anomalous diffraction of synchrotron radiation." Science 254(5028): 51-8.

Hendrickson, W. A., J. R. Horton, et al. (1990). "Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure." Embo J 9(5): 1665-72.

Hertz-Fowler, C., C. S. Peacock, et al. (2004). "GeneDB: a resource for prokaryotic and eukaryotic organisms." Nucleic Acids Res 32(Database issue): D339-43.

Hollstein, M., D. Sidransky, et al. (1991). "p53 mutations in human cancers." Science 253(5015): 49-53.

Holm, L. and P. Rosenstrom (2010). "Dali server: conservation mapping in 3D." Nucleic Acids Res 38(Web Server issue): W545-9.

Horvath, M. P., V. L. Schweiker, et al. (1998). "Crystal structure of the Oxytricha nova telomere end binding protein complexed with single strand DNA." Cell 95(7): 963-74.

Huda, N., H. Tanaka, et al. (2009). "DNA damage-induced phosphorylation of TRF2 is required for the fast pathway of DNA double-strand break repair." Mol Cell Biol 29(13): 3597-604.

Humpal, S. E., D. A. Robinson, et al. (2009). "Marks to stop the clock: histone modifications and checkpoint regulation in the DNA damage response." Biochem Cell Biol 87(1): 243-53.

Jacob, C., W. Maret, et al. (1998). "Control of zinc transfer between thionein, metallothionein, and zinc proteins." Proc Natl Acad Sci U S A 95(7): 3489-94.

Jones, D. T. (1999). "Protein secondary structure prediction based on position-specific scoring matrices." J Mol Biol 292(2): 195-202.

Kabsch, W. (2010). "Xds." Acta Crystallogr D Biol Crystallogr 66(Pt 2): 125-32.

Kanter-Smoler, G., A. Dahlkvist, et al. (1994). " Improved method for rapid transformation of intact Schizosaccharomyces pombe cells." Biotechniques 16(5):798-800.

Keller, R. B., K. E. Gagne, et al. (2012). "CTC1 Mutations in a patient with dyskeratosis congenita." Pediatr Blood Cancer 59(2): 311-4.

Kim, S. H., P. Kaminker, et al. (1999). "TIN2, a new regulator of telomere length in human cells." Nat Genet 23(4): 405-12.

Kipling, D. and H. J. Cooke (1990). "Hypervariable ultra-long telomeres in mice." Nature 347(6291): 400-2.

Kirwan, M., T. Vulliamy, et al. (2009). "Defining the pathogenic role of telomerase mutations in myelodysplastic syndrome and acute myeloid leukemia." Hum Mutat 30(11): 1567-73.

Klobutcher, L. A., M. T. Swanton, et al. (1981). "All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus." Proc Natl Acad Sci U S A 78(5): 3015-9.

Knight, S. W., N. S. Heiss, et al. (1999). "Unexplained aplastic anaemia, immunodeficiency, and cerebellar hypoplasia (Hoyeraal-Hreidarsson syndrome) due to mutations in the dyskeratosis congenita gene, DKC1." Br J Haematol 107(2): 335-9.

Koonin, E. V., S. F. Altschul, et al. (1996). "BRCA1 protein products ... Functional motifs." Nat Genet 13(3): 266-8.

Krissinel, E. and K. Henrick (2007). "Inference of macromolecular assemblies from crystalline state." J Mol Biol 372(3): 774-97.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature 227(5259): 680-5.

Lamzin, V. S. and K. S. Wilson (1997). "Automated refinement for protein crystallography." Methods Enzymol 277: 269-305.

Larkin, M. A., G. Blackshields, et al. (2007). "Clustal W and Clustal X version 2.0." Bioinformatics 23(21): 2947-8.

Larrivee, M., C. LeBel, et al. (2004). "The generation of proper constitutive G-tails on yeast telomeres is dependent on the MRX complex." Genes Dev 18(12): 1391-6.

Laskowski, R. A. (2001). "PDBsum: summaries and analyses of PDB structures." Nucleic Acids Res 29(1): 221-2.

Lei, M., E. R. Podell, et al. (2003). "DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA." Nature 426(6963): 198-203.

Lei, M., E. R. Podell, et al. (2004). "Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection." Nat Struct Mol Biol 11(12): 1223-9.

Lendvay, T. S., D. K. Morris, et al. (1996). "Senescence mutants of Saccharomyces cerevisiae with a defect in telomere replication identify three additional EST genes." Genetics 144(4): 1399-412.

Leonardi, J., J. A. Box, et al. (2008). "TER1, the RNA subunit of fission yeast telomerase." Nat Struct Mol Biol 15(1): 26-33.

Li, B., S. Oestreich, et al. (2000). "Identification of human Rap1: implications for telomere evolution." Cell 101(5): 471-83.

Lingner, J., T. R. Cech, et al. (1997). "Three Ever Shorter Telomere (EST) genes are dispensable for in vitro yeast telomerase activity." Proc Natl Acad Sci U S A 94(21): 11190-5.

Lingner, J., T. R. Hughes, et al. (1997). "Reverse transcriptase motifs in the catalytic subunit of telomerase." Science 276(5312): 561-7.

Liu, D., A. Safari, et al. (2004). "PTOP interacts with POT1 and regulates its localization to telomeres." Nat Cell Biol 6(7): 673-80.

Loayza, D., H. Parsons, et al. (2004). "DNA binding features of human POT1: a nonamer 5'-TAGGGTTAG-3' minimal binding site, sequence specificity, and internal binding to multimeric sites." J Biol Chem 279(13): 13241-8.

Longhese, M. P., S. Anbalagan, et al. (2012). "The role of shelterin in maintaining telomere integrity." Front Biosci 17: 1715-28.

Lovell, S. C., I. W. Davis, et al. (2003). "Structure validation by Calpha geometry: phi,psi and Cbeta deviation." Proteins 50(3): 437-50.

Luderus, M. E., B. van Steensel, et al. (1996). "Structure, subnuclear distribution, and nuclear matrix association of the mammalian telomeric complex." J Cell Biol 135(4): 867-81.

Lundblad, V. and J. W. Szostak (1989). "A mutant with a defect in telomere elongation leads to senescence in yeast." Cell 57(4): 633-43.

Marcand, S., E. Gilson, et al. (1997). "A protein-counting mechanism for telomere length regulation in yeast." Science 275(5302): 986-90.

Makarov, V. L., Y. Hirose, et al. (1997). "Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening." Cell 88(5): 657-66.

Marcand, S., D. Wotton, et al. (1997). "Rap1p and telomere length regulation in yeast." Ciba Found Symp 211: 76-93; discussion 93-103.

Martin, V., L. L. Du, et al. (2007). "Protection of telomeres by a conserved Stn1-Ten1 complex." Proc Natl Acad Sci U S A 104(35): 14038-43.

Mason, J. M., R. C. Frydrychova, et al. (2008). "Drosophila telomeres: an exception providing new insights." Bioessays 30(1): 25-37.

Matsumoto, T., K. Fukui, et al. (1987). "Identification of healed terminal DNA fragments in linear minichromosomes of Schizosaccharomyces pombe." Mol Cell Biol 7(12): 4424-30.

McClintock, B. (1938). "The Production of Homozygous Deficient Tissues with Mutant Characteristics by Means of the Aberrant Mitotic Behavior of Ring-Shaped Chromosomes." Genetics 23(4): 315-76.

McClintock, B. (1941). "The Stability of Broken Ends of Chromosomes in Zea Mays." Genetics 26(2): 234-82.

Meyne, J., R. J. Baker, et al. (1990). "Distribution of non-telomeric sites of the (TTAGGG)n telomeric sequence in vertebrate chromosomes." Chromosoma 99(1): 3-10.

Miller, K. M., M. G. Ferreira, et al. (2005). "Taz1, Rap1 and Rif1 act both interdependently and independently to maintain telomeres." Embo J 24(17): 3128-35.

Miller, K. M., O. Rog, et al. (2006). "Semi-conservative DNA replication through telomeres requires Taz1." Nature 440(7085): 824-8.

Mitchell, J. R., E. Wood, et al. (1999). "A telomerase component is defective in the human disease dyskeratosis congenita." Nature 402(6761): 551-5.

Miyake, Y., M. Nakamura, et al. (2009). "RPA-like mammalian Ctc1-Stn1-Ten1 complex binds to single-stranded DNA and protects telomeres independently of the Pot1 pathway." Mol Cell 36(2): 193-206.

Miyoshi, T., J. Kanoh, et al. (2008). "Fission yeast Pot1-Tpp1 protects telomeres and regulates telomere length." Science 320(5881): 1341-4.

Morris, R. J., A. Perrakis, et al. (2003). "ARP/wARP and automatic interpretation of protein electron density maps." Methods Enzymol 374: 229-44.

Moser, B. A., Y. T. Chang, et al. (2011). "Tel1ATM and Rad3ATR kinases promote Ccq1-Est1 interaction to maintain telomeres in fission yeast." Nat Struct Mol Biol 18(12): 1408-13.

Moyzis, R. K., J. M. Buckingham, et al. (1988). "A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes." Proc Natl Acad Sci U S A 85(18): 6622-6.

Murnane, J. P., L. Sabatier, et al. (1994). "Telomere dynamics in an immortal human cell line." Embo J 13(20): 4953-62.

Murshudov, G. N., A. A. Vagin, et al. (1997). "Refinement of macromolecular structures by the maximum-likelihood method." Acta Crystallogr D Biol Crystallogr 53(Pt 3): 240-55.

Murzin, A. G. (1993). "OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences." Embo J 12(3): 861-7.

Nakamura, T. M., G. B. Morin, et al. (1997). "Telomerase catalytic subunit homologs from fission yeast and human." Science 277(5328): 955-9.

Nandakumar, J. and T. R. Cech (2012). "DNA-induced dimerization of the single-stranded DNA binding telomeric protein Pot1 from Schizosaccharomyces pombe." Nucleic Acids Res 40(1): 235-44.

Nishikawa, T., H. Okamura, et al. (2001). "Structure of the DNA-binding domain of human telomeric protein, TRF1 and its interaction with telomeric DNA." Nucleic Acids Res Suppl(1): 273-4.

Noble, J. R., Z. H. Zhong, et al. (2004). "Alterations in the p16(INK4a) and p53 tumor suppressor genes of hTERT-immortalized human fibroblasts." Oncogene 23(17): 3116-21.

O'Connor, M. S., A. Safari, et al. (2004). "The human Rap1 protein complex and modulation of telomere length." J Biol Chem 279(27): 28585-91.

Ohi, M. D. and K. L. Gould (2002). "Characterization of interactions among the Cef1p-Prp19p-associated splicing complex." Rna 8(6): 798-815.

Okazaki, K., N. Okazaki, et al. (1990). "High-frequency transformation method and library transducing vectors for cloning mammalian cDNAs by trans-complementation of Schizosaccharomyces pombe." Nucl. Acids Res. 18:6485-6489.

Olovnikov, A. M. (1973). "A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon." J Theor Biol 41(1): 181-90.

Palm, W. and T. de Lange (2008). "How shelterin protects mammalian telomeres." Annu Rev Genet 42: 301-34.

Pardue, M. L. and P. G. DeBaryshe (2008). "Drosophila telomeres: A variation on the telomerase theme." Fly (Austin) 2(3): 101-10.

Perrakis, A., T. K. Sixma, et al. (1997). "wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models." Acta Crystallogr D Biol Crystallogr 53(Pt 4): 448-55.

Peska, V., E. Sykorova, et al. (2008). "Two faces of Solanaceae telomeres: a comparison between Nicotiana and Cestrum telomeres and telomere-binding proteins." Cytogenet Genome Res 122(3-4): 380-7.

Pich, U., J. Fuchs, et al. (1996). "How do Alliaceae stabilize their chromosome ends in the absence of TTTAGGG sequences?" Chromosome Res 4(3): 207-13.

Pich, U. and I. Schubert (1998). "Terminal heterochromatin and alternative telomeric sequences in Allium cepa." Chromosome Res 6(4): 315-21.

Polvi, A., T. Linnankivi, et al. (2012). "Mutations in CTC1, encoding the CTS telomere maintenance complex component 1, cause cerebroretinal microangiopathy with calcifications and cysts." Am J Hum Genet 90(3): 540-9.

Prescott, J. and E. H. Blackburn (1997). "Functionally interacting telomerase RNAs in the yeast telomerase complex." Genes Dev 11(21): 2790-800.

Price, C. M., K. A. Boltz, et al. (2010). "Evolution of CST function in telomere maintenance." Cell Cycle 9(16): 3157-65.

Raffa, G. D., L. Ciapponi, et al. (2011). "Terminin: a protein complex that mediates epigenetic maintenance of Drosophila telomeres." Nucleus 2(5): 383-91.

Ramachandran, G. N., C. Ramakrishnan, et al. (1963). "Stereochemistry of polypeptide chain configurations." J Mol Biol 7: 95-9.

Rogan, E. M., T. M. Bryan, et al. (1995). "Alterations in p53 and p16INK4 expression and telomere length during spontaneous immortalization of Li-Fraumeni syndrome fibroblasts." Mol Cell Biol 15(9): 4745-53.

Rossmann M. G. and Blow D. M. (1962). "The detection of sub-units within the crystallographic asymmetric unit." Acta Cryst. 15, 24-31

Sancar, A., L. A. Lindsey-Boltz, et al. (2004). "Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints." Annu Rev Biochem 73: 39-85.

Sarthy, J., N. S. Bae, et al. (2009). "Human RAP1 inhibits non-homologous end joining at telomeres." Embo J 28(21): 3390-9.

Sasaki, T. and H. Fujiwara (2000). "Detection and distribution patterns of telomerase activity in insects." Eur J Biochem 267(10): 3025-31.

Savage, S. A., N. Giri, et al. (2008). "TINF2, a component of the shelterin telomere protection complex, is mutated in dyskeratosis congenita." Am J Hum Genet 82(2): 501-9.

Shampay, J., J. W. Szostak, et al. (1984). "DNA sequences of telomeres maintained in yeast." Nature 310(5973): 154-7.

Shay, J. W., O. M. Pereira-Smith, et al. (1991). "A role for both RB and p53 in the regulation of human cellular senescence." Exp Cell Res 196(1): 33-9.

Shay, J. W. and W. E. Wright (2011). "Role of telomeres and telomerase in cancer." Semin Cancer Biol 21(6): 349-53.

Shiloh, Y. (2003). "ATM and related protein kinases: safeguarding genome integrity." Nat Rev Cancer 3(3): 155-68.

Smogorzewska, A. and T. de Lange (2004). "Regulation of telomerase by telomeric proteins." Annu Rev Biochem 73: 177-208.

Spink, K. G., R. J. Evans, et al. (2000). "Sequence-specific binding of Taz1p dimers to fission yeast telomeric DNA." Nucleic Acids Res 28(2): 527-33.

Stewart, J. A., F. Wang, et al. (2012). "Human CST promotes telomere duplex replication and general replication restart after fork stalling." Embo J 31(17): 3537-49.

Stewart, S. A., I. Ben-Porath, et al. (2003). "Erosion of the telomeric single-strand overhang at replicative senescence." Nat Genet 33(4): 492-6.

Strahl-Bolsinger, S., A. Hecht, et al. (1997). "SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast." Genes Dev 11(1): 83-93.

Sugawara N. Cambridge, MA: Harvard University; 1988. DNA Sequences at the Telomeres of the Fission Yeast S. pombe. Ph.D. Dissertation.

Sun, J., Y. Yang, et al. (2011). "Structural bases of dimerization of yeast telomere protein Cdc13 and its interaction with the catalytic subunit of DNA polymerase alpha." Cell Res 21(2): 258-74.

Sun, J., E. Y. Yu, et al. (2009). "Stn1-Ten1 is an Rpa2-Rpa3-like complex at telomeres." Genes Dev 23(24): 2900-14.

Surovtseva, Y. V., D. Churikov, et al. (2009). "Conserved telomere maintenance component 1 interacts with STN1 and maintains chromosome ends in higher eukaryotes." Mol Cell 36(2): 207-18.

Sykorova, E., K. Y. Lim, et al. (2003). "Telomere variability in the monocotyledonous plant order Asparagales." Proc Biol Sci 270(1527): 1893-904.

Szostak, J. W. and E. H. Blackburn (1982). "Cloning yeast telomeres on linear plasmid vectors." Cell 29(1): 245-55.

Tarkanyi, I. and J. Aradi (2008). "Pharmacological intervention strategies for affecting telomerase activity: future prospects to treat cancer and degenerative disease." Biochimie 90(1): 156-72.

Taylor, G. (2003). "The phase problem." Acta Crystallogr D Biol Crystallogr 59(Pt 11): 1881-90.

Terwilliger, T. (2004). "SOLVE and RESOLVE: automated structure solution, density modification and model building." J Synchrotron Radiat 11(Pt 1): 49-52.

Terwilliger, T. C. (2000). "Maximum-likelihood density modification." Acta Crystallogr D Biol Crystallogr 56(Pt 8): 965-72.

Terwilliger, T. C. (2003). "Improving macromolecular atomic models at moderate resolution by automated iterative model building, statistical density modification and refinement." Acta Crystallogr D Biol Crystallogr 59(Pt 7): 1174-82.

Terwilliger, T. C. (2003). "SOLVE and RESOLVE: automated structure solution and density modification." Methods Enzymol 374: 22-37.

Terwilliger, T. C. (2003). "Statistical density modification using local pattern matching." Acta Crystallogr D Biol Crystallogr 59(Pt 10): 1688-701.

Terwilliger, T. C. and J. Berendzen (1999). "Automated MAD and MIR structure solution." Acta Crystallogr D Biol Crystallogr 55(Pt 4): 849-61.

Tomaska, L., S. Willcox, et al. (2004). "Taz1 binding to a fission yeast model telomere: formation of telomeric loops and higher order structures." J Biol Chem 279(49): 50764-72.

Tomita, K. and J. P. Cooper (2008). "Fission yeast Ccq1 is telomerase recruiter and local checkpoint controller." Genes Dev 22(24): 3461-74.

Tomita, K., T. Kibe, et al. (2004). "Fission yeast Dna2 is required for generation of the telomeric single-strand overhang." Mol Cell Biol 24(21): 9557-67.

Touzot, F., I. Callebaut, et al. (2010). "Function of Apollo (SNM1B) at telomere highlighted by a splice variant identified in a patient with Hoyeraal-Hreidarsson syndrome." Proc Natl Acad Sci U S A 107(22): 10097-102.

Trujillo, K. M., J. T. Bunch, et al. (2005). "Extended DNA binding site in Pot1 broadens sequence specificity to allow recognition of heterogeneous fission yeast telomeres." J Biol Chem 280(10): 9119-28.

Tsakiri, K. D., J. T. Cronkhite, et al. (2007). "Adult-onset pulmonary fibrosis caused by mutations in telomerase." Proc Natl Acad Sci U S A 104(18): 7552-7.

Ueno, M., R. Kurokawa, et al. (2001). "Schizosaccharomyces pombe taf1+ is required for nitrogen starvation-induced sexual development and for entering the dormant GO state." Curr Genet 38(6): 307-13.

Vallee, B. L. (1995). "The function of metallothionein." Neurochem Int 27(1): 23-33.

Vaziri, H. and S. Benchimol (1998). "Reconstitution of telomerase activity in normal human cells leads to elongation of telomeres and extended replicative life span." Curr Biol 8(5): 279-82.

Vonrhein, C., E. Blanc, et al. (2007). "Automated structure solution with autoSHARP." Methods Mol Biol 364: 215-30.

Vulliamy, T., R. Beswick, et al. (2008). "Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita." Proc Natl Acad Sci U S A 105(23): 8073-8.

Vulliamy, T., A. Marrone, et al. (2002). "Association between aplastic anaemia and mutations in telomerase RNA." Lancet 359(9324): 2168-70.

Vulliamy, T., A. Marrone, et al. (2001). "The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita." Nature 413(6854): 432-5.

Walker, J. R. and X. D. (2012). Zhu "Post-translational modifications of TRF1 and TRF2 and their roles in telomere maintenance." Mech Ageing Dev 133(6): 421-34.

Walne, A. J. and I. Dokal (2008). "Dyskeratosis Congenita: a historical perspective." Mech Ageing Dev 129(1-2): 48-59.

Walne, A. J., T. Vulliamy, et al. (2008). "TINF2 mutations result in very short telomeres: analysis of a large cohort of patients with dyskeratosis congenita and related bone marrow failure syndromes." Blood 112(9): 3594-600.

Walne, A. J., T. Vulliamy, et al. (2007). "Genetic heterogeneity in autosomal recessive dyskeratosis congenita with one subtype due to mutations in the telomerase-associated protein NOP10." Hum Mol Genet 16(13): 1619-29.

Wan, M., J. Qin, et al. (2009). "OB fold-containing protein 1 (OBFC1), a human homolog of yeast Stn1, associates with TPP1 and is implicated in telomere length regulation." J Biol Chem 284(39): 26725-31.

Wang, F., E. R. Podell, et al. (2007). "The POT1-TPP1 telomere complex is a telomerase processivity factor." Nature 445(7127): 506-10.

Wang, L., S. R. Dean, et al. (2002). "Oligomerization of the telomerase reverse transcriptase from Euplotes crassus." Nucleic Acids Res 30(18): 4032-9.

Watson, J. D. (1972). "Origin of concatemeric T7 DNA." Nat New Biol 239(94): 197-201.

Watson, J. M. and K. Riha (2010). "Comparative biology of telomeres: where plants stand." FEBS Lett 584(17): 3752-9.

Webb, C. J. and V. A. Zakian (2008). "Identification and characterization of the Schizosaccharomyces pombe TER1 telomerase RNA." Nat Struct Mol Biol 15(1): 34-42.

Webb, C. J. and V. A. Zakian (2012). "Schizosaccharomyces pombe Ccq1 and TER1 bind the 14-3-3-like domain of Est1, which promotes and stabilizes telomerase-telomere association." Genes Dev 26(1): 82-91.

Wellinger, R. J., A. J. Wolf, et al. (1992). "Use of non-denaturing Southern hybridization and two dimensional agarose gels to detect putative intermediates in telomere replication in Saccharomyces cerevisiae." Chromosoma 102(1 Suppl): S150-6.

Wellinger, R. J., A. J. Wolf, et al. (1993). "Structural and temporal analysis of telomere replication in yeast." Cold Spring Harb Symp Quant Biol 58: 725-32.

Wenz, C., B. Enenkel, et al. (2001). "Human telomerase contains two cooperating telomerase RNA molecules." Embo J 20(13): 3526-34.

Willson, J., S. Wilson, et al. (1997). "Isolation and characterization of the Schizosaccharomyces pombe rhp9 gene: a gene required for the DNA damage checkpoint but not the replication checkpoint." Nucleic Acids Res 25(11): 2138-46.

Wright, W. E., O. M. Pereira-Smith, et al. (1989). "Reversible cellular senescence: implications for immortalization of normal human diploid fibroblasts." Mol Cell Biol 9(7): 3088-92.

Xin, H., D. Liu, et al. (2007). "TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase." Nature 445(7127): 559-62.

Yaghmai, R., A. Kimyai-Asadi, et al. (2000). "Overlap of dyskeratosis congenita with the Hoyeraal-Hreidarsson syndrome." J Pediatr 136(3): 390-3.

Yamaguchi, H., G. M. Baerlocher, et al. (2003). "Mutations of the human telomerase RNA gene (TERC) in aplastic anemia and myelodysplastic syndrome." Blood 102(3): 916-8.

Yamaguchi, H., R. T. Calado, et al. (2005). "Mutations in TERT, the gene for telomerase reverse transcriptase, in aplastic anemia." N Engl J Med 352(14): 1413-24.

Yamazaki, H., Y. Tarumoto, et al. (2012). "Tel1(ATM) and Rad3(ATR) phosphorylate the telomere protein Ccq1 to recruit telomerase and elongate telomeres in fission yeast." Genes Dev 26(3): 241-6.

Ye, J. Z., D. Hockemeyer, et al. (2004). "POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex." Genes Dev 18(14): 1649-54.

Zakian, V. A. (1995). "Telomeres: beginning to understand the end." Science 270(5242): 1601-7.

Zaug, A. J., E. R. Podell, et al. (2010). "Functional interaction between telomere protein TPP1 and telomerase." Genes Dev 24(6): 613-22.

Zhang, X., S. Morera, et al. (1998). "Structure of an XRCC1 BRCT domain: a new protein-protein interaction module." Embo J 17(21): 6404-11.

Zhong, F., S. A. Savage, et al. (2011). "Disruption of telomerase trafficking by TCAB1 mutation causes dyskeratosis congenita." Genes Dev 25(1): 11-6.

Zhong, Z., L. Shiue, et al. (1992). "A mammalian factor that binds telomeric TTAGGG repeats in vitro." Mol Cell Biol 12(11): 4834-43.

Zinsser, H. (1906). "A Simple Method for the Plating of Anaerobic Organisms." J Exp Med 8(4): 542-6.

Curriculum Vitae

Cian Stutz, PhD

Rue Daniel-Jeanrichard 43
CH-2300 La Chaux-de-Fonds
Tel.: +41 79 261 07 87
E-Mail: cian.stutz@gmail.com

Personal details

Date and place of birth: December 16, 1983 in Basel, Switzerland
Nationality: Swiss, Irish
Marital status: married

Education and Work Experience

Oct 2012-current	Research Officer in Antibody Engineering at Glenmark Pharmaceuticals in La Chaux-de-Fonds, Switzerland. Tasks: Biophysical and biochemical characterization as well as structure-guided optimization of novel therapeutic antibodies. Humanization of mouse antibodies.
13. March 2013	PhD degree in Structural Biology, University of Basel / Friedrich Miescher Institute for Biomedical Research in Basel, Switzerland, in the group of Dr. Nicolas Thomä (with Prof. Dr. Susan Gasser, Prof. Dr. Marc Bühler and Prof. Dr. Peter Baumann as members of my thesis committee).
2008-2012	PhD Student, Friedrich Miescher Institute for Biomedical Research of NIBR Novartis in Basel in the group of Dr. Nicolas Thomä Project: Towards the Crystal Structure of the *S. pombe* Shelterin Complex Tasks: Cloning, expression, purification, crystallization and structure determination of the proteins in the Shelterin complex as well as the biochemical and functional characterization thereof.
2006-2008	Master in Structural Biology, Biozentrum, University of Basel in the group of Prof. Dr. Olga Mayans Project: Evolution of enzymes in the Tryptophan biosynthesis pathway Tasks: Crystallization and 3D structure determination of a protein which evolved by a gene duplication event and acquired an unknown function
2003-2006	Biology studies, Biozentrum, University of Basel Graduation: Bachelor in Molecular Biology
1997-2002	Gymnasium Kirschgarten, Basel Graduation: Matura

Method knowledge

Molecular biology

PCR
Primer design
Construction of plasmids
Sequencing

Mutagenesis (Point mutations, deletions, fusions)
Transformation
Transfection (Baculovirus for insect cell expression, Mammalian cells)
Yeast (*S. pombe*) *in vivo* genome manipulation (deletion, integration and modification)

Protein chemistry

Expression of proteins:	Bacteria:	*E. coli*
	Insect cells:	*Sf9, Hi-5*
	Mammalian cells:	HEK

Purification of proteins:
Affinity-chromatography (His-tag, GST, Strep-tag, Ig affinity)
Size exclusion chromatography
Ionic exchange chromatography
Hydrophobic interaction chromatography
FPLC (ÄKTA), HPLC, HPLC-MS

Biophysical methods

UV-VIS
Native- and SDS-PAGE
Western-Blot analysis
SEC-MALS
DLS
LC-MS
SPR (Biacore)
Differential Scanning Calorimetry
Microscale Thermophoresis (Nanotemper)

Structural biology

Crystallization:
Vapor-diffusion: sitting and hanging drop
Crystallization robotics (Phenix)
RockImager for automated crystallization tray management

Crystallography:
Data collection (In-house and Synchrotron)
Processing of diffraction data
Structure determination and refinement
Virtual Ligand-Receptor docking using various Ligand-Docking software
(based on volume complementarity and energy minimization)

Qualifications

Languages:
German (native language), English (native language), French (intermediate knowledge),
Russian (beginner)

Computer skills:
Microsoft Windows, Linux, OSX, MSOffice, Photoshop, Illustrator, Maple
VectorNTI, CLC Workbench and servers for protein and DNA sequence analysis, Software packages for structural biology

Publications

Cian Stutz, Patrick Babinger, Darko Skegro, Olga Mayans. "Crystal structure of YbiB – Insight into the evolution of novel protein functions".
(unpublished)

Cian Stutz, Daniel Hess, Heinz Gut, Peter Baumann, Nicolas Thomä. "The Crystal structure of the *S. pombe* Shelterin components Poz1-Tpz1".
(in preparation)

References

Dr. Darko Skegro, Team leader, Head of Bispecific Antibodies Platform, Glenmark Pharmaceuticals, La Chaux-de-Fonds, Schweiz
Telefon: ++41 76 732 32 98, darkos@glenmarkpharma.com

Dr. Nicolas Thomä, Friedrich Miescher Institute for Biomedical Research, Basel, Schweiz
Telefon: ++41 61 697 86 30, nicolas.thoma@fmi.ch

Prof. Dr. Olga Mayans, Institute of Integrative Biology, University of Liverpool, United Kingdom
Telefon: ++44 151 795 4472, olga.mayans@liverpool.ac.uk