Published in final edited form as:

Methods Enzymol. 2016 ; 573: 403–419. doi:10.1016/bs.mie.2016.04.006.

Crystallographic Studies of Telomerase

H. Hoffman and E. Skordalakes¹

The Wistar Institute, Philadelphia, PA, United States

Abstract

Telomeres are nucleoprotein complexes that maintain the ends of our chromosomes thus providing genomic stability. Telomerase is a ribonucleoprotein reverse transcriptase that replicates the short tandem repeats of DNA known as telomeres. The telomeric DNA is specifically associated with two major complexes, the shelterin and CST complexes both of which are involved in telomere length regulation and maintenance along with telomerase. Obtaining structural information on these nucleoprotein complexes has been a major bottleneck in fully understanding the mechanism of action of telomeric nucleoproteins for over two decades. The recent advances in molecular and structural biology have enabled us to obtain atomic resolution structures of telomeric proteins alone and in complex with their nucleic acid substrates transforming the field and our understanding and interpretation of this unique biological pathway. Here we report our approach to obtain the structure of the *Triobolium castaneum* catalytic subunit of telomerase TERT (tcTERT) in its apo- and substrate-bound states.

1. INTRODUCTION

Telomerase replicates the ends of our chromosomes by adding multiple identical repeats of DNA known as telomeres. Even though telomeres are noncoding parts of the chromosome, they are critically important for the integrity of our genome as indicated by a host of diseases associated with dysfunctional telomerase, telomerase deregulation, and short telomeres. Some of the most common maladies associated with dysfunctional telomeric complexes are cancer, aplastic anemia, and dyskeratosis congenital, just to mention a few (Calado & Young, 2009; Kim et al., 1994; Vulliamy et al., 2002). In fact, telomerase is known to be active in approximately 85% of human cancers where the continuous replication of the chromosome ends is required, a prerequisite for cell proliferation as it allows the cells to overcome their Hayflick limit and to continue dividing indefinitely (Hayflick, 1965; Kim et al., 1994).

The core telomerase holoenzyme is a ribonucleoprotein complex, consisting of a protein subunit (telomerase reverse transcriptase, TERT) that catalyzes the reaction of nucleotide addition at the ends of chromosomes via phosphodiester bond formation. TERT belongs to the family of reverse transcriptases and consists of four or five domains depending on species. These include the TEN (telomerase essential N-terminal), TRBD (telomerase RNA binding domain), fingers, palm, and thumb domains (reverse transcriptase nomenclature)

¹Corresponding author: skorda@wistar.org.

(Gillis, Schuller, & Skordalakes, 2008). The TRBD, fingers, palm, and thumb domains are universally conserved and essential for telomerase function, while the TEN domain, which comprises the N-terminal portion of the protein is absent in insects and possibly in worms (Fig. 1A) (Osanai et al., 2006). The TEN domain contributes to telomerase recruitment to telomeres but does not affect telomerase activity (Schmidt, Dalby, & Cech, 2014). One possibility is that the TEN domain in the later organisms exists as an independent polypeptide. This hypothesis is supported by data showing that the fully active enzyme can be reconstituted by transiently adding the TEN domain to the TEN-truncated telomerase in vitro (Robart & Collins, 2011).

The telomerase RNA (TER) is an integral component of the holoenzyme and carries the template TERT uses during telomere replication (Fig. 1B). The TER template usually consists of 1.5 telomeric repeats (Lee & Blackburn, 1993; Lingner, Hendrick, & Cech, 1994; Shippen-Lentz & Blackburn, 1990). For example, in humans the telomeric repeat is TTAGGG with the RNA template being CUAACCCU (Lee & Blackburn, 1993; Lingner et al., 1994; Shippen-Lentz & Blackburn, 1990). Interestingly, while TERT is conserved across species, TER is vastly divergent. One of the smallest known TERs belongs to the ciliate *Tetrahymena paravorax* and is 147 nucleotide long (Ye & Romero, 2002), while the fungi TERs are the largest known with *Mycosphaerella graminicola* (2425 dNTPs) (Qi et al., 2013) being the largest identified to date (http://telomerase.asu.edu/sequences_tr.html). To add to the complexity of this system, the proper telomerase holoenzyme assembly requires the enzymatic contribution of a number of proteins with most of them acting on the RNA component of the enzyme. In vertebrates, the enzyme dyskerin modifies the 3'-end of TER, a process required for its stability and proper assembly of telomerase (Mitchell, Wood, & Collins, 1999; Shay & Wright, 1999). In ciliates, the enzyme p65 is required to bind and fold the RNA for proper TERT, TER assembly (Witkin & Collins, 2004). The complexity of the holoenzyme is demonstrated in the recent EM structure of *Tetrahymena thermophila* telomerase (Jiang et al., 2013).

Replicating telomeres is only part of chromosome end protection and overcoming the Hayflick limit. Chromosome end maintenance is carried out by telomere binding complexes such as CST and shelterin (de Lange, 2005; Grandin, Reed, & Charbonneau, 1997; Grandin, Damon, & Charbonneau, 2000, 2001; Lin & Zakian, 1996; Miyake et al., 2009). Both are essential complexes as the deletion of any of their components has deleterious effects for cells, usually leading to telomere length deregulation, chromosome end uncapping, and genomic instability associated with senescence, apoptosis, and in some cases cell immortalization and carcinogenesis. Here, we will focus primarily on the CST named after the three proteins that constitute this complex (cell division cycle 13 (Cdc13)/Ctc1, suppressor of cdc thirteen (Stn1), and Ten1). The CST complex specifically binds the telomeric overhang, and although it is commonly known as a telomere-capping complex, its function at the ends of chromosomes is far more complex. It was believed that this complex was unique to yeast until 2013 when the Ishikawa group identified the human CST complex, which was followed by the discovery of CST in plants, and most recently in ciliates (Lin & Zakian, 1996; Miyake et al., 2009; Wan et al., 2015).

Although the degree of functional conservation of the CST complex across species is currently unclear, there are a few things one should take into consideration when thinking of this nucleoprotein assembly. The trimeric CST complex binds single-stranded telomeric DNA (ssDNA) with high affinity and specificity, and caps the ends of chromosomes. The Stn1 and Ten1 subcomplex structurally resembles the replication protein A complex (containing 32 and 14 kDa subunits), binds DNA nonspecifically, and localizes to various parts of the chromosome where it regulates DNA replication (Bryan et al., 2013; Grandin et al., 2001, 1997; Sun et al., 2009). Yeast (*Saccharomyces cerevisiae*) CST recruits telomerase to telomeres via binding to the Est1 protein (Pennock, Buckley, & Lundblad, 2001), a process that most likely requires significant structural rearrangement of the trimeric complex (Mason & Skordalakes, 2010). In contrast, current data suggest that the human CST complex, and in particular Ctc1, interacts with TPP1, a component of the shelterin complex, which recruits telomerase to telomeres and downregulates telomere replication (Chen, Redon, & Lingner, 2012; Tejera et al., 2010; Wang et al., 2007).

At this point, it is well established that telomeres are replicated and maintained by large nucleoprotein complexes consisting of multiple domains involved in protein–protein and nucleic acid binding. As previously stated, TERT consists of four or five domains containing a large integral RNA, which, as we now know, makes extensive contacts with almost all domains of TERT. Cdc13, the core subunit of the yeast CST complex, consists of five domains, each of which contributes to the function of this complex. Four of these domains are OB (oligonucleotide/oligosaccharide binding) folds with some being involved in Cdc13 dimerization while others are important for Cdc13 ssDNA binding (Bourne et al., 1998; Greetham et al., 2015; Mason & Skordalakes, 2010; Mason et al., 2013; Mitchell et al., 2010a; Nugent et al., 1996; Qi & Zakian, 2000). With this wide array of intricate protein complexes serving multiple essential functions, it is not surprising that structural data were limited to only a domain of Cdc13 (DNA-binding domain, DBD) and the TEN domain of telomerase until the mid-2000s (Jacobs, Podell, & Cech, 2006; Mitton-Fry et al., 2004).

Our understanding of telomeric complexes has been greatly enhanced by the recent structural breakthroughs in the field. In this chapter, we examine strategies utilized in recent years to obtain high-resolution structures of telomerase, breakthroughs that have transformed not only the telomere field and our understanding of telomere biology but also its implications in human disease.

2. CRYSTAL STRUCTURE DETERMINATION OF THE CATALYTIC SUBUNIT OF *TRIOBOLIUM CASTANEUM* TELOMERASE (tcTERT)

2.1 tcTERT Protein Isolation

Considering the size and complexity of telomerase, it is not surprising that efforts to obtain high-resolution atomic data on this enzyme have been met with limited success. NMR structures of fragments of the RNA component of telomerase were accessible early on, mostly because of their small size. One can still glean much information from individual domains, especially when combined with biochemical studies. The concerted effort from a confluence of laboratories resulted in the high-resolution structures of several fragments of

the RNA component of telomerase such as the pseudoknot, the TBE, and the activation domain of TER (Theimer & Feigon, 2006). In the case of TERT, the structures of the TEN and TRBD domains became available in 2006 and 2007, respectively (Jacobs et al., 2006; Rouda & Skordalakes, 2007).

The catalytic subunit of telomerase, TERT, posed a different challenge. A major limitation to obtaining structural information on TERT was associated with the lack of accessible active, soluble protein in sufficient quantities for crystallographic studies. Efforts to obtain soluble TERT by a number of prominent labs were met with limited success. These issues were no exception to our lab. Initially, we tested TERT genes from a variety of organisms, primarily focusing on vertebrate, fungi, and ciliates. These groups of organisms had been studied extensively, and hence there was a wealth of biochemical data to draw from when studying these systems. Unfortunately, this approach did not yield any suitable protein for further studies. When expression was possible, the protein was either degraded or insoluble. At this point, we decided that coexpression of TERT with its integral RNA component TER may address this problem. We tried this approach both in *Escherichia coli* and in baculovirus, but the results were no better than with TERT alone. At this point it was clear that we needed to think outside of the box and go after TERT genes from organisms that had not been looked at previously; a comprehensive list of TERT genes can be found in the online available telomerase database (<http://telomerase.asu.edu>). Although there are significant evolutionary differences among telomerases from different species (Fig. 1), the core mechanism of telomere replication appears to be highly conserved and as such one can still garner a wealth of information from their structural studies. A survey of the telomerase database revealed that the size of the catalytic subunit of telomerase varies considerably in size. TERTs in birds are in the range of 150–160 kDa, in yeasts, around 90–110 kDa, and in worms and insects around 70–80 kDa, while human TERT is 127 kDa. Identifying a TERT gene that lacks all the nonconserved insertion loops could be the answer to addressing this difficult crystallographic problem. With this notion in mind, we decided to test TERT of some of the smaller genes, including the insect *T. castaneum* (70 kDa) and the nematode *Caenorhabditis elegans* (66 kDa).

The next step was to identify the correct expression system and conditions so that we could obtain the protein in large, stable, pure, and active quantities required for crystallographic studies. Expressing proteins in *E. coli* is inexpensive and efficient, plus it eliminates all of the posttranslational modifications that could interfere with structural studies as they can be quite heterogeneous. However, expressing eukaryotic proteins in *E. coli* can be challenging, and therefore the baculovirus system provides a good alternative if necessary. To maximize our yields and quality of protein obtained, we screened protein expression in a range of *E. coli* cell lines (BL21 (pLysS), RIPL, Rosetta (DE3) pLysS, *E. coli* ScarabXpress-1 (T7Lac) cells, and SHuffle T7 Competent *E. coli*), all of which are commercially available (EMD Millipore (Novagen), Billerica, MA, Scarab Genomics, Madison, WI, and New England BioLabs Inc., Ipswich, MA). Many of these cell lines have been modified to be compatible with eukaryotic protein expression by the introduction of rare codons required for mRNA translation. Alternatively, one can obtain the codon-optimized gene that is compatible with the prokaryotic protein expression system being used, although this can be rather expensive, especially when dealing with large genes such as TERT, which spans the length of 1800–

3500 nucleotides, depending on organism. For tcTERT overexpression studies, we used the synthetic gene (GeneArt and ThermoFisher Scientific) codon optimized for *E. coli* overexpression cloned into a modified version of the pET28b vector containing a cleavable hexahistidine tag at its N-terminus. tcTERT protein overexpression tests showed that the *E. coli* Rosetta (DE3) pLysS (EMD Millipore (Novagen)) was the ideal cell line for obtaining this protein in large quantities required for structural studies.

Identifying the appropriate cell line is important as it is essential to identify the correct expression conditions, which will define the rate of cell growth and protein expression. Frequently, overexpression of proteins in *E. coli* or baculovirus leads to protein misfolding, which in turn forces the protein into inclusion bodies or to be degraded by the cell. One can improve the protein quality and eliminate aggregation by slowing down the rate and time at which the protein is expressed. One way to achieve this is by varying the temperature at which the cells are grown and the protein expressed. Using temperatures of 16–20°C is common practice, although temperatures ranging from 37°C to 4°C have been used successfully. For tcTERT, constant temperatures in the range of 37–16°C produced mostly small amounts of protein that allowed for a limited number of crystal studies. However, this procedure was expensive and cumbersome so optimization was essential. Interestingly, our over-expression tests indicated that a slow transition in temperature (from high to low) after cell induction with IPTG showed promise worth pursuing. With this in mind, we screened several temperature ranges (37–30°C, 37–20°C, 37–16°C, and 30–20°C). Through this process, we found that a drop in temperature to 30°C from 37°C when the cell density (OD) was 0.4, followed by cell induction when the OD reached 0.6, was ideal for obtaining maximum yields of protein. The period of cell induction/protein expression is also critical. At 37–30°C, 2–4 h overexpressions are commonly used, and this was also true in the case of tcTERT.

The purification of the tcTERT was a three-step process. The protein was overexpressed with a hexahistidine tag at its N-terminus, so we used a Ni-NTA resin for the initial purification of the protein from cell lysate, which resulted in 90% pure protein. However, for crystallographic studies one needs a highly pure and monodisperse protein, so this step alone was not sufficient. At this point, we took advantage of the high pI of the protein (9.37; Wilkins et al., 1999) as it allows for tight binding to cation exchange resin, such as HS poros (Applied Biosystems). Loading the tcTERT to the resin at high salt concentration (500 mM KCl) allows for contaminant proteins coming from the host cell (*E. coli* Rosetta (DE3) pLysS (EMD Millipore (Novagen))) to flow through the column while tcTERT binds to the resin tightly. Following loading of the protein to the resin, the column was washed thoroughly (~100 mL) with 50% buffer B (25 mM Tris-HCl, pH 7.5, 1 M KCl, 1 mM DTT, 5% glycerol) to remove any contaminant proteins still associated with the sample. The sample was then eluted off the HS poros using a salt gradient (50–100% buffer B, which translates to 500 mM–1 M KCl), and the fractions were run on a gel to determine the purity of the sample. The fractions, clean of any contaminants, were pulled together and concentrated to less than 1 mL for the final purification step involving size exclusion chromatography (Sephedex S200). Size exclusion chromatography is an essential step for crystallographic studies as it allows for the separation of protein aggregates generated primarily due to the over-expression of the protein and which can significantly hinder

protein crystallization. We eluted the protein from the S200 column in 25 mM Tris-HCl, 500 mM KCl, 1 mM MTCEP, and 5% glycerol, pH 7.5, which we refer to as “storage buffer.” We found that most of the proteins we work with are relatively stable under these buffer conditions for a few days after purification. At this point, the protein (tcTERT) was >99% pure and therefore ready to be used for crystallographic studies.

2.2 tcTERT Crystallization

For crystallographic studies, there are several points of concern one should consider prior to setting trays. The protein concentration and pI are critical for identifying one or more crystallization conditions. The temperature at which the crystallization is carried out can also be significant as it can affect the stability but most importantly the solubility of the protein. The concentration of the protein should also be taken into account, and our experience indicates that the smaller the protein, the higher the concentration required for crystallization, primarily due to crowding reasons. However, there is no hard rule as to the concentration of the protein sample since the protein solubility factors into this process. Our experience has shown that for large proteins such as ~300 kDa, even 1 mg/mL protein is sufficient in some cases to obtain crystals.

Crystal screening is usually done using a range of sparse matrix conditions commercially available from a number of sources such as Hampton Research (<https://www.hamptonresearch.com>). The goal of this step is to screen a wide range of precipitants, salts, buffers, pHs, and additives that when added together facilitate the orderly arrangement of the protein molecules leading to a diffracting crystal. For this reason and for the sparse matrix conditions to be effective, the high concentration of buffer, salt, and glycerol of the “storage buffer” needs to be replaced with much lower concentration of buffer, salt, and in most cases glycerol must be removed. The gentlest way to exchange existing buffer conditions of a macromolecule is through dialysis. Dialysis allows for the gradual exchange of buffers, thus providing a window of time for the protein to adjust to the new buffer conditions. There are several devices available for this purpose; we used Slide-A-Lyzer MINI Dialysis Units from ThermoFisher Scientific. We dialyzed the stock tcTERT into 5 mM Tris-HCl, 100 mM KCl, 1 mM MTCEP (pH 7.5) prior to crystallization trials. Crystallization trials were carried out at room temperature and 4°C. As mentioned previously, temperature is an important factor as it can affect the solubility and therefore the stability and rate of crystallization of the protein. We found that tcTERT is far more soluble at 4°C than at room temperature, which resulted in a large number of clear crystallization drops even at high concentrations of protein that ultimately did not produce any useful crystals. We addressed this issue by performing all additional crystallization trials at room temperature. However, neither temperature produced any crystal hits. At this point, we decided that cocrystallization of the tcTERT with a natural substrate such as ssDNA or RNA may facilitate crystallization. We prepared the tcTERT-ssDNA sample by mixing the dialyzed (5 mM Tris-HCl, 100 mM KCl, 1 mM MTCEP (pH 7.5)) tcTERT protein with 1.2-fold excess ssDNA. The ssDNA consisting of two or three telomeric repeats was purchased from Integrated DNA Technologies (IDT) and prepared in 5 mM Tris-HCl and 50 mM KCl, pH 7.5. Mixing the dialyzed tcTERT resulted in heavy precipitation of the protein as indicated by a thick white solution; it was unclear at this point, if the protein was useful for

crystallization but we proceeded anyway. Interestingly, we found that mixing the protein with various sparse matrix crystallization conditions resulted in resolubilization of the protein and clear crystallization drops. Our approach subsequently resulted in two crystal hits (a) 100 mM Tris pH 8.5, 1.5 M NaNO₃ and (b) 100 mM Tris (pH 8.0) and 2 M (NH₄)₂SO₄.

Crystal optimization is standard and is usually achieved with adjustment of the concentration and pH of the reagents that produced the crystal hit, as well as the addition of a wide range of additives or detergents. Additive and detergent screens, sampling a wide range of metal ions, small organics, and detergents are commercially available and used widely by the crystallographic community. Crystals grown in 100 mM Tris, pH 8.5, and 1.5 M NaNO₃ were small and diffracted to 4 Å resolution at best, so improvement of crystal quality for better resolution was essential. Using the current crystal growth conditions as a guide, we screened a wide range of buffers, pHs, salts, and precipitant (NaNO₃) concentrations. Through this process we found that the best quality crystals, diffracting to better than 3 Å, grew in 50 mM HEPES (pH 7.0), 1.2–1.3 M NaNO₃ and 1 mM TCEP at room temperature. Crystals grown in the (NH₄)₂SO₄ condition did not diffract better than 3.25 Å even after extensive optimization, so we focused solely on the NaNO₃-dependent crystals of tcTERT plus ssDNA.

2.3 tcTERT Structure Determination

Data collection was carried out at NSLS, beam line X6A. The best crystals grown in NaNO₃, diffracted to 2.71 Å resolution, belonged to the orthorhombic space group $P2_12_12_1$ and contained two molecules in the asymmetric unit. Phasing the structure proved challenging. We initially tried preparing the selenomethionine protein for phasing, but the protein expression levels from minimal media growths were disappointingly low. At this point, we decided that metal derivatization of the crystals may be a good alternative. This can be done either by use of a covalent binder, such as a mercury compound that will coordinate free cysteines or by a noncovalent binder, such as NaBr. We found that soaking the crystals with high concentrations of NaBr caused significant damage to the crystals and impaired diffraction. At this point, we took a step back, examined the protein sequence, found that tcTERT consists of 13 cysteines, and therefore the use of mercury derivatization was a possible option. Soaking the crystals with methyl mercury(II) chloride CH₃HgCl or HgCl₂ solution (1–10 mM) even for a few minutes resulted in poorly diffracting crystals suggesting that a gentler approach was needed for effective derivatization. To address this issue, we decided to try the following two approaches: the first involved the addition of a very small amount of solid CH₃HgCl into the crystallization drop containing the crystals of interest. The idea behind this approach was to allow for the relatively insoluble mercury to slowly seep into the crystals, thus allowing for slow and specific derivatization of solvent-accessible cysteines without damaging the crystal lattice. The second approach involved the addition of 5 mM MeHgCl₂ solution incrementally (0.5 mM every 5 min). We found that the second method of slowly adding CH₃HgCl to the crystals resulted in derivatized crystals that diffracted to 3.5 Å resolution.

We calculated phases using the method of single isomorphous replacement with anomalous signal (SIRAS). More specifically, we used one native and two isomorphous CH₃HgCl derivative datasets collected at wavelengths Hg1—1.00850 Å and Hg2—1.00800 Å. We first used SOLVE (<https://solve.lanl.gov>) (Terwilliger, 2000; Terwilliger & Berendzen, 1999) to identify the high occupancy sites (12 heavy atom sites were located), which we subsequently refined and used to calculate new phases with MLPHARE (Collaborative Computational Project No. 4, 1994). We then used the MLPHARE improved phases to calculate an anomalous difference map, which allowed us to identify the remaining heavy atom sites (22 in total) to 3.5 Å resolution. MLPHARE phases obtained using all the heavy atom sites were then used in DM with twofold NCS and phase extension using the high-resolution native data (2.71 Å) to calculate starting experimental maps and build the model.

The structure revealed a ring configuration that closely resembles the polymerase domain of HIV RTs and viral RNA polymerases (Gillis et al., 2008). A striking difference between TERT and RTs as well as viral RNA polymerases is the presence of the N-terminal extension, which comprises the TRBD and is unique to telomerase. The structure reveals that this domain adopts an almost all-helical structure that makes extensive interactions with the C-terminal (thumb domain) extension, which facilitates the formation of the closed ring configuration observed in the TERT structure. Subsequent EM studies of the *T. thermophila* and human TERT (Jiang et al., 2013; Sauerwald et al., 2013) showed that the structure of TERT is conserved across species.

3. CRYSTAL STRUCTURE DETERMINATION OF A PARTIAL TELOMERASE ELONGATION COMPLEX

3.1 Nucleic Acid Substrate Design

We then asked how TERT assembles with TER and the telomeric overhang to form an active elongation complex. While we have the tcTERT protein in large quantities, the RNA component (tcTER) of *T. castaneum* telomerase is still unknown. For this reason, we decided to determine the structure of tcTERT in complex with the RNA template and telomeric DNA. As we mentioned earlier, the TER template is usually 1.5 telomeric repeats. In the case of *T. castaneum* telomerase, the telomeric repeat is TCAGG, so the putative RNA template should be (5'-rCrUrGrArCrCrU-3') (Mitchell et al., 2010b). We designed the RNA–DNA hairpin so that the RNA and DNA strands are linked together for stability with a stable tetraloop (TTCG). We also designed the hairpin to contain a three-nucleotide overhang (5'-rCrUrGrArCrCrGTTCGAGGT-3') so that we can (a) test the enzymatic activity of tcTERT in the presence of the hairpin and (b) trap the enzyme in its catalytic state by cocrystallizing with dNTPs and Mg²⁺ ions.

3.2 Activity Assays

To ensure tcTERT assembles with the designed RNA–DNA hairpin in a productive manner, we tested its activity in the presence of the hairpin. Since the hairpin does not contain the full-length TER required for repeat addition processivity, we only expected the addition of three nucleotides to the end of the telomeric DNA associated with the hairpin. We carried out a standard reverse transcriptase assay in a buffer containing 50 mM Tris–HCl, 100 mM

KCl, 1.25 mM MgCl₂, 5 mM DTT, 5% (w/v) glycerol, pH 8, 100 μM dNTPs (dATP, dTTP, and dGTP), 10 μM [³²P]dCTP (80 Ci/mmol), 5 μM RNA–DNA hairpin, and 1 μM recombinant TERT. The results clearly indicated the enzyme is indeed active under these conditions and allowed us to move forward in attempting to capture its catalytic state.

3.3 Complex Crystallization

Trapping telomerase in its catalytic state can be challenging as the active enzyme can hydrolyze the dNTPs during either the sample preparation or crystallization. To avoid this issue, we purchased nonhydrolyzable dNTPs (dNTPαS—Jenna Biosciences) and prepared the sample by mixing the protein–RNA–DNA hairpin at a 1:1.2 ratio in the presence of 5 mM dNTPαS and 5 mM MgCl₂.

Crystallization of tcTERT in the presence 5′-rCrUrGrArCrCrUTTCGAGGT-3′, dNTPαS, and MgCl₂²⁺ was carried out at room temperature and 4°C using the commercially available sparse matrix conditions we used for the apo-tcTERT. Co-crystallization of tcTERT with the above hairpin, dNTPαS and Mg²⁺ did not produce any diffracting crystals, so we decided to modify the length of the linker that connects the RNA template with the telomeric DNA to facilitate crystallization. To this end, we designed two additional hairpins consisting of the RNA template (5′-rCrUrGrArCrCrU-3′), the complementary telomeric DNA (5′-AGGT-3′), the tetraloop (5′-TTCG-3′), and two or six additional nucleotides (Fig. 2). Of all three hairpins only the one with the additional six nucleotides (highlighted in bold-italic, 5′-rCrUrGrArCrCrGrGrArCTTCGGTCAGGT-3′) produced crystals amenable to structural studies. Interestingly, several crystal hits appeared within a week, however, all of them with the exception of one gave the rod-like crystals we observed previously with the tcTERT protein alone, suggesting that these crystals did not contain nucleic acid. We confirmed the absence of RNA–DNA in these crystals by dissolving them and running them on an agarose gel in the presence of ethidium bromide. Inspection of the crystallization conditions revealed that crystals that did not contain the hairpin had high salt concentration as a precipitant, which suggested that the contacts between the protein and the nucleic acid were mostly hydrophilic in nature and involved the backbone of the RNA–DNA hairpin.

Crystals containing the hairpin grew in 0.1 M HEPES (pH 7.5), 0.2 M KCl, and 12% 1,6-hexanediol or PEG 4K. Under these crystallization conditions, small crystals appeared within a couple of days and grew to useful diffracting size (200 × 50 × 20 μm) within 2 weeks. For diffraction studies, we transferred these crystals into cryoprotectant containing 0.1 M HEPES (pH 7.5), 15% (w/v) 1,6-hexanediol or PEG 4K, 15% (w/v) glycerol, 0.2 M KCl, and 1 mM TCEP and harvested them by flash-freezing in liquid nitrogen. The crystals ultimately diffracted to 2.7 Å resolution and belong to the monoclinic space group *P*2₁.

3.4 Structure Determination

We determined the structure using the method of molecular replacement and the apotcTERT (PDB ID: 3DU6) structure as a search model. The structure revealed one molecule in the asymmetric unit with clear density for the nucleic acid in the interior cavity of the TERT ring. Close inspection of the nucleic acid density revealed three additional nucleotides at the 3′-end of the telomeric DNA (Fig. 3), which suggested that the active tcTERT had extended

the DNA in the crystallization drop. The nucleotide located at the 3'-end of the DNA occupied the active site of the enzyme thus providing partial evidence of the mechanism of dNTP addition by telomerase to the end of our chromosomes.

This structure revealed that during telomere elongation TERT binds the RNA template and telomeric DNA in a similar fashion to that of HIV RTs revealing new insights into the mechanism of telomere elongation by telomerase.

4. CONCLUSIONS

The main obstacle to obtaining the high-resolution structure of telomerase has been the isolation of the active TERT in large quantities required for crystallographic studies. Screening a large number of TERT genes from different organisms paid off in identifying a TERT gene, which lacked most of the nonconserved regions of the protein. However, several other factors contributed to the successful completion of this project. We found that using a synthetic TERT gene, codon-optimized for bacterial overexpression and identifying the appropriate cell strain for protein overexpression were essential to the overall success of this project.

Another finding of note was that using fresh protein and avoiding flash-frozen samples was critical for effective crystallization with this being especially true for the protein and RNA samples, while DNA samples appeared relatively stable. Not only was the design of a substrate important, but an understanding of the biochemistry of the assembly in general is indispensable to the success of any project.

These experiments provide a view of the catalytic subunit of telomerase alone and in complex with its RNA template and telomeric DNA. Telomerase has evolved over time, which explains the diversity observed both in the protein and the RNA component of telomerase across species. It is clear that the core TERT enzyme is highly conserved across species as are the RNA motifs implicated in telomerase-dependent telomere replication. The information obtained from these studies has been instrumental in the interpretation and design of many experiments and has shaped the way we view telomerase in the field.

Acknowledgments

The research was funded by the NIGMS (5 R01 GM088332) and The Wistar Cancer Center Support Grant (P30 CA10815).

ABBREVIATIONS

Cdc13	cell division cycle 13
Stn1	suppressor of cdc thirteen
Ten1	telomeric pathways in association with Stn1 number 1
TER	telomerase RNA
TERT	telomerase reverse transcriptase

References

- Bourns BD, et al. Sir proteins, Rif proteins, and Cdc13p bind *Saccharomyces* telomeres in vivo. *Molecular and Cellular Biology*. 1998; 18(9):5600–5608. [PubMed: 9710643]
- Bryan C, et al. Structure of the human telomeric Stn1-Ten1 capping complex. *PloS One*. 2013; 8(6):e66756. [PubMed: 23826127]
- Calado RT, Young NS. Telomere diseases. *The New England Journal of Medicine*. 2009; 361(24):2353–2365. [PubMed: 20007561]
- Chen LY, Redon S, Lingner J. The human CST complex is a terminator of telomerase activity. *Nature*. 2012; 488:540–544. [PubMed: 22763445]
- Collaborative Computational Project No. 4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallographica. Section D, Biological Crystallography*. 1994; 50(Pt. 5):760–763. [PubMed: 15299374]
- de Lange T. Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes & Development*. 2005; 19(18):2100–2110. [PubMed: 16166375]
- Gillis AJ, Schuller AP, Skordalakes E. Structure of the *Tribolium castaneum* telomerase catalytic subunit TERT. *Nature*. 2008; 455(7213):633–637. [PubMed: 18758444]
- Grandin N, Damon C, Charbonneau M. Cdc13 cooperates with the yeast Ku proteins and Stn1 to regulate telomerase recruitment. *Molecular and Cellular Biology*. 2000; 20(22):8397–8408. [PubMed: 11046137]
- Grandin N, Damon C, Charbonneau M. Ten1 functions in telomere end protection and length regulation in association with Stn1 and Cdc13. *The EMBO Journal*. 2001; 20(5):1173–1183. [PubMed: 11230140]
- Grandin N, Reed SI, Charbonneau M. Stn1, a new *Saccharomyces cerevisiae* protein, is implicated in telomere size regulation in association with Cdc13. *Genes & Development*. 1997; 11(4):512–527. [PubMed: 9042864]
- Greetham M, et al. The telomere binding protein Cdc13 and the single-stranded DNA binding protein RPA protect telomeric DNA from resection by exonucleases. *Journal of Molecular Biology*. 2015; 427:3023–3030. [PubMed: 26264873]
- Hayflick L. The limited in vitro lifetime of human diploid cell strains. *Experimental Cell Research*. 1965; 37:614–636. [PubMed: 14315085]
- Jacobs SA, Podell ER, Cech TR. Crystal structure of the essential N-terminal domain of telomerase reverse transcriptase. *Nature Structural & Molecular Biology*. 2006; 13(3):218–225.
- Jiang J, et al. The architecture of Tetrahymena telomerase holoenzyme. *Nature*. 2013; 496(7444):187–192. [PubMed: 23552895]
- Kim NW, et al. Specific association of human telomerase activity with immortal cells and cancer. *Science*. 1994; 266(5193):2011–2015. [PubMed: 7605428]
- Lee MS, Blackburn EH. Sequence-specific DNA primer effects on telomerase polymerization activity. *Molecular and Cellular Biology*. 1993; 13(10):6586–6599. [PubMed: 8413255]
- Lin JJ, Zakian VA. The *Saccharomyces* CDC13 protein is a single-strand TG1-3 telomeric DNA-binding protein in vitro that affects telomere behavior in vivo. *Proceedings of the National Academy of Sciences of the United States of America*. 1996; 93(24):13760–13765. [PubMed: 8943008]
- Lingner J, Hendrick LL, Cech TR. Telomerase RNAs of different ciliates have a common secondary structure and a permuted template. *Genes & Development*. 1994; 8(16):1984–1998. [PubMed: 7958872]
- Mason M, Skordalakes E. Insights into Cdc13 dependent telomere length regulation. *Aging (Albany NY)*. 2010; 2(10):731–734. [PubMed: 20975208]
- Mason M, et al. Cdc13 OB2 dimerization required for productive Stn1 binding and efficient telomere maintenance. *Structure*. 2013; 21(1):109–120. [PubMed: 23177925]
- Mitchell JR, Wood E, Collins K. A telomerase component is defective in the human disease dyskeratosis congenita. *Nature*. 1999; 402(6761):551–555. [PubMed: 10591218]

- Mitchell M, et al. Cdc13 N-terminal dimerization, DNA binding, and telomere length regulation. *Molecular and Cellular Biology*. 2010a; 30(22):5325–5334. [PubMed: 20837709]
- Mitchell M, et al. Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nature Structural & Molecular Biology*. 2010b; 17(4):513–518.
- Mitton-Fry RM, et al. Structural basis for telomeric single-stranded DNA recognition by yeast Cdc13. *Journal of Molecular Biology*. 2004; 338(2):241–255. [PubMed: 15066429]
- Miyake Y, et al. RPA-like mammalian Ctc1-Stn1-Ten1 complex binds to single-stranded DNA and protects telomeres independently of the Pot1 pathway. *Molecular Cell*. 2009; 36(2):193–206. [PubMed: 19854130]
- Nugent CI, et al. Cdc13p: A single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science*. 1996; 274(5285):249–252. [PubMed: 8824190]
- Osanai M, et al. Identification and characterization of the telomerase reverse transcriptase of *Bombyx mori* (silkworm) and *Tribolium castaneum* (flour beetle). *Gene*. 2006; 376(2):281–289. [PubMed: 16793225]
- Pennock E, Buckley K, Lundblad V. Cdc13 delivers separate complexes to the telomere for end protection and replication. *Cell*. 2001; 104(3):387–396. [PubMed: 11239396]
- Qi H, Zakian V. The *Saccharomyces* telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase alpha and the telomerase-associated est1 protein. *Genes & Development*. 2000; 14(14):1777–1788. [PubMed: 10898792]
- Qi X, et al. The common ancestral core of vertebrate and fungal telomerase RNAs. *Nucleic Acids Research*. 2013; 41(1):450–462. [PubMed: 23093598]
- Robert AR, Collins K. Human telomerase domain interactions capture DNA for TEN domain-dependent processive elongation. *Molecular Cell*. 2011; 42(3):308–318. [PubMed: 21514196]
- Rouda S, Skordalakes E. Structure of the RNA-binding domain of telomerase: Implications for RNA recognition and binding. *Structure*. 2007; 15(11):1403–1412. [PubMed: 17997966]
- Sauerwald A, et al. Structure of active dimeric human telomerase. *Nature Structural & Molecular Biology*. 2013; 20(4):454–460.
- Schmidt JC, Dalby AB, Cech TR. Identification of human TERT elements necessary for telomerase recruitment to telomeres. *eLife*. 2014; 3
- Shay JW, Wright WE. Mutant dyskerin ends relationship with telomerase. *Science*. 1999; 286(5448):2284–2285. [PubMed: 10636790]
- Shippen-Lentz D, Blackburn EH. Functional evidence for an RNA template in telomerase. *Science*. 1990; 247(4942):546–552. [PubMed: 1689074]
- Sun J, et al. Stn1-Ten1 is an Rpa2-Rpa3-like complex at telomeres. *Genes & Development*. 2009; 23(24):2900–2914. [PubMed: 20008938]
- Tejera AM, et al. TPP1 is required for TERT recruitment, telomere elongation during nuclear reprogramming, and normal skin development in mice. *Developmental Cell*. 2010; 18(5):775–789. [PubMed: 20493811]
- Terwilliger TC. Maximum-likelihood density modification. *Acta Crystallographica. Section D, Biological Crystallography*. 2000; 56(Pt. 8):965–972. [PubMed: 10944333]
- Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. *Acta Crystallographica. Section D, Biological Crystallography*. 1999; 55(Pt. 4):849–861. [PubMed: 10089316]
- Theimer CA, Feigon J. Structure and function of telomerase RNA. *Current Opinion in Structural Biology*. 2006; 16(3):307–318. [PubMed: 16713250]
- Vulliamy T, et al. Association between aplastic anaemia and mutations in telomerase RNA. *Lancet*. 2002; 359(9324):2168–2170. [PubMed: 12090986]
- Wan B, et al. The *Tetrahymena* telomerase p75-p45-p19 subcomplex is a unique CST complex. *Nature Structural & Molecular Biology*. 2015; 22:1023–1026.
- Wang F, et al. The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature*. 2007; 445(7127):506–510. [PubMed: 17237768]
- Wilkins MR, et al. Protein identification and analysis tools in the ExPASy server. *Methods in Molecular Biology*. 1999; 112:531–552. [PubMed: 10027275]

- Witkin KL, Collins K. Holoenzyme proteins required for the physiological assembly and activity of telomerase. *Genes & Development*. 2004; 18(10):1107–1118. [PubMed: 15131081]
- Ye AJ, Romero DP. Phylogenetic relationships amongst tetrahymenine ciliates inferred by a comparison of telomerase RNAs. *International Journal of Systematic and Evolutionary Microbiology*. 2002; 52(Pt. 6):2297–2302. [PubMed: 12508900]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1.

(A) Primary structure and sequence identity of TERT proteins from a diverse group of organisms (Human, *T. thermophila*, *S. cerevisiae*, *E. cuniculi*, and *T. castaneum*). (B) Schematic of the vertebrate telomerase RNA; conserved motifs are shown in color. The activation domain (CR4/5) and the pseudoknot are highlighted with a blue, dashed line.

Fig. 2.

Primary structure and sequence of the RNA–DNA hairpins used in cocrystallization trials with tcTERT. The RNA template and complementary telomeric DNA are shown in *black letters*; the linker in *blue* and the stable tetraloop in *red*.

Fig. 3.
(A) Simulated annealed omit map of the RNA (*magenta stick*)–DNA (*yellow stick*) hairpin cocrystallized with tcTERT at 1.0σ contour level. (B) tcTERT surface charge representation, showing the RNA–DNA hybrid (*stick*) docked in the interior cavity of the TERT ring.