Structural studies of human telomerase complexes by cryo-electron microscopy

Jin, Yibo

2020

Jin, Y. (2020). Structural studies of human telomerase complexes by cryo-electron microscopy. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/139941>

<https://doi.org/10.32657/10356/139941>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 20:29:10 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**STRUCTURAL STUDIES OF HUMAN TELOMERASE
COMPLEXES BY CRYO-ELECTRON MICROSCOPY**

YIBO JIN

SCHOOL OF BIOLOGICAL SCIENCES

2020

**STRUCTURAL STUDIES OF HUMAN TELOMERASE
COMPLEXES BY CRYO-ELECTRON MICROSCOPY**

YIBO JIN

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological
University in partial fulfilment of the requirement for the
degree of Doctor of Philosophy

2020

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

24.01.2020

Date

A handwritten signature in black ink, consisting of stylized, overlapping loops and lines, positioned above the printed name.

Yibo Jin

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

24.01.2020

Date

A handwritten signature in black ink, appearing to read 'Daniela Rhodes', written in a cursive style.

Prof. Daniela Rhodes

Authorship Attribution Statement

This thesis **does not** contain any materials from papers published in peer-reviewed journals or from papers accepted at conferences in which I am listed as an author.

24.01.2020

Date

A handwritten signature in black ink, consisting of stylized, overlapping loops and lines, positioned above a horizontal baseline.

Yibo Jin

Acknowledgement

First, I would like to thank my supervisor Professor Daniela Rhodes for her guidance and support in the development of my scientific career. It was a wonderful experience to take on this challenging project, through which there were tears as well as joys. Without her help, it would have not been possible to work through such a tough project. I would also like to thank my co-supervisor Sara Sandin, for her helpful suggestions and guidance on electron microscopy work. Importantly, I need to express my gratitude for Dr David Gill and Dr Simon Lattman for training me to work as a scientist. I would also like to thank my thesis advisory committee, Peter Droge and Shashi Bhushan for their generous help throughout my PhD life.

I wish to thank all my colleague in NTU. I would like to express my special thanks to Ziqi Long, who I spent cheerful five years together as friend and colleague. I wish to thank Qin Gong, Chenrui Xu for useful scientific and general discussions. I wish to thank Dr Andrew Wong and Dr Emmanuel Smith for critical discussion on EM techniques. I would also like to thank all my friends for the emotional supports received. Special thanks to Yishen Wang who helps me to get out of my miserable times.

Above all, I would like to thank my parents for their upbringing and support when I chose to study biology during my bachelors and starting a PhD journey far from home. Without their support, I wouldn't be able to follow my passion for life. Lastly, I would like to thank Nanyang Technological University (NTU) and the Ministry of Education (MOE) for financial support during my course.

Contents

List of abbreviations	ix
Summary.....	x
Chapter 1 Introduction.....	1
1.1 Telomerase is a solution to the end replication problem	1
1.1.1 The telomere is at the linear end of eukaryotic chromosomes.....	1
1.1.2 Semi-conservative replication and End replication problem	3
1.1.3 Telomere binding complexes	6
1.1.4 Telomerase biological function and regulation.....	11
1.1.5 Human telomerase biogenesis and recruitment.....	15
1.1.6 Telomerase recruitment and the catalytic cycle	19
1.1.7 Telomerase structure analysis	26
1.2 The era of “Resolution Revolution” in biological cryo-electron microscopy	32
1.2.1 History of transmission electron microscope.....	32
1.2.2 Electron microscope objective system enhances image contrast	35
1.2.3 Development of detectors in electron microscopes.....	36
1.2.4 Electron microscopy for biological specimens	38
1.2.5 Cryo-EM sample preparation.....	41
1.2.6 Single-particle cryo-EM and processing.....	45
1.2.7 Cryo-electron tomography – directly studying proteins from cell.....	51
1.2.8 Resolution revolution.....	54
1.3 Research aim	56
Chapter 2 Materials and methods	58
2.1 Biochemical methods	58
2.2 Electron microscopy methods	65
Chapter 3 Overexpression and purification of human telomerase complexes from the mammalian expression system.....	69
3.1 Introduction	69
3.2 New telomerase expression system	71
3.3 Telomerase expression and cell line selection.....	73
3.4 Telomerase purification protocol optimisation.....	80
3.5 Substrate-based human telomerase purification	84
3.6 Purification of the telomerase complex containing HSP90.....	86
3.7 Sucrose gradient characterisation of the human telomerase.....	91
3.8 Characterisation of the purified human telomerase complex	94
3.9 Telomerase activity	96

3.10	Shelterin proteins increase telomerase processivity	99
3.11	Conclusions and discussions	102
Chapter 4	Structural analysis of human telomerase by using electron microscopy	107
4.1	Introduction	107
4.2	Optimisation of grid preparation to capture human telomerase sample on grids	108
4.3	Cryo-EM analysis of human telomerase on graphene oxide grids	110
4.4	Cryo-EM analysis of human telomerase on holey grids.....	113
4.5	Negative stain EM analysis of oligo-pull down human telomerase	118
4.6	Cryo-EM analysis of substrate-based purified human telomerase on holey grids with phase plate	120
4.7	Electron tomography of human telomerase.....	122
4.8	Cryo-EM analysis of HSP90 bound human telomerase	125
4.9	Comparison between EM densities of human telomerase.....	130
4.10	Conclusions and discussions	133
Chapter 5	Discussion.....	136
5.1	Composition and oligomeric state of human telomerase.....	136
5.2	The conformational heterogeneity of human telomerase limits the resolution of structural analysis by cryo-EM.....	141
5.3	Cryo-EM development and the future	146
5.4	Current difficulties in cryo-EM	148
5.5	Future work and outlook	152
Reference		155

List of abbreviations

ALT	Alternative lengthening of telomeres
ATM	Ataxia-Telangiectasia mutated
ATR	Ataxia telangiectasia and Rad3-related protein
CMV	Cytomegalovirus
CPU	Central processing unit
DAT	Dissociates activities of telomerase
DKC	Dyskeratosis congenita
DNA	Deoxyribonucleic acid
DQE	Detective quantum efficiency
ds	Double strand
DSB	Double strand break
EM	Electron Microscopy
ETS	E-twenty-six
FIB	Focused ion beam
FSC	Fourier shell correlation
G4	G quadruplex
GPU	Graphic processing unit
HEK	Human embryonic kidney
HSP	Heat shock protein
MW	Molecular weight
Ni-NTA	Nitrilotriacetate Nickel
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
POT1	Protection of telomeres 1
PUA	Archaeosine-specific transglycosylase
RAP1	Ras-proximate-1
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RT	Reverse transcriptase
SDS	Sodium dodecyl sulfate
SNR	Signal to Noise Ratio
ss	Single strand
TERC	Telomerase reverse transcriptase RNA
TERT	Telomerase reverse transcriptase
TIN2	TERF1-interacting nuclear factor 2
TRBD	Telomerase RNA binding domain
TRF1/2	Telomeric repeat-binding factor 1/2
VPP	Volta phase plate
WT	Wild type

Summary

Chromosomal DNA ends shorten after each round of replication due to the end-replication problem and post replication processing. The ends of linear eukaryotic chromosomes are capped and protected by protein-DNA complexes forming a structure called the telomere. The telomere length is crucial as critically short telomeres activate the DNA damage response pathway leading to cell senescence. The telomere is extended by a protein-RNA enzyme complex with reverse transcriptase function named telomerase. The process by which telomerase extends telomeres must be tightly regulated. Human telomerase is inactive in normal human somatic cells but activated by various mechanisms in 85~90% of cancer cells. Mutations in telomerase binding proteins that cause reduced telomerase activity *in vivo* lead to developmental disorders such as Dyskeratosis Congenita. Although many studies were carried out to study human telomerase, the molecular mechanisms behind telomerase biogenesis, recruitment and elongation process are still not fully understood. Thus, a high-resolution structure of human telomerase would provide molecular details of human telomerase catalytic cycle and recruitment. Historically, overexpression of full-length human telomerase has been shown to be difficult and does not give sufficient material for structural studies and biochemical analysis. In this thesis, I present optimised human telomerase expression and purification methods from mammalian cells, which routinely yield around 500µg of active human telomerase sample, allowing further studies of human telomerase structure by cryo-EM. The purified human telomerase sample was biochemically characterised to contain only the catalytic subunit hTERT and RNA subunit hTERC and is shown to be fully active and processive. Previously, Daniel Rhodes' group had determined the structure of an active full-length human telomerase dimer to 25Å resolution by negative stain EM. More recently the Collins group determined the structure of monomeric human telomerase with other binding factors complex to 8Å resolution by cryo-EM. In my studies, I

have obtained a 15Å structure of human telomerase dimer constituting only hTERT and hTERC by cryo-EM. In addition, HSP90/P23 bound human telomerase was purified and studied using cryo-EM. I determined a low-resolution structure of HSP90 bound human telomerase to 25Å resolution and the structure of HSP90/P23 sub-complex to 4.8Å resolution. Despite efforts spent to improve the resolution of the complex, data suggested the structural heterogeneity in the sample is the key limiting factor of the resolution of the structure, a large amount of data may compensate for the problem. Through biochemical and structural analysis, I concluded that human telomerase is likely to function as a dimer containing hTERT and hTERC alone. Finally, this thesis describes the outlook in studying human telomerase structure and its functional role at telomeres by using cryo-electron tomography (cryo-ET).

Chapter 1 Introduction

1.1 Telomerase is a solution to the end replication problem

This thesis focuses on the structural investigation of human telomerase, an enzyme that extends the eukaryotic chromosome ends named telomeres. This chapter aims to introduce the biological background of telomere structure and the end replication problem, which subsequently lead to the proposal and discovery of the telomerase enzyme. At the same time, the protein/nucleic acid complexes telomerase and shelterin, which are located at the chromosomal ends protecting and regulating the telomeres, will be introduced, and the interaction between telomerase and shelterin will be discussed. Finally, the biological importance and structure information of the human telomerase will be introduced in detail.

1.1.1 The telomere is at the linear end of eukaryotic chromosomes

There are several differences between the prokaryotic and eukaryotic genomes such as genome size, DNA packing, and perhaps the biggest difference is their shape. Prokaryotes contain circular genomes whereas eukaryotes contain linear DNA bound to histones and packaged in highly organised structures named chromosomes. The linear nature of the eukaryotic chromosome means that it contains two ends per chromatid, and these ends are named telomeres. Telomeres are essential for genomic stability and cell viability of eukaryotic cells from ciliates to vertebrates. Telomeres contain highly conserved G-rich DNA repeats, but the repeat sequence varies somewhat across species; for example, TTGGGG in *Tetrahymena*, $G_{(2-3)}TG_{(1-6)}T$ in *Saccharomyces cerevisiae* and TTAGGG in vertebrates and human (Flint et al. 1997). The telomere ends consist of a double-stranded DNA region and G-rich 3'overhang region which are bound by protein factors (De Lange 2009; Palm and de Lange 2008).

The exposed telomere structure resembles one half of a DNA double-strand break (DSB), and the overhang region may be recognised as a single-strand break. The DSB may be recognised by ATM triggering DNA damage response, which may result in chromosome fusion through end-joining mechanisms (Weaver 1998). To prevent triggering of the DNA damage response, both double and single strand DNA binding proteins bind to the chromosomal DNA ends forming a cap-like protective structure named shelterin.

The physiological structure of the telomere is still under debate. The classical view suggests the structure is linear and includes proteins that bind to the single and double-stranded region of the telomere (Figure 1.1.1A).

Figure 1.1.1 Three views of the structures at telomeric ends. (A) The classical view suggests the telomeric ends is linear. (B) The T loop and D loop model suggest telomere ends are packed in a duplex lariat structure folded back forming loops. (C) The G4 model suggest that the 3' overhang may form G4 structure, and with ssDNA binding complex as regulatory factors.

In 1999, electron microscopy revealed that mouse and human telomeres are organised in a duplex lariat structure, the so-called T-loop (Figure 1.1.1B) (Griffith et al. 1999). In this model, the G-overhang folds back and forms base pairs with the C-rich strand in the double-stranded telomeric DNA region folding into a D-loop. The closed configuration of the T-loop

defines the telomere region and prevents chromosome fusion. Although the model is effective for explaining end-protection, there is no implicit *in vivo* evidence for its existence and regulation.

Over 30 years ago, Klug suggested that a four-stranded DNA structure called G-quadruplex (G4) should have a functional role *in vivo*, due to its spontaneous formation *in vitro* in physiological salt conditions (reviewed in Rhodes & Lipps, 2015). Potential G-quadruplexes forming sequences are found at various locations in the genome, but they are most commonly found at the telomere. The G-rich nature of telomeric sequence meaning that it can form G-quadruplexes, demonstrated by various *in vitro* studies (Lipps, Grissem, and Prescott 1982; Sundquist and Klug 1989) and *in vivo* studies (Paeschke et al. 2005; Schaffitzel et al. 2001). The G4 formation at telomere ends are shown to have a regulatory role on telomerase activity (Wang et al. 2011) and may protect the telomere ends (Figure 1.1.1C). Despite the strong evidence from many light microscopy experiments showing the existence of G4 at telomeres, the resolution limit means it's unclear whether the G4 structures are formed at the telomeric ends or sub-telomeric regions in human (Rhodes and Lipps 2015).

1.1.2 Semi-conservative replication and End replication problem

In both prokaryotic and eukaryotic cells, DNA replication follows a semi-conservative replication model which was first proposed after the discovery of the DNA structure by Watson and Crick in 1953 (Watson and Crick 1953) and subsequently proved by the Meselson-Stahl experiment in 1958 (Meselson and Stahl 1958). The process involves the separation of parental DNA into single strands, and the newly synthesised strand is based paired to the parental strand of DNA. In double-stranded DNA, the two DNA strands are in opposite directions (antiparallel), thus the mechanism of replication for each stand is

different. DNA synthesis can only process from 5'- 3' direction as DNA polymerase requires a 3-OH group for the addition of nucleotides. DNA polymerases cannot initiate DNA replication *de novo* and the process is initiated by RNA primers. At the replication fork, as two strands of DNA separate, the leading strand unwinds in the 3'- 5' direction, allowing the continuous synthesis of the new strand from the 5'- 3' direction from a single primer at the replication fold. However, the lagging strand unwinds in the 5'- 3' direction, which poses a problem. The cell deals with this by employing multiple RNA primers which are placed in 100-200bp gaps to allow synthesis of DNA fragments. The RNA primers on lagging strands are later removed leaving fragments of DNA, named Okazaki fragments. These fragments are joined by DNA ligase. The complicated process requires additional enzymes and close regulation at different levels.

Figure 1.1.2 The end replication problem. (A) During the semi-conservative DNA replication, the removal of RNA primer (red) at 5' end leaving an unfilled gap which leads to gradual shortening of telomeres after each round of replication. (B) A nuclease carries out post-processing of replication products and creates 3' G-overhang on both ends of a chromosome.

The replication mechanism requires the coordination of multiple enzymes, including DNA polymerase, helicase, and PCNA (clamp). The replication starts with the synthesis of RNA primers complementary to template DNA, which is carried out by the DNA Polymerase α /Primase complex. The RNA-DNA duplex is bound by replication factor C (RFC) which displaces the Polymerase α /Primase complex. The ring-shaped PCNA complex is then recruited next to RFC forming a DNA clamp structure. The binding of PCNA helps the binding of polymerase epsilon (Pol ϵ) to the leading strand, and Pol δ to the lagging strand (Tsurimoto, Melendy, and Stillman 1990). As the replication fork progresses, DNA synthesised by Pol δ displaces the RNA primer from the lagging strand into a single strand RNA flap, and the short flaps are removed by FEN1, whereas the long flaps are first coated by RPA and then removed by RNaseHII (Finger et al. 2012). This leaves gaps in the backbone between Okazaki fragments, which are joined together via DNA ligase 1 (Turchi et al. 1994).

Since the eukaryotic chromosome is linear, following this replication mechanism, when the replication fork reaches the linear end of the chromosome, it leaves a gap. This problem is called the end-replication problem (Levy et al. 1992; Wynford-Thomas and Kipling 1997) and was first described in studies of linear phage DNA replication (Tsurimoto et al. 1990). On the lagging strand, the addition of the primer becomes impossible at the end of chromosomes, which means the strands covered by an RNA primer at the 5' end of both the leading and lagging strands are lost each time a chromosome is replicated (Figure 1.1.2). This means that in human somatic cells, the DNA ends of chromosomes progressively become shorter after each cell division. On estimate, telomeres in human somatic cells shorten by 50-200 bp per cell cycle. Interestingly, the rate of telomere shortening in human and mice is faster than predicted by the end replication problem. Data suggest that after each round of replication, the 5' terminal of DNA is further processed by a nuclease (Makarov, Hirose, and

Langmore 1997) and regulated by POT1(Hockemeyer et al. 2005). The post-process leads to the production of a 3' G-overhang on both leading and lagging strand ends.

To prevent loss of important genetic information, the end of eukaryotic DNA consists of non-coding sequence repeats, and in human, the repeat sequence is (TTAGGG) n and has a typical size of 15-20kb at birth. For somatic cells, the number of cell divisions is limited as the telomeres progressively become shorter, and eventually, when the telomeres are critically short, they trigger the cell damage response leading to replication, senescence, or apoptosis causing cell death (Blasco et al. 1997). However, in germ cells and cancer cells, the length of the telomere is maintained or extended to prevent cell death caused by telomere shortening. The extension of the telomere is carried out by an enzyme called telomerase, which will be further described in this chapter.

1.1.3 Telomere binding complexes

Due to the importance of telomere ends for cell survival, they must be well organised, protected, and regulated. The telomeres of the mammalian chromosome ends are bound by six-protein complexes – the Shelterin complex (de Lange 2005). Together they are responsible for the protection and capping of the telomere, recruitment and regulation of telomerase, and hence the extension and maintenance of telomeres. In vertebrates, the Shelterin components are TRF1, TRF2, POT1, TPP1, TIN2, and RAP1 (Figure 1.1.3) (Hu et al. 2017). The domain and interaction map of all six shelterin proteins is shown on the next page.

Figure 1.1.3 Domains and interaction map of the six Shelterin components: POT1, TPP1, TRF1, TIN2, TRF2 and RAP1. (A) Domains details of the six Shelterin components. (B) Crystal structure of DNA binding domains and interacting domains of shelterin components available, with dotted line mapping the interactions. Figure adapted from (Hu et al. 2017) (Chapman, personal communication).

Three components of the shelterin complex bind specifically and directly to telomeric DNA repeats. TRF1 and TRF2 bind dsDNA whereas POT1 binds to ssDNA (Bianchi 1997; Chen et al. 2008b). Other components, hTPP1, hTIN2, and Rap1 were shown to have no direct interaction with telomeric DNA, but instead, they form the Shelterin complex through interacting with the DNA binding proteins. Structurally, hTRF1 and hTRF2 have a highly conserved TRF homology (TRFH) domain which interacts with the TBM domain from hTIN2 (Chen et al. 2008b; Court et al. 2005). More importantly, they contain nearly identical

Myb motifs that recognise and directly bind to the dsDNA TTAGGG repeats (Court et al. 2005). However, whilst TRF1 and TRF2 have the same structure, their surfaces are chemically distinct forming different complexes in order to carry out different functions. Structurally, the N-terminal of TRF1 consists of an acidic region whereas TRF2 contains a basic region. In addition, only TRF2 has an RBM domain that interacts with the RCT domain from RAP1, and TRF1 shows no such interaction. The formation of a complex with TRF2 is essential for the localization of RAP1 to telomeres and its stability (Li, Oestreich, and De Lange 2000). Otherwise, human RAP1 is poorly characterised.

The ssDNA telomeric G-overhang is protected by the hPOT1-hTPP1 complex. The POT1-TPP1 complex is very important in telomeric protection, as the deletion of the complex leads to chromosomal fusion (Grill et al. 2019). It has been suggested that the complex between TPP1 and POT1 is important for telomere localisation (Pike et al. 2019). The deletion of TPP1 causes POT1 to lose its localisation to telomeres, although it has a high binding affinity specifically for telomeric repeats. Some data suggest hTPP1 can bind to the junction of ssDNA repeats and dsDNA repeat region at telomere ends (Erdel et al. 2017).

TIN2 is a very important mediator protein that bridges the dsDNA binding protein complex TRF1/TRF2-RAP1 to the ssDNA binding protein POT1. Domain analysis shows the TIN2 TBM interacts with the TRF1/2 TRFH domain, and the TIN2 TRFH domain interacts with the TPP1 TBM domains. It also plays critical roles in both telomere maintenance and telomere length regulation through two different mechanisms (Ye et al. 2004). TRF1 can be inhibited by the poly (ADP-ribose) polymerase (PARP) activity of its interacting partner tankyrase 1, which abolishes its DNA binding activity and inhibits TRF1 localisation to the telomeres *in vivo*. TIN2 is a modulator of tankyrase and stabilises TRF1's binding to telomeres (Ye and De Lange 2004). Also, TIN2 competes with SCFFBX4 for TRF1 binding, thus preventing degradation of TRF1 through the ubiquitination pathway (Zeng et al. 2010).

Thus, loss of TIN2 would cause activation of DDRs via both pathways, highlighting the importance of TIN2 in telomere end protection. Furthermore, TIN2 interacts with three shelterin proteins and TIN2 deletion compromises the integrity of shelterin components at telomeres (Ye et al. 2004). The precise roles of TIN2-mediated interactions in telomere protection are currently unclear. Mutational studies have given some insights into the interaction between TIN2 and TPP1 (Frescas and de Lange 2014; Kim, Kaminker, and Campisi 2000). The interactions of TIN2 with TRF1 and TRF2 appear to be important for preventing non-homologous end-joining mediated chromosomal fusion (Arnoult and Karlseder 2015).

Unprotected telomeres resemble DNA breaks, and exposed DNA triggers DNA damage response through the activation of ATR and ATM. ATM is specific for sensing double-strand breaks, whereas ATR is activated by a single-strand break (Ciccio and Elledge 2010).

Physically, the shelterin complex prevents telomeric DNA from being exposed, and sterically hinders the binding of DNA damage detectors ATM and ATR. Therefore, shelterin components TRF2 and POT1 actively suppress DNA damage response by downregulating the ATM and ATR pathways respectively (Rai et al. 2010; Sfeir and de Lange 2012).

In summary, all shelterin components have their integrated role in the protection of telomeres and regulation of telomere length. The protein complex that is formed in the ssDNA region of the telomere is TPP1-POT1. The complex is bridged to the dsDNA binding complex TRF2-RAP1 and TRF1 by TIN2.

Besides its function in telomere protection and maintenance, the shelterin components also play a crucial role in telomerase recruitment and regulation. The N-terminal OB-fold of hTPP1 does not contribute to telomere specific binding but has been shown to play an important role in telomerase recruitment and regulation (Grill, Tesmer, and Nandakumar 2018). In Dyskeratosis Congenita (DKC) patients with a TIN2-R282H mutation, localisation

and binding to other shelterin components is unaffected and thus is not causing activation of DDR. It has been shown the telomerase elongates telomeres at a reduced frequency in TIN2-R282H mutant cells, which suggest the role of TIN2 in telomerase recruitment and regulation. Data suggest that both hTIN2 and hTPP1 are essential for telomerase recruitment, through direct interaction between the N-terminal of hTPP1 and TEN domains from hTERT (Abreu et al., 2010; Grill et al., 2018, Anne Hempel, personal communication) Similar observations have also been made in ciliates in which the hTPP1 homologue TEBP β was shown to be essential for telomerase recruitment to telomeres (Paeschke et al. 2008). The exact mechanism by which TIN2 regulates telomerase activity is unknown. However, recruitment and regulation of telomerase may be through the coordination of TPP1 and TIN2. One hypothesis is that TPP1 may have a direct role in telomerase recruitment, and TIN2 may change the conformation of telomeres, making them more accessible to telomerase resulting in upregulation of telomerase activity (Pike et al. 2019).

The interaction between G4 DNA and shelterin components also have been shown to have a regulatory role. TRF2 binding to telomeric DNA inhibits the formation of G-quadruplex *in vitro* (Lipps and Rhodes 2009; Mendez-Bermudez et al. 2018). The formation of G4 at telomere ends are shown to be important for human telomerase regulation (Moye et al. 2015). The regulation of human telomerase by G4 may not be through direct interaction, instead, through shelterin components POT1 which unfold G4 (Zaug, Podell, and Cech 2005). Currently, only limited information is available on the telomerase recruitment process to telomeres. Due to the close interaction between human telomerase and shelterin components, structural information on the telomerase and its recruitment complexes may provide insights into the functional role of shelterin components during telomerase recruitment and regulation process.

1.1.4 Telomerase biological function and regulation

Figure 1.1.4 History of telomere and telomerase research and telomerase structure. A timeline from 1938 to 2018, highlighting key events in the telomere research history. The illustration shows telomerase structure consisting of protein subunit TERT and RNA subunit TERC.

The discovery of the telomerase enzyme and its function is the results of decades of work from many great scientists (Figure 1.1.4). The existence of a capping structure at linear chromosomal ends was first proposed in 1938 by Muller, long before the discovery of the structure of DNA in 1953 by Watson and Crick. The discovery of DNA structure (Watson and Crick 1953), together with DNA polymerase by Kornberg in 1958 (Lehman et al. 1958) gave rise to the question of what is happening at the telomeric ends. In 1961, Hayflick found that cells can only divide 40-60 times before they stop dividing, named after him as the Hayflick limit (Hayflick and Moorhead 1961). Following this observation, Watson and Olovnikov independently proposed there must be a polymerase-like enzyme to “repair” telomere ends. Later work by several scientists in 1970 to establish the link between telomere length and cell senescence, lead to the hypothesis that Hayflick limit is related to the

telomere length, and there must be an enzyme that lengthens the telomere ends. In 1981, Blackburn and Szostak (Katzen, Cann, and Blackburn 1981; Shampay, Szostak, and Blackburn 1984) started work on sequencing of telomere ends which lead to the result that the telomere is essential in stabilising genomes. In the same year, Klobutcher showed that the DNA sequence of the telomeric region of hypotrichous ciliates consists of a sequence repeat and also first described the G- rich strand extends in a 3'direction forming 3'overhang (Klobutcher et al. 1981). In 1985, Blackburn and Greider first discovered the RNA subunit of tetrahymena telomerase enzyme and reported its function (Greider and Blackburn 1985). This key discovery was later awarded the Nobel Prize in Physiology or Medicine in 2009 (shared with Szostak). However, the protein catalytic subunit was not discovered until 1997, when Lingner and Lundblad identified telomerase the protein catalytic subunit in *Euplotes aediculatus* and *Saccharomyces cerevisiae* respectively (Lingner et al. 1997; Morris and Lundblad 1997). This sets the beginning of telomerase related research, and nowadays, telomerase is considered an excellent therapeutic target for medical researchers in many areas.

The new findings suggest that telomerase is a reverse transcriptase that carries its own RNA template in the RNA subunit TERC that adds TTAGGG repeats to the end of telomeres through the action of its catalytic subunit TERT (Figure 1.1.4). Mammalian cells have a complicated mechanism for regulation of the cell lifespan. Normal somatic cells have limited growth potential and undergo senescence after a fixed number of cell divisions (the Hayflick limit). Surprisingly, telomerase activity is only detected in embryonic stem cells and germ cells; but low or no telomerase activity is detected from the majority of normal stem cells regardless of their proliferative potential (reviewed in Hiyama & Hiyama, 2007). In contrast, >85% of cancer cells maintain telomere length by activation and overexpression of telomerase. The other 15% of cancer cells maintain telomere length by alternative

lengthening of telomeres (ALT) pathway. Either telomere maintenance pathway enables cells to have unlimited proliferation potential and leads to as “immortalised” cells.

Since the main functional role of the telomerase is to maintain telomerase length, it must be closely regulated to prevent any unwanted consequences. A reduced telomerase activity leads to shortened telomeres causing pre-mature ageing diseases, while an increase in telomerase activity leads to uncontrolled cell proliferation, a key characteristic of cancer cells.

Under normal circumstance, human telomerase activity can only be detected from germline cells and embryonic stem cells but not in normal adult tissues (Wright et al. 1996). Germline cells and stem cells during early developmental stage undergoes constant replication thus the telomere length need to be maintained to counteract telomere shortening caused by the end-replication problem. Failing to do so would lead to cell senescence or apoptosis.

Dyskeratosis Congenita (DKC) is a well-known developmental disease that is related to reduced human telomerase activity. The disease is caused by functional mutations in the gene *DKC1* encoding the Dyskerin protein which binds to TERC at the H/ACA box (Figure 1.1.4). Patient mutations usually contain a single amino acid substitution at the N-terminal domain or archaeosine-specific transglycosylase (PUA) domain (Garcia, Wright, and Shay 2007), reducing the binding to TERC. Reduced or abolished binding of *DKC1* to TERC reduced the stability of TERC leading to a reduction in human telomerase activity. Patients with severe DKC phenotype also have very short telomeres, and those with longer telomeres have less physiological effects. Thus reduced telomerase activity and shorter telomeres are both required for DKC to manifest itself (Goldman et al. 2005).

Since TERT is not expressed in most normal human cells, it is widely used as a potential cancer biomarker. Interestingly, the mutations of hTERT that contribute to cancer are not common in the coding region but mostly in the promoter region of the *TERT* gene. Multiple point mutations in the *TERT* promoter region have been identified to contribute to cancer by

genome sequencing. These mutations generate binding motifs for the E-twenty-six (ETS) transcription factor in the *TERT* promoter region, leading to the upregulation of TERT mRNA expression (Bell et al. 2015; Li et al. 2015). *TERT* promoter mutations are currently the most common non-coding somatic mutations in cancer and are present in many types of cancers, including melanoma (67%), ovarian clear cell carcinoma (15.9%), thyroid cancer (10%), and bladder cancer (59%) (Horn et al. 2013; Huang et al. 2013; Killela et al. 2013). Regulation of *TERT* expression is through transcription factors but as well as epigenetic modifications. DNA methylation has often been associated with gene silencing, however, the level of DNA methylation in the *TERT* promoter region is positively correlated with TERT expression levels (Lai, Serene 2005). Alternative splicing has also been shown to be linked to human telomerase upregulation (Wong, Wright, and Shay 2014).

Figure 1.1.5 How cells become immortalised by activation of telomerase (adapted from (Okamoto and Seimiya 2019)). Oncogenes activated or tumour suppressor inactivated cells which also has activated telomerase activation would lead to cell immortalisation.

Although telomerase is overexpressed by various pathways in cancer cells, telomere length in cancer is usually shorter than normal somatic cells. This suggests the telomerase overexpression does not cause cancer directly but is a key contribution to cell immortalisation. As cancer cells undergo uncontrolled proliferation, as a result, their telomeres wear out very quickly and cells enter a senescence state through the DDR pathway. When mutations occur in tumour suppressor genes such as p53, which allow these cells to escape senescence and continue to proliferate until the telomere is critically short (Figure 1.1.5)(Okamoto and Seimiya 2019), and many of the cells would undergo chromosomal

fusion and trigger apoptosis (Hayflick 1965). Cells that have re-expression of TERT or ALT may maintain the telomere length, allowing the cell to become immortalised (Chin et al. 1999). Data suggest that TERT expression and telomerase activation is too low after the acquisition of TERT promoter mutations to extend all telomeres (Hemann et al. 2001; Ouellette et al. 2000). The telomerase is preferentially recruited to the shortest telomeres, and as a result, the average length of all telomeres gradually shortens. When telomeres are all critically short, TERT expression and telomerase activity are increased through unknown mechanisms, which stabilise telomere length (Ouellette et al. 2000; Zhu et al. 1998) and hence cell proliferation.

1.1.5 Human Telomerase biogenesis and recruitment

Human telomerase is a ribonucleoprotein complex and consists of protein catalytic subunit TERT and RNA subunit TERC that contains the template region for telomere repeat synthesis (Figure 1.1.4). Since the protein component TERT is synthesised by ribosomes in the cell cytoplasm, it needs to be transported and assembled with its RNA component TERC in the nucleus. TERC is directly transcribed from a coding region, but it needs to undergo sophisticated processing before it can be assembled together with TERT. *In vitro* reconstituted human telomerase from rabbit reticulocyte extracts showed significant less telomerase activity compared to purified telomerase from cellular extracts (Bachand and Autexier 1999; Weinrich et al. 1997). This suggests that human telomerase assembly may happen at the same time as the synthesis of TERC, or there are specific protein factors that are required for the proper folding of telomerase that is missing from rabbit reticulocyte. The complete mechanism of the human telomerase biogenesis pathway is still unclear, but biochemical data has provided some useful information.

The TERT protein biogenesis is poorly studied due to its intrinsically low abundance in cells. Data suggest overexpression of hTERT disrupts its subcellular localization, which means the observations under the overexpression system may not represent the true physiological nature (Zhong et al. 2012). Through current understanding, hTERT is synthesised in the cytoplasm. hTERT contains a putative nuclear localization signal (NLS) and nuclear export signal (NES), suggesting that TERT crossing the nuclear membrane is likely to be involved in telomerase assembly or activity regulation (Chung, Khadka, and Chung 2012). Interestingly, a mutation in the NLS has little effect on telomerase activity and cell immortalisation and poses questions on the importance of nuclear trafficking of TERT. TERT has been reported to be bound by the chaperone HSP90/P23 as well as the AAA+ ATPases Pontin and Reptin (Mao and Houry 2017). Knocking down those components leads to the reduction of telomerase activity, but the role of each cofactor remains unclear. Although HSP90/P23 is a chaperone protein and is commonly thought to have a function of assisting protein folding, data suggest it may play a more important role. Immunoblotting of HSP90/P23 shows it copurifies with telomerase, suggesting this form is preferentially bound to the mature/active form of human telomerase. Further studies suggest two forms for telomerase, the HSP90/P23 bound, and non-bound form, and the activity of these two forms is indistinguishable (Forsythe et al. 2001; Holt et al. 1999). This proposes the possibility of HSP90/P23's involvement in both human telomerase biogenesis and regulation. The other components Pontin and Reptin seem to be associated with hTERT that is not incorporated into an enzymatically active telomerase RNP. Pontin and Reptin have also been shown to interact with hTERC, which will further be explained.

Figure 1.1.6 Human telomerase biogenesis and recruitment pathway. hTERT is synthesised in the cytoplasm and transported into the nucleus with HSP90 and p23. hTERT is assembled with telomerase RNA TERC in the nucleus and accumulate in Cajal bodies before being recruited to telomeres. (adapted from Jafri, Ansari, Alqahtani, & Shay, 2016)

The hTERC subunit encodes the template for telomeric DNA repeat synthesis. The biogenesis of hTERC is much more complicated compared to hTERT. The mature form of telomerase RNA is a 451bp noncoding RNA that contains functional features: the template for reverse transcription, the pseudoknot domain, a stem-loop that interacts with TERT, and a 3' element required for RNA stability (Webb and Zakian 2016; Zvereva, Shcherbakova, and Dontsova 2011). However, the function of the rest of the RNA is poorly understood. Unlike scaRNAs and snoRNAs, whose precursors are contained in intronic regions of mRNAs, TERC is transcribed from its own promoter by RNA polymerase II, with 5' capped and processed and 3' controlled by H/ACA (Feng et al. 1995). TERC contains an H/ACA box that is bound of H/ACA box associated factors: dyskerin, NOP10, NHP2 and NAF1 (GAR1 in humans). The 3' end of TERC is not polyadenylated like normal mRNA, instead, it's cleaved by exonuclease up to the boundary of the H/ACA, where further cleavage is

prevented by binding of dyskerin. The binding of the H/ACA box is co-transcriptional, and the rapid binding prevents 3' degradation of telomerase RNA allowing the accumulation of hTERC in the Cajal body (Darzacq et al. 2006; Fu and Collins 2003). Pontin and Reptin have also been shown to transiently associate with hTERC during biogenesis, but not in the maturation form. Like other scaRNAs, TERC also contains the Cajal body localization signal (CAB box) which binds to TCAB1. TCAB1 is essential for telomerase localisation to the Cajal body, however, the role of TCAB1 in telomerase biogenesis is unknown (Stern et al. 2012).

Also, the essential function unclear. It is thought the Cajal body may act as a distribution centre to deliver matured telomerase to telomeres. Knockdown of TCAB1 does not affect the overall telomerase activity *in vivo*, but instead significantly reduces the number of telomerase molecules recruited to the telomere (Venteicher and Artandi 2009). The hTERC subunit encodes the template for telomeric DNA repeat synthesis. The biogenesis of hTERC is much more complicated compared to that of hTERT. The mature form of human telomerase RNA is a 451bp noncoding RNA that contains functional features: the template for reverse transcription, the pseudoknot domain, a stem-loop that interacts with TERT, and a 3' element required for RNA stability (Webb and Zakian 2016; Zvereva et al. 2011). However, the function of the rest of the RNA is poorly understood. Unlike scaRNAs and snoRNAs, whose precursors are contained in intronic regions of mRNAs, hTERC is transcribed from its own promoter by RNA polymerase II, with 5' capped and processed and 3' controlled by H/ACA (Feng et al. 1995). TERC contains an H/ACA box that is bound of H/ACA box associated factors: dyskerin, NOP10, NHP2 and NAF1 (GAR1 in humans) (Figure 1.1.6). The 3' end of TERC is not polyadenylated like normal mRNA, instead, it's cleaved by exonuclease up to the boundary of the H/ACA, where further cleavage is prevented by binding of dyskerin. The binding of the H/ACA box is co-transcriptional, and the rapid binding prevents 3'

degradation of telomerase RNA allowing the accumulation of hTERC in the Cajal body (Darzacq et al. 2006; Fu and Collins 2003). Pontin and Reptin have also been shown to transiently associate with hTERC during biogenesis, but not in the maturation form. Like other scaRNAs, TERC also contains the Cajal body localization signal (CAB box) which binds to TCAB1. TCAB1 is essential for telomerase localisation to the Cajal body, however, the role of TCAB1 in telomerase biogenesis is unknown (Stern et al. 2012; Zhong et al. 2011). It is thought the Cajal body may act as a distribution centre to deliver matured telomerase to telomeres. Knockdown of TCAB1 does not affect the overall telomerase activity *in vivo*, but instead significantly reduces the number of telomerase molecules recruited to the telomere (Venteicher and Artandi 2009).

1.1.6 Telomerase recruitment and the catalytic cycle

The human telomerase is accumulated in the Cajal body and carries out its function at telomeres. Most abundant enzymes, such as ribosomes and metabolic enzymes, find their substrates by simple diffusion. The low copy number of both telomerase and telomere means telomerase must be accurately and efficiently recruited to telomeres that need to be elongated. It has been estimated that HEK293T and HeLa cells have around 100-250 copies of telomerase (Xi and Cech 2014), and during S phase, there is a total of 184 telomeres, although not all need to be lengthened at the same time. The mechanism of how telomerase is recruited to the telomeres remains unclear, but some ideas are acquired through the mapping of protein interactions.

The recruitment mechanism was first identified through the DAT (dissociates activities of telomerase) mutants. The DAT mutated telomerase was active *in vitro* but inactive *in vivo*. A number of DAT mutations are primarily clustered in the TEN domains of hTERT, which is

thought to play an important role in telomerase recruitment. Since telomeres are naturally protected by shelterin components, a link was drawn between the shelterin component TPP1 and the DAT mutants (Armbruster et al. 2001, 2004). Cell biological experiments including immunofluorescence and fluorescence in situ hybridization (IF–FISH) to telomeres as well as telomere chromatin immunoprecipitation (ChIP) to hTERT confirmed TPP1 recruits telomerase to the telomere (Abreu et al. 2010). Experimental evidence suggests hTERT TEN domain directly interacts with hTPP1's N-terminal OB-fold (Anne Hempel, personal communication). As mentioned previously, the TPP1-POT1 complex stimulates telomerase activity *in vivo*. More specifically, mutational studies revealed a patch of amino acids (termed the TEL patch) in the OB-fold of TPP1 that is essential for telomerase recruitment to telomeres *in vivo* and telomerase RAP stimulation *in vitro* (Dalby, Hofr, and Cech 2015). Experimental evidence shows that the fusion of one of the N-DAT mutants to POT1 rescues activity *in vivo*, suggesting that the telomerase DAT mutants were inactive *in vivo* because they were not recruited to telomeres. More recently, in a compensatory charge reversal experiment, DAT mutation (K78E) of hTERT and point mutation (E215K) in TPP1 which individually would disrupt telomerase stimulation *in vitro* and telomerase recruitment *in vivo* are rescued by each other, proving the direct interaction between hTERT TEN and hTPP1 without a bridging factor (Schmidt, Dalby, and Cech 2014). However, these experiments only prove that the hTPP1-hTERT TEN domain is essential for telomerase recruitment, but they do not prove this interaction alone is sufficient and that other factors may also have a role in telomerase recruitment such as TCAB1 mentioned above.

Recently, studies of TIN2 mutations in Dyskeratosis Congenita have revealed the possibility of its involvement in telomerase recruitment. Such mutations affect telomerase activity without disrupting TIN2 interaction with other shelterin proteins, suggesting TIN2 involvement in telomerase recruitment to the telomere, either through direct interaction or

through a mediator protein. Studies have also shown binding of heterochromatin protein 1 (HP1) to an HP1-binding motif (PTVML) in the region of TIN2 which contains mutations linked to dyskeratosis congenita-associated mutation, suggesting HP1 could be another regulatory protein (Canudas et al. 2011).

Due to the functional role of telomere maintenance and oncogenic potential of telomerase, its regulation is cell cycle dependent. Interestingly, telomerase is accumulated at the Cajal body most of the time, and only recruited to telomeres during the S-phase of the cycle, coinciding to when DNA replication occurs (Ten Hagen et al. 1990; Wright et al. 1999). Many studies have demonstrated telomerase localisation to telomeres only occurs in the S-phase of the cell division cycle (Paeschke et al. 2008).

It is perhaps not surprising since the telomere is most accessible during S-phase.

Posttranslational modifications are usually involved in cell cycle-dependent regulation of other enzymes. Data shows that the phosphorylation site in the OB-fold domain of TPP1 (S111) could be responsible for the cell cycle dependency (Zhang et al. 2013). However, the experimental results in HeLa cells and embryonic stem cells give contradictory results (Sexton et al. 2014). This may be due to the fundamental physiological difference between cancer cells and stem cells. More studies are required to characterise the regulation clearly. In support of the argument, regulation of telomerase through post-translational modification exists in fission yeast, where phosphorylation of the telomere component Ccq1 is required for assembly with the telomerase holoenzyme (Armstrong et al. 2018). This is clearly evident in ciliates in which phosphorylation by CDK2 of the TPP1 homologue is required for telomerase recruitment to telomeres in late S-phase (Paeschke et al. 2008).

Once the telomerase is recruited to the telomere, the repeat addition process is initiated. The current model of telomere elongation is initiated by the recognition and binding of telomeric DNA repeats by the template region of telomerase RNA, TERT and the anchor site (Figure

1.1.7). The 3' end of the telomeric DNA forms a hybrid with the 3' RNA template region, whereas the 5' region of the DNA is postulated to interact with two proposed anchor sites (Greider and Blackburn 1985; Patrick et al. 2019; Zaug, Podell, and Cech 2008). Next, the polymerase activity of hTERT extends the 3' telomeric DNA until the 5' boundary of the template region of hTERC. When the boundary is reached, telomerase undergoes translocation via an unknown mechanism, realigning the 3' telomeric DNA end to the first repeat of the 3' template region (Figure 1.1.7). Then, another round of addition is initiated, for multiple rounds until the substrate dissociated from the enzyme. The number of repeats added is regulated by a collection of interactions: RNA-DNA, TERT-DNA, TERT-RNA, and TERT-TERT interactions and possibly by an intrinsic nuclease activity as well as telomerase-associated proteins (Lewis and Wuttke 2012).

Figure 1.1.7 Human telomerase extension cycle. The substrate binds to the template region of telomerase RNA. The protein subunit extends the telomere to the boundary of the template region. Telomerase translocates with the help of anchor sites, allowing a new round of extension to take place.

The TERT protein architecture contains a reverse transcriptase domain that catalyses nucleotide addition to a substrate's 3' hydroxyl group and forms a product–template complex. Like other polymerases, telomerase nucleotide addition is metal-dependent, specifically through Mg^{2+} (Steitz et al. 1996). Accurate telomeric repeat synthesis is controlled by the template region within TERC. The process in which telomerase extends the telomere is through a progressive manner, and the Repeat Addition Processivity (RAP) measures the number of repeats added before the product dissociates from the enzyme (Wu, Tam, & Collins, 2017). The processive nature of telomerase enzyme means it must undergo translocation which involves breaking and forming double-strand bonds. While the helix is broken, to prevent the dissociation of the enzyme from the product, another “anchor” site is required. Recently, by biophysical approach experiment, the Schmidt group suggested that the fully assembled human telomerase RNP that is functional at the telomere contains a secondary DNA binding site but does not demonstrate the molecular nature of the binding

site (Patrick et al. 2019). Other publications suggest the TEN domain contain regions responsible for the anchoring region, due to its spatial proximity to the template binding region when in the catalytic ring formed by TRBD, RT and CTE of TERT. Biophysical data shows an anchor site dissociation constant and the Michaelis-Menten constant of telomerase, meaning instead of the hTERC template site, the anchor site is the primary binding site of DNA during telomerase catalysis. While the anchor site does not move, it has been suggested the extended telomere folds back to form a loop and subsequently released (Patrick et al. 2019). As mentioned previously, hTPP1-POT1 complexes act as a positive regulator by interacting with telomerase in a coordinating manner to increase telomerase affinity for its substrate, thus increasing the processivity (Figure 1.1.8).

Figure 1.1.8 Model of human telomerase extension regulation by TPP1. This model shows TPP1 binding prevents substrate dissociation for the human telomerase, aiding translocation of during the catalytic cycle, increasing the processivity of human telomerase (Schmidt and Cech 2015).

Negative regulators of telomerase terminate repeat synthesis when a certain number of repeats are already added. On average, about 60 nucleotides are added to the telomeric 3' overhang per telomere extension cycle. Data suggest the maximum number of repeats synthesised is controlled by interaction with the CST complex, consisting of Ctc1, Stn1 and

Ten1 (Chen, Redon, and Lingner 2012). Upon reaching a certain length, the telomeric primer is released by the association of CST in late S-phase/early G2-phase. Also, CST interacts with POT1-TPP1 and may inhibit TPP1 mediated stimulation of RAP and telomeric localization of telomerase (Chen et al. 2012; Zhao et al. 2011).

The mechanism of telomerase dissociation from the telomere in humans is unknown. Two popular hypotheses rising from other organisms provide insights on the dissociation process of human telomerase: the protein-counting model and replication fork model. The protein-counting model originates from observation in *S. cerevisiae*, where the repeat addition length was shown to be independent of telomeric repeats. Mutations in Rap1, the main telomeric DNA binding protein in yeast, and knock-out of Rap1 interacting partner RIF1 and RIF2 lead to abnormally long telomeres (Li et al. 2000; Wotton and Shore 1997). The repeat length addition is reduced when the number of Rap1p molecules bound to the telomere end increases. Since the telomere-binding proteins TRF1, TRF2, POT1, and TIN2 cause excessive telomere elongation, the mammalian telomerase may also adopt a similar protein counting mechanism.

The replication fork model arises from observations in ciliates. In *Oxytricha*, telomerase associated with the “replication band” and travels with the replication forks during the S phase (Fang and Cech 1995). In this model, telomerase repeat addition is linked to the movement of the replication fork, and telomerase travels with the replication fork and must reach the telomere ends to carry out its catalytic function. The telomere binding protein encountered along the way would increase the probability of telomerase dissociation. The replication fork model explains why telomerase is preferentially recruited to shorter telomeres, and how telomerase-binding proteins function as negative regulators of telomerase (Buonomo et al. 2009). Interestingly, this explains why hTRF1 and hTRF2 are negative regulators of telomerase activity whilst the terminal binding TPP1-POT1 complex has a

positive regulatory effect. However, the molecular detail of telomerase recruitment and elongation remains an interesting area of future research.

1.1.7 Telomerase structure analysis

Telomerase is a ribonucleoprotein complex that consists of a protein subunit (TERT) and an RNA component (TERC). The template region of TERC base pairs with the TTAGGG repeat acting as a template for telomere elongation, and together with the catalytic function of protein subunit TERT, adds TTAGGG repeats in telomere ends when activated (Figure 1.1.7). The full-length human TERT subunit has a molecular weight of 127kDa (Lingner et al. 1997). Telomerase has four conserved domains across various species: N-terminal TEN domain, TERC binding TRBD domain, the catalytic domain RT domain, and C-terminal CTE domain. Both the N-terminal domain and TRBD domain contains ssDNA binding folds, which stabilise telomerase - DNA binding in addition to TERC-DNA binding of TRBD domain. As described previously, the hTEN domain is flexibly linked to the catalytic domain, which serves as an anchor site for the DNA substrate (Figure 1.1.7). The human TEN domain's interaction with TPP1 is important during telomerase recruitment and elongation. (Zhong et al. 2012)

RNA binding motifs in the TRBD region are responsible for the specific binding of TERT to the TERC CR4/5 loop, and the complete assembly of TERT and TERC is essential for telomerase function. (Zhang, Kim, and Feigon 2011) As the name suggests, the reverse transcriptase domain shares structural and functional homology to the viral protein reverse transcriptase, carrying out the catalytic activity of nucleotide addition. Mutations in the catalytic site of reverse transcriptase domain of human telomerase results in a loss of activity (Sandin and Rhodes 2014) (Figure 1.1.9). The TRBD and RT domain of TERT are mostly conserved across different species due to their importance in assembly and activity, but the

rest of TERT shows great variation in sequence and size, suggesting a diversion in evolution. The structure of full-length human telomerase atomic-resolution is yet to be determined. Currently, limited structural information comes from crystal structures of homologous species and medium-resolution EM structures.

Figure 1.1.9 TERT protein domains and predicted TERC structure (adapted from (Sandin and Rhodes 2014)) The TERT consists of four domains, TEN, TRBD, RT and CTE. The TERC contains the core domain which includes the template region, CR4/CR5 domain, H/ACA and scaRNA-like domain.

The human telomerase RNA TERC component is 451 nucleotides long (Feng et al. 1995). The size for TERC varies greatly across species, from ~150bp in ciliates to ~1300 bp in fungi. Although the size variation is huge across species, the core region, transactivation domain region and the terminal region are mostly conserved due to their functional importance (Podlevsky and Chen 2016). The core domain contains the template region, which is complementary to the telomere repeat sequence. It is required for sequence recognition of telomeric ends as it functions as a template for repeat synthesis. Experiments demonstrated that both the transactivation region and the core domain region contribute to the interaction with the TERT protein subunit (Zhang et al. 2011). The terminal domain region of

human TERC contains an H/ACA scnRNA like structure, which binds to multiple H/ACA binding proteins. As described in the previous chapter, the H/ACA binding proteins are thought to play an important role in telomerase biogenesis and mutation of H/ACA binding protein DKC1 prevents telomere accumulation *in vivo*. The high-resolution structure of full-length human TERC is not available, and the only structural information comes from several NMR structures of important human TERC sub-domains (Chen and Greider 2004; Feigon et al. 2010; Theimer et al. 2008).

The minimal telomerase complex required for telomere repeat synthesis consists of hTERT and hTERC. Full assembly of the hTERT and hTERC is required to carry out its catalytic function. Previously, only the EM structure of *Tetrahymena* telomerase and the crystal structure of Beetle TERT in complex with short DNA-RNA helix were available (Mitchell et al. 2010; Theimer et al. 2008). However, these structures only provide very limited structural information on human telomerase. Only recently and during the period of my thesis work, more information becomes available through the two published EM structures of human telomerase (Nguyen et al. 2018; Sauerwald et al. 2013).

The physiological oligomerisation state of the human telomerase remains a debate. Whether telomerase functions as a monomer or dimer remains a puzzle. The existence of functional dimeric telomerase has been reported for human, yeast, and *E.crassus* telomerase. However, *Tetrahymena* telomerase was shown to function as a monomer. Telomerase was first proposed to be a dimer by S. Cohen in 2007 (Cohen et al. 2007). Data suggested that telomerase extracted from various immortalised cell lines sediments on density gradients at approximately the size of 670kDa which correspond to the size of two copies of TERT, TERC and DKC1 *in vivo* (Cohen et al. 2007). Both TERT and TERC were shown to have the ability to dimerise alone (Ly 2003; Wu et al. 2015). Two distinct domains region of RNA have been shown to have the ability to dimerise *in vitro*, however, there is no *in vivo* data for

this (Ly 2003; Ren et al. 2003). Collins group has demonstrated by single-molecule imaging, that telomerase exists in both monomeric and dimeric states (Wu et al. 2015). The recent cryo-EM structure of the human telomerase at 8Å resolution is of a monomeric complex that contains additional cofactors to TERT and TERC (Nguyen et al. 2018).

From the evidence in the literature, it can be concluded that is likely that the human telomerase exists in a dynamic state switching between monomeric and dimeric states. Two hypotheses suggest the human telomerase may function as a dimer at telomeric ends. One suggests that telomerase dimerization allows a parallel extension of a pair of telomeres on sister chromatids (Sauerwald et al. 2013). Another alternative suggests two monomers of the human telomerase acting in a co-operative manner with one monomer serves as the “anchor” during translocation while the other monomer is responsible for extension (Beattie et al. 2001). The physiological oligomeric state of human telomerase *in vivo* at telomeres remains to be uncovered.

The first structure of active human telomerase was published in 2013 by the Rhodes group (Sauerwald et al. 2013). The 25 Å resolution structure was solved by negative stain EM. The structure reveals telomerase to be a dumbbell-shaped dimer, but intriguingly, the conformation of two lobes is different. Although the structure is at low resolution, thus making the model fitting hard, nano-gold labelling experiments allow determination of the approximate location of TERT and TERC (Figure 1.1.10). TERC density was assigned to the stalk region, whereas TERT density is located at two ends. The holes in both lobes are thought to be the catalytic pocket of human telomerase. The previous publication showed that functional human telomerase must contain two copies of telomerase RNA (Wenz 2001). In addition, the Rhodes group presented experimental data showing that human telomerase has to function as a dimer biochemically and structurally (Sauerwald et al. 2013). The functional

implication of two different conformations of the two lobes in the telomerase dimer is unclear.

Figure 1.1.10 Negative stain structure of human telomerase at 20 Å (adapted from (Sauerwald et al. 2013)). (A)(B)(C) Front, side and top view of telomerase open lobe, dark blue indicates TRBD density and light blue indication RT domain (D) Electron micrographs of individual negatively stained telomerase containing His-tagged TERT subunit in complex with Ni-NTA-5nm colloidal gold (E) Distribution plot showing the distance between the Ni-NTA-gold and the dimer interface (y-axis), and the number of gold particles (x-axis) (F) Cartoon representation showing the interpretation of the EM density map.

Figure 1.1.11 Cryo-EM structure of human telomerase at 8 Å (adapted from (Nguyen et al. 2018)). (A) Schematic of subunit arrangements. (B) Front (left) and back (right) views of the cryo-EM reconstructions for the H/ACA lobe at 8.2 Å and the catalytic core at 7.7 Å, with fitted subunits colour-coded as indicated.

Later in 2018, the cryo-EM structure of human telomerase was published at 8 Å by the Collins and Nogales group (Nguyen et al. 2018). The EM density has been interpreted as monomeric telomerase consisting of one hTERT and one hTERC where hTERC is bound by two dyskerin, two NHP2, two NOP1, GAR1 and TCAB1. The structure shows the catalytic ring formation of human telomerase and allowed mapping of the position of the template region of hTERC. However, the fitting of the H/ACA region is ambiguous, there is little biological explanation of fitting of two copies of Dyskerin, and density for other proteins are obtained through simulation.

When comparing the EM density of the Rhodes negative stain structure and the Collins cryo-EM structure, it is clear there is a mismatch suggesting that they are structurally different complexes. Biochemically, since they contain different components, they must be different (Nguyen et al. 2018; Sauerwald et al. 2013). This suggests the human telomerase exists as multiple complexes in the cell, perhaps unsurprisingly due to its complicated biogenesis pathway as described previously. This poses an interesting question of what the physiological complex functioning at the telomeres is. This question is difficult to answer though due to the low abundance of telomeres and telomerase complexes, but perhaps the development of cryo-ET may aid to solve this question. After all, the atomic resolution structure of human telomerase remains to be solved.

1.2 The era of “Resolution Revolution” in biological cryo-electron microscopy

During my PhD, the main method used to investigate the structure of human telomerase was electron microscopy. Human telomerase is a ribonucleoprotein complex, which has been shown by previous work from the group to be 670kDa dimer consisting of two TERT protein and two TERC RNA subunits. Due to difficulties in producing a high concentration of human telomerase required for crystallography as well the rapid development in cryo-EM, the most suitable method for structure determination human telomerase is cryo-EM

This section of the chapter aims to give an overview of the basic principles of electron microscopy and how cryo-EM has evolved from the “blobology” era to producing near-atomic resolution structures of proteins, that in the past was only possible by using X-ray crystallography.

1.2.1 History of transmission electron microscope

The development of electron microscopes dates back to 1926 when Busch first researched the trajectories of charged particles in axially symmetric electric and magnetic fields and showed that such fields could focus electron beams and act as electron gun lenses (Busch 1926).

Electrons exhibit properties of both particles and waves, a property known as the wave-particle duality. French scientist de Broglie researched the concept of corpuscular waves, which define the frequency and wavelength of charged particles. The combination of both ideas initiated the development of electron microscopes.

The aim of developing electron microscopes is to observe material with better resolution than light microscopes. The first electron microscope was constructed in 1931 with a nominal magnification of 16x. In 1932, Knoll and Ruska tried to estimate the resolution limit of the electron microscope. Assuming the resolution limit formula of the light microscope was still

valid for electrons, the wavelength of electrons accelerating at a Voltage of 75 kV was used in the calculation. The resulted theoretical resolution limit was 0.22nm or 2.2Å. This provided a fundamental reason to drive the construction of electron microscopes. In 1933, Ruska built an improved 3-lens electron microscope with a resolution reaching 50nm surpassing the resolution of light microscopes by ten times. Their breakthrough led to the development of the first commercial microscope by Metropolitan-Vickers EM1 in the UK in 1936 (Ruska 1987). However, commercial electron microscopes became widely available by Siemens and Halske in Germany in 1939.

There are two main types of electron microscopes, the Transmission Electron Microscope (TEM) and the Scanning Electron Microscope (SEM). For biological specimens related to this dissertation, only TEM will be discussed in detail. TEMs became widely available from several other sources (Hitachi, JEOL, Philips, etc) after the Second World War. And after decades of development, the quality of objective lenses was improved, the spherical aberrations were reduced, and commercially available electron Microscopes could now achieve more than 100 000x magnification.

Figure 1.2.1 Schematic diagram of a modern electron microscope. The electron beam is generated by the electron gun, focused on the sample by condenser lenses. The objective lens system project the image onto an image recording system.

A typical modern electron microscope consists of an electron gun, a condenser lens, an objective lens, two projector lenses, and a fluorescent screening or charge-coupled device (CCD) camera that captures the images. In the old days, images were collected on film instead of by detectors. Customised additional components may also be added including sample autoloader or an energy filter which eliminates low energy electrons from scattering. The most recent development includes image recording with direct electron detectors, that allow accurate detection of signals from individual electrons improving the signal to noise ratio.

1.2.2 Electron microscope objective system enhances image contrast

Unlike the microscope, the objective system in electron microscopes is important for generating image contrast. In TEM, spherical aberrations and defocus of the objective lens is used to enhance the contrast of biological specimens. These aberrations distort the TEM images of specimens, which generate distortion in a contrast transfer function (CTF). The distorted contrast transfer function needs to be computationally corrected in order to recover the original signal. Instead of an objective aperture, a phase plate is inserted in the back focal plane of the objective lens to modulate the scattered wave and non-scattered wave with an additional phase difference of $\sim\pi/2$. In TEM, such modulation could convert the CTF from a sine-shape function to a cosine-shape function, so low-frequency signals are preserved better and the image contrast is greatly enhanced.

Figure 1.2.2 Types of electron scattering through a specimen. Black lines illustrate scattering types with no energy loss that can generate usable signals. Red lines illustrate scattering types with energy loss that produce signals that are not usable. (adapted from (Russo and Passmore 2016))

The Volta phase plate (VPP) is a thin film without holes that uses electro-potential charging properties that generate a Volta potential. This allows smaller protein particles to be judged and picked from raw micrographs, which otherwise have a contrast that is too weak to be picked by human eyes (Danev, Tegunov, and Baumeister 2017; Wang and Fan 2019).

1.2.3 Development of detectors in electron microscopes

Since the electrons are not visible by eye, it requires special detectors to produce an image.

The most basic electron detector is a fluorescent screen, where a phosphor is excited by electrons and emits visible light, providing light and dark contrast. However, more advanced detectors are required to preserve signal and capture high-resolution images.

Detector performance is very important when dealing with radiation-sensitive samples such as biological samples studied in electron cryo-microscopy (cryo-EM). Biological samples are damaged when they are struck by electrons, and the electron dose must be kept within a reasonable range ($<100\text{e}/\text{\AA}^2$) to preserve the natural state of the sample. Low dose means the images will have an inherent low signal to noise ratio. The amount of additional noise added by a detector is measured by its detective quantum efficiency (DQE) which is defined as the square of the ratio of the output signal-to-noise, SNR_o , to that of the input, SNR_i , i.e.,

$$\text{DQE} = \text{SNR}_o^2 / \text{SNR}_i^2.$$

Theoretically, a perfect detector has a DQE value of 1.

Figure 1.2.3 The DQE of different detectors. (adapted from (Faruqi, Henderson, and McMullan 2015; McMullan et al. 2014)) The DQE of K2, Falcon II, DE-20, Film and FalconIII at different fractions of Nyquist frequency.

In the best practical cases, a typical CCD camera with 15 μm pixel size can record usable signals up to 2/3 Nyquist frequency (Chen et al. 2008a). Unfortunately, new CCD cameras with smaller pixel sizes have not resulted in improved resolution, suggesting that current CCD technology has reached its limit.

While the K2 camera has a very high internal frame rate (400 frames/s), it's usually operated at a dose rate of around $5\text{--}8 \text{ e}/\text{\AA}^2$, because coincidence loss becomes significant at dose rates above $\sim 8 \text{ e}/\text{pixel}^2/\text{s}$ (Li et al. 2013). The electron counting mode has helped to push the resolution limits of cryo-EM, and many 3D structures are produced near-atomic resolution including the recent map of β -galactosidase with an estimated resolution of $2.2 \text{ \AA}$ (Bartesaghi et al. 2015).

Compared to the commonly used detectors (Figure 1.2.3), the K2 Summit gives the highest DQE at low spatial frequencies. With increasing spatial frequency, the DQE of the K2 Summit falls roughly as the MTF^2 but the DQE of the DE-20 and Falcon II stays constant up

to $1/2$ the Nyquist frequency, due to parallel falls in their respective noise power spectra. The DQE falls for all detectors from $1/2$ Nyquist towards the Nyquist frequency due to increasing contributions from an alias in the noise power spectra. Beyond $3/4$ of the Nyquist frequency, the DQE of the Falcon III becomes the highest of all the detectors. However, at below $1/2$ Nyquist, K2 remains the best detector, but above $1/2$ Nyquist, DQE of Falcon III is much higher than K2. Higher DQE/SNR gives better constant and more 2D information of specimens, which in turn leads to higher resolution data. The improvement of detector DQEs made a great contribution to the resolution improvement of electron density maps obtained by cryo-EM (Bammes et al. 2012).

1.2.4 Electron microscopy for biological specimens

Electron microscopy has been widely used by material scientists to investigate the structure of hard matter. However, biological specimens are much more sensitive to radiation damage from the electron beams and have far significantly lower contrast. The molecular details of biological samples can only be resolved using high-resolution imaging, which can only be achieved by using electron microscopes.

Porter published the first picture of an intact cell taken with an electron microscope in 1945 (Porter 1945). The cell was fixed with osmium tetroxide. Their TEM image magnified the specimen 1600 times and their first electron micrograph of a cell revealed the mitochondria.

Figure 1.2.4 Electron micrograph of a portion of a cultured cell (Porter 1945). Details of outline and cytoplasmic content are clear except in the central portion where the density and thickness of the cell have prevented penetration of electrons.

These problems were first addressed by new techniques for dehydration and freezing substitution, and electron beam damage was reduced through absorption by staining with heavy metals (Glaeser 2008). Heavy metal stains such as urinal acetate act as a fixative which preserves the structure and enhances the contrast of biological samples. However, osmium is a heavy metal that contributes to electron density when bound to cellular structures. In addition, since the infiltration and diffusion of stains and fixatives into the cells is low, fixation artefacts would be introduced.

For biological samples of macromolecular complexes (virus, protein), negative stain methods of biological macromolecular particles were first clearly described by Brenner and Horne (Brenner and Horne 1959). Since the contrast of an electron micrograph is generated based on the “thickness” of a biological specimen, biological macromolecules that are less than

10nm thick is difficult to see. When biological macromolecules are stained with a heavy metal salt solution and layered on to TEM grids, the heavy metal binds to the surface of the macromolecules giving them contrast on a light background (Thompson et al. 2016).

Figure 1.2.5 Positive and negative staining using heavy metal salts. (A) In negative staining, the stain fully envelops the macromolecular complex; in the micrograph, the complex appears white on a dark background. (B) Positive staining results in a small amount of stain forming a thin shell around the molecule, meaning the sample appears dark against a light background.

However, the resolution obtained by this method still does not give atomic details. The effect of heavy metal staining severely affects the accurate determination of macromolecule structures. Stains can cause structural changes in macromolecules such as shrinkage due to dehydration, and also bind to the outer surface and add additional thickness. Stains can also lead to more inelastic electron scattering, causing signal loss and reducing the resolution even further. In combination, these effects limit the resolution of negatively stained EM to about 1-2nm, while in the past, the resolution of structures obtained via this method was around 5nm.

Figure 1.2.6 FEI Vitrobot and a cryostat device developed by the Laboratory of Molecular Biology (LMB) (Russo, Scotcher, and Kyte 2016). The FEI Vitrobot is a semi-automated plunge freeze device. The cryostat designed by LMB keeps the liquid ethane in a constant temperature range to prevent solidification.

A breakthrough was made in 1981, when in order to preserve the nature of the sample and protect the sample from radiation damage, Dubochet and Alasdair McDowell presented a method for producing a thin-film of non-crystalline samples on a specimen grid (Dubochet and McDowell 1981). The liquid sample was rapidly frozen in 90k, bypassing the formation of crystalline ice, and creating vitrified ice. This technique allowed biological samples to be imaged in their native state without staining artefacts, which simplifies the image interpretation process. This method laid the foundation of modern cryo-EM as a tool for structure investigation at atomic resolution.

1.2.5 Cryo-EM sample preparation

Over the years, cryo-EM sample preparation has become more standardised (Figure 1.2.7). In order to avoid damage, samples are frozen at cryogenic temperature. The sample preparation process involves multiple steps. A protein sample is purified from a source and applied onto the EM grids. The grids are then blotted with filter paper to form a thin layer of liquid and

plunge frozen in liquid ethane. The speed of freezing has to be very rapid, and temperature must be kept under -137°C to avoid the formation of ice crystals. (Dubochet et al. 1982; Fernandez-Leiro and Scheres 2016). If the temperature is above -160°C , areas of cryo-grids can be severely compromised by crystalline ice, resulting in poor quality micrographs (Fernandez-Leiro and Scheres 2016). Grids must be frozen extremely rapidly, at a rate of 10^6 $^{\circ}\text{C}$ per second. Liquid ethane is used instead of liquid nitrogen because the specific heat capacity allows rapid freezing of grids. Liquid nitrogen, however, due to its low specific heat capacity, when in contact with a grid at room temperature is affected by laidenfrost effect, and liquid nitrogen immediately evaporates forming a layer of gaseous LN2, which has an insulating effect and slows down the freezing process causing the formation of crystalline ice. Liquid ethane and propane mix have also been shown to work for this purpose. Currently, the process is semiautomated with the development of the FEI Vitrobot and Leica plunger, while more automated sample handlers are constantly being developed.

Figure 1.2.7 The workflow of structure determination by cryo-EM. (Fernandez-Leiro and Scheres 2016) A protein sample is applied to EM grids and plunge frozen. The grids containing the protein is imaged by an electron microscope. Multiple images of the EM grids are recorded and processed to give rise to the 3D structure of the target protein.

The grid and support itself can affect particle distribution on grids, which may affect the final resolution. This was best characterised by Passmore and Russo. The most common material grids are made from is copper or gold. The most common grid support foil material is amorphous carbon. Grid supports contain holes (holey grids) or include an additional support film covering the holes. The surface used may hugely affect the particle density, distribution and orientation.

Figure 1.2.8 Design of cryo-EM specimen supports (adapted from (Russo and Passmore 2016)). Top view and section diagrams of typical specimen support geometries, comprising a perforated foil on a metal mesh grid.

Chris Russo (Russo and Passmore 2016) was the first to test the character of gold as a material for grid support foil. During imaging, beam-induced motion occurs due to high energy electrons hitting the surface, and different support materials perform differently when hit by electrons. Gold supports show the least beam-induced movement.

Gold and graphene support was developed by Chris Russo from MRC, LMB. The surface can be functionalised to have graphene and graphene oxide. Methods to produce reliable graphene oxide (Palovcak et al. 2018; Pantelic et al. 2010) and graphene grids (Han et al. 2019; Naydenova, Peet, and Russo 2019) has been described. Further developments include the functionalisation of graphene (D’Imprima et al. 2019) and graphene oxide surfaces (Wang et al. 2019). Both methods have been shown to modify the surface properties and have different effects on the binding and distribution of different samples. The change in distribution may be subtle but may have a profound effect on particle data processing in later stages. However, the interaction between the protein sample and grid surface cannot be

accurately predicted. The optimal support for different proteins is different and still requires a process of trial and error.

1.2.6 Single-particle cryo-EM and processing

After successful grid preparation capturing macromolecular complexes either by negative stain or freezing, the following step is to acquire images. Single-particle analysis describes the 3D reconstruction of an object based on 2D images of individual particles in different orientations. This method was first used to solve the structure of the tail of T4 bacterial phage by De Rosier and Klug in 1968 (De Rosier and Klug 1968). Later, the development of computer programs and more powerful computers has allowed combining images with similar orientation together (classification) to enhance the contrast.

Both software development and hardware development allow faster and easier processing of single-particle data. Software development including improved algorithms and utilisation of graphics processing unit (GPU) achieved a processing speed of 10-50 times faster than previous algorithms using central processing unit (CPU).

In the “old days”, images captured on films were scanned for processing. Technology developments resulted in film first being replaced by CCD camera images and now surpassed by newly developed direct electron detectors. As described in the previous chapter, current detectors capture images that have better SNR and allow better processing.

The first step in EM data processing is to correct image motion induced by microscopes.

When images are captured by camera/detectors, the captured images have local translational movements due to microscope instability and/or beam-induced movement. The latest technologies allow data acquisition in movie mode, which split the exposure duration into frames, and the translational movement between frames can be measured and re-aligned by a

computer algorithm (Zheng et al. 2017), creating sharper corrected images compared to the uncorrected images. (Figure 1.2.9) Scheres has recently introduced a new algorithm based on a Bayesian approach for motion correction (Zivanov, Nakane, and Scheres 2019).

Figure 1.2.9 Translational alignment of movie frames to reduce blurring in images affected by beam-induced motion. Adapted from (Brilot et al. 2012) Aligned cryo-EM images of viral particles before (A) and after (B) motion correction by a computer program.

Once the images are obtained and corrected, then the particles may be picked from the micrographs. The particles are “cut” or extracted out from the micrographs. Traditionally, particle picking was done manually, where individual particles were recognised by eye and picked. Recent advancements in computation methods allow picking to be done more accurately and automatically. Two approaches are predominating; template-based, and generic Gaussians based methods. Template-based autopicking requires an input of (Gautomatch, Relion picker) 2D classes, and the program picks particles with similar size and orientations from the micrographs. Non-template picking requires an input of particle size, and the program generates a generic model based on “difference of Gaussians”, where a broad 2D Gaussian function of negative amplitude is subtracted from a narrow 2D Gaussian of positive amplitude. In generic turns, the program picks regions with high SNR and the

same size. However, this method is less efficient because frequently ice/ethane contamination and carbon regions that have high SNR would also be recognised as particles in such methods. Recently, a new particle (cryolo) picking algorithm based on deep learning has also shown to be effective (Wagner et al. 2019). Instead of requiring an input of 2D projections, only input of particles is required, and 100 particles are usually enough. Typically, 20-100 particles may be obtained per micrograph depending on the size of protein and magnification used. The collection of particles is saved as “particle stacks”.

As introduced in Chapter 1.2.2, an objective defocus is routinely applied during the collection process in order to enhance the image SNR, and together with the spherical and chromatic aberrations, they contribute to the distortion of the CTF from the micrograph. Thus, during data processing, the distorted CTF needs to be accounted for and corrected. The CTF is measured and the spectrum is compared and fitted with a spectrum corresponding to different theoretical values. This process is known as CTF correction and the following step uses the information of the image’s Fourier transform.

Figure 1.2.10 Schematic of single-particle reconstruction Protein purification. Negative stain is useful to check the sample before going to cryo-EM for data collection (Carroni and Saibil 2016)

The basic principle of cryo-EM data processing is illustrated in Figure 1.2.11. From the images, different 2D projections of the specimen are obtained. Then different 2D projections are mathematically transformed into 2D Fourier space. Each 2D projection in Fourier space is

a section for a 3D model in Fourier space, and with the aid of computational algorithms described later in this chapter, 2D Fourier transforms are orientated and compiled together to build a model in 3D Fourier space. The 3D Fourier model undergoes a process called Fourier inversion, and as a result, a 3D density map in real space is obtained.

Figure 1.2.11: Principle of single-particle electron microscopy reconstruction (adapted from (Milne et al. 2013)) The 2D projected images are converted into signals by Fourier transforms. The transformed images are used to build densities in 3D Fourier space, which is converted back to real space densities by Fourier inversion.

The basic principle may appear to be straight forward, however, there are difficulties during each step. Both the low SNR of individual cryo-EM images and the distortion due to CTF effects make it difficult to evaluate the images in a particle stack. Multiple 2D images of the same projection are clustered together into a class and represented by a class average, and the addition of signal makes the protein features more visible. In a high-resolution dataset, secondary-structure elements such as alpha-helices can be seen from the 2D classes.

The clustering of similar particle images was first introduced by van Heel and Frank (Van Heel and Frank 1981) and the problem has received much attention; for a recent comparison of methods, see Zhao and Singer (Zhao and Singer 2013). The notable method, which is heavily used in this thesis, is the maximum-likelihood 2D classification implemented in Relion (Scheres 2012). The program starts with a set of random ‘reference’ images. Then for each particle image, probabilities are computed with respect to its rotation, translation and degree of matching to each reference. Then the translational and rotational parameter with maximum probabilities is identified and applied to that particle image. The aligned images and their 2D Fourier transform together with CTF weightings are combined to form the overall average images. The resulted aligned images have their signal reinforced, and therefore represent a much clearer view than that of individual particles.

The 3D classification algorithm is similar in principle to the 2D classification of images. However, in 3D space, there is an additional parameter of the Euler axis and angle, requiring greater computational power. A 3D initial model input is required by most programs. The model may be generated by many means: volume with similar dimension filled with Gaussian noise, density extracted from tomogram, previous reconstruction from negatively stained EM, or most recent development allows the building of ab-initio model from the same dataset (Reboul et al. 2018). The model is used as a “reference”, the set of probabilities computed for each particle image that also includes the probability that the image arises from each of several different 3D densities.

After multiple iterations of refinement, the translational and angular parameters of each particle image related to a 3D model is usually unambiguous. The same result should be obtained from reconstructions of images from separately assigned particle stacks. However, the 3D classification may not work well due to the influence of the 3D models given, and many structures often fail to be discovered. In practice, one usually performs multiple rounds

of classification to obtain a population of particle images yielding the best-resolution structure. Recently software developments also allow masking, a technique to “hide” a certain region while processing others, which allows focused alignment of the unmasked region, and in some cases, this method can have a profound effect in the final resolution of the EM density map (Galej et al. 2016; Nguyen et al. 2015, 2016; Yan et al. 2015).

After obtaining the 3D reconstruction, the model can be reprocessed in Relion by polishing individual particles that are refined and reextracted from the raw movie files. Although the micrographs have been corrected during the first steps, this step allows a finer correction on each individual particle used in the final reconstruction and further improve SNR of individual particles. Then the final EM density map is exported.

The final map resolution estimate is based on the Fourier shell correlation (FSC). To measure the FSC, two volumes are the result of two three-dimensional reconstructions, each based on half of the available data set (Liao and Frank 2010). FSC of unmasked particles is affected by the real-space voxels that arise from the solvent background, leading to worse resolution estimates than the protein region (protein only occupies <10% of the volume).

Relion and cisTEM (Grant, Rohou, and Grigorieff 2018; Scheres 2012) follow two alternative ways to address this problem. In Relion, a mask is used which flattens the solvent, and phase randomization is performed to minimise artificial correlations due to the mask. In cisTEM, no mask is used. Instead, the FSC curve is computed between two unmasked volumes and then scaled mathematically to “counter-act” the fact that many voxels are just solvent. The resolution is an estimated RSC that uses the correlation coefficient cut off at 0.5 or the Scheres’ suggestion of a Golden standard cut of at 0.143 for a more reliable estimate. Programs designed to estimate the local resolution of densities are proven to be helpful in cases where the flexible part of macromolecule affects the overall resolution estimated of the whole density (Scheres and Chen 2012).

1.2.7 Cryo-electron tomography – directly studying proteins from cell

Figure 1.2.12 Structural Analysis of the Translocon in situ (Beck and Baumeister 2016). (A) and isosurface rendering (B) showing isolated ER vesicles that assume a spherical topology in isolation. (C) Cryo-EM structure of the translocon complex together with the translocon associated protein complex. Tomographic slice (D) and isosurface rendering (E) of the outer nuclear membrane observed *in situ sensu strictu*. Ribosomes form clearly recognizable linear polysome chains on the low curvature membrane.

One important aspect of cell biology is to study the localisation and interaction of macromolecular assemblies in cellular systems *in situ* and at different scales of resolution. Much has been learned in the identification and characterization of the molecular structures involved in cellular processes through biochemical experiments, but little is known about the spatial distribution and organisation of macromolecules within the cell.

At present, cryo-ET is the only 3D imaging method that can image cells and their organelles in their native state at molecular resolution (Gruska et al. 2008; Nickell et al. 2007; Sartori et al. 2007). Important applications of cryo-ET include the 3D structure of the proteasome; the 70S ribosome (Ortiz et al. 2006) and the 3D molecular mapping proteins in an entire cell. The

Briggs group produced many density maps of viruses at a resolution of 8-10Å using cryo-ET to uncover the molecular details of viral infection mechanisms. (multiple citations)

Figure 1.2.13 In-situ cryo-lamella milling (Rigort and Plitzko 2015). (a)(d) Illustration showing targeting of a region of interest to make it accessible for electron cryo-tomography. Milling from (b) top and (c) bottom of the lamella. (e)(f) lamella after milling

A typical eukaryotic cell has a thickness of $\sim 10\ \mu\text{m}$. In TEM, electrons can penetrate samples with a thickness of only 100~500nm without significant loss of signal due to inelastic scattering. The focused ion beam (FIB) system was developed to produce thin cell lamellar and provided a solution to this problem. While operating in a cryogenic temperature, FIB technology can be used to produce sections (lamella) in situ, with a thickness suitable to be analysed by TEM.

Figure 1.2.14 Illustration of the influence of the tilt range and tilt increment on the quality of tomographic reconstructions (Nudelman, de With, and Sommerdijk 2011). As the tilt limit reduce, the blurriness of the images caused by missing wedge issue increases.

Once lamella is obtained from FIB and transferred to TEM grids, they are ready for collection. The data collected are called tomograms, where the stage is tilted at an angle with angular increments to obtain information from all dimensions. However, the most predominant problem is the missing wedge. Also, at a very high tilt angle, the sections imaged becomes extremely thick and the image quality drops significantly, thus the range of usable tilt angles is limited. The highest tilt angle recommended is around ± 60 degrees. The missing 2-D information at tilt angle greater than 60 degrees means the 3D information is incomplete. This is then solved by sub-tomogram averaging as illustrated, where complementary information for other orientations is added to give a full final reconstruction. Internal symmetry of particles helps to identify the missing information and produce a better reconstruction; hence viral structures readily give high-resolution densities from cryo-ET.

Usually, sub tomogram averages of sufficiently abundant complexes typically reach resolutions in the range of 10–30 Å (Bykov et al. 2017; Guo et al. 2018). At this resolution, the conformational states can be determined with the possibility of discovering new interaction partners. The highest resolution obtained by cryo-ET of a non-virus protein target is 9 Å of COPI reported by Briggs, whereas the highest resolution obtained by cryo-ET of a virus is 3.4 Å for HIV, also obtained by Briggs group (Turoňová et al. 2017).

In this dissertation, cryo-ET is primarily used to characterise the distribution of telomerase particles in ice as used in Carragher’s work (Noble et al. 2018). However, the power of cryo-ET would provide a good perspective for the future of the project.

1.2.8 Resolution revolution

The term “resolution revolution” is often used to describe the development of the cryo-EM field from 2013-2019. The improvement of cryo-EM technologies has aided in the determination of hundreds of structure of protein complexes close to atomic resolution, which was previously not thought to be possible. In the past, structural determination from EM and cryo-EM was limited to large complexes or low resolution (20-30Å). The term “blobology” was used to refer to the density obtained by EM by crystallographers.

Around 2012, it was generally thought only large protein complexes (>500kDa) would be the prime target of cryo-EM because smaller proteins were not expected to produce enough contrast for analysis. The first structures obtained by cryo-EM with a resolution around 3.5 Å were published in 2013 (Campbell et al. 2012; McMullan et al. 2014) and the first structure obtained at a resolution at 2.8 Å maps in as the 20S proteasome in 2015 (Campbell et al. 2015) which allowed unambiguous fitting of the side chains.

By 2017, cryo-EM could produce atomic resolution structures of various proteins, as small as 95kDa (Merk et al. 2016). To date, a 52 kDa streptavidin (SA) is the smallest protein that has been resolved at a near-atomic resolution of 3.1Å using cryo-EM (Fan et al. 2019).

A combination of factors mentioned in this chapter has contributed to this huge leap of resolution in recent years. A new generation of detectors with lower DQE enables the recording of high-quality images. New software tools allow correction for microscope or bead induced motion during image acquisition. The development of the phase plate greatly boosts the contrast of small protein that would be difficult to be observed. New development in computational software allows more accurate and rapid processing of EM data. Densities generated at near-atomic resolution enables model building de novo, providing a powerful tool to solve unknown protein structures.

Figure 1.2.15 Number of cryoEM structures became available from 2005-2019. The data shows the number of structures deposited to PSB obtained using EM. The number of structures deposited per year shows a very rapidly increased from 2015 onwards. (Data from the PDB website (rcsb.org)(Berman 2000))

A cryo-EM review in 2015 mentioned that obtaining a high-resolution structure for small complexes is still difficult (Bai, McMullan, and Scheres 2015a). However, in early 2017,

near-atomic resolution structures of small macromolecular complexes (<500 kDa) are readily produced. In 2019, more than 500 structures at near-atomic resolution solved by cryo-EM have been deposited. This demonstrates that cryo-EM technology has rapidly evolved beyond expectation in the past few years. The main actors, Henderson, Dubochet and Frank, responsible for the “resolution revolution” in cryo-EM were recognised by the award of the Nobel Chemistry Prize in 2017.

1.3 Research aim

Telomerase is an important enzyme involved in developments, ageing and cancer. However as mentioned above, many aspects of human telomerase are still poorly understood. The molecular details of human telomerase biogenesis, recruitment and regulation remain to be discovered. Uncovering those mysteries may potentially provide key information on controlling cancer cell immortalisation as well as rescue telomerase activity deficiency-related diseases such as Dyskeratosis Congenita.

The project aims to answer the question from a structural approach. Despite many previous studies and work done within the Rhodes group, the yield of telomerase from overexpression was still very poor. Therefore, the first focus of my project was to establish and optimise an expression and purification protocol to obtain a concentrated sample of full-length human telomerase for biochemical and structural characterisation. The next focus was to characterise the composition of the purified human telomerase complexes by biophysical and biochemical methods. The third and most important aim was to obtain a near-atomic resolution structure of human telomerase using cryo-EM methods. Determining the structure should uncover details of the mechanism of the telomere elongation cycle, and further help us understand the translocation mechanism. Furthermore, by assembling the TPP1-POT1-telomerase complex, we hoped to understand the molecular mechanism behind TPP1-POT1 induced positive

regulation of human telomerase. Understanding the molecular structure of telomerase also provides insights on novel methods of telomerase inhibition by targeting the assembly and recruitment of the human telomerase.

This project aims to obtain the three-dimensional structure of human telomerase by using cryo-EM, which is a newly developed and rising method for structural determination.

Structural determination of protein using cryo-EM is becoming routine for structurally homogeneous complexes, yet it remains difficult to obtain the structure of flexible proteins or complexes. This project also focusses on the utilization of functionalised EM grid supports, and collection and processing methods that will help establish protocols for obtaining structures of flexible proteins or complexes.

Chapter 2 Materials and methods

2.1 Biochemical methods

2.1.1 Molecular cloning

Mammalian expression vector encoding hTERT and hTERC was created by multiple rounds of restriction and ligation using multiple cloning sites following the standard protocol (Sambrook, Fritsch, and Maniatis 1989). The vector encodes 2xStrepII-EGFP-TERT-3xFLAG construct with TEV and HRV 3C site as shown below. TERC is expressed under U3 promoter in the opposite direction. The vector also contained the puromycin resistance gene. The final plasmid is named as pSTL1.

2.1.2 Transfection

Human embryonic kidney HEK293T cell lines were maintained in Dulbecco's modified Eagle's medium supplemented with 15% fetal calf serum (FCS), 2mM L-glutamine and 1 mM sodium pyruvate at 37° C in a humidified 10% CO₂ incubator. Transient transfections were performed with Lipofectamine 3000 (Invitrogen) according to the manufacturer's instructions. Transfected cells were cultured for 24–36 h prior to testing for transgene expression.

2.1.3 Cell line selection

Stable cell lines over-expressing telomerase were obtained by transfecting HEK293T cells with the pSTL1 plasmid. Polyclonal populations of telomerase-expressing cells were obtained by puromycin selection 48 h post-transfection. Cells were then further screened for enhanced telomerase activity and TERT overexpression. To generate monoclonal lineages of telomerase-over-expressing HEK293T cells, the polyclonal cell lines were further subjected to single-cell isolation using fluorescence-activated cell sorting (FACS) and EGFP-tagged TERT as a fluorescent marker. Following expansion, each monoclonal stable cell line was assessed for enhanced telomerase activity and TERT over-expression.

2.1.4 Cell line growth for telomerase expression analysis

HEK293T cells stably over-expressing telomerase was grown in 15cm TC treated dishes (Nunc), with DMEM high glucose (Gibco) supplemented with 15% FBS (Gibco). Cells were washed with ice-cold PBS (Gibco) and then lifted by rinsing with PBS (Gibco). Cells were spin down at 300xg for 10mins at 4°C, then flash-frozen in liquid nitrogen, stored at - 80°C before use.

2.1.5 Large scale cell culture with HYPERFlask

HEK293T cells stably overexpress telomerase were cultured in 15cm TC treated dishes (Nunc), with DMEM high glucose (Gibco) supplemented with 15% FBS (Gibco). When grown to full confluency, the cells were passaged into HYPERFlask (Corning) following standard protocol. Cells were grown until fully confluent, usually around 6 to 7 days before harvest. The HYPERflask were first washed with 50ml PBS at room temperature, then 200ml of ice-cold PBS (Gibco) was added and gently shaken to release the cell from the surface. The detachment of cells from the growth surface can be easily observed. PBS containing cells were spun down at 300xg for 10min, flash-frozen in liquid nitrogen, and kept at -80 °C for future use.

2.1.6 Human telomerase purification protocol

Cell pellets obtained from six HYPERflasks (~20g in total) were lysed in 100ml of Lysis buffer (50mM HEPES-OH pH8.0, 300mM KCl, 2mM MgCl₂, 10% sucrose, 1% TritonX-100 (Sigma), 1mM DTT, and Roche cOmplete™ Protease Inhibitor Cocktail) tablet on ice for 30mins. The cell lysate was clarified by spinning at 40,000 xg for 20mins. The clarified lysate was incubated with 3ml volume of StrepTactin (GE) resin (Sigma) for 5hours on a roller at 4°C to capture telomerase carrying Strep II tag. Resin was washed with 100ml wash buffer (50mM HEPES-OH pH8.0, 300mM KCl, 2mM MgCl₂, 10% sucrose, 0.1% TritonX-100 (Sigma), 1mM DTT) to remove non-bound material. The telomerase was then eluted by

incubation with 10ml elution buffer (Wash buffer +50mM Desthiobiotin) for 2hours. The eluted sample was diluted to have a final KCl of 100mM and loaded on pre-equilibrated 1ml of heparin beads for 30mins. Beads were washed with 30ml wash buffer and eluted with 3ml of elution buffer (50mM HEPES-OH pH8.0, 1M KCl, 2mM MgCl₂, 10% sucrose, 0.1% TritonX-100 (Sigma), 1mM DTT). The eluent was immediately diluted to 300mM KCl and concentrated using a spin concentrator (20ml, MWCO 100kDa) to a final volume of 500µl before size fractionation using a sucrose gradient. 10-50% sucrose gradient (Wash buffer with 10% or 50% sucrose as a base buffer) was made using gradient maker (Biocomp Gradientmaster 109) in 4ml SW60 Ti UltraClear tubes (Beckman) and kept undisturbed in a cold room for at least 1 hour. For cryo-EM work, the sample was crosslinked by adding glutaraldehyde to a final concentration of 0.4% in the 50% sucrose buffer using GraFix protocol (Stark 2010). A volume of 700µl was removed from the top of the tube; 500µl of telomerase sample was loaded on to the gradient. The gradient was spun at 215,000xg for 18 hours and then fractionated into 200µl fractions manually. Peak fractions F11 to F13 were pooled and concentrated using a spin concentrator (500µl, MWCO 100kDa). For cryo-EM grid preparation, the sample was buffer exchanged by diluting using grid buffer (50mM HEPES-OH pH8.0, 150mM KCl, 2mM MgCl₂, 1mM DTT) in a 1:10 ratio for a total of 4 times. Finally, the telomerase sample was concentrated to a final volume of 30µl to 50µl at 0.01mg/ml concentration.

2.1.7 Affinity purification of telomerase using a telomeric substrate

After the heparin ion exchange step, the telomerase sample was purified using substrate-based affinity purification. 5nM desthiobiotinylated DNA ligand (Table below) was added to the telomerase sample for binding on ice for 2 hours, the mixtures were then incubated with pre-equilibrated 5ul neutravidin beads (Thermofisher) for capturing of the DNA bound telomerase on beads for 4 hours. The beads were washed with 1.4 ml of wash buffer for 4

times. Telomerase was released from the beads by competitive elution with 500ul buffer containing biotin. After the substrate-based purification step, the telomerase was further processing using GraFix as described before.

Primer name	Sequence (5' - 3')
C Strand	GACTATGAGTACGCGCTTTGCACGAGAG/ideSBioTEG/T
G Strand	CTCTCGTGCAAAGCGCGTACTCATAGTCAGTG TTAGGGTTAGGGUUAGGGUUAG (U = 2'-O-Methyl RNA Bases)

2.1.8 Purification of the HSP90/Telomerase complex

Cell pellets from six HYPERflask (~20g in total) were lysed in 100ml of Lysis buffer (50mM HEPES-OH pH8.0, 50mM KCl, 2mM MgCl₂, 10% sucrose, 1% TritonX-100 (Sigma), 1mM DTT, Roche cOmplete™ Protease Inhibitor Cocktail) tablet on ice for 30mins. The cell lysate was clarified by spinning at 40,000 xg for 20 mins. The clarified lysate was incubated with 3ml StrepTactin (GE) resin (Sigma) for 5 hours on a roller at 4°C. The resin was washed with 100ml wash buffer (50mM HEPES-OH pH8.0, 50mM KCl, 2mM MgCl₂, 10% sucrose, 0.1% TritonX-100 (Sigma), 1mM DTT), then eluted by adding TEV protease to 10 ml of wash buffer to a final concentration of 50µg/ml of for 4 hours. The eluted sample was concentrated using a spin concentrator (20ml, MWCO 100kDa) to a final volume of 500µl before loading onto a sucrose gradient. 10-50% sucrose gradient (Wash buffer with 10% or 50% sucrose as a base buffer) was made using gradient maker (Biocomp Gradientmaster 109) in SW60 Ti UltraClear tubes (Beckman) and kept still in a cold room for at least 1 hour. For cryo-EM work, the sample was crosslinked by adding glutaraldehyde to a final concentration

of 0.4% in the 50% sucrose buffer. 700µl was removed from the top of the tube; 500µl of sample is loaded on to the gradient. The gradient was spun at 215,000xg for 18 hours and manually broken into 200µl fractions. F11 to F13 were pooled and concentrated using spin concentrator (500µl, MWCO 100kDa), the sample was buffer exchanged by diluting using grid buffer (50mM HEPES-OH pH8.0, 50mM KCl, 2mM MgCl₂, 1mM DTT) in a 1:10 ratio for a total of 4 times, concentrating to a final volume of 30µl to 50µl.

2.1.9 SDS-PAGE analysis

Denatured sample of protein was made by mixing protein sample with homemade 4XSDS sample buffer and boiled for 5 mins. SDS samples were run with 4-12% Bis-Tris PAGE gel (Nupage) in MOPS SDS buffer at 200V for 50mins. The gel was either stained with Coomassie blue or silver stain (Thermofisher) following the standard protocol. Images are captured using a CCD camera (Biorad Chemidoc Touch) in optimal exposure mode.

2.1.10 Western blot analysis

Denatured protein samples were separated by running gel electrophoresis with 4-12% Bis-Tris PAGE gel (Nupage) in MOPS SDS buffer at 200V for 50mins. After gel electrophoresis, the protein was transferred on to a PVDF membrane (Biorad) using wet transfer (XCell II™ Blot Module, Life technologies). The membrane was blocked with 3% BSA in TBST buffer (20mM Tris pH 7.5, 150mM NaCl, 0.1% Tween 20) at room temperature for 1 hour. The membrane was incubated with primary antibodies for 1 hour at room temperature, subsequently washed with 3 changes of TBST at 10-minute intervals. The membrane was then probed with the corresponding secondary antibodies (1:5000 dilution), for 1 hour, and washed with TBST as described. The membrane was developed by addition of ECL substrate (Advansta) and imaged on a CCD camera (Biorad Chemidoc Touch) in optimal exposure mode. List of antibodies used:

Antibody Name	Catalogue Number	Dilution Factor
hTERT	Sc393013	1:500
TCAB1	Ab99376	1:2000
DKC1	Ab156877	1:3000
NOP10	Ab133726	1:2000
NHP2	Ab172481	1:2000
Strep II	Ab76949	1:1000
Anti-goat HRP	NA935	1 :5000
Anti-rabbit HRP	NA9340	1 :5000
Anti-mouse HRP	NA9310	1 :5000

2.1.11 qTRAP Telomerase activity assay

The telomerase activity in whole-cell extracts and in the enriched telomerase fractions was analysed by the quantitative TRAP assay (qTRAP, (Herbert et al. 2006)). An aliquot of 250 ng protein from cell extracts or 2 µl of the enriched telomerase fractions was incubated (30 min, 30° C) in 25 µl qTRAP reaction buffer (20 mM Tris-Cl (pH 8.3), 1.5 mM MgCl₂, 63 mM KCl, 0.05% (v/v) Tween-20; 1 mM EGTA) supplemented with 50 µM dNTPs, 0.35 µM (each) forward TS and reverse ACX primers (Table below), 1 U/µl murine RNase inhibitor, 0.4 mg/ml BSA, 10 µM SYTO™ 82 fluorescent nucleic acid intercalating dye and 0.04 U/µl hot start Taq DNA polymerase (NEB). The telomerase extension reaction was terminated by 30-s heat inactivation at 95° C. The telomerase-extended TS primers were amplified (50 cycles: 30 s at 95° C, 30 s at 55° C, 30 s at 68° C) on a Bio-Rad CFX96 real-time PCR detection system. The relative telomerase activity was determined using a standard curve and a linear equation model (Herbert et al. 2006).

Primer name	Sequence (5' - 3')
TS primer:	AATCCGTCGAGCAGAGTT
ACX primer:	GCGCGGCTTACCCTTACCCTTACCCTAACC
NT primer:	ATCGCTTCTCGGCCTTTT
TSNT primer:	AATCCGTCGAGCAGAGTTAAAAGGCCGAGAAGCGAT

2.1.12 RNA analysis by quantitative (RT-qPCR)

For quantitative TERC RNA analyses, total RNA was extracted with TRIzol reagent according to the manufacturer's instructions. Reverse transcription was performed using the iScript™ cDNA Synthesis Kit (Bio-Rad). Quantitative PCR reactions were performed in technical duplicates using the SsoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad), according to the manufacturer's instructions, on a Bio-Rad CFX96 real-time PCR detection system and analysed with the CFX Manager 3.1 Software (Bio-Rad). Relative transcript levels were determined using the 2-ΔCt method. For each primer pair (Table below), the efficiency of amplification was determined to be equal or superior to 1.8. Control reactions lacking the reverse transcriptase or template RNA confirmed the specificity of the amplification reactions.

Primer name	Sequence (5' - 3')
264_Fw_TRAP-TS	AATCCGTCGAGCAGAGTT
267_Rv_TRAP_ACX	GCGCGGCTTACCCTTACCCTTACCCTAACC
361_Fw_TERC	TCTAACCCTAACTGAGAAGG
362_Rv_TERC	GTTTGCTCTAGAATGAACGG

2.2 Electron microscopy methods

2.2.1 Grid preparation for negative stain EM

For negative stain EM, grids were either homemade by depositing 30nm carbon on 300 copper meshes grids or purchased from Electron Microscopy Science (CF300-Cu). Prior to sample application, grids were glow-discharged for 60 seconds in a plasma chamber (0.2mBar and 40mA). 4µl of the sample was applied to the grids and incubated for 1 min at room temperature. The grids were manually blotted using 3MM filter paper, and 4µl of 2% UA was immediately added for 1 minute. The UA was blotted off with 3MM filter paper, the grids were air-dried and kept in a humidity-controlled room before viewing.

2.2.2 Graphene oxide grids preparation

Graphene oxide coated grids were prepared for cryo-EM. GO solution (Sigma 777676, 4mg/ml) was diluted using MiliQ water to 0.4mg/ml. 3ul of the diluted GO solution was added to Quantifoil 1.2/1.3 grids and was blotted off using 3MM filter paper. They were then washed with 10ul of H₂O twice and blotted off with 3MM paper. The grids dried for at least overnight before use. The grids were glow discharged for 30 seconds before use to prepare cryo-EM grids sample.

2.2.3 Grids preparation for Cryo-EM

Cryo-EM sample on holey grids with no surface, the Quantifoil R1.2/1.3 300Au grids were glow discharged for 90 seconds in a plasma chamber(0.2mBar and 40mA). 5µl of human telomerase sample at ~0.01mg/ml was added to the grid in 100% humidity chamber at 4°C. The grid was blotted for 2 seconds with blot force -1, then plunge frozen in liquid ethane, using a Vitrobot™ Mark IV (FEI).

For Cryo-EM samples on Graphene oxide grids, Precoated Quantifoil R1.2/1.3 300Au grids were glow discharged for 30 seconds plasma chamber (0.2mBar and 40mA). 5µl of human

telomerase sample $\sim 0.01\text{mg/ml}$ was added to the grid in 100% humidity chamber at 4°C . The grid was blotted for 2 seconds with blot force -1, then plunge frozen in liquid ethane, using Vitrobot™ Mark IV (FEI).

2.2.4 Grid preparation for cryo-electron tomography

To prepare telomerase for tomography analysis, 1ml of 5nm nanogold solution was resuspended by vortexing for 10mins. The resuspended nanogold solution was added to the purified human telomerase sample at 0.01mg/ml in 1:7 ratio. Quantifoil R1.2/1.3 300 Au grids were glow discharged for 90 seconds. $5\mu\text{l}$ of the sample was applied to the grid in a 100% humidity chamber at 4°C . The grid was blotted for 2 seconds with blot force -1, and plunge frozen in liquid ethane, using Vitrobot™ Mark IV (FEI).

2.2.5 Data collection of single particle cryo-electron microscopy

Micrographs were collected using a 300kV Titan Krios Microscope (FEI) in EFTEF mode with a K2 detector. The collection was under movie mode with 40 frames with an objective aperture inserted. The nominal magnification used was 105000x ($1.41\text{ \AA}$ per pixel) or 135000x ($1.1\text{ \AA}$ per pixel) and 105000x ($1.41\text{ \AA}$ per pixel), a defocus range of -1.5 to $-4\mu\text{l}$, for different datasets, total dose rate $\sim 5\text{e}/\text{\AA}^2/\text{s}$, with 12 second exposure time, achieving a total dose $\sim 60\text{e}/\text{\AA}^2$. Micrographs were automatically collected using EPU (Thermo Scientific). For phase plate datasets, the Volta phase plate was inserted instead of the objective capture, and the defocus used was $-0.7\mu\text{m}$, with the other settings kept the same.

2.2.6 Data processing of micrograph collected on human telomerase on graphene oxide coated grids

A dataset of 5285 images was collected on graphene oxide grids using Titan Krios as described. The raw movie files were motion-corrected using MotionCorr2 (Zheng et al. 2017) following the standard protocol. The motion-corrected images were then CTF corrected using

CTFFIND4 (Rohou and Grigorieff 2015), and incorrectly fitted images and empty images were removed by manual inspection. From this dataset, 760,533 particles were picked with Gautomatch using default settings. Further data processing was carried out using Relion 2.1, and particles were classified using 2D classification, resulting in many aligned classes. The good 2D classes showing bilobal shape with a total of 94,117 particles were further classified by 3D classification. The 3D classification yielded 4 classes which had even distribution and a resolution about 17Å.

2.2.7 Data processing of micrograph collected on human telomerase on holey grids

In this dataset 11672 images were collected on holey carbon grids, and the raw movie files were motion-corrected using MotionCorr2 (Zheng et al. 2017) following standard protocols. The motion-corrected images were then CTF corrected using CTFFIND4 (Rohou and Grigorieff 2015), and incorrectly fitted images and empty images were removed by manual inspection and 8166 micrographs were selected for further processing. 1,340,635 particles were picked with Gautomatch using default settings. Further data processing was carried out using Relion 2.1, and the particles were classified using 2D classification. This resulted in many aligned classes and good 2D classes showing bilobal shape with a total of 269, 693 particles. These particles were further classified by many rounds of 3D classification. The 3D classification yielded 4 classes which had even distribution and have a resolution about 15Å.

2.2.8 Processing of micrograph collected on HSP90 bound human telomerase

1321 micrographs of unfixed sample of HSP90 bound human telomerase were motion-corrected using Motioncorr2 and CTF corrected using CTFFIND4. 215,483 particles were picked with cryolo picker and 2D classified. In this case, the well-aligned 2D classes show HSP90 alone resembles the shape of a key, and the 2D classes show clear secondary features were selected including 56,207 particles for 3D classification with the best 3D classes

selected. Eventually, 24,924 particles were used for 3D masked refinement which gives an HSP90 structure at 4.8Å resolution using FSC golden standard cut-off.

For fixed sample of HSP90 bound human telomerase, 1292 micrographs were collected. The micrographs were motion-corrected using Motioncorr2 and CTF corrected using CTFFIND4. 101,476 particles were picked with Gautomatch using previous human telomerase 2D classes as a reference and particles 2D classified using Relion. However, the resulted 2D classes are surprisingly bad, the blurry 2D classes indicate a poor alignment of the molecules. Around 69,112 particles in 2D classes that resemble HSP90 bound human telomerase were 3D classified and the best class containing 29,291 particles were 3D refined using previous obtained 3D density as the initial model. This result in a 3D density ~21Å resolution

2.2.9 Cryo-electron tomography data collection and processing

For tomography data collection. micrographs were collected using 300kV Titan Krios Microscope (FEI) in EFTEF mode with K2 detector. The nominal magnification used was 105000x (1.41 Å per pixel) defocus range of -2.5 to -3.5µm. Data were collected using a single tilt angle from +60 to -60 degrees, with 2-degree increments using continuous collection scheme. A total of 61 images was collected for each tomogram stacks and the total dose used was 120e/Å².

Tomograms with not enough fiducial or burnt ice were discarded. Tomograms were processed using etomo from the IMAGIC package, following the default settings for fiducial alignment as described on the user manual (van Heel and Keegstra 1981). The final aligned tomogram was manually inspected to identify the water-air interface. Subregions of the 3D volume were extracted using etomo. The density map of individual particles was extracted using Chimera (Pettersen et al. 2004) following the user manual.

Chapter 3 Overexpression and purification of human telomerase complexes from the mammalian expression system

3.1 Introduction

The aim of the work presented in this chapter was to improve the expression and purification of human telomerase to enable cryo-EM analysis and hopefully determine the high-resolution structure. The catalytic protein subunit TERT consists of 1131 amino acids with a MW=127kDa and the RNA subunit TERC consists of 451 nucleotides with a MW=151kDa. On estimation, cancer cells contain only 150~200 copies of human telomerase per cell (Akincilar et al. 2015; Xi and Cech 2014), therefore extraction of endogenous human telomerase would unlikely provide enough sample for structural work. Thus, overexpression was necessary for the purification of sufficient quantities of human telomerase to enable structural analysis. Previous work from the Rhodes group determined the low-resolution structure of human telomerase at 25Å resolution by negative stain EM (Sauerwald et al. 2013). The human telomerase sample was purified from transiently transfected HEK293 suspension cells. Since the biogenesis of human telomerase is complicated and potentially involves many chaperones only present in mammalian cells, to date, the mammalian expression system is the best system to use. However, in the early Rhodes work, only picomolar amount of telomerase could be purified from transient transfection. Therefore, in order to enable structure studies by cryo-EM and to reach atomic resolution, a higher concentration of human telomerase is required.

Because large-scale transient transfection was found to be irreproducible and the expression level varies between batches, we set out to develop an expression system that could reliably generate higher amounts of pure human telomerase. When I joined the group in 2016, David Gill had initiated work on creating a stable cell line using the lentivirus system. Lentivirus promotes the integration of hTERT and hTERC genes into the genome of the host cells. First,

lentiviruses containing the hTERT and hTERC gene were generated separately by transfecting WT HEK293T cells. A stable HeLaS3 cell line was achieved by first transfecting with lentivirus containing *TERT* and followed by transfection by lentivirus containing *TERC*. Cells with the highest TERT expression were selected by single clone selection as assessed by western blot against TERT, and the clone selected was named Sub J. The cell line expresses approximated 20 times more of the TERT subunit and has approximately 30 times more activity than the host cell-line.

Initially, I spent a considerable amount of time optimising the purification protocol for the telomerase from the HeLaS3 (SubJ) cell line. The construct contained an N-terminal Protein A tag, a FLAGx3 tag and a TEV cleavage site developed based on published work (Sauerwald et al. 2013). The purification strategy was to capture the telomerase on IgG affinity beads and release the telomerase by TEV cleavage, and further purified the complex by ion-exchange heparin column chromatography followed by size exclusion on a sucrose gradient as described in previous work from the Rhodes group (Sauerwald et al. 2013). A large portion of telomerase was forming aggregation during the sucrose gradient step, and only large amorphous aggregates were observed in negative-stain EM of the purified sample. After many experiments it was found that the plasmid that carried the TERC contained a design error; there was no transcription stop site between the GFP reporter and the TERC gene. Although the TERC 3' end should be regulated and truncated in a sequence-specific manner, it was suspected that some TERC expression was not terminated correctly, hence some TERC may contain additional RNA sequence encoding GFP which resulted in misassembly of the telomerase complex and aggregation. The lentiviral infection may also lead to physiological changes in cells, affecting telomerase biogenesis. Hence, our old cell lines were discarded, and a new cell line was established using a new plasmid construct and method.

3.2 New telomerase expression system

Since our first strategy using lentivirus overexpression system failed to improve the yield of purified human telomerase, a new expression system was designed in collaboration with Simon Lattmann. In the new system, the protein subunit TERT and the RNA subunit TERC genes were cloned together in the same plasmid (pSTL1), that should permit integration of the two genes together in the cellular genome.

First, we searched for the promoters that would give the best overexpression. The hTERT gene was placed downstream of a CMV promoter, which is a strong promoter commonly used for mammalian overexpression systems. On the 3' end of the hTERT gene, we cloned an SV40 polyA sequence which aims to increase the stability of mRNA hence aiding the overexpression. Furthermore, a gene for fluorescent protein tag EGFP was placed at the N-terminus of the hTERT gene to act as a marker for successful transfection and expression.

For overexpression of the telomerase RNA subunit TERC, the endogenous promoter is weak and subject to regulation. Therefore, based on previous publications, we replaced this with cloned in the U1 or U3 promoter (Cristofari and Lingner 2006; Wong and Collins 2006) to improve the overexpression of hTERC. During the previous work, we had suspected that TERC expression is the limiting factor in human telomerase assembly. Therefore, three copies of the TERC gene were cloned together with one copy of protein subunit TERT gene on the same plasmid. The final plasmids were constructed using multiple cloning site as described in Chapter2 based on the modification of a vector previously used by Lattman. The map of the plasmid is as shown in Figure 3.1. The DNA sequences of the constructed plasmid were verified by DNA sequencing.

Figure 3.1 Vector (pSTL1) for co-expression of human TERT and TERC in mammalian cells. The vector map shows hTERT gene fused to EGFP the under the CMV promoter and hTERC gene under the U3 promoter.

The overexpression vector contains a puromycin resistance gene (puroR) under an internal ribosome entry site (IRES). The puroR gene is included to allow antibiotic resistance selection of transfected cells, enhancing expression. The gene is expressed under IRES instead of a separate promoter in order to ensure the expression of puroR is expressed with the hTERT gene.

The CMV promoter drives strong expression of human telomerase when the plasmid enters the cell by transfection. Most of the cells are expected to express the genes encoded on the plasmid for a short amount of time (transient) because the genetic material is not integrated into the genome and the plasmid would degrade over time. In a small subpopulation of the cells, the viral sequence of the CMV promoter assists permanent integration of genes carried

on the plasmid into the host genome. The cells would continuously express the genes and pass them on to future cell generations. We exploit this nature by applying a selection pressure such as antibiotics, only cells that have the vector stably integrated into the genome could be selected, resulting in a constitutive telomerase overexpression.

3.3 Telomerase expression and cell line selection

As described above, due to the poor and unreliable overexpression of human telomerase using transient transfection, we set out to enhance the expression levels by creating a stable cell line overexpressing both hTERT and hTERC genes. The process involves the selection of cells that had the genes integrated into the cellular genome followed by cloning out the strongest overexpressing cells.

To create new cell lines, first, the mammalian expression plasmid was amplified in *E. coli* and the plasmids were purified using a commercial Medi-prep kit (QIAGEN). WT HEK293T cells were transiently transfected with the purified plasmid using Lipofectamine 3000 and cells grown adherently in 6-cm dishes. The transfection efficiency was judged by the GFP signal, and transfected cells were incubated ~48 hours to allow expression of the genes before further processing.

To select cells with plasmid integration, the cells were first subjected to antibiotic selection using 8, 16, 32 and 64 mg/ml of puromycin. Dead cells were removed after 48 hours post-treatment and new medium was added to the cells and the remaining cells were allowed to amplify. Cells that survived higher concentration of puromycin are expected to have a stronger GFP signal since GFP is fused to the *TERT* gene. However, a large variation in GFP intensity within the same cell population was observed. The discrepancy between GFP signal and antibiotic resistance suggested even though the puroR gene was expressed under IRES,

perhaps due to incomplete plasmid integration into the genome, some cells expressing the *puroR* gene were not overexpressing human telomerase at the same level. This showed antibiotic resistance selection alone was insufficient for selection of telomerase expressing cells. Thus, a further selection step was required to segregate cells that were overexpressing human telomerase. Since the *TERT* gene is fused to an EGFP tag, the level of fluorescence signal should correlate to the level of expression of human telomerase. Fluorescence-activated cell sorting (FACS) was used as the second selection step. However, the GFP intensity was still surprisingly low, as a consequence of the low expression level of human telomerase. The cells were segregated into three fractions of fluorescence level and the fraction with the highest fluorescence signal was used for further selection.

In order to further improve human telomerase expression, the cells were further selected using single clonal selection. In this approach, cells were diluted to a single clone following standard protocols, and then single clones were allowed to amplify in 6 well plates. Next, the cells were analysed by measuring the telomerase expression level by SDS-PAGE and western blot analysis with an antibody against hTERT. Well-expressing clones were further analysed by checking the telomerase activity in crude extracts. The final cell clones were selected based on high protein expression levels and high telomerase activity (Figure 3.2).

To verify that the human telomerase produced in the new overexpression is active, a quick method for testing telomerase activity was required. Traditionally, the direct telomerase assay or the TRAP assay is overnight processing involving the use of radioactive labelling ^{32}P and visualisation of telomere synthesis on long denaturing polyacrylamide gels (Slatko and Albright 2001). Since the human telomerase is an unstable protein/RNA complex and is required to be in pristine condition before going for EM studies, a rapid activity assay was necessary to check the sample quality. To aid the work, Simon Lattman developed a fast fluorescent-based TRAP, providing a quick way to judge sample quality before going on with

EM analysis. This activity assay quantifies telomerase activity by using a DNA binding dye to measure the amount of extension product formed during the telomerase extension reaction. To enable fast analysis, the reaction master mix (Chapter 2) was aliquoted and frozen. Activity tests were conducted by adding telomerase sample to the master mix and running a qPCR protocol, which allows verification of human telomerase activity in just 2 hours instead of more than 10 hours for standard direct telomerase assay.

Figure 3.2 The relative activity of selected single clones. The hTERT expression level was quantified by western blot and ranked from highest lowest from left to right. The bar graph is plotted with a logarithm scale showing the relative activity of single clones normalised to untransfected WT HEK293T cells. The red line marks the telomerase activity of the cell population before single cloning was carried out.

The hTERT expression was quantified using western blot analysis against the Strep II tag at the N-terminus of the TERT protein, which detects overexpressed telomerase. The human telomerase activity of single clones was quantified by qTRAP assay and normalised to untransfected WT HEK293T cells. To select the best cell clones for telomerase production, we need to consider both expression and activity. We could not simply select the clones based on the level of expression of protein subunit TERT or RNA subunit TERC alone

because telomerase required both subunits to be properly assembled to function. Consequently, expression levels may not correlate with the degree of assembly and enzymatic activity. Thus, we need to consider the level of TERT expression in combination with the resulting telomerase activity of each cell clone. Cell clones were ranked in order of highest TERT expression from left to right in Figure 3.2. The results show that all single clones had higher telomerase activity compared to wild type HEK293T cells, however, only some clones had higher activity levels compared to the cell population before single cell selection. Unexpectedly, many clones had lower activity, which was puzzling because we expected that single cloning process would enhance the activity of human telomerase. In the first instance, we selected cell lines with the highest hTERT expression that also had higher activity than post-FACS cells for further analysis. Thus, we selected clones 683 and 731 for further analysis. Clone number 721 and 716 had the highest telomerase activity, but since the activity was not significantly higher than that of 683 and 731, and with the fact that the activity levels did not correlate with the TERT expression level, those clones were not picked for further analysis.

Figure 3.3 Sucrose gradient as a size exclusion step. Protein samples were loaded on top of the sucrose gradient, the proteins segregate in the gradient according to their MW after centrifugation at 215,000xg for 18 hours. The graphs show the distribution of standard protein size markers on a sucrose gradient, detected by the Bradford reagent.

Since one key indication of correct telomerase assembly is the MW of the human telomerase complex as described by Rhodes group (Sauerwald et al. 2013), the MW of telomerase in the new cell lines was estimated by sucrose gradient fractionation. Samples of crude cell extract from clone 683 and 731 (Figure 3.4) were loaded onto a sucrose gradient that had been calibrated with size markers of known molecular weight. After ultracentrifugation, the 4ml 10-50% sucrose gradient was fractionated manually into 20 fractions labelled F1-F20, where F1 was the top of the gradient (Figure 3.1). In this gradient, aldolase with a MW=158kDa migrates as a peak in fractions F7-9, ferritin with a MW=440kDa as a peak in fractions F10-F12. The distribution of 669kDa thyroglobulin is rather broad but mainly migrated in fractions F11-13. The 10-50% sucrose gradient used in later chapters was made and fractionated in the same way as described here, thus this size calibration is a valid guide of MW in subsequent sucrose gradients.

Figure 3.4 The distribution of telomerase from crude cell extract on sucrose gradient for clone number 683 and 731. Crude cell extracts from clone 683 and 731 were fractionated by sucrose gradient. The presence of telomerase in each fraction was detected using western blot against TERT. The distribution profile shows both clone 683 and 731 are expressing correctly assembled telomerase (F11-F13).

The presence of human telomerase in the fractionated sucrose gradient was analysed by western blots as described in materials and methods. As the fractions of the sucrose gradient correspond to different molecular weights, properly assembled telomerase dimer containing two copies of TERT (150kDa) and two copies of TERC (151kDa) should migrate in fractions F11-13 corresponding to a MW

~ 600kDa. For unassembled human telomerase, the free TERT should migrate approximately in fractions F7-F9 which corresponds to MW=150kDa, whereas aggregated proteins should migration in fraction F14 and higher, corresponding to a MW=700kDa and above. A good amount of human telomerase is present in fractions F11-13 for all clones analysed, indicating that the new cell lines produced assembled human telomerase with a molecular weight of a dimer. Since for clone 731, there was a clear peak visible for correctly assembled telomerase in fractions F11 – F13 for assembled telomerase, this clone was selected for telomerase production and future studies.

Based on literature and standard protocols for cryo-EM grid preparation, it was believed that at least 4 μ l of human telomerase at $\sim$ 0.1mg/ml concentration (400 μ g) is required to prepare a single cryo-EM grid. Small scale overexpression trails with one 15-cm dish and purification only yielded less than 10 μ g of telomerase based on rough estimation from silver-stained gels. From this, it was estimated that 60-90 15-cm dishes would be required to purify sufficient human telomerase for cryo-EM analysis from this cell line. Since the HEK293T cell line is adherent, large scale growth is difficult and required many manipulations. In order to reduce the risk of contamination and achieve an easier bulk harvest, I decided to culture the cells using corning HYPERflask (Figure 3.5).

Figure 3.5 The workflow with corning HYPERflask allows the growth of many cells. Cells are grown in multi-layered flasks and harvested by shaking in ice-cold PBS solution followed by spinning down the at 300xg.

Each HYPERflask contains ten air-permeable surfaces optimised for cell growth with a surface area equivalent of ten 15-cm dishes. The adherent cells were expanded in the HYPERflask, and cells can be harvested by gentle shaking in ice-cold PBS. The cells in

suspension were spun down at 300xg and the resulting cell pellets were flash frozen and stored in -80°C freezers for future use.

3.4 Telomerase purification protocol optimisation

On estimate, six HYPERflask could produce ~500µg of telomerase. The low expression levels of human telomerase already pose a great challenge, but the purification process poses further challenges due to the instability of the telomerase enzyme. Human telomerase contains RNA, and it is widely known that RNA is extremely sensitive to nuclease degradation, thus the purification steps need to be fast to obtain a pristine sample.

Figure 3.6 The human telomerase construct (in collaboration with Simon Lattman). The New construct shows the current construct used for overexpression and purification of human telomerase. The construct contains N-terminal 2xStrepII tag and EGFP, separated from the hTERT by TEV cleavage site and a linker region. On the C-terminal, it contains a 3xFLAG tag linked to hTERT by 3C cleavage site.

Based on the Sauerwald telomerase purification protocol (Sauerwald et al. 2013), many experiments were carried out to optimise the purification protocol, however, some issues were raised during the optimisation tests. First, the human telomerase had very low binding to many types of affinity beads including IgG, Ni-NTA, GST, FLAG. By estimating the relative amount of telomerase by western blot analysis, I concluded that purification using these beads only captured about 1% of the human telomerase present in the cell extract, with 99% being in the flow-through, which is extremely inefficient. Since the amount of telomerase captured was significantly less than the maximum binding capacity of the beads used, this suggested that the N-terminal tag present in TERT protein may not be exposed to binding due to steric hindrance. Attempt to improve the exposure of tags by extending the linker was of

little success. The reason behind the lack of binding remained unclear, but one possible explanation is that overexpression of telomerase may cause aggregation *in vivo*. In addition to the low binding of telomerase on affinity beads, the telomerase elution efficiency from the beads is also low, with only 30~50% of the telomerase eluting from the beads according to the recommended protocol.

In the end, amongst the several affinity tags tested, human telomerase can only be sufficiently captured by the Strep-II tag, but even then, the capture rate was only 10%~20% as compared to the total expression. Only ion-exchange chromatography using heparin beads has been shown to be an efficient second step to capture >90% of human telomerase. Due to the low amount of human telomerase, and despite the RNA having a high UV absorbance at A_{260} , the overall signal was still too low to use for reliable concentration determinations. Size exclusion chromatography with gel filtration column was even more troublesome. The telomerase sample spread over too many fractions and there was no obvious resolution of sizes, presumably due to non-specific interactions with the column material, this makes obtaining a homogenous sample more difficult. Since sucrose gradient was used to separate human telomerase by molecular weight by the Rhodes group and others (Cohen et al. 2007; Sauerwald et al. 2013; Schnapp et al. 1998), thus, the sucrose gradient was used as the final size exclusion step. Eventually, after many tests, the current protocol of purification was established that can produce sufficient telomerase for structural analysis by cryo-EM.

Figure 3.7 Telomerase purification protocol and SDS-PAGE analysis of purified human telomerase. (A) A flow chart summarising the purification protocol. (B) SDS-PAGE analysis of purified telomerase samples stained with silver stain. The gel shows the sample after Strep II competitive elution (SE) is not pure, along with 152kDa TERT band, there are many other bands present. The purity of the sample after heparin elution (HE) has improved with only a few bands still present. (C) SDS-PAGE of protein sample before loading onto the gradient (HE conc.). The gel shows a predominant TERT band and some other faint bands. (D) SDS-PAGE and Coomassie Blue stain of unfixed human telomerase sample fractions (F11-13) that correspond to 600kDa, showing the TERT band. The fractions are subsequently pooled and used for EM analysis.

The established protocol for purification of human telomerase is shown in Figure 3.7A. The clarified crude cell lysate was first incubated with streptactin beads in suspension, which binds to the Strep II tag on telomerase with nanomolar affinity. The beads were thoroughly washed using wash buffer to remove any unbound material. Despite the construct contains a TEV cleavage site (Figure 3.7), in order to shorten the incubation time, bound telomerase was released from the affinity beads by competitive elution using 50mM desthiobiotin for two hours, whereas TEV cleavage would take at least four hours on ice. The protein composition of the eluent was analysed by SDS-PAGE stained with silver stain. A strong band at around 150kDa (Figure 3.7BC) corresponding to TERT (127kDa hTERT + 27kDa GFP) can be seen. However, many other bands were also present in addition to TERT (Figure 3.7). In order to purify the telomerase sample further, the eluent from the affinity purification was purified by ion-exchange chromatography using heparin beads. Heparin beads function as a cation

exchanger, which bind to positively charged surfaces. Elution of telomerase was carried out using a buffer containing 1M KCl, which is considered a harsh condition for protein/RNA complex. However, this step was established in the purification protocol from the Rhodes group and shown to produce active telomerase (Sauerwald et al. 2013). To reduce the possibility of any dissociation of the complex, the eluted sample was immediately diluted to 300mM KCl concentration. The analysis of the heparin eluent by SDS-PAGE showed the purity was improved because there are fewer protein bands visible in the SDS-PAGE (Figure 3.7B) compared to the eluent from the Strep II affinity purification. The eluent from the heparin beads was concentrated 20-fold to permit an analysis by Coomassie blue SDS-PAGE. This analysis (Figure 3.7C) shows a strong band for TERT at 150kDa with only a few faint bands, suggesting the sample was relatively pure. The concentrated sample was loaded on a sucrose gradient serving as a size exclusion step. After ultracentrifugation, the gradient was fractionated into 20 fractions containing 200µl each, and the fractions were analysed by SDS-PAGE (Figure 3.7D). After this final purification step, the predominant protein band is that of hTERT.

At the end of the purification, the concentration of human telomerase is still too low to be quantified accurately. Based on estimation using Coomassie blue and Bradford, the final concentration was around 0.01mg/ml with a sample volume ~30µl. For sample used for cryo-EM work, the sample was crosslinked by the addition of glutaraldehyde in the sucrose gradient step following the GraFix (Stark 2010). Fractions F11 and F13 containing telomerase dimer were pooled, concentrated and buffer exchanged to remove sucrose. Sucrose contamination hinders cryo-EM analysis. The purified sample was immediately used for grid preparation for cryo-EM analysis. The EM results are discussed in Chapter 4.

3.5 Substrate-based human telomerase purification

Inspection of micrographs from cryo-EM analysis of the human telomerase purified using the protocol described in Section 3.4, suggested the purified telomerase complex was heterogeneous. One explanation could be that the purified telomerase was not sufficiently pure. In order to obtain a purer sample, an additional purification step was introduced. Figure 3.9A shows the flow chart of the new purification strategy, where the purification protocol is identical to the one described in Section 3.4, but with an additional substrate-based affinity purification step added after the heparin purification step.

Figure 3.8 Principle of substrate-based purification. (A) Telomerase sample is incubated with desthiobiotinlayted DNA substrate. (B) Substrate bound telomerase is captured by neutravidin beads (C) and non-bound material including unassembled free TERT is removed through successive washes.

The telomerase sample obtained after heparin elution following the standard purification protocol was concentrated to around 500 μ l. The biotinylated telomeric G-overhang substrate was first incubated with the telomerase sample for 2 hours to allow binding to the template region of the telomerase RNA component TERC through base pairing (Figure 3.8A). G-overhang bound telomerase was then immobilised with neutravidin beads for 4 hours (Figure

3.8B). The beads were extensively washed to remove unassembled telomerase and other contaminants (Figure 3.8C). The bound telomerase was released from the beads through competitive elution by incubation with a buffer containing biotin. This strategy allows telomerase to be purified through the pull-down of TERC, eliminating unassembled telomerase, thus telomerase obtained by this purification must have the ability to bind telomeric substrate.

Figure 3.9 Substrate-based human telomerase purification and SDS-PAGE analysis. (A) shows a flow chart of the modified purification protocol including the oligo pulldown step. (B) SDS-PAGE analysis of protein samples during the purification. The heparin elution (HE) is incubated with the telomeric substrate and neutravidin beads, which is not very pure containing many other bands. By comparing flow-through material (OFT) and HE, which shows only ~10% of human telomeres is captured by the beads. OB fraction shows the protein sample that is bound to the beads. OE shows the eluted human telomerase from this purification is relatively pure.

The SDS-PAGE analysis of the sample shown in Figure 3.9 shows the sample from substrate-based purification had fewer contamination bands than the telomerase sample after heparin purification. However, the yield suffered from a heavy loss of ~70% of the sample. The amount of human telomerase obtained after the oligo purification was very little estimated to be 100µg. About a tenth of the sample was used for this SDS-PAGE analysis. After this

purification, the TERT band at 150kDa was prominent. Further analysis of purity was not possible due to the low concentration of the telomerase sample obtained. The eluent containing telomerase was fractionated same as before by sucrose gradient fractionation containing glutaraldehyde for crosslinking, and fractions corresponding the molecular weight of dimeric telomerase (F11-F13) were pooled, concentrated, and buffer exchanged to give final volume ~30 μ l. This concentration was too low to give a correct estimation of telomerase concentration, nevertheless, the sample was analysed by both negative stain EM and cryo-EM. The results of the EM analysis are shown in Section 4.5.

3.6 Purification of the telomerase complex containing HSP90

The purification protocol described in Section 3.4 and 3.5 yielded human telomerase that contains hTERT and hTERC subunits only. After analysing the EM results from samples obtained from previous sections, I was unable to obtain a high-resolution structure of human telomerase (Chapter 4). My suspicion is that human telomerase enzyme consisting of hTERT and hTERC alone is conformationally heterogeneous which hindered structure determination of the structure at high resolution. During the course of my PhD, a cryo-EM structure of human telomerase at 8Å resolution was published (Nguyen et al. 2018). This telomerase complex compared to our telomerase contains additional telomerase cofactors such as H/ACA box proteins (DKC1, NOP10, GAR1, NHP2) and TCAB1 that bind to the RNA subunit. This led to the thought that additional binding partners of human telomerase may improve stability and reduce the structural flexibility of the sample, hence improving the resolution of cryo-EM analysis. Based on these observations, I modified my purification protocol intending to obtain a telomerase complex with additional binding partners.

In our standard purification protocol described in Section 3.4, the buffer used in the first extraction step contained 300 mM KCl. Since this “high-salt” step could cause disruption of protein-protein or protein-RNA interactions, the salt concentration in the extraction step was reduced. In trial experiments, cells were lysed with lysis buffer containing from 25 to 300mM KCl. The clarified cell lysate was mixed with streptactin beads, and telomerase was immobilised to the beads through Strep II affinity tag. The protein immobilised on streptactin beads was thoroughly washed to remove any non-bound material. Then SDS sample buffer was added to the resin and boiled to extract proteins immobilised on the beads. The denatured protein samples were analysed on SDS-PAGE stained with Coomassie blue. As shown in Figure 3.10, in addition to the band migrating at 150kDa corresponding to the TERT subunit, there is a band at 90kDa where the immobilisation of telomerase was carried out in 25 to 150 mM of KCl. The 90kDa band is not visible in the sample where immobilisation was carried out in 300mM of KCl. Importantly, the sample prepared in 300mM KCl contained fewer bands than at lower salt concentration which shows the purity of telomerase is higher at higher salt concentration. In order to strike a balance between obtaining a relatively pure telomerase sample and preserving the protein-protein or protein-RNA interactions between human telomerase and cofactors, I decided to carry out the purification at 50mM KCl concentration. 25mM KCl was not chosen because the additional contamination band at 250kDa is present in high amount.

Figure 3.10 Analysis of the effect of salt concentration on purified telomerase composition. SDS-PAGE analysis showing the immunoprecipitated telomerase sample carried out in 50 to 300 mM KCl. TERT band is at 150kDa, and the presence of HSP90 band at 90kDa is only visible up to KCl concentration of 150mM.

Since the heparin purification step in the standard purification protocol involved a 300mM KCl wash step and a 1M KCl elution step, which would disrupt the interaction of telomerase cofactors based on previous results, the heparin purification step was omitted from the purification protocol. In the standard purification protocol, telomerase is released from the streptactin beads by competitive elution because it is faster compared to cleavage elution by TEV protease. In principle, the sample obtained through complete elution would be less pure, because other biotinylated proteins present in the cell lysate would also be bound and eluted together with telomerase. The loss of purity by using competitive elution was compensated by having an additional ion-exchange heparin step in the standard protocol. However, since the removal of the heparin purification step would reduce the purity of the telomerase. To improve on sample purity, I decided to use the TEV cleavage site within the construct, to release human telomerase from streptactin beads. The TEV protease specifically cleaves the TEV cleavage site which is in the TERT subunit, hence only releasing telomerase

and proteins bound to telomerase. This improves the purity of the telomerase sample and proteins that co-eluted with telomerase in this step would be interacting with telomerase. Because TEV protease has a MW=27kDa and migrates in different fractions as telomerase, thus it was removed by the final sucrose gradient fractional step.

Same as for the previous purification, clarified cell lysate in 50mM KCl was incubated with streptactin beads, immobilising the telomerase containing Strep II tag on the beads. After washing to remove the non-bound material, instead of complete elution, the telomerase was released from the beads by incubation with homemade TEV protease for 4 hours on ice. The sample was subsequently concentrated using a spin concentrator with MW cut-off of 100kDa, which partially removed the 27kDa TEV protease. SDS-PAGE analysis shows that the sample after TEV elution contained a strong hTERT band at 150kDa and co-factor band at 90kDa. The co-elution suggests the co-factor is bound to the human telomerase rather than non-specifically bound to the beads. The band at 90kDa was cut out and analysed by mass spectrometry which identified HSP90 with a high confidence level. As discussed in the introduction, HSP90 is expected to be bound to active human telomerase to form a complex (Forsythe et al. 2001), thus the co-purification of HSP90 in this protocol is unsurprising and expected.

The eluted sample was concentrated and loaded onto a sucrose gradient for fractionation by molecular weight. The samples in the fractions were analysed by SDS-PAGE (Figure 3.11B) showing the profile of hTERT distribution across the gradient. The telomerase activity of each fraction was also analysed. The gel analysis clearly shows the hTERT band at 150kDa and HSP90 band at ~90kDa. The co-migration of HSP90 on sucrose gradient suggest that HSP90 forms a stable complex with telomerase. HSP90 naturally forms a homodimer with an MW=180kDa (Wayne and Bolon 2007). The intensity of the HSP90 band with respect to that of TERT is consistent with HSP90 being a dimer in this complex.

The active telomerase bound to HSP90 was migrating in fractions F12-F14 which has an estimated molecular weight of 550~750kDa. Normally, active dimeric human telomerase is expected to migrate in fractions F11-F13 (450~670kDa), thus the shift in migration position suggested the human telomerase complex has gained mass due to bound HSP90 dimer as expected. The HSP90 bound human telomerase dimer is expected to have an MW of ~780kDa, which is consistent with the migration position of this complex.

Figure 3.11 Purification of the HSP90 bound human telomerase. (A) The purification protocol was modified by using a lower ionic strength (50mM KCl) buffer and omitted the harsh heparin purification step. (B) SDS-PAGE analysis of protein fractions from the sucrose gradient shows the distribution of telomerase coupled with relative telomerase activity of each fraction. The combined result shows the HSP90 is co-migrating with hTERT on a sucrose gradient and migrate as MW ~780kDa.

As the SDS-PAGE gel analysis shows, the sample is not very pure with many other bands. Nevertheless, I decided to investigate the sample by cryo-EM before spending more time in optimising the purification protocol. Since the non-fixed telomerase sample in fraction F12 and F13 have the highest telomerase activity, these two fractions were pooled and concentrated and then analysed by cryo-EM. In another purification, the sample was fixed

using Grafix described before, and fractions F12 and F13 were pooled and also analysed by cryo-EM. The results of EM analysis are shown in Section 4.6.

3.7 Sucrose gradient characterisation of the human telomerase

Throughout the work presented in this thesis, the purified human telomerase was tracked mainly using the SDS-PAGE and western blot detecting the TERT subunit, and the telomerase activity assays to analyse the migration of active telomerase in the sucrose gradient fractionation. Tracking telomerase through the RNA subunit is difficult due to the limitation caused by the low concentration of human telomerase obtained after purification. To understand the composition and oligomeric state of purified telomerase, the fractions of purified telomerase obtained from the sucrose gradient fractionation following the standard purification protocol were analysed for the presence of TERT, TERC and telomerase activity. Since multiple samples need to be processed rapidly and at the same time, the analyses were carried out in collaboration with Simon Lattman.

Figure 3.12 Analysis of sucrose gradient fractionated human telomerase. Human telomerase sample is fractionated on a sucrose gradient for characterisation. The western-blot against TERT shows the distribution of TERT across the sucrose gradient. The presence of telomerase RNA in gradient fractions was quantified by qRT-PCR. The telomerase activity was detected using qTRAP activity assay. The figure shows active telomerase containing telomerase RNA is located in fractions F11-13, corresponding to a MW of ~ 600kDa, the size of dimeric telomerase.

First, the sucrose gradient fractions were separated using SDS-PAGE and the presence of telomerase protein subunit was detected by western blot against hTERT. An equal amount of each fraction was loaded on the gel to allow direct comparison of TERT amounts between different fractions. The western blot analysis (Figure 3.12) shows the TERT distribution across the gradient. The distribution of TERT has a widespread profile from fractions F6 to F16, with high amounts of TERT present in fractions F6 to F10 which corresponds to the MW of TERT alone (150kDa). The lower amounts of TERT present in fractions F11 to 13 corresponding to the size of dimeric human telomerase (600kDa). Trace amounts of TERT are also present in fractions higher than F14, that correspond to a molecular weight >700kDa. Inspection of western blot analysis revealed the presence of unassembled TERT subunit in the purified telomerase sample. This could arise from the incomplete assembly of telomerase during overexpression or dissociation of telomerase during purification. Nevertheless, the sucrose gradient was effective in the separation and removal of the unassembled telomerase.

The presence of the RNA subunit TERC cannot be accurately detected by direct methods due to the low concentration, thus qRT-PCR was used to amplify the RNA. The telomerase RNA was extracted from the sample using Trizol, then reverse transcribed using cDNA synthesis kit following the manufacturer's protocol. In order to quantify TERC RNA, quantitative PCR reactions were performed on the reverse transcribed sample using primers corresponding to DNA sequences flanking the full-length TERC gene. The qRT-PCR results show that the telomerase RNA was only present in fraction F10 and above, with a sharp peak at fractions F11 to F13, which corresponds to the migration of dimeric human telomerase. Some TERC was also present in fraction F14 and above corresponding to high molecular weight >700kDa, which could be higher-order oligomers or aggregates of the telomerase. As expected, the TERC subunits were found only in the fractions corresponding to the MW of dimeric telomerase and no unassembled TERC was observed as the purification strategies were targeting the protein subunit TERT.

Lastly, the activity of human telomerase was analysed using qTRAP assay as described earlier in Section 3.2. The activity started to emerge in fraction F10, peaked at fraction F11 and F12 and started to drop in F13, dropping to very low from fractions F14 onwards. The fractions with the highest telomerase activity coincide with the molecular weight of dimeric human telomerase. As expected, the samples from fractions containing higher molecular weight complexes and hence likely aggregated telomerase, show little telomerase activity. The strong correlation between the distribution of TERC and telomerase activity was expected because only properly assembled telomerase containing the RNA subunit should be active. The fact that TERC was mainly distributed in fraction F11-F13 corresponding to the where the telomerase dimer (MW= 670kDa) migrates provides further evidence that human telomerase is assembled as a dimer in its active form. Trace amounts of TERC and telomerase activity were detected in fractions F9-F10 which corresponds to a MW of

~300kDa, possibly indicating the presence of a telomerase monomer consisting of one copy of TERT and one copy of TERC alone.

In conclusion, by comparing the distribution of telomerase against molecular weight markers in the sucrose gradient, we found that active human telomerase is co-migrating with ferritin (440kDa) and thyroglobulin (670kDa). TERT+TERC has a MW=300kDa; 2xTERC+2xTERC has a MW=600kDa. Furthermore, the telomerase activity peaks at in fractions F11 to F13, which corresponds to a MW~450-670kDa. This provides strong evidence that our purified human telomerase is active in telomere repeat synthesis as a dimer. However, the wide distribution of unassembled TERT subunit in the sucrose gradient suggests that some unassembled TERT subunit is likely to be present in fractions F11 to F13 and hence it might be a contaminant in our cryo-EM analysis.

3.8 Characterisation of the purified human telomerase complex

In 2018, during the course of my thesis work the cryo-EM structure of the monomeric human telomerase in complex other cofactors (DKC1, GAR1, NHP2, NOP10, TCAB1) was published by the Collins group at 8Å resolution (Nguyen et al. 2018). This telomerase sample was purified using a milder purification protocol consisting of isolation by substrate binding following by affinity tag purification, with no ion exchange or size exclusion step. This led to the isolation of monomeric human telomerase complexes containing additional components. This inspired me to modify our purification protocol using milder extraction conditions resulting in a telomerase complex containing HSP90 (Section 3.4). In the Collins fitted structure cofactors (DKC1, GAR1, NOP10, NHP2, TCAB1) were fitted into the density map in stoichiometric amounts. The presented data by the Collins group indicates the presence of additional co-factors is important for the catalytic activity of human telomerase.

This poses a question of the validity of the human dimeric telomerase complex purified by us and further has important implications for the interpretation of EM maps.

Figure 3.13 Western blot and SDS-PAGE analysis of purified human telomerase show our sample contains a negligible amount of telomerase cofactors. The purified telomerase sample was analysed by (A) western blot against TERT, TCAB1, DKC1, NHP2 and (B)(C) SDS-PAGE (B)(C). The combined result shows our purified telomerase sample contains predominantly TERT and negligible amount of telomerase cofactors. The red arrow marks the position of TERT band, the three orange arrows mark the position of TCAB1, DKC1 and NHP2 from top to bottom. Green arrows mark two predominant bands present in the sample, that were identified by mass-spectrometry as HSP70 and tubulin which are non-specific contaminants as a result of purification.

Although in our work, protein analysis of the SDS-PAGE fractionation of our purified telomerase sample showed that our telomerase did not contain a stoichiometric amount of additional factors, the SDS-PAGE and western blot analyses of telomerase sample after the sucrose gradient fractionation suggested the main components of our purified telomerase are TERT and TERC. The western blot analysis of the composition of our purified telomerase after 1st step (Strep II) and 2nd step (Heparin) shows that H/ACA box proteins TCAB1, DKC1 and NHP2 and TCAB1 are lost during the purification (Figure 3.13). Two predominant bands present in our telomerase sample we identified by mass-spectrometry as HSP70 and tubulin which are non-specific contaminants purification. In summary, the combined results of western blot and SDS-PAGE stained with Coomassie blue showed that our telomerase

purification results in an active telomerase complex that is compositionally different from the one produced by the Collins group. These results also further validate that the telomerase purified by me consisted of the protein catalytic subunit TERT and RNA subunit TERC only and negligible amounts of other cofactors. This observation coincides with previous conclusions by the Rhodes group showing their purified human telomerase contains TERT and TERC only (Sauerwald et al. 2013).

In the published work by the Collins group (Nguyen et al. 2018), the presented experiments also suggest the binding of telomerase cofactors are required for telomerase activity. This clearly contradicts my results that only TERT and TERC are required for telomerase activity. Any experiment demonstrating the activity of telomerase would require precise knowledge of the concentration of telomerase in the sample. However, in my purification, this was not possible as the concentration of telomerase is too low to be accurately estimated using optical absorbance. To conclude that my purified sample is fully functional and active, further telomerase activity analysis was required to address this question.

3.9 Telomerase Activity

In the work presented in this thesis, so far, I have shown that our purified human telomerase is active. Due to the small quantities of telomerase enzyme produced, it was difficult to establish its specific activity. Our collaborator Scott Cohen volunteered to estimate the specific activity of our telomerase sample using a direct telomerase activity assay that he has developed (Cohen and Reddel 2008) and also wished to compare the activity of our human telomerase preparation with the one he produces using a different purification protocol. This could help to explain, first, whether the additional cofactors are necessary for telomerase activity, and second, whether our purified telomerase is in a pristine condition.

Figure 3.14 Principle of telomerase activity assay. (A) Telomerase is incubated with the substrate in the reaction mixture. The sample is incubated at 37 °C to allow telomerase extension. The resulting extended products have different numbers of telomeric repeats. (B) The extension products are separated using urea denaturing PAGE gel and detected by ^{32}P incorporated G nucleotide. Each band represent extension products with a certain number of repeats. The bands at higher molecular weight indicate longer extension products. An increased amount of longer product produced indicate an increase in telomerase processivity.

The direct telomerase activity assay was carried out based on Cohen's published protocol (Cohen and Reddel 2008). Since human telomerase is a processive enzyme, this assay can measure both the activity and processivity of the enzyme. Human telomerase is incubated with reaction mixtures containing radiolabelled guanines [^{32}P αGTP], and as telomerase extends the telomeric substrate, the radiolabelled G gets incorporated into the product. The product was analysed on a denaturing urea PAGE gel (Figure 3.14), which allows the separation of product according to the number of nucleotides added. Each band in Figure 3.14 represents a dissociated product that has been extended by human telomerase, and the bands are in a ladder with 6 nucleotides gap which is the number of nucleotides extended by telomerase per extension cycle. Longer extension products are represented by bands with high molecular and vice versa. Higher intensity of bands at higher molecular weight product

shows high telomerase processivity. The activity is quantified by addition of the intensity of all the bands in the gel.

Figure 3.15 Comparison of telomerase activity of samples purified by us and S. Cohen (work of Scott Cohen). Direct telomerase activity assay was performed and products were analysed using 10% denaturing PAGE (A) Results of direct activity assay, the pattern of bands indicate telomerase activity. On the left, the activity of a serial dilution of our collaborator's sample (CMRI) is shown compared to our sample at strepII elution (NTU B), and at heparin elution (NTU C). On the bottom is the northern blot against hTR to normalise the amount of human telomerase. (B) Quantified data represents hTR amount and activity normalised against CMRI sample. The result shows NTUB sample has the same activity but more hTR, and NTUC sample has a similar amount of hTR and similar activity compared to the CMRI sample, verifying our telomerase sample is fully active and comparable to their sample.

In his activity assay, the telomerase concentration was first established by spotting telomerase sample on to filter paper and probing with radiolabel oligonucleotides complementary to the template region of the human telomerase RNA subunit, which allows rough quantification of the relative amount of human telomerase (Figure 3.15). The spot with similar intensity indicates a similar amount of human telomerase, which later allows normalisation of activity per moles of human telomerase. The urea PAGE gel analysis in Figure 3.15A shows a clear ladder of bands demonstrating purified telomerase by Cohen and us can achieve both base additional and processivity. Further quantification in Figure 3.15B shows the relative specific

activity of the telomerase sample after first step purification (SE) is lower than the sample of Cohen after the initial purification step, and the telomerase sample after the heparin step (HE) has the same specific activity as their sample (CMRI Lane1). Interestingly, the gel analysis showed higher intensities for bands representing longer telomerase synthesis products for my telomerase sample, which suggest my purified telomerase sample is more processive. After normalising the activity based on the absolute amount of human telomerase, the unit activity of our purified human telomerase is the same as theirs (Figure 3.15B). By personal communication, our collaborator has verified that his sample only contained DKC1 in addition to TERT subunit and no other additional factors were found. Considering the results of this activity assay together with our findings, I could conclude that the human telomerase consisting of hTERT and hTERC only is fully active and does not require additional cofactors for activity. The results also gave confidence that the telomerase sample I purified is a fully active and properly assembled human telomerase complex.

3.10 Shelterin proteins increase telomerase processivity

Human telomerase is reported to interact with the shelterin components at the telomeres (Nandakumar and Cech 2013; Paeschke et al. 2008). Since the shelterin complex was available in the Rhodes group, we were interested in investigating the effect of full shelterin complex on human telomerase activity. This experiment would also further demonstrate that our purified telomerase is active and functional. Emmanuel Smith from our group kindly provided the purified shelterin complexes TIN2-POT1-TPP1 (TPT), TRF2-RAP1 (TR) and TRF2-RAP1-TIN2-TPP1-POT1 (TRTPT). As illustrated in Figure 1.1.4, the TRF2-RAP1 complex is bound to the dsDNA region of the telomere, the TPP1-POT1 complex is bound to the ssDNA G-overhang region of the telomere and TIN2 bridges between the dsDNA and ssDNA binding complex. The TPP1-POT1 complex is thought to be involved in human

telomerase recruitment to the telomeres (Schmidt et al. 2014). From previous publications, it was expected that the shelterin complex containing TPP1-POT1 would increase the processivity of human telomerase, whereas TRF2-RAP1 was not expected to directly interact with human telomerase (de Lange 2005). For this experiment, human telomerase purified following the protocol presented in section 3.4 including the heparin step and concentration was incubated with the telomeric substrate to allow binding. Subsequently, different shelterin components, TR, TPT and TRTPT, were added to the reaction mixture and incubated at 37C° to allow substrate extension and the product analysed by denaturing gel electrophoresis. The ladder of bands represents telomeric repeat additions where each band shows a repeat through the incorporation of P³² nucleotides. As described in the previous section, a band would represent a dissociated product that has been extended by human telomerase. Bands with more nucleotides represent longer extensions, and vice versa. Higher intensity of bands of the longer extension products shows high telomerase processivity.

Figure 3.16 Telomerase activity stimulation by human shelterin components (work of Simon Lattman). Direct telomerase activity was performed, and the product was analysed on a 10% denaturing PAGE gel showing the human telomerase sample without shelterin components (-), with TRF2-RAP1 (TR), with TPP1-POT1-TIN2 (TPT 5-100nm), and with TRF2-RAP1-TPP1-POT1-TIN2 (TR-TPR 5-100nm). The gel shows telomerase processivity is stimulated by incubation with shelterin complexes containing TPP1-POT1-TIN2 but not TRF2-RAP1.

In Figure 3.16, the gel-electrophoresis analysis shows that the addition of TPT alters the pattern of bands compared to purified telomerase alone. As TPT concentration increase, the bands representing longer extension products increase in intensity, while the intensity of the bands representing shorter extension products decrease. This experiment demonstrates that longer extension products are preferentially formed upon addition of TPT in a concentration-dependent manner. The experiments demonstrate that telomerase processivity is regulated by TPT and that telomerase becomes more processive when interacting with TPT. The addition of TRTPR shows a similar effect as the addition of TPT complex to the reaction mixture increasing the telomerase processivity also in a concentration-dependent manner.

Interestingly, the level of increased processivity after addition of 12 nmole of TRTPR is

similar to increase achieved by adding 18 nmole of TPT, suggesting the full TRTPT complex increase telomerase processivity more than the TPT complex itself.

The telomerase activity analysis (Figure 3.16) provided additional evidence that the human telomerase sample we purified is fully active and its processivity can be increased by telomere end-binding proteins. The lane to the left and the right on the gel shows the effect on the telomerase activity by shelterin proteins. Whilst all shelterin proteins TRF2, RAP1, TIN2, POT1 and TPP1 increase processivity, it can be concluded that the proteins primarily responsible for the increased activity is the TIN2-POT1-TPP1 complex.

Our results are in agreement and go beyond previously published results showing that the addition TPP1 alone increases the human telomerase processivity (Zaug et al. 2010). It has been proposed that the binding of shelterin components to human telomerase increases its affinity for G-overhang substrate preventing substrate dissociation and aids translocation (Figure 1.1.4) hence achieving increased telomerase processivity. These results not only demonstrate that our purified human telomerase is active in telomeric repeat synthesis but it may interact with telomere binding protein directly resulting in an increased processivity. This further proves our purified human telomerase containing hTERT and hTERC alone is fully active and functional.

3.11 Conclusions and discussions

To summarise, optimisation of the expression and purification protocol for human telomerase has been very challenging. Despite using several strategies to overexpress human telomerase, the expression level was still very low compared to the overexpression obtained for other novel protein complexes. An important point to note, I am trying to overexpress a complex formed by a protein subunit and an RNA subunit together, the processing pathway for the two components are very different, while others usually co-express complexes containing

multiple proteins. It was only possible to produce sufficient material for cryo-EM analysis by increasing the number of cells grown. The large-scale growing system in HYPERflasks was introduced which provides a convenient way to handle large scale growth of adherent cell cultures. Later, other groups' publications confirmed our observations, as they were only able to produce enough human telomerase for structural work by using ~100 dishes of cells (Nguyen et al. 2018).

Published purification protocols of the human telomerase resulted in very small yields of human telomerase thus making the structural analysis very challenging. Previous work from the Rhodes group also demonstrated establishing a purification protocol for human telomerase was difficult, purifying only ~10 pmol of telomerase from 10g of cells pellets transiently expressing the TERT and TERC genes. With some optimisations, I have successfully established a protocol for purifying human telomerase for cryo-EM analysis, although the yield is still very low. Eventually, with the established protocol I was able to produce 30~50µl of human telomerase sample at 0.01mg/ml for analysis by cryo-EM. Typically, 4µl of protein sample is required per cryo-EM grids, and hence with the limited amount of sample, I could only make ~10 grids per purification. This is a great hindrance in testing different EM grids preparation conditions, therefore I had to invest considerable time in repeatedly purify human telomerase and use it immediately for cryo-EM analysis.

Figure 3.17 SDS-PAGE and activity assay of human telomerase from the Collins group (Nguyen et al. 2018). (A) The SDS page of samples containing elution fractions from their purification. (B) Activity assay from corresponding samples.

Biochemical characterisation of the human telomerase is also complicated by the low concentration of human telomerase obtained. By sucrose gradient fractionation we demonstrated that human telomerase we purify migrate as a dimer with MW ~600kDa. The published work by the Collins group reports that their purified human telomerase is monomeric and contains other binding proteins in addition to TERT and TERC (Nguyen et al. 2018). The density map was fitted with stoichiometric amounts to telomerase with a ratio of 2x DKC1: 1x GAR1: 2x NHP2: 2x NOP10: 1x TCAB1: 1x TERT: 1xTERC. However, the SDS-PAGE analysis did not contain bands with stoichiometric intensity in agreement with their fittings (Figure 3.17). Furthermore, our collaborator purified telomerase using a different protocol containing DKC alone but no other proteins (Cohen, personal communication). After verification by western blot, I concluded that the telomerase sample produced by me consists of TERT and TERC alone. The different human telomerase composition purified by three groups is may be explained by considering that different overexpression and purification strategies were used. In the expression system used by the

Collins group, the hTERC gene was fused to the HDV gene which codes for hepatitis delta virus ribozyme, shown to increase hTERC accumulation at Cajal bodies (Vogan et al. 2016). This abnormal accumulation of hTERC may affect the human telomerase biogenesis causing H/ACA proteins to remain bound to human telomerase and inhibiting the human telomerase dimerization process. Our overexpression strategy and purification protocol were similar to Scott Cohen's purification strategy, and neither of us purified a stoichiometric amount of NHP2, TCAB1, GAR1 and NOP10 in our sample. Whether a different overexpression strategy directly caused the difference in composition of purified human telomerase complex is unclear.

An additional difference is that the Collins group argued that human telomerase is only active with the additional cofactors bound (Figure 3.17B). However, our activity assays of purified telomerase and addition of shelterin demonstrate that this is not necessarily true. It is evident from all of the analysis presented here that human telomerase consisting of only the protein catalytic subunit hTERC and RNA subunit hTERC is fully active. While testing telomerase activity, we could not meaningfully compare the activity of our human telomerase to published work, thus it was not possible to deduce whether the additional co-factor affects telomerase activity at all. But we demonstrated our purified human telomerase had increased processivity when shelterin proteins were added, which simulates the interaction of telomerase at telomeres, suggesting telomerase does not require other co-factors for activity.

In conclusion, the aim of the work presented in this chapter was achieved. The expression and purification protocol established and optimised readily yield human telomerase samples for structural analysis, although in less than satisfactory quantities. In addition, a protocol to purify the HSP90 bound human telomerase complex was also established. Furthermore, the purified telomerase sample by us was biochemically characterised to show that it contains TERT and TERC subunits only with no additional cofactors and is fully active. Also, the

purified telomerase could interact with telomere binding proteins in agreement with previous literature (Grill et al. 2018). All these data provide evidence that the human telomerase sample I purified is suitable for structural analysis.

Chapter 4 Structural analysis of human telomerase by using electron microscopy

4.1 Introduction

The Rhodes group determined the first human telomerase structure was to 25Å resolution by using negative stain EM. However, the thickness of the heavy metal stain sets the resolution limit of negative stain EM at about 10~15Å. In order to obtain an atomic-resolution structure of telomerase, cryo-EM was chosen as the method for structural analysis. In addition, due to the low yield of human telomerase (~500µg per large scale purification), carrying out crystallography trials was not possible at these limited quantities of protein sample. The recent spectacular developments in cryo-EM have made it the method of choice for analysing the structure of big complexes. Having obtained small quantities of pure and active telomerase, we went onto cryo-EM structure analysis. Despite all the effort described in the previous chapter, the highest amount of telomerase obtained was 50µl sample at 0.01mg/ml. Typically, in order to obtain grids with a good distribution of particles, a concentration of 0.1~4mg/ml is used. This means only ten grids may be made from the sample obtained from a single purification which is a bottleneck for testing different grid conditions. The first challenge we faced is sample capture on cryo-EM grids. The preparation step includes a blotting step, where blotting paper comes in contact with the sample droplet blotting off the excess liquid to form a thin layer of ice upon freezing. During this step, the majority of the protein is removed due to charge interaction with filter paper. Given the particularly low concentration of human telomerase I obtained, the production of cryo-EM grids suitable for data collection remained very challenging. Upon successful production of cryo-EM grids with nice particle distribution, one can then go on to data collection and processing to achieve the ultimate aim of obtaining the high-resolution structure of human telomerase.

To establish a sample capture protocol on EM grids, I used the Talos Arctica to record images and assessed the distribution and number of particles. Grids that contained too few or no particles were immediately discarded, whereas grids containing good particle distribution of human telomerase were saved and used for data collection using the high-end microscope 300kV Titan Krios. Before April 2018, the only electron microscope available for cryo-EM work at NTU was the 200kV Talos Arctica with the Falcon II detector. The Titan Krios with K2 detector only became available in April 2018 which allowed a rapid collection of high-resolution datasets.

This chapter describes the cryo-EM analysis performed on purified telomerase samples using the protocol described in Chapter 3. By obtaining the structure of various human telomerase complexes, we hope to gain molecular insights on the functions of human telomerase.

4.2 Optimisation of grid preparation to capture human telomerase sample on grids

To produce cryo-EM grids, a protein sample must be loaded onto a support foil. The support or otherwise referred to as EM grids come in many variations. Then, the loaded grids must be blotted to remove excess liquid before freezing to produce a thin layer of vitreous ice containing protein particles. The blotting process is usually carried out using the semi-automated FEI Vitrobot machine with adjustable blot force and time. A typical grid used for negatively stained EM contains ~20nm thick of carbon, which would not be suitable for cryo-EM analysis due to its high contrast. Quantifoil grids are carbon grids with evenly spaced holes, often referred to as holey grids. An additional layer of support may be added on the holes in order to enhance sample capture. However, there is no universal guide on the optimal support type for grids preparation as each sample behaves differently. Thus, this step

becomes a trial and error process and it is particularly challenging when working with samples of low concentration such as human telomerase.

With the low concentration of human telomerase sample, I had no choice but to optimise the blotting conditions attempting to capture telomerase on EM grids. In the beginning, the standard Vitrobot settings yielded empty cryo-EM grids, thus tests were carried out to identify the optimal blotting conditions for sample capture and optimal ice thickness. The ice thickness contributes to the signal of the background, hence thick ice would reduce the micrograph contrast which hinders the picking of particles and reduce the maximum resolution achievable. After the optimising blotting conditions to obtain optimal ice thickness, I moved on to improve telomerase sample capture. During the process of trial and error, holey grids and carbon-coated grids were tested using different blotting conditions. Unfortunately, many combinations of grids support and blotting conditions were not able to produce good grids, as either no or little particles were observed during the screening process. Thus, I decided to explore layering grids with graphene oxide which has a lower background versus carbon coating and more importantly has been shown to improve sample capture (Bai et al. 2015b).

Figure 4.1 Hexagonal diffraction pattern of monolayer graphene oxide on Quantifoil grids.

Graphene oxide grid was imaged using 120kV T12 microscope in diffraction mode. The image shows only six spots are visible per ring shows the monolayer graphene oxide coating of grids is successful.

Graphene oxide grids were produced by the addition of a graphene oxide solution droplet to Quantifoil R1.2/1.3 Au300 grids followed by washes with MiliQ water. The grids were then air-dried, and glow discharged before use. The coverage of graphene oxide was visually examined for every grid on the T12 electron microscope. The micrographs revealed the graphene oxide coverage was around 60~70%. Some regions have an evenly distributed monolayer, while some regions contain folded or multilayers of graphene oxide. The condition of graphene layers is hard to judge due to its high electron transparency, however, graphene oxide has a hexagonal diffraction pattern as shown in Figure 4.1, the presence of single hexagonal pattern indicates the deposition of monolayered graphene oxide.

4.3 Cryo-EM analysis of human telomerase on graphene oxide grids

Using graphene oxide coated grids, it was finally possible to capture human telomerase on cryo-EM grids. Figure 4.2 shows the raw micrograph. Human telomerase particles are clearly visible in the micrographs with even distribution. In some regions of the micrographs, the

wrinkles produced from the graphene oxide sheet is also clearly visible. Inspection of the micrograph suggests that the telomerase has been captured in different orientations on the grid. The side views have a bilobal shape which resembles the solved negative structure by the Rhodes group. The length of the bilobal particles is estimated to be around 25-28nm (Figure 4.2). Particles with smaller size are also present in the micrographs, which could be the top views of dimeric telomerase or dissociated human telomerase particles. Nonetheless, the distribution and particle quality appear to be good and based on this I proceeded to collect a dataset.

Figure 4.2 Electron micrographs of telomerase captured on graphene oxide grids. Micrographs collected using 300kV Titan Krios electron microscope at a defocus of $-3\mu\text{m}$ show (A) micrograph with flat GO (B) micrograph with wrinkled GO. The red circles highlight examples of human telomerase particles.

A dataset was collected on the 300kV Titan Krios with K2 detector using movie mode with 40 frames per second and with energy filter inserted. The nominal magnification used was 105,000x which resulted in a pixel size of $1.31\text{\AA}/\text{pixel}$. The exposure time was 15secs with a total dose $\sim 60\text{ e}/\text{\AA}^2$. The defocus range used was -2.5 to $-4\mu\text{m}$. The dataset was collected automatically using a computer program supplied by the manufacturer (FEI-EPU).

Figure 4.3 Classification of telomerase particles from the GO dataset Particles were classified in Relion. (A) Representative 2D classes showing side views of dimeric human telomerase, the box size is 400 Å, the length of the particles is ~280 Å (B) Four classes of human telomerase densities after 3D classification containing ~23,000 particles each. Results from 2D and 3D classification show telomerase have a bilobal shape.

5285 images were collected, and raw movie files were motion-corrected using MotionCorr2 following the standard protocol (Zheng et al. 2017). The motion-corrected images were then CTF corrected using CTFFIND4 (Rohou and Grigorieff 2015), and incorrectly fitted images and empty images were removed by manual inspection. 760,533 particles were picked with Gautomatch (Zhang 2018) using default settings. The micrographs and particles coordinates were imported and processed in Relion 2.1 (Scheres 2012). The particles were classified using 2D classification and resulted in many aligned classes. Figure 4.3 shows some of the good 2D classes of bilobal shaped particles as found by negative structure analysis (Sauerwald et al. 2013). The 2D classification clearly shows particles with dimeric size (280 Å in length). The good 2D classes were selected and combined with a total of 94,117 particles, which were further classified by 3D classification. Several rounds of 3D classification using Relion yielded class sums that look similar to each other and with even particle distribution. The shape is clearly bilobal with hinge region bridging the two lobes, the two lobes appear to be of different sizes in all classes with dimensions of 280 Å x 150 Å x

150Å, in agreement with previous observations by the Rhodes group (Sauerwald et al. 2013). The hinge region is where the hTERC is expected to be located. Class 1 (grey) and Class 3 (cyan) have the hinge region more extended and seem to be making two contacts which each lobe. The hinge region in Class 4 (purple) appears to be disconnected, which could represent telomerase particles that are damaged and dissociated. Class 2 (yellow) have a single hinge connecting two lobes, which best resembles the solved negative stain structure by the Rhodes group. A more detailed structure analysis is shown later in this chapter. However, the best resolution of the class sums was limited to ~18Å resolution. The result is puzzling, as the data were collected under conditions that should give a high-resolution structure. Further attempts to process the data using different programs (cryoSparc, cisTEM) were unsuccessful. At this point I decided to make grids with another grid support material.

4.4 Cryo-EM analysis of human telomerase on holey grids

One explanation for the low resolution obtained of the telomerase could be due to deformation of telomerase caused by the interacting with the graphene oxide surface. Although graphene oxide is mostly electron transparent, it still contributes to some image contrast. Holey grids are expected to have lower background and a higher signal to noise ratio, which may produce an image with higher contrast. Also, because there is no additional support, protein particles are frozen freely in a thin layer of ice, meaning they are closer to their native conditions and conformations.

Therefore, I went back to try and capture human telomerase in holey grids again. With more focused and thorough tests, I was able to capture human telomerase molecule with nice distribution in holey grids by glow-discharging for an atypically long time of 100 seconds. A raw micrograph is shown in Figure 4.4. The contrast of the image is better than graphene

oxide grids and particles can be more clearly identified (Figure 4.4). However, the particles were not evenly distributed in the holes, and instead of in most regions, the particles either cluster in a circle in the middle of the holes or form a ring-shaped pattern. The particles look very distinct and clear, and surprisingly, the raw images seem to contain detailed features, which may be secondary structure elements in TERT or TERC. The particle length and shape appear to correspond to these on graphene oxide grids, with bilobal shaped particles with a length of 25-28nm. Smaller particles are also present. From the visual inspection of the raw micrographs, the particles are with better contrast, which appears to be promising. Therefore, I went on to collect a dataset for further analysis.

Figure 4.4 Electron micrographs of human telomerase sample from holey grids. Micrographs collected using 300kV Titan Krios electron microscope at a defocus of $-3\mu\text{m}$ show (A) telomerase particles form a circular cluster in the middle or (B) the telomerase particles form a ring shape. Telomerase particles are clearly visible on the micrographs.

The dataset was collected on the Titan Krios with the same setting as described in Section 4.3, except the defocus range. The defocus ranged was changed to $-1.5\mu\text{m}$ to $-3\mu\text{m}$. Lower defocus value was used because the image contrast was better, hence particles remain clearly

visible even when the defocus value was reduced. Images captured closer to focus have better preserved high-resolution information.

Initially, a dataset of 2523 micrograph was collected and processed using the same method as described in Section 4.3. The 2D classification resulted in good dimer classes. Initially, processing resulted in 3D classes having a resolution of $\sim 20\text{\AA}$. The limited resolution was an indication that the number of suitable particles was insufficient to give high-resolution structures.

At the time, it started to emerge that many published high-resolution structures were solved from a dataset containing more than 1 million particles, for example, the spliceosomes (Nguyen et al. 2016). It was emerging that huge datasets with more than 10,000 micrographs were collected, and autopicking resulted in more than 1 million particles, however, only around 30,000 particles ($<3\%$) were used in the final reconstruction giving high-resolution structures with $3\sim 4\text{\AA}$. This observation led me to think that the poor resolution was because I did not have enough particles with high-resolution information that contributed to the high-resolution structure. Therefore, I went on to collect a large dataset containing a total number of 9459 micrographs, using the same collection strategy as described earlier in this section.

The micrographs from all dataset were combined and were motion-corrected using MotionCorr2 (Zheng et al. 2017), and subsequently, CTF corrected using CTFFIND4 (Rohou and Grigorieff 2015). The corrected micrographs were visually inspected to remove empty and bad micrographs, resulting in 8166 good micrographs. Particles were picked using template-based picking with Gautomatch using bilobal 2D classes obtained from Section 4.4. This resulted in around 1,340,635 particles being picked. Then the classification and refinements were carried out using Relion 2.1 (Scheres 2012). The picked particles were first classified using 2D classification. Figure 4.5 shows representative 2D classes, showing

clearly the bilobal side view of human telomerase. From the resulting classes, well-aligned bilobal classes and some top view classes were selected and combined resulting in a total of 269, 693 particles. The particles were classified using 3D class generated from Section 3.4 as the initial model which produced 3 classes with dimeric human telomerase with a resolution around $\sim 15\text{\AA}$. The particle distribution across 3 classes is even. The telomerase map obtained at $15\text{\AA}$ has a bilobal shape, which is similar to published Rhodes structure. Despite the lack of atomic-level details, the map has improved resolution compared to the published telomerase structure at $25\text{\AA}$ resolution by Rhodes group (Sauerwald et al. 2013).

Figure 4.5 2D classes and 3D classes from the combined dataset of 8166 micrographs. (A) Representative 2D classes show the side view of human telomerase, with length $\sim 280\text{\AA}$. (B) Class 1-3 reconstructed from good particles from 2D classification, containing $\sim 90,000$ particles each. The density has two lobes and a hinge region. (C) The class 2 from 3D classification was used for the fitting. Two lobe densities are fitted with two beetle TERT (PDB: 3KYL) and a section of TERC (PDB: 2MHI) fitted into the hinge region.

So why is the resolution stuck at $15\text{\AA}$ despite the huge dataset being used? Initially, I thought the low resolution is a result of particle flexibility as similar problems were reported by many published works (Nguyen et al. 2016; Rawson et al. 2016; Sauerwald et al. 2013; Truman et al. 2017). In some of these complexes, the problem of flexibility was overcome by using

computational masking in Relion (Scheres 2012). Masking allows focused alignment on a specific region of the map, hence reduces the misalignment caused by the flexible region of the protein. Therefore a mask against the bigger lobe was created to mask out a single lobe of the telomerase using Chimera (Pettersen et al. 2004) and used for masked classification carried out in Relion 2.1. Particles used in 3D classification were pooled and classified into 3 classes using the mask. The resulted 3D classes are shown in Figure 4.6. Despite the mask, the resolution of the single lobe was still limited to 15Å resolution with particles evenly distributed across these classes. The ring in the density could be the pocket formed by hTERT domains, where the substrate is expected to bind. At this resolution, the map is hard to be interpreted, and since masking did not improve the resolution suggesting that there is an internal alignment issue in addition to the particle flexibility.

Figure 4.6 3D classes using masked 3D classification focusing on alignment on one lobe. The result 3D classes have an estimated resolution ~15 Å, with even distribution of particles.

Disappointingly, despite a large number of micrographs and particles picked, the resolution did not improve. One possible explanation is poor particle alignment. The poor alignment can either be a result of poor signal to noise ratio or particle heterogeneity. Since in the micrographs the particles are visible to human eyes, the poor signal to noise can, therefore, be ruled out as a contributing cause. Thus, the problem is likely caused by particle heterogeneity. Particle heterogeneity could arise during grid preparation that may have

caused particle damage. This has been reported for other complexes in the Rhodes group in which well characterised complexes fall apart during grid preparation and plunge freezing.

4.5 Negative stain EM analysis of oligo-pull down human telomerase

Since we were suspecting that freezing causes damage induced heterogeneity as reported by many other scientists, I went on to perform negative stain EM analysis on the substrate-based purified telomerase sample prepared as described in Section 3.5. Telomerase purified using the new method has improved purity (Figure 3.9) and substrate-binding ability which should, in theory, improve homogeneity.

The negative stain EM grids were prepared using commercial carbon-coated grids. EM analysis following the standard protocol as described in Chapter 2. Micrographs were collected on a Tecnai T12, with a nominal magnification of 78,000x, representing a pixel size of 2.11Å/pixel. The total exposure time is 2 seconds giving a total dose of 40e/Å². The defocus used was -2.5 µm to -3 µm. Around 150 micrographs were collected manually.

The raw micrographs were CTF corrected using CTFFIND4, and a total number of 26,940 particles were picked manually. The 2D classification was performed in Relion 2.1 (Scheres 2012). In the raw micrographs, we could already observe the particle heterogeneity, particles with bilobal shape are clearly visible, however, some particles appear to be very non-uniform. 26,940 particles were classified using 2D classification, the class sums are shown in Figure 4.7. Three distinct conformations of human telomerase may be observed, classes highlighted in red are small (~120Å), which may be the monomeric or top view of human telomerase. Classes highlighted in yellow are medium-sized (~200Å), this may be bent human telomerase particles due to the flexibility in the hinge region. Classes highlighted in blue are dimeric

human telomerase classes, which is consistent with our previous observations in cryo-EM and previous negative stain structure by the Rhodes group (Sauerwald et al. 2013).

Figure 4.7 Negative stain EM analysis for substrate-based purified telomerase sample.

Micrographs of negative stain sample were collected using 120kV T12. (A) Representative micrographs of the sample. The sample appears to be heterogeneous in size. (B) Representative 2D classes showing the heterogeneity of the sample. Red highlight classes that are small ($\sim 120\text{\AA}$), which may be the monomer or top view of human telomerase. Yellow highlight classes that are medium-sized ($\sim 200\text{\AA}$), this may be bent human telomerase particles due to the flexibility in the hinge region. Red highlights dimeric human telomerase classes, consistent with previous observations.

This analysis shows that sample heterogeneity was not due to freezing, but already present within the sample. Despite the purity of the sample, the negative stain data looks very different from previous data obtained by the Rhodes group (Sauerwald et al. 2013), where it was reported that most particles visible on EM grids were bilobal. Although it is difficult to explain this difference as the composition and the molecular weight of telomerase obtained in the sample is the same. However, the samples were prepared using different expression and purification system.

4.6 Cryo-EM analysis of substrate-based purified human telomerase on holey grids with phase plate

Having failed to obtain a high-resolution structure of human telomerase using cryo-EM, I decided that the combination of using phase plate to get better contrast and telomerase sample obtained through substrate-based purification might result in a structure of higher resolution.

A Volta phase plate had been installed in our Titan Krios microscope making this data collection possible. Phase plate has been shown to enhance the signal to noise ratio effectively, which should improve the alignment of particles during 2D classification, allowing the high-resolution structure of smaller proteins to be solved (Chua et al. 2016). Thus, I decided to collect a dataset with a phase plate with the hope that it could improve the alignment of particles which should improve the resolution.

Cryo-EM grids were prepared using Quantifoil R1.2/1.3 Au300 grids glow discharged for 100 seconds in a plasma chamber. The blotting was done using the Vitrobot with standard blotting conditions. Because this substrate-based purified sample was dilute, a 60 seconds incubation time was added before blotting to improve particle capture. The dataset was collected using the 300kV Titan Krios with Volta phase plate inserted and adjusted to phase shift between 0.3 - 0.5 radian (Danev et al. 2014). The micrographs were imaged with a K2 detector using movie mode with 40 frames/s and with the energy filter inserted. The nominal magnification used was 135,000x which resulted in a pixel size of 1.1Å/pixel. The exposure time was 15secs with a total dose $\sim 60 \text{ e}/\text{\AA}^2$. The defocus used was -0.7 μm which is standard for data collection using a phase plate. The dataset was collected automatically using protocol included in FEI EPU.

Figure 4.8 Cryo-EM analysis of substrate-based purified telomerase. (A) Representative micrograph collected using phase plate, the particles form ring shapes, the sample contrast is much better comparing to non-phase plate data (Figure 4.4). (B) Representative 2D classes showing the side view of dimeric human telomerase, the particles have a length of 280 Å and a bilobal shape (C) 3D classes obtained from classification with a bilobal shape, and the resolution is about ~26 Å.

The micrograph (Figure 4.8) shows that the particle contrast is significantly better than the contrast for previous micrographs (Figure 4.4). In this instance, some of the particles have small extensions protruding out from particles. This could be the telomerase substrate that remained bound to the telomerase after the purification.

1473 of micrographs were collected, and motion-corrected using MotionCorr2(Zheng et al. 2017) and subsequently CTF corrected as described in Chapter 2. Micrographs were manually inspected; empty and bad micrographs were removed. 109,096 particles were picked using Gautomatch using a template from the previous 2D classifications. The classification of particles was carried out in Relion 2.1 (Scheres 2012). Representative 2D classes are shown in Figure 4.8B. The aligned 2D classes show a clear bilobal structure with a length of around 280 Å, which is the same size as previously shown in Section 4.4. However, the protruding substrate is not seen in the 2D classes, suggesting that the substrate is flexible and thus not aligned during the 2D classification process. 14,831 particles were

then classified into two 3D classes with resolution $\sim 25\text{\AA}$. The low resolution (below $20\text{\AA}$) is an indication that this dataset is not capable of producing an atomic resolution structure of telomerase. The shape of 3D density is clearly a bilobal shape, which is consistent with all previous findings.

Despite using a purer sample and imaging with phase plate, the structure was still at a low resolution. The problem that hindered our structure work is still not yet understood. The most likely explanation is conformational flexibility with the human telomerase sample. So far, our EM analysis shows the human telomerase is a dimer with dimensions similar to previously published work by the Rhodes group (Sauerwald et al. 2013).

4.7 Electron tomography of human telomerase

During the course of this thesis, it was concluded that protein particles sit at the water-air interface are damaged resulting in an increased sample heterogeneity (Noble et al. 2018). Therefore, I decided to investigate whether the low structural resolution observed was due to damage of telomerase particles located at the water-air interface was performed.

As introduced in Chapter 1.9, cryo-electron tomography is a technique that allows 3D data collection through a series of tilted images with different angles by physically tilting the cryo-EM grid. This means 3D information of a single macromolecular complex may be obtained. By analysing the tomogram acquired using cryo-electron tomography, I should be able to establish the position of protein particles in ice and more importantly also to obtain a density map of individual telomerase particles. The results from this analysis should help us to establish whether the heterogeneity we observed for telomerase is due to damage or different conformations. The methods should also provide some insights on different conformations of

human telomerase which can serve as a better initial model for 3D classification for our single-particle work.

Nanogold is added to the sample to serve as a marker for accurate alignment between tomograms. The telomerase sample used for cryo-ET data collection was prepared as described in section 3.4 and mixed with 5nm nanogold particles at a 7:1 ratio before sample application. The nanogold-containing sample was then applied to Quantifoil 1.2/1.3 Au300 grids, blotted and plunge using Vitrobot frozen following the standard protocol.

Data was collected using 300kV Titan Krios in tomography mode with the energy filter and phase plate inserted. Individual micrographs were acquired using K2 in linear mode. The nominal magnification used was 64,000x with a pixel size of 2.26 Å/pixel. The grid was tilted with a tilt angle between -60 and +60 degrees with 2-degree increment between images, following the continuous dose distribution scheme. The dose rate was adjusted to $2e/\text{Å}^2/\text{s}$, and the total dose was $120 e/\text{Å}^2$. The defocus range was -0.5 to -1 µm. The data was collected automatically using FEI Tomography software.

Twelve tomograms were collected at different positions in the same grid. The tomogram stacks were processed using etomo from the IMOD package (Kremer, Mastronarde, and McIntosh 1996). The images were manually inspected and micrographs with an insufficient number of fiducials or melted ice were discarded. The fiducial based alignment was done using nanogold as the reference to correct for translational movement between tomogram. then the 3D tomogram was reconstructed using the IMOD (Kremer et al. 1996). Telomerase particle densities were extracted using Chimera as described from previous literature (Pettersen et al. 2004).

Figure 4.9 Cryo-ET analysis on human telomerase sample captured in holey grids.

(A) Illustration of tomogram reconstruction shows the human telomerase particles are not located at the water-air interface. (B) An extracted particle of human telomerase at $\sim 40\text{\AA}$ resolution. The particle is $\sim 290\text{\AA}$ in length and has two lobes and a hinge. The side view is deformed due to the missing wedge problem.

By inspecting the slices through the reconstructed tomogram from Z height, the boundary of tomogram which corresponds to the air-water interface can be assigned. From this, it can be concluded that the human telomerase is not located at the water-air interface in holey carbon grids (Figure 4.9A), but instead embedded in the ice. Therefore, the heterogeneity we observed in single-particle cryo-EM was unlikely to be a result of damage at the water-air interface. The extracted density shows the density of a single human telomerase molecule, which has a bilobal shape with a hinge region in the middle. The dimension of the shown density is around $280\text{\AA} \times 150\text{\AA} \times 120\text{\AA}$, which matches the dimension of 3D density obtained through single-particle analysis. However, as expected the extracted densities show a high degree of deformation across one dimension (Figure 4.9B). This is the typical effect of the missing wedge problem in tomography since the tilt angle was limited between $+60$ and -60 degrees. By knowing that our human telomerase particles are not located at the water-air interface, this suggests the main problem that hindered of structural determination is due to the telomerase sample. Given that the biochemical analysis showed the telomerase we purified is fairly homogenous and fully active, I concluded the main contribution to the lack

of high resolution during structural determination was due to the conformational flexibility of the human telomerase.

4.8 Cryo-EM analysis of HSP90 bound human telomerase

As described in Section 3.7, in order to obtain a human telomerase sample with lower flexibility, I attempted to purified human telomerase bound to additional binding partners or cofactors. Using a modified purification protocol, I was able to purify HSP90 bound human telomerase.

Both unfixed and fixed HSP90 bound human telomerase samples were prepared and cryo-EM grids were prepared as described in Section 4.3. Datasets of unfixed and fixed telomerase sample were collected using a 300kV Titan Krios with K2 detector using movie mode with 40 frames/s with energy filter inserted. The nominal magnification used was 135,000x repressing a pixel size of 1.1Å/pixel. The exposure time was 15secs with a total dose ~60 e/Å². The defocus range used was -1.5 to -3 µm. Datasets were collected automatically using the protocol included in the computer program (FEI EPU). 1321 micrographs were collected for non-fixed telomerase sample, and 1292 micrographs were collected for the fixed sample.

The micrographs from unfixed HSP90 bound telomerase sample were motion-corrected using Motioncorr2 and CTF corrected using CTFFIND4. 215,483 particles were picked with cryolo picker (Wagner et al. 2019) and 2D classified. The well-aligned 2D classes (Figure 4.10) show HSP90 alone resembles the shape of a key and clear secondary features can be seen. The 2D classes showed the particle is symmetrical and maybe an oligomer. The observation fits with published work that HSP90 functions as a homodimer with a MW=180kDa (Wayne and Bolon 2007). Interestingly, additional clouds of densities can be seen in the 2D classes of HSP90, which may come from human telomerase. The good 2-D

classes were selected including 56,207 particles for 3D classification with the best 3D classes selected. Eventually, 24,924 particles were used for 3D masked refinement which gives an HSP90 density at 4.8Å resolution using Fourier Shell Correlation (FSC) golden standard cut-off (Scheres and Chen 2012). Local resolution was estimated using Resmap, which showed that the local resolution of most of the density is 4~5Å, but some small surface regions have a lower resolution of 8Å which does not affect the interpretation of the map. The map shows a homodimer of HSP90 protein with alpha-helices clearly visible, the fitting of crystal structures of HSP90 shows the map is unambiguous. There is an area of additional density is present in the final reconstruction of HSP90, by comparing the map with the available crystal structure of HSP90/P23 (Ali et al. 2006), the additional density was identified as P23. This is in agreement with previously published work that HSP90-P23 has a regulatory role in telomerase activity (Holt et al. 1999). Although the structure of HSP90-P23 was determined to high resolution, the more important structure is the complexes of HSP90 bound form of human telomerase. In the unfixed sample, I was not able to obtain any clear 2D classes containing human telomerase, thus, I moved to process the dataset collected on the crosslinked sample.

Figure 4.10 Cryo-EM analysis of HSP90 alone. The data processing was carried out using Relion (A) Representative 2D classes representing views of HSP90 dimer, the length is $\sim 160\text{\AA}$. (B) The final 3D reconstruction of HSP90 at a global resolution of $4.8\text{\AA}$, fitted with the available crystal structure (PDB: 2CG9) (C) The FSC curve of the final reconstruction showing at 0.143 cut off, the resolution of the density is at $4.8\text{\AA}$ (D) Resmap of the final 3D reconstruction showing local resolution of the density map. Most part of the density is at $4\sim 5\text{\AA}$, but some surface regions are less ordered and have a lower resolution of $8\text{\AA}$.

The micrographs from the fixed sample were motion-corrected using Motioncorr2 (Zheng et al. 2017) and CTF corrected using CTFFIND4 (Rohou and Grigorieff 2015). 101,476 particles were picked with Gautomatch (Zhang 2018) with previous human telomerase 2D classes as a reference and particles 2D classified using Relion (Scheres 2012). However, the resulted 2D classes are surprisingly bad, the blurriness in 2D classes indicates a poor alignment of the molecules. From the 2D classes, with difficulties, the shape of HSP90 was still visible. Therefore, I decided to do further processing using 3D classification hoping to obtain a low-resolution density of HSP90 bound human telomerase. Around 69,112 particles in classes that resemble HSP90 bound human telomerase were 3D classified. The best class containing 29,291 particles were 3D refined using previous obtained 3D density as the initial model. This result in a 3D density at about 22Å resolution. The dimension of the densities corresponded to the densities obtained in Section 3.4 for telomerase alone, suggesting the density is of human telomerase. The bilobal shape suggests this is a complex of human dimeric telomerase, with extra density at the hinge region of the telomerase density. However, since the density is at a very low resolution, the fitting would be difficult. On the bright side, the distinct shape of HSP90 can still be identified in the low-resolution density. Therefore, I have fitted the density map with the 4.8Å HSP90/P23 structure obtained by me into the hinge region as shown in Figure 4.11C. The model fitting suggests HSP90 makes extensive contact to the hinge region of human telomerase. The density suggests that HSP90 may directly contact the telomerase RNA subunit TERC since the hinge region was fitted with TERC in the negative stain structure of telomerase by the Rhodes group (Sauerwald et al. 2013). Although previous work suggests HSP90 is involved in the assembly and regulation of human telomerase by binding to TERT only (Forsythe et al. 2001), HSP90 has been shown to be interacting with other RNA domains as well (Huang et al. 2012). My results, therefore, suggest HSP90 interacts with TERT and TERC together and could play a

vital role in human telomerase biogenesis. Previous work has also shown that HSP90 is associated with active human telomerase (Lagadari et al. 2016), and this was also shown by my telomerase activity assay of HSP90 bound telomerase complex (Figure 3.11). My observations suggest binding of HSP90 may stabilise the dimeric structure in order for it to carry out catalytic function at telomeres. However, due to the very poor resolution, we were unable to map the detailed interactions between human telomerase and HSP90 at the molecular level. Nonetheless, by improving the resolution of this complex, it should be possible to gain structural insights into the interaction between HSP90 and human telomerase.

Figure 4.11 Cryo-EM analysis of HSP90 bound human telomerase. (A) Representative 2D classes from the dataset show very poor alignments, the shape of HSP90 is observable, and human telomerase is blurry. (B) 3D classes reconstructed from good particles with a resolution of $\sim 22\text{\AA}$. The density has a similar dimension as dimeric human telomerase density in Figure 4.5. (C) Fitting of HSP90 density into the density of HSP90 bound human telomerase which was possible because of the distinct shape of HSP90.

4.9 Comparison between EM densities of human telomerase.

Despite all the attempts to obtain a high-resolution cryo-EM structure of human telomerase, the best resolution obtained by us was $\sim 15\text{\AA}$. The recently published telomerase monomer structure at $8\text{\AA}$ by Collins group (Nguyen et al. 2018) suggests human telomerase contains one copy of TERT and one copy of TERC as well as additional binding factors (TCAB1, DKC1, NOP10, GAR1, NHP2). This is in clear contradiction to our bilobal human telomerase density maps showing our telomerase is a dimer containing two TERT subunits and two TERC. Our cryo-EM structure is in agreement with previously solved human telomerase structure at $25\text{\AA}$ by Rhodes group using negative stain EM (Sauerwald et al. 2013). Trying to understand the difference in telomerase structures, I decided to carry out a comparative analysis of the structure of the human telomerase obtained. Whilst all telomerase densities have a similar dimension of approximately $280\text{\AA} \times 150\text{\AA} \times 150\text{\AA}$, the shape of the densities has some differences (Figure 4.12).

Figure 4.12 Comparison of density maps of human telomerase. (A) Comparison between densities obtain by our group in Singapore, the Rhodes group (EMD-2310) and Collins group (EMD-7521). (B) Comparison between telomerase density alone and HSP90 bound telomerase density obtained by me.

Densities obtained by me is similar to previously published Rhodes structure by negative stain, both densities have a clear bilobal shape with a hinge region in the middle (Figure 4.12A). In both cases, biochemical studies showed that the telomerase sample gave rise to the density map should contain two TERT and two TERC subunit. The telomerase density obtained by me has an improved resolution of 15Å compared to Rhode previous structure at

25Å by the negative stain. More details of the folds on the surface of the lobes and the twisted RNA region can be seen. In agreement with Rhodes group's density, the two lobes from my density also appear to be conformationally different, with one lobe being more "open" than the other one. Although the overall shape is similar, there are some differences between the two densities. Also, my telomerase density has an extensive hinge region composed of TERC connecting the two lobes. The hinge region of Rhodes group's negative stain density occupied more volume, which is due to the flexible RNA region not being well resolved as they suggested in the paper (Sauerwald et al. 2013). The densities of two lobes from my telomerase appear to be closer to each other in comparison to the negative stain density. I attempted to fit the two lobe densities with beetles TERT with the ring-shaped catalytic pocket positioned on a "hole" from the density (Figure 4.12).

Collins group's density is also bilobal, however, the overall shape, as well as individual lobes, have a very different shape compared to my density. They have fitted one lobe with TERT, while the other lobe was fitted with the cofactors, in contrast, we have fitted both lobal density with TERT which means we must be investigating a different telomerase structure (Figure 4.12A). The density of the hinge region obtained by Collins group occupied less space than my density, as they have only fitted one copy of TERC, while I expect two TERC to be fitted in the same region. Although the length is about 280 Å in both densities, the width of our telomerase is slightly longer than Collins, 150Å vs 120Å. The total volume occupied by telomerase from my density is also greater than Collin's density, due to a larger molecular weight and the presence of additional TERC.

The density of HSP90 bound telomerase has a similar dimension compared to the density of telomerase alone. The HSP90 bound human telomerase also has a bilobal shape with an additional density of HSP90 in the hinge region, also suggesting the human telomerase is a dimeric complex (Figure 4.12B). However, the two lobes in the density of HSP90 bound

telomerase is closer to each other compared to the density of telomerase alone, this could be due to interaction with HSP90.

In conclusion, the 3D-densities obtained by me suggests human telomerase functions as a dimer. The two lobes are structurally different from each other, suggesting they may have different functions during the catalytic cycle. The biochemical difference of telomerase complexes purified by our group and Collins group is addressed in detail Section 5.1.

4.10 Conclusions and discussions

At the beginning of the work presented in this thesis, it was thought the biggest hindrance in obtaining a high-resolution cryo-EM structure of human telomerase was the sample purification. The cryo-EM procedure was perceived as a straightforward process once a high-quality human telomerase sample was purified. However, as the project unfolded, I have encountered many difficulties during different stages of the cryo-EM workflow.

Despite efforts spent to optimise telomerase expression and purification protocol, the concentration of purified telomerase sample was still not ideal for structural analysis. With no other choices, I was forced to optimise grids preparation conditions in order to achieve a good sample capture. On estimation, around 99.9% of protein particles are lost during the grids blotting step (Kemmerling et al. 2012). This issue did not gain too much attention from the field as the loss could be easily compensated by increasing the concentration of protein samples. However, this was not possible with human telomerase due to the low yield. As many researchers want to use cryo-EM to obtain structures of less abundant proteins or complexes, there are several ongoing developments addressing this issue which will be discussed in Section 5.4. Eventually, I successfully prepared cryo-EM grids with good particle distribution.

From the first cryo-EM dataset obtained by capturing telomerase on graphene oxide coated grids, nice bilobal 2D classes were obtained which correspond to the size of dimeric human telomerase as established in the previous work by Rhodes group (Sauerwald et al. 2013).

However poor alignment in the 3D classification resulted in the low-resolution reconstruction of human telomerase at $\sim 15\text{\AA}$ showing a bilobal shape. First, it was suspected that the graphene oxide interferes with the processing, thus I switched to capture the sample in holey carbon grids. The 3D reconstruction of particles from the new dataset remained at around $15\sim 20\text{\AA}$ with no obvious improvement. Attempting to improve the resolution by having more particles, I decided to collect a large dataset, totalling around 12, 000 images.

Disappointingly, even with large amounts of data, the best resolution obtained was around $15\text{\AA}$. At this resolution, model fitting is ambiguous. As we were very puzzled by the result and want to exclude the possible damage to telomerase particles through grids preparation, I went back to analyse the sample by negative stain EM. This revealed that the sample is comprised of three different shapes, which suggest the purified human telomerase sample is conformationally heterogeneous. The cause of heterogeneity was then subjected to further investigation.

Damage of protein particles at the air-water interface is an emerging issue in the field, which will be addressed in detail in Chapter 5.4. This prompted a further investigation to whether my telomerase is affected by this. Through the cryo-ET analysis, the water-air interface was identified, and the data showed that telomerase does not gather at the water-air interface and hence the heterogeneity must come from other causes. A low-resolution density was also extracted from the tomograms. This density has a bilobal shape in agreement with my previous 3D reconstructions from single-particle analysis.

Attempt to solve the heterogeneity issue included stabilising human telomerase with additional binding partners and the HSP90/telomerase complex was purified. This complex

gave rise to a high-resolution structure of HSP90/P23 to 4.8Å using cryo-EM. However, the HSP90/P23 bound telomerase did not refine well and only at low resolution (~ 20 Å) even when the HSP90 density was clearly visible. The results of this experiment provide evidence that my cryo-EM sample preparation and processing is capable of obtaining a high-resolution structure. Thus, the problem that hindered me from obtaining the high-resolution structure of the human telomerase likely to arise from the high flexibility of human telomerase dimer.

In conclusion, I have successfully optimised conditions for grid preparation step to successfully capture human telomerase with good particle distribution in ice. Despite many attempts in trying to obtain the high-resolution structure of the human telomerase, I was only able to obtain a density of human telomerase to a resolution of ~ 15 Å which is an improvement on the 25 Å structure of human telomerase solved by negative stain. However, at this resolution, it is still not possible to do an accurate model fitting. Preliminary attempts to fit TERT and TERC into the density shows that the telomerase likely contains two TERT and two TERC subunits. Structural comparison between the low-resolution structure obtained from different dataset consistently showed densities of a bilobal shape, which indicates the dimeric nature of the complex. The composition and oligomeric state of human telomerase are discussed in detail in Section 5.1. Since different ways of purifying telomerase have been tried and yet we did not see any improvement of resolution, I highly suspect the fundamental problem arises from the structural heterogeneity of the human telomerase sample, which will be further discussed in Section 5.2.

Chapter 5 Discussion

5.1 Composition and oligomeric state of human telomerase

As introduced in Chapter 1, the biogenesis pathway of the human telomerase is a complex, multistep process. This involves the synthesis of hTERT protein in the cytoplasm by ribosomes, transport into the nucleus and the Cajal body for assembling with the RNA subunit TERC and the H/ACA box protein. Furthermore, human telomerase has been shown to be associated with over 30 proteins (reviewed in (Yang, Song, and Johnson 2016)). Some interacting partners such as TCAB1, HSP90, TPP1 have been shown to regulate telomerase activity, whereas others such as dyskerin, pontin and reptin have been shown to aid the assembly and maturation of human telomerase (Gallardo and Chartrand 2008). The function of other binding partners has not been well studied. These observations suggest the human telomerase can exist in complex with different proteins depending on the location in the cells. Potentially, there could be many human telomerase complexes with different binding partners or cofactors with different cellular localisation and functions.

The question of the natural composition and oligomeric state of active human telomerase is poorly understood and remains to be answered. Several studies regarding the oligomeric state of human telomerase report contradictory findings. The protein catalytic subunit hTERT (127kDa) and the RNA subunit hTERC (153kDa) are the key components of telomerase. In the early days, researchers of the field developed methods to reconstitute human telomerase actively by translating protein subunit TERT in presence of *in vitro* transcribed TERC in rabbit Reticulocytes lysate (RRL). Studies showed reconstituted telomerase using RRL is active suggesting only hTERT and hTERC are necessary and sufficient for activity *in vitro* (Weinrich et al. 1997). Linger's group showed that *in vitro* transcribed human telomerase complex contains two TERC by quantifying the amount of TERC per telomerase (Wenz 2001). The first evidence that physiological human telomerase exists as a dimer come from

Cohen, where he shows endogenous human telomerase purified from various cancer cell lines have a size of a dimer (~670kDa) on a sucrose gradient (Cohen et al. 2007). Later, Rhodes' group demonstrated that overexpressed and purified human telomerase is a dimer both structurally and biochemically. They used nano-gold particles to identify the location of protein and RNA subunit (Sauerwald et al. 2013). The first experiment showed two nanogold particles were binding to the two lobal ends of human telomerase dimer with His-tagged hTERT, showing that complex must contain two copies hTERT. The second experiment used nanogold bound to monovalent streptavidin, acting as a probe for biotin labelled G-overhang substrate pre-bound to human telomerase dimer. This data shows that up to two nanogold could be bound per human telomerase particle. As each RNA subunit contains only one substrate binding side, this shows there must be two copies of hTERC. They also biochemically showed that co-expression of wild type and a TERT active site mutant of human telomerase in 1:1 ratio caused human telomerase to lose its activity completely. This proved that human telomerase must be functional as an oligomer and since the molecular weight of the isolated human telomerase corresponds to that of two hTERT and two TERC it was concluded that human telomerase is a dimer and functions as a dimer (Sauerwald et al. 2013).

By contrast, telomerase from various of other species such as yeast and tetrahymena has been shown to be monomeric (Bryan, Goodrich, and Cech 2003; Shcherbakova et al. 2009), thus some researchers argue that human telomerase must also be monomeric. By biophysical approach, Balasubramanian group showed the human telomerase reconstituted from RRL is a monomer (Alves et al. 2008). Whereas later, by single-molecule fluorescent microscopy, the Collins group showed that human telomerase purified from mammalian cells is also monomeric (Wu et al. 2015). The most profound evidence comes from the published cryo-EM structure of monomeric human telomerase by Collins group (Nguyen et al. 2018). The

8Å map, whilst having a bilobal shape similar to the Rhodes structure, was fitted with one hTERT, one hTERC and various additional binding proteins as shown in Chapter 1. They also show some experimental evidence that human telomerase is more catalytically active in the presence of the H/ACA box proteins. However, they have not reported the molecular weight of their human telomerase complex. They have fitted their monomeric telomerase density map with a stoichiometric amount of additional cofactors to telomerase. In clear contrast to their SDS-PAGE analysis, they have not demonstrated that the cofactors are present in a stoichiometric amount as fitted in the density map (discussed in detail below).

The work I presented in this thesis argues that human telomerase is functional as dimer consisting of hTERT and hTERC alone. First, I have shown that the human telomerase complex I have purified migrates at a molecular weight of 450~670kDa in the sucrose gradient corresponding to the MW of a dimer, consisting only 2x TERT and 2xTERC with a combined Mw = 600kDa. Furthermore, we demonstrate that only telomerase migrating as a dimer in the sucrose gradient is active. Then, through analysing the protein composition by western blot against the H/ACA box protein and silver stain, I showed that my purified human telomerase sample contains almost no H/ACA box protein and consist of TERT and TERC alone. This observation coincides with the previous work showing that only TERT and TERC are required for activity *in vitro* (Bachand and Autexier 1999). Going a step further, we showed that the shelterin binding complex TIN2-TPP1-POT1 increases the processivity of our purified dimeric human telomerase complex containing TERT and TERC alone. This functional interaction suggests that the human telomerase complex recruited to telomeres may be a dimer consisting of twoTERT and two TERC.

Although I only obtained a low-resolution 3D density map of human telomerase at about a resolution of 15Å, which prevents an accurate and independent interpretation, the density

map still provides strong evidence on the oligomeric state of human telomerase. First, a bilobal shaped density was consistently obtained in 3D classifications of sample purified with different methods as shown in Chapter 4. The bilobal shape agrees with the previous published dimeric structure of human telomerase from negative stain EM (Sauerwald et al. 2013). Since biochemical analysis shows the sample used for cryo-EM analysis contains only TERT and TERC, the dimensions of the density corresponding to the size of two TERT and two TERC. Also, the 2D classes clearly show human telomerase classes with bilobal shape. Interestingly, when I attempted to purify human telomerase bound to additional cofactors by using a different extraction condition, the purified sample contained stoichiometric amounts of HSP90, which was not reported by either Rhodes or Collins groups. The interaction between HSP90 and human telomerase could not be fully characterised due to the limited resolution obtained. Initially, work by others suggested HSP90 is involved in the biogenesis of human telomerase (Viviescas, Cano, and Segatto 2018), whilst other groups have suggested HSP90 has a regulatory role on human telomerase (Lagadari et al. 2016). The purified HSP90 bound form of human telomerase was shown to be catalytically active by me (Figure 3.11) and others (Forsythe et al. 2001), suggesting this complex exist *in vivo* and may be involved in the telomerase recruitment process (Lagadari et al. 2016).

It is important to note that the results obtained using purified telomerase only represent how telomerase functions *in vitro*, the nature of the telomerase complex at telomeres *in vivo* is unknown. Although we showed the addition of shelterin component increases the processivity of the human telomerase enzyme, I was not able to co-purify any of shelterin components with human telomerase in pull-down studies. This may be explained by the tight regulation of telomerase recruitment; the hTPP1 components of the shelterin complex may only interact with human telomerase under certain conditions during the S-phase of the cell

cycle (Tomlinson et al. 2006). This was also biochemically demonstrated for ciliate telomerase *in vivo* (Paeschke et al. 2008).

In my opinion, human telomerase may exist both as monomeric and dimeric complex in cells, depending on the cellular location and stage of the cell cycle. There is no conclusive evidence to show which form is more physiologically relevant. This may be answered by the continuously developing EM field in the future. The cellular cryo-Electron Tomography is under rapid development, and new tools such as cryo-Fib and dual-axis tilt have been introduced (Schaffer et al. 2019). A cryo-FIB lift-out technique enables molecular-resolution cryo-ET within cells and native *Caenorhabditis Elegans* tissue. This technique allows direct observation of proteins complexes *in situ*. Currently, the cryo-ET is primarily used for large and abundant proteins complexes because the densities are obtained from sub-tomograms averaging from large particle sets. From a single protein molecule, the best resolution obtained so far is only around 35Å and suffers hugely from the missing wedge problem (Paavolainen et al. 2014). The missing wedge problem arises from the limitations of tilting angles, the imaged ice thickness increases as tilt angle increase, which reduces the image contrast and imaging of tilt angle above 60 degrees is generally not recommended. At the resolution limit of 35Å, fitting of the densities is not possible, it may provide useful information about size, shape and conformation of the telomerase complexes. This could be used to identify the density of human telomerase bound at telomeres. With further improvement in resolution, we may learn about the telomerase recruitment process. Once a clear telomerase structure at the telomeres is solved, it would provide many more insights into the telomere extension mechanism and telomerase's interaction with shelterin components. This would be very exciting and eventually end the debate on what is the biologically relevant composition of the telomerase and telomeres.

5.2 The conformational heterogeneity of human telomerase limits the resolution of structural analysis by cryo-EM

From the work presented in this thesis, I conclude that it is the structural heterogeneity of the dimeric human telomerase that prevented us from obtaining a higher resolution structure. The structural heterogeneity likely rises from the conformationally flexible nature of human telomerase. Flexible proteins are often involved in the regulatory mechanism, especially the ones interacting with DNA. The link has been drawn between conformational flexibility and the regulation in transcription factors (Khan and Kumar 2009). As suggested in the introduction, human telomerase is tightly regulated and needs to undergo significant conformational change during its catalytic cycle during telomere extension. Furthermore, telomerase is expected to undergo conformational changes during the process of recruitment. In addition, human telomerase is an RNP particle and its RNA component that may change its shape during telomere synthesis or in different complexes adding additional complexity to the conformational flexibility of telomerase.

Cryo-EM provided a solution to solve structures of macromolecular complexes that cannot be crystallised or are of low abundance. Crystallisation relies on large quantities of sample and a rigid protein forming a regular lattice structure, which is difficult if the protein contains flexible or unstructured domains, thus truncations of domains are very common. The fundamental of EM processing relies heavily on the classification process, which is based on computational image alignments. Macromolecular complexes that are captured in fixed conformations can be classified very well with this algorithm. However, many studies show that the structure determination of highly flexible macromolecular targets remains difficult (Nwanochie and Uversky 2019).

EM structures solved in the early days were large macromolecular complexes like viruses, filaments and ribosomes. Macromolecular complexes that are flexible but may be restricted

to certain conformations with a small substrate have been resolved after considerable efforts to high resolution by cryo-EM. Such structures are represented by the ribosome and spliceosome. In addition, structures of some DNA and RNA binding proteins that have multiple conformations can be determined by cryo-EM. For example, William's worked on ATM kinase which has been resulting in a low-resolution structure for many years, until a change in sample preparation and improved processing enabled the structure to be solved to 5.7Å resolution in the closed state and 11.5Å resolution in the open state (Truman et al. 2017). Another group working on ATM kinase was only able to obtain a high-resolution structure by using the solved structure as an initial model. In Lori Passmore's Cleavage and Polyadenylation Factor (CPF/CPSF) work, this problem is further highlighted. The published work shows attempts to solve the structure of the complex by making many different constructs, but the resolution was still limited to 25Å (Hill et al. 2019). Also, Passmore's recent structure of the monoubiquitin ligase complex to ~5Å was also hindered for many years previously (Shakeel et al. 2019). Computational techniques provided a partial solution to conformation flexibility problem through the introduction of masking and multi-body refinement which allowed more protein structures to be solved to high resolution. For example, the secretin complex of *Thermus thermophilus* has been solved to ~7Å with the multibody refinement protocol, which was previously limited to only ~15Å resolution (D'Imprima et al. 2017).

In theory, the structure of proteins and complexes that are flexible and have many conformations should still be amenable to be determined using cryo-EM, since obtaining the structure of cryo-EM relies on aligning molecule with the same conformation. The question now becomes the numbers of particles possessing the exact same conformation in the dataset. Taking spliceosome as a prime example, it is an RNP particle that could undergo many conformational changes. In this work, the structure of three conformations of human

activated spliceosomes was solved to 4~5Å (Zhang et al. 2018). During the final 3D refinement only around 3% of the total number of picked particles were used for the reconstruction of the high-resolution structure. The protein sample may contain a small percentage of the spliceosome in different conformation, and accurate alignment of those in the same conformation is the important step in producing high-resolution structure.

In Collins human telomerase work, more than 1million particles were picked, and less than 3% of particles were used for the final reconstruction (Nguyen et al. 2018), with rest of the particles discarded during the processing step. However, the composition or conformation of discarded particles was overlooked during the process. As highlighted in red in Figure 5.1, some of Collins telomerase 3D classes look very similar to 3D classes of our human telomerase with a bilobal shape connected with a single hinge region. These discarded classes contained 21.8% of the total number of particles (Figure 5.1). The highlighted class in blue also looks similar to our 3D class of human telomerase, with a bilobal shape connected by two hinge regions connecting each lobe, and those classes that were discarded contained 31.6% of the total number of picked particles. It is perhaps surprising and alarming that a huge proportion of particles in their 3D classes (53.4%) is in agreement with our work. However, the structure they have chosen to present was very different. Collins did not address the discarded classes. Furthermore, all highlighted discarded classes have a longer length than the classes picked, which suggest these classes may be structurally and compositionally different from their obtained structure. We could only speculate that they discarded these classes because they could not obtain a high-resolution structure from the particles in these classes.

Figure 5.1 The processing workflow leads to the 8Å structure of human telomerase monomer by the Collins group (Nguyen et al. 2018). The workflow shows the processing steps involved to obtain the structure of telomerase. The black box shows the selected class from the original publication. Red boxes highlight clear bilobal densities that were not selected (21.8% of total). Blue boxes highlight classes with bigger dimensions than solved structures which were also not selected (31.6% of total).

Looking back at our human telomerase datasets, despite trying many different purification and processing strategy, we were still unable to obtain a high-resolution structure of human telomerase. The problem could be explained by the dimeric telomerase existing in many different conformations. The conformational flexible means that the computational algorithm could not sufficiently align particles into 2D or 3D classes. The cause may either be that the computational algorithm could not differentiate between similar conformations, or because

the existence of many different conformations, there is an insufficient number of particles in each conformation to be classified. If the latter is the case, then this problem may potentially be solved by collecting an even bigger dataset with ten times the number of particles.

In contrast to the purified dimeric human telomerase, the process of obtaining a high-resolution structure of HSP90/P23 dimer was a straightforward process. Although the protein sample analysed contained a mixture of HSP90 and human telomerase it was possible to obtain a high-resolution HSP90/P23 structure alone from a sample containing both HSP90 and telomerase. In addition, it is important to note that the HSP90/P23 complex is only ~230kDa which is considered a small target for structural studies by cryo-EM, but despite this by setting appropriate parameters in the autopicking software cryolo (Wagner et al. 2019), the program efficiently picked the particles resembling HSP90/P23. This was demonstrated by the fact that in the subsequent step of 2D and 3D classification, the alignment of the particles was excellent, and even secondary structures are visible in the 2D classes. The entire process of obtaining the structure of the HSP90/P23 complex was straightforward without further complications. This showed if a protein has a rigid and well-defined shape, the protein is a good target for studies using cryo-EM. The result suggests it may be possible to specifically pick particles by using auto picking from a relative “crude” sample containing a mixture of proteins if the protein is rigid and have a defined shape and size. The classification algorithm is very powerful in distinguishing between completely different protein complexes, and in classifying particles with very distinct conformations. This also provided a control for the validity of the correctness of the data processing.

Due to the impressive developments occurring in the EM field, the process of obtaining a high-resolution structure of a well-behaved homogenous macromolecule complex is straightforward. However, if the protein undergoes many conformational changes and the flexibility occurs on a global level, then the structure of the protein may still be difficult to determine by

cryo-EM. The problem may be compensated by the collection of a huge dataset with a large number of particles (>1million). Or in the future, the development of computational algorithms may be able to track the changes between conformations and compensate for flexibility, thus providing a solution to this problem.

5.3 Cryo-EM development and the future

The cryo-EM of biological samples has undergone a dramatic change during 2012-2019, and this is often referred to as the “resolution revolution”. Coincidentally, the timeline of cryo-EM development overlapped with the period of my PhD project and happened as my projects unfolded. The first major improvement came with the introduction of high energy (300kV) and a more stable electron microscopy instrument released by FEI, named Titan Krios. The number of high-end 300kV microscope has largely increased, and more users have access to the technology. More users mean more attention cryo-EM methods gains, which means more research groups/companies conduct related studies to improve the technology constantly. At the beginning of my PhD in 2016, there were less than 20 FEI Titan Krios available around the world, and now there are more than 80. The direct electron detector K2 was first developed in 2014 which allowed a more accurate recording of signals, and together with computer programs allowed high throughput data collection. The FEI Titan Krios installed at NTU that I performed data collection on was only commissioned to use in April 2018. Before this, the only in-house microscope available was a 200kV Tecnai Arctica. The instalment of Titan Krios allowed me to obtain micrographs with better contrast and rapid collection of large datasets, which helped my structural work. Beam-induced motion correction software first became available in 2012, but the improved and widely used version only become available in 2017, which had a huge contribution to improving the resolution of structure solved. Also, in 2017, the commercial release of the Volta phase plate allowed imaging of

small protein particles (<300kDa) with high contrast. When the Volta phase plate first became accessible inhouse in 2017, I was optimising the sample capture on cryo-EM grids. The phase greatly enhanced particle contrast on the 200kV Arctica microscope allowed me to unambiguously assess the grid quality, which would be difficult using standard objective lenses. The development of various cryo-EM data processing software is a very active field and is ongoing, and many versions were released during the course of my PhD, each time improving the algorithm for alignment and clarity of interface for uses. Notably, the method of multibody refinements in Relion 3 (Nakane et al. 2018) greatly helps to solve multiprotein complexes that have defined regions which improves the local resolution of huge complexes. However, this method was not suitable for my use, as the human telomerase cannot be split into well-defined regions due to lack of clear boundaries.

A future exciting prospect is the rapid emergence of the cryo-electron tomography field. The main advantage of cryo-ET is that it allows direct observation of protein and complexes *in situ* in their native cellular conditions, best preserving the spatial arrangements.

Traditionally, tomography was performed on fixed plastic sections of cells or tissue, which produced images with sufficient resolution to see cellular organelles. In 2002, the first application of cryo-ET was used to view the architecture of cells, and the resolution was limited to 5-6nm (Medalia 2002). Later, the method was widely used in microbiology and virology studies. The main difficulties with cryo-ET were that biological samples suffered from radiation damage and using a low dose of electrons resulted in low contrast images. The development of the phase plate in 2017 provided a great solution to this problem. Together with the other improvements on cryo-EM during the resolution revolution, scientists can observe proteins in frozen native cells at ~20Å resolution (Albert et al. 2017). At this resolution, the assembly, functional state, and binding state may be differentiated.

The future of the cryo-EM field is moving towards faster data acquisition as well as higher resolution. The first structures produced from cryo-EM was a virus structure, large chaperon, and ribosomes. At the beginning of my Ph.D., it was widely believed complexes with MW>500kda would be a good target for cryo-EM work. But recently, published work shows an atomic resolution structure of 52kDa may be obtained through cryo-EM. This only sets the limits for now, but with further technology developments, it is conceivable that even smaller proteins may be solved using cryo-EM. Considering there are potential improvements in many more areas of cryo-EM, the application could become wider. I believe the resolution of cryo-ET will be further improved in the next few years, and cryo-ET may be used to study many more biological problems. Cryo-ET may become a novel approach to discover protein binding partners in cells and investigate the spatial location of different binding states aiding cell biology studies. Even more interesting, by using gene-editing techniques, cell biologists may obtain direction information on the cellular effects of mutations and knockouts.

5.4 Current difficulties in cryo-EM

The development of EM hardware and software has enormously helped scientists to obtain high-resolution structures of proteins. As the technique develops, many underlining problems start to emerge which were previously not discovered or considered due to previous limitations of EM. There are three major issues that I have encountered during my project and have also been reported by other scientists.

One major problem is what happens in the sample blotting step during grid preparation. As introduced, during cryo-EM grids preparation, the sample is applied to the grid support and blotted by filter paper and subsequently plunge frozen. This process may be carried out manually or automatically. Currently, the most reproducible and reliable way of grid

preparation is using a semi-automated machine such as FEI Vitrobot, Leica EM GP (and GP2) and Gatan Cryoplunge 3. After blotting, because of the very small volume of liquid on grids, any evaporation would alter the concentration of buffer composition, which may drastically affect the stability of protein complexes. Another major issue is sample loss during the blotting step when the blotting papers come in contact with the sample droplet. Either the physical interaction or charge interaction leads to a major loss of protein particles in the final grids. This issue is well known and commonly solved by using high concentrations of samples (0.5 mg- 4mg/ml). However, this creates a limitation for samples that are difficult to purify to a high concentration. Human telomerase is a prime example where only around 300~500µg of protein can be obtained from six HYPERflask of culture. Therefore, the rate of testing grid preparation condition was greatly limited by the amount of protein required, although the required protein amount is still far less than required by crystallography. Other challenging targets include the spliceosome, which is also hard to overexpress and shares the same technical difficulties as telomerase. In order to create enough sample to solve the spliceosome structure, 60L of yeast culture had to be inoculated. So far, such a problem can only be resolved by increasing the resources used. Blotting also can be controlled by manipulation through the old-fashioned manual blotting, which was successful for many published works; however, the reproducibility remains low. As more scientists become aware of the issue, many research groups are presenting solutions to eliminate the blotting process during grid preparation. For example, Vitrojet is a robot that utilises pin-painting, which “paints” a small amount of protein, followed by an automated jet-freezing process (Ravelli et al. 2019). The Spotiton or Chameleon robot functions by sparing a small volume of protein on grids (Dandey et al. 2018), which produces sufficiently thin ice directly without the need for blotting. Such sophisticated machines eliminate the process of blotting and provide an automated platform for sample freezing. Potentially this provides an ultimate solution to the

sample blotting issue and allows rare and low concentration protein complexes to be studied using cryo-EM. Furthermore, the automated freezing and transfer platform may increase the reproducibility and reduce the number of failed grids caused by human error. Hopefully, the introduction of the new tools will allow easier cryo-EM sample preparation with human telomerase.

Another major issue is the “preferred orientation” problem (Drulyte et al. 2018). When the macromolecules are frozen in ice, theoretically, they may be orientated 360 degrees in all directions. However, in many cases, particles are found to orientated in one, two or few preferred orientations. The reason behind it is difficult to understand, but reasons may include effects by the physical shape and charge distribution of the complex. As introduced before, the 3D reconstruction process relies on information from 2D classes, and each orientation contributes to the 3D slice in Fourier space, thus more orientations mean better completeness of the density. Having a few orientations may be sufficient for a 3D-reconstruction, but if only one or two preferred orientation is present, then the information is limited to those orientations only. The reconstruction will be missing 3D-information from many other directions, and the computational algorithms attempt to create a density by assigning false densities to the missing dimensions. Luckily, this problem can be easily spotted by mapping the orientation distribution or by analysing the density directly. Proteins that have a heavily charged surface or located at the water-air interface are particularly susceptible to this problem. Currently, there are no computational methods that can compensate for this problem. Researchers attempt to solve this issue by using different grid support materials and treatments, which aim to change the charge interaction with the grid surface. Alternatively, collecting datasets at tilt may help obtain information for other orientations (Tan et al. 2017). However, the data acquisition process is difficult due to the limitation of the tilt axis and data processing may present challenges due to CTF distortions created by tilting the stage.

Currently, there are no reliable methods that can provide a solution to this problem. The human telomerase was not affected by this problem; however, the human telomerase particles were not evenly distributed across the grids, which potentially could be solved by changing the surface.

In addition to the above, the third most commonly recognised issue arises from the “water-air interface problem”. This problem has been long-standing (Glaeser and Han 2017), but only well addressed recently. In recently published work, 51 protein samples were tested and characterised using tomography methods (D’Imprima et al. 2019; Noble et al. 2018), and more than 30 proteins were found to locate at the water-air interface. Different proteins that sit at the water interface can behave very differently, some may be unaffected, while some may sustain partial denaturation, complete denaturation, or adopt a preferred orientation.

When this problem was discovered, the first method proposed to overcome the problem was the use of detergents which form a physical barrier that stops proteins from being in contact with the water/air interface (Glaeser et al. 2016). Others have attempted to solve this problem by modifying the properties of the grid-support. Most common cryo-EM grids are coated with amorphous carbon which is difficult to functionalise or modify. Russo and other groups established a protocol to produce graphene and graphene oxide grids aimed to improve sample capture and also lower the background in EM images (Russo and Passmore 2016). Later on, other groups exploited the chemical properties of graphene and graphene oxide to produce a modified surface (D’Imprima et al. 2019; Wang et al. 2019). The modified surface can have a positive or negative charge, or even further modified to have a specific affinity to DNA or purification tags. All these modifications aim to attract the proteins to a surface and prevent movement to the water-air interface. However, the effectiveness of such methods is still inconsistent across different protein targets. Even though human telomerase particles were not located at the water-air interface, it could still be damaged by coming in

contact with the air due to fast diffusion (Israelachvili and Israelachvili 2011). The frequent contact with the water-air interface may cause damage to the protein and hence hindering the structural work.

In summary, the outlined major technical issues may hinder structural studies of many proteins. The cryo-EM community is actively seeking solutions to these problems. While the blotting issue is the closest to be resolved, the other two issues require significantly more work due to the complexity of chemical interactions.

5.5 Future work and outlook

The high-resolution structure of the human telomerase is yet to be determined. Currently, the best available human telomerase structure is at $\sim 8\text{\AA}$ resolution (Nguyen et al. 2018). At this resolution, the fitting of protein and RNA is ambiguous, and it is impossible to deduce any molecular mechanism in detail. My findings suggest that human telomerase may have many different conformations and forming different complexes *in vivo*, however, the most physiological relevant one is yet to be identified. Below I present my future strategy for obtaining a high-resolution structure of human telomerase and a method to investigate the physiologically relevant form of human telomerase.

In order to enhance the overexpression of the human telomerase, a permeant cell line was selected and cultured in a period of 2-3 months. However, if a different constructor cell line needs to be tested, it would add significant time to the process since it is very time-consuming. I would first establish a protocol to transiently transfect mammalian cells with high efficiency of over-expression. This would allow quick testing of constructs with mutations or truncations, and even constructs that co-express telomerase with other proteins.

The grid preparation step has huge potential to be improved. The functionalisation of support surfaces provides a possible way to create affinity grids. Since the telomerase construct contains a Strep-II tag, it is possible to modify the surface of the grid to be covered by streptavidin which binds Strep-II tag with nanomolar affinity. Or it may be possible to attach telomeric DNA substrates to the grid surface, that selectively bind to properly assembled and matured telomerase. Both methods would increase the specific capture of telomerase sample on grids and physically preventing telomerase from contacting the water-air interface. Also, in the hope to make telomerase distribute more evenly across the surface of the grid.

Since I concluded that the major problem that hinders high-resolution determination of telomerase is its conformational heterogeneity. This may be addressed by collecting an even larger dataset (>30000 micrographs) and manually classifying telomerase particles of different conformations into different classes. This method potentially allows a more detailed analysis of multiple telomerase conformational states through the catalytic cycle. However, it would be difficult to judge subtle changes in conformational states by human eyes, thus some computational aid will also be required.

Biochemically, it is important to investigate the nature of the human telomerase complex at the telomere ends. By trapping cells with short telomeres in the S-phase of the cell cycle, telomerase is expected to localise to the telomeres. By freezing the cells and using cryo-ET to image the telomere directly *in situ* would provide insights on interactions between human telomerase and telomeres. At the resolution limit of cellular tomography, different conformational states of human telomerase trapped in different stages of the catalytic cycle may also be observed. Moreover, the cryo-ET provides a tool to study the spatial arrangements of Cajal bodies and their functional role in human telomerase recruitment.

Another interesting aspect is the role of shelterin proteins in the telomerase recruitment process. I managed to produce results which are in agreement with past literature showing shelterin components TPP1 has a stimulatory effect on human telomerase activity and processivity (Zhong et al. 2012). However, I was not able to copurify any shelterin component with telomerase complex, thus I was unable to obtain any structural information of interactions between shelterin proteins and telomerase. Potentially, overexpressing a tethered construct linking hTPP1 and hTERT together may allow structural studies on this complex bound to a motel telomere. The structure may provide insights into the recruitment and regulation of human telomerase.

Once the atomic resolution structure of human telomerase is obtained, direct inhibitors of telomerase activity and allosteric inhibitors may be developed based on the structure. The structural and spatial effects of telomerase inhibitors may be studied using both single-particle EM and tomography techniques. This potentially could aid the development of anticancer and antiaging therapies.

Reference

- Abreu, E., E. Artonovska, P. Reichenbach, G. Cristofari, B. Culp, R. M. Terns, J. Lingner, and M. P. Terns. 2010. "TIN2-Tethered TPP1 Recruits Human Telomerase to Telomeres In Vivo." *Molecular and Cellular Biology* 30(12):2971–82.
- Akincilar, Semih Can, Kee Chung Low, Chia Yi Liu, Ting Dong Yan, Asami Oji, Masahito Ikawa, Shang Li, and Vinay Tergaonkar. 2015. "Quantitative Assessment of Telomerase Components in Cancer Cell Lines." *FEBS Letters* 589(9):974–84.
- Albert, Sahradha, Miroslava Schaffer, Florian Beck, Shyamal Mosalaganti, Shoh Asano, Henry F. Thomas, Jürgen M. Plitzko, Martin Beck, Wolfgang Baumeister, and Benjamin D. Engel. 2017. "Proteasomes Tether to Two Distinct Sites at the Nuclear Pore Complex." *Proceedings of the National Academy of Sciences* 114(52):13726–31.
- Ali, Maruf M. U., S. Mark Roe, Cara K. Vaughan, Phillipe Meyer, Barry Panaretou, Peter W. Piper, Chrisostomos Prodromou, and Laurence H. Pearl. 2006. "Crystal Structure of an Hsp90–Nucleotide–P23/Sba I Closed Chaperone Complex." *Nature* 440(7087):1013–17.
- Alves, David, Haitao Li, Rosalind Codrington, Angel Orte, Xiaojun Ren, David Klenerman, and Shankar Balasubramanian. 2008. "Single-Molecule Analysis of Human Telomerase Monomer." *Nature Chemical Biology* 4(5):287–89.
- Armbruster, B. N., S. S. R. Banik, C. Guo, A. C. Smith, and C. M. Counter. 2001. "N-Terminal Domains of the Human Telomerase Catalytic Subunit Required for Enzyme Activity in Vivo." *Molecular and Cellular Biology* 21(22):7775–86.
- Armbruster, B. N., C. M. Linardic, T. Veldman, N. P. Bansal, D. L. Downie, and C. M. Counter. 2004. "Rescue of an HTERT Mutant Defective in Telomere Elongation by Fusion with HPot1." *Molecular and Cellular Biology* 24(8):3552–61.
- Armstrong, Christine A., Vera Moiseeva, Laura C. Collopy, Siân R. Pearson, Tomalika R. Ullah, Shidong T. Xi, Jennifer Martin, Shaan Subramaniam, Sara Marelli, Hanna Amelina, and Kazunori Tomita. 2018. "Fission Yeast Ccq1 Is a Modulator of Telomerase Activity." *Nucleic*

Acids Research 46(2):704–16.

- Arnoult, Nausica, and Jan Karlseder. 2015. “Complex Interactions between the DNA-Damage Response and Mammalian Telomeres.” *Nature Structural & Molecular Biology* 22(11):859–66.
- Bachand, François, and Chantal Autexier. 1999. “Functional Reconstitution of Human Telomerase Expressed in *Saccharomyces Cerevisiae*.” *Journal of Biological Chemistry* 274(53):38027–31.
- Bai, Xiao-chen, Greg McMullan, and Sjors H. W. Scheres. 2015a. “How Cryo-EM Is Revolutionizing Structural Biology.” *Trends in Biochemical Sciences* 40(1):49–57.
- Bai, Xiao-chen, Chuangye Yan, Guanghui Yang, Peilong Lu, Dan Ma, Linfeng Sun, Rui Zhou, Sjors H. W. Scheres, and Yigong Shi. 2015b. “An Atomic Structure of Human γ -Secretase.” *Nature* 525(7568):212–17.
- Bammes, Benjamin E., Ryan H. Rochat, Joanita Jakana, Dong Hua Chen, and Wah Chiu. 2012. “Direct Electron Detection Yields Cryo-EM Reconstructions at Resolutions beyond 3/4 Nyquist Frequency.” *Journal of Structural Biology* 177(3):589–601.
- Bartesaghi, A., A. Merk, S. Banerjee, D. Matthies, X. Wu, J. L. S. Milne, and S. Subramaniam. 2015. “2.2 Å Resolution Cryo-EM Structure of β -Galactosidase in Complex with a Cell-Permeant Inhibitor.” *Science* 348(6239):1147–51.
- Beattie, T. L., W. Zhou, M. O. Robinson, and L. Harrington. 2001. “Functional Multimerization of the Human Telomerase Reverse Transcriptase.” *Molecular and Cellular Biology* 21(18):6151–60.
- Beck, Martin, and Wolfgang Baumeister. 2016. “Cryo-Electron Tomography: Can It Reveal the Molecular Sociology of Cells in Atomic Detail?” *Trends in Cell Biology* 26(11):825–37.
- Bell, Robert J. A., H. Tomas Rube, Alex Kreig, Andrew Mancini, Shaun D. Fouse, Raman P. Nagarajan, Serah Choi, Chibo Hong, Daniel He, Melike Pekmezci, John K. Wiencke, Margaret R. Wrensch, Susan M. Chang, Kyle M. Walsh, Sua Myong, Jun S. Song, and Joseph F. Costello. 2015. “The Transcription Factor GABP Selectively Binds and Activates the Mutant TERT Promoter in Cancer.” *Science* 348(6238):1036–39.

- Berman, H. M. 2000. "The Protein Data Bank." *Nucleic Acids Research* 28(1):235–42.
- Bianchi, Alessandro. 1997. "TRF1 Is a Dimer and Bends Telomeric DNA." *The EMBO Journal* 16(7):1785–94.
- Blasco, M. A., H. W. Lee, M. Prakash Hande, Enrique Samper, Peter M. Lansdorp, R. A. DePinho, and Carol W. Greider. 1997. "Telomere Shortening and Tumor Formation by Mouse Cells Lacking Telomerase RNA." *Cell* 91(1):25–34.
- Brenner, S., and R. W. Horne. 1959. "A Negative Staining Method for High Resolution Electron Microscopy of Viruses." *Biochimica et Biophysica Acta* 34:103–10.
- Brilot, Axel F., James Z. Chen, Anchi Cheng, Junhua Pan, Stephen C. Harrison, Clinton S. Potter, Bridget Carragher, Richard Henderson, and Nikolaus Grigorieff. 2012. "Beam-Induced Motion of Vitrified Specimen on Holey Carbon Film." *Journal of Structural Biology* 177(3):630–37.
- Bryan, Tracy M., Karen J. Goodrich, and Thomas R. Cech. 2003. "Tetrahymena Telomerase Is Active as a Monomer." *Molecular Biology of the Cell* 14(12):4794–4804.
- Buonomo, Sara B. C., Yipin Wu, David Ferguson, and Titia De Lange. 2009. "Mammalian Rif1 Contributes to Replication Stress Survival and Homology-Directed Repair." *Journal of Cell Biology* 187(3):385–98.
- Busch, H. 1926. "Berechnung Der Bahn von Kathodenstrahlen Im Axialsymmetrischen Elektromagnetischen Felde." *Annalen Der Physik* 386(25):974–93.
- Bykov, Yury S., Miroslava Schaffer, Svetlana O. Dodonova, Sahrada Albert, Jürgen M. Plitzko, Wolfgang Baumeister, Benjamin D. Engel, and John AG Briggs. 2017. "The Structure of the COPI Coat Determined within the Cell." *ELife* 6.
- Campbell, Melody G., Anchi Cheng, Axel F. Brilot, Arne Moeller, Dmitry Lyumkis, David Veesler, Junhua Pan, Stephen C. Harrison, Clinton S. Potter, Bridget Carragher, and Nikolaus Grigorieff. 2012. "Movies of Ice-Embedded Particles Enhance Resolution in Electron Cryo-Microscopy." *Structure* 20(11):1823–28.

- Campbell, Melody G., David Veessler, Anchi Cheng, Clinton S. Potter, and Bridget Carragher. 2015. "2.8 Å Resolution Reconstruction of the Thermoplasma Acidophilum 20S Proteasome Using Cryo-Electron Microscopy" edited by S. H. W. Scheres. *ELife* 4:e06380.
- Canudas, Silvia, Benjamin R. Houghtaling, Monica Bhanot, Ghadir Sasa, Sharon A. Savage, Alison A. Bertuch, and Susan Smith. 2011. "A Role for Heterochromatin Protein 1 at Human Telomeres." *Genes & Development* 25(17):1807–19.
- Carroni, Marta, and Helen R. Saibil. 2016. "Cryo Electron Microscopy to Determine the Structure of Macromolecular Complexes." *Methods* 95:78–85.
- Chen, Dong-Hua, Joanita Jakana, Xiangnan Liu, Michael F. Schmid, and Wah Chiu. 2008a. "Achievable Resolution from Images of Biological Specimens Acquired from a 4k×4k CCD Camera in a 300-KV Electron Cryomicroscope." *Journal of Structural Biology* 163(1):45–52.
- Chen, J. L., and Carol W. Greider. 2004. "An Emerging Consensus for Telomerase RNA Structure." *Proceedings of the National Academy of Sciences* 101(41):14683–84.
- Chen, Liuh Yow, Sophie Redon, and Joachim Lingner. 2012. "The Human CST Complex Is a Terminator of Telomerase Activity." *Nature* 488(7412):540–44.
- Chen, Yong, Yuting Yang, Megan Van Overbeek, Jill R. Donigian, Paul Baciú, Titia De Lange, and Ming Lei. 2008b. "A Shared Docking Motif in TRF1 and TRF2 Used for Differential Recruitment of Telomeric Proteins." *Science* 319(5866):1092–96.
- Chin, Lynda, Steven E. Artandi, Qiong Shen, Alice Tam, Shwu-Luan Lee, Geoffrey J. Gottlieb, Carol W. Greider, and Ronald A. DePinho. 1999. "P53 Deficiency Rescues the Adverse Effects of Telomere Loss and Cooperates with Telomere Dysfunction to Accelerate Carcinogenesis." *Cell* 97(4):527–38.
- Chua, Eugene Y. D., Vinod K. Vogirala, Oviya Inian, Andrew S. W. Wong, Lars Nordenskiöld, Juergen M. Plitzko, Radostin Danev, and Sara Sandin. 2016. "3.9 Å Structure of the Nucleosome Core Particle Determined by Phase-Plate Cryo-EM." *Nucleic Acids Research*

44(17):8013–19.

Chung, Jeeyun, Prabhat Khadka, and In Kwon Chung. 2012. “Nuclear Import of Htert Requires a Bipartite Nuclear Localization Signal and Akt-Mediated Phosphorylation.” *Journal of Cell Science* 125(11):2684–97.

Ciccia, Alberto, and Stephen J. Elledge. 2010. “The DNA Damage Response: Making It Safe to Play with Knives.” *Molecular Cell* 40(2):179–204.

Cohen, S. B., M. E. Graham, G. O. Lovrecz, N. Bache, P. J. Robinson, and R. R. Reddel. 2007. “Protein Composition of Catalytically Active Human Telomerase from Immortal Cells.” *Science* 315(5820):1850–53.

Cohen, Scott B., and Roger R. Reddel. 2008. “A Sensitive Direct Human Telomerase Activity Assay.” *Nature Methods* 5(4):355–60.

Court, Robert, Lynda Chapman, Louise Fairall, and Daniela Rhodes. 2005. “How the Human Telomeric Proteins TRF1 and TRF2 Recognize Telomeric DNA: A View from High-resolution Crystal Structures.” *EMBO Reports* 6(1):39–45.

Cristofari, Gaël, and Joachim Lingner. 2006. “Telomere Length Homeostasis Requires That Telomerase Levels Are Limiting.” *The EMBO Journal* 25(3):565–74.

D’Imprima, Edoardo, Davide Floris, Mirko Joppe, Ricardo Sánchez, Martin Grininger, Werner Kühlbrandt, Edoardo D’Imprima, Davide Floris, Mirko Joppe, Ricardo Sánchez, Martin Grininger, and Werner Kühlbrandt. 2019. “Protein Denaturation at the Air-Water Interface and How to Prevent It.” *ELife* 8:1–18.

D’Imprima, Edoardo, Ralf Salzer, Ramachandra M. Bhaskara, Ricardo Sánchez, Ilona Rose, Lennart Kirchner, Gerhard Hummer, Werner Kühlbrandt, Janet Vonck, and Beate Averhoff. 2017. “Cryo-EM Structure of the Bifunctional Secretin Complex of *Thermus Thermophilus*.” *ELife* 6.

Dalby, Andrew B., Ctirad Hofr, and Thomas R. Cech. 2015. “Contributions of the TEL-Patch Amino Acid Cluster on TPP1 to Telomeric DNA Synthesis by Human Telomerase.” *Journal of*

Molecular Biology 427(6):1291–1303.

Dandey, Venkata P., Hui Wei, Zhening Zhang, Yong Zi Tan, Priyamvada Acharya, Edward T. Eng, William J. Rice, Peter A. Kahn, Clinton S. Potter, and Bridget Carragher. 2018. “Spotiton: New Features and Applications.” *Journal of Structural Biology* 202(2):161–69.

Danev, Radostin, Bart Buijsse, Maryam Khoshouei, Jürgen M. Plitzko, and Wolfgang Baumeister. 2014. “Volta Potential Phase Plate for In-Focus Phase Contrast Transmission Electron Microscopy.” *Proceedings of the National Academy of Sciences of the United States of America* 111(44):15635–40.

Danev, Radostin, Dimitry Tegunov, and Wolfgang Baumeister. 2017. “Using the Volta Phase Plate with Defocus for Cryo-EM Single Particle Analysis.” *ELife* 6:1–9.

Darzacq, Xavier, Nupur Kittur, Sujayita Roy, Yaron Shav-Tal, Robert H. Singer, and U. Thomas Meier. 2006. “Stepwise RNP Assembly at the Site of H/ACA RNA Transcription in Human Cells.” *Journal of Cell Biology* 173(2):207–18.

Drulyte, Ieva, Rachel M. Johnson, Emma L. Hesketh, Daniel L. Hurdiss, Charlotte A. Scarff, Sebastian A. Porav, Neil A. Ranson, Stephen P. Muench, and Rebecca F. Thompson. 2018. “Approaches to Altering Particle Distributions in Cryo-Electron Microscopy Sample Preparation.” *Acta Crystallographica Section D Structural Biology* 74(6):560–71.

Dubochet, J., J. Lepault, R. Freeman, J. A. Berriman, and J. -C Homo. 1982. “Electron Microscopy of Frozen Water and Aqueous Solutions.” *Journal of Microscopy* 128(3):219–37.

Dubochet, J., and A. W. McDowell. 1981. “VITRIFICATION OF PURE WATER FOR ELECTRON MICROSCOPY.” *Journal of Microscopy* 124(3):3–4.

Erdel, Fabian, Katja Kratz, Smaranda Willcox, Jack D. Griffith, Eric C. Greene, and Titia de Lange. 2017. “Telomere Recognition and Assembly Mechanism of Mammalian Shelterin.” *Cell Reports* 18(1):41–53.

Fan, Xiao, Jia Wang, Xing Zhang, Zi Yang, Jin Can Zhang, Lingyun Zhao, Hai Lin Peng, Jianlin Lei,

- and Hong Wei Wang. 2019. “Single Particle Cryo-EM Reconstruction of 52 KDa Streptavidin at 3.2 Angstrom Resolution.” *Nature Communications* 10(1):1–11.
- Fang, Guowei, and Thomas R. Cech. 1995. “Telomerase RNA Localized in the Replication Band and Spherical Subnuclear Organelles in Hypotrichous Ciliates.” *Journal of Cell Biology* 130(2):243–53.
- Faruqi, A. R. R., Richard Henderson, and Greg McMullan. 2015. “Progress and Development of Direct Detectors for Electron Cryomicroscopy.” Pp. 103–41 in *Advances in Imaging and Electron Physics*. Vol. 190. Elsevier Inc.
- Feigon, J., N. K. Kim, R. D. Peterson, Z. Wang, and Q. Zhang. 2010. “Structurally Conserved Five Nucleotide Bulge Determines the Overall Topology of the Core Domain of Human Telomerase RNA.” *Proceedings of the National Academy of Sciences* 107(44):18761–68.
- Feng, J., W. Funk, S. Wang, S. Weinrich, A. Avilion, C. Chiu, R. Adams, E. Chang, R. Allsopp, J. Yu, and E. Al. 1995. “The RNA Component of Human Telomerase.” *Science* 269(5228):1236–41.
- Fernandez-Leiro, Rafael, and Sjors H. W. Scheres. 2016. “Unravelling Biological Macromolecules with Cryo-Electron Microscopy.” *Nature* 537(7620):339–46.
- Finger, L. David, John. M. Atack, Susan Tsutakawa, Scott Classen, John Tainer, Jane Grasby, and Binghui Shen. 2012. *The Eukaryotic Replisome: A Guide to Protein Structure and Function*. Vol. 62. edited by S. MacNeill. Dordrecht: Springer Netherlands.
- Flint, J., G. P. Bates, K. Clark, A. Dorman, D. Willingham, B. A. Roe, G. Micklem, D. R. Higgs, and E. J. Louis. 1997. “Sequence Comparison of Human and Yeast Telomeres Identifies Structurally Distinct Subtelomeric Domains.” *Human Molecular Genetics* 6(8):1305–13.
- Forsythe, Heidi L., Jennie L. Jarvis, John W. Turner, Lynne W. Elmore, and Shawn E. Holt. 2001. “Stable Association of Hsp90 and P23, but Not Hsp70, with Active Human Telomerase.” *Journal of Biological Chemistry* 276(19):15571–74.

- Frescas, David, and Titia de Lange. 2014. "Binding of TPP1 Protein to TIN2 Protein Is Required for POT1a,b Protein-Mediated Telomere Protection." *Journal of Biological Chemistry* 289(35):24180–87.
- Fu, Dragony, and Kathleen Collins. 2003. "Distinct Biogenesis Pathways for Human Telomerase RNA and H/ACA Small Nucleolar RNAs." *Molecular Cell* 11(5):1361–72.
- Galej, Wojciech P., Max E. Wilkinson, Sebastian M. Fica, Chris Oubridge, Andrew J. Newman, and Kiyoshi Nagai. 2016. "Cryo-EM Structure of the Spliceosome Immediately after Branching." *Nature* 537(7619):197–201.
- Gallardo, Franck, and Pascal Chartrand. 2008. "Telomerase Biogenesis: The Long Road before Getting to the End." *RNA Biology* 5(4):212–15.
- Garcia, Christine Kim, Woodring E. Wright, and Jerry W. Shay. 2007. "Human Diseases of Telomerase Dysfunction: Insights into Tissue Aging." *Nucleic Acids Research* 35(22):7406–16.
- Glaeser, Robert M. 2008. "Retrospective: Radiation Damage and Its Associated 'Information Limitations.'" *Journal of Structural Biology* 163(3):271–76.
- Glaeser, Robert M., and Bong-Gyoon Han. 2017. "Opinion: Hazards Faced by Macromolecules When Confined to Thin Aqueous Films." *Biophysics Reports* 3(1):1–7.
- Glaeser, Robert M., Bong-Gyoon Han, Roseann Csencsits, Alison Killilea, Arto Pulk, and Jamie H. D. Cate. 2016. "Factors That Influence the Formation and Stability of Thin, Cryo-EM Specimens." *Biophysical Journal* 110(4):749–55.
- Goldman, Fred, Rachida Bouarich, Shashikant Kulkarni, Sara Freeman, Hong Yan Du, Lea Harrington, Philip J. Mason, Arturo Londoño-Vallejo, and Monica Bessler. 2005. "The Effect of TERC Haploinsufficiency on the Inheritance of Telomere Length." *Proceedings of the National Academy of Sciences of the United States of America* 102(47):17119–24.
- Grant, Timothy, Alexis Rohou, and Nikolaus Grigorieff. 2018. "CisTEM, User-Friendly Software for Single-Particle Image Processing." *ELife* 7.

- Greider, Carol W., and Elizabeth H. Blackburn. 1985. "Identification of a Specific Telomere Terminal Transferase Activity in Tetrahymena Extracts." *Cell* 43(2 PART 1):405–13.
- Griffith, Jack D., Laurey Comeau, Soraya Rosenfield, Rachel M. Stansel, Alessandro Bianchi, Heidi Moss, Titia de Lange, Jack D. Griffith, Laurey Comeau, Soraya Rosenfield, Rachel M. Stansel, Alessandro Bianchi, and Heidi Moss. 1999. "Mammalian Telomeres End in a Large Duplex Loop." *Cell* 97(4):503–14.
- Grill, Sherilyn, Kamlesh Bisht, Valerie M. Tesmer, Adrienne Niederriter Shami, Saher S. Hammoud, and Jayakrishnan Nandakumar. 2019. "Two Separation-of-Function Isoforms of Human TPP1 Dictate Telomerase Regulation in Somatic and Germ Cells." *Cell Reports* 27(12):3511–3521.e7.
- Grill, Sherilyn, Valerie M. Tesmer, and Jayakrishnan Nandakumar. 2018. "The N Terminus of the OB Domain of Telomere Protein TPP1 Is Critical for Telomerase Action." *Cell Reports* 22(5):1132–40.
- Gruska, Manuela, Ohad Medalia, Wolfgang Baumeister, and Andrew Leis. 2008. "Electron Tomography of Vitreous Sections from Cultured Mammalian Cells." *Journal of Structural Biology* 161(3):384–92.
- Guo, Qiang, Carina Lehmer, Antonio Martínez-Sánchez, Till Rudack, Florian Beck, Hannelore Hartmann, Manuela Pérez-Berlanga, Frédéric Frottin, Mark S. Hipp, F. Ulrich Hartl, Dieter Edbauer, Wolfgang Baumeister, and Rubén Fernández-Busnadiego. 2018. "In Situ Structure of Neuronal C9orf72 Poly-GA Aggregates Reveals Proteasome Recruitment." *Cell* 172(4):696–705.e12.
- Ten Hagen, K. G., D. M. Gilbert, H. F. Willard, and S. N. Cohen. 1990. "Replication Timing of DNA Sequences Associated with Human Centromeres and Telomeres." *Molecular and Cellular Biology* 10(12):6348–55.
- Han, Yimo, Xiao Fan, Haozhe Wang, Fang Zhao, Christopher G. Tully, Jing Kong, Nan Yao, and Nieng Yan. 2019. "High Yield Monolayer Graphene Grids for Near-Atomic Resolution Cryo-

- Electron Microscopy.” *BioRxiv* 827089.
- Hayflick, L. 1965. “The Limited in Vitro Lifetime of Human Diploid Cell Strains.” *Experimental Cell Research* 37(3):614–36.
- Hayflick, L., and P. S. Moorhead. 1961. “The Serial Cultivation of Human Diploid Cell Strains.” *Experimental Cell Research* 25(3):585–621.
- Van Heel, Marin, and Joachim Frank. 1981. “Use of Multivariate Statistics in Analysing the Images of Biological Macromolecules.” *Ultramicroscopy* 6(2):187–94.
- van Heel, Marin, and Wilko Keegstra. 1981. “IMAGIC: A Fast, Flexible and Friendly Image Analysis Software System.” *Ultramicroscopy* 7(2):113–29.
- Hemann, Michael T., Margaret A. Strong, Ling-Yang Hao, and Carol W. Greider. 2001. “The Shortest Telomere, Not Average Telomere Length, Is Critical for Cell Viability and Chromosome Stability.” *Cell* 107(1):67–77.
- Herbert, Brittney-Shea Shea, Amelia E. Hochreiter, Woodring E. Wright, and Jerry W. Shay. 2006. “Nonradioactive Detection of Telomerase Activity Using the Telomeric Repeat Amplification Protocol.” *Nature Protocols* 1(3):1583–90.
- Hill, Chris H., Vytautė Boreikaitė, Ananthanarayanan Kumar, Ana Casañal, Peter Kubík, Gianluca Degliesposti, Sarah Maslen, Angelica Mariani, Otilie von Loeffelholz, Mathias Girbig, Mark Skehel, and Lori A. Passmore. 2019. “Activation of the Endonuclease That Defines mRNA 3’ Ends Requires Incorporation into an 8-Subunit Core Cleavage and Polyadenylation Factor Complex.” *Molecular Cell* 73(6):1217-1231.e11.
- Hiyama, E., and K. Hiyama. 2007. “Telomere and Telomerase in Stem Cells.” *British Journal of Cancer* 96(7):1020–24.
- Hockemeyer, Dirk, Agnel J. Sfeir, Jerry W. Shay, Woodring E. Wright, and Titia De Lange. 2005. “POT1 Protects Telomeres from a Transient DNA Damage Response and Determines How Human Chromosomes End.” *The EMBO Journal* 24(14):2667–78.

- Holt, Shawn E., Dara L. Aisner, Joseph Baur, Valerie M. Tesmer, Marife Dy, Michel Ouellette, James B. Trager, Gregg B. Morin, David O. Toft, Jerry W. Shay, Woodring E. Wright, and Michael A. White. 1999. "Functional Requirement of P23 and Hsp90 in Telomerase Complexes." *Genes & Development* 13(7):817–26.
- Horn, Susanne, Adina Figl, P. Sivaramakrishna Rachakonda, Christine Fischer, Antje Sucker, Andreas Gast, Stephanie Kadel, Iris Moll, Eduardo Nagore, Kari Hemminki, Dirk Schadendorf, and Rajiv Kumar. 2013. "TERT Promoter Mutations in Familial and Sporadic Melanoma." *Science* 339(6122):959–61.
- Hu, Chunyi, Rekha Rai, Chenhui Huang, Cayla Broton, Juanjuan Long, Ying Xu, Jing Xue, Ming Lei, Sandy Chang, and Yong Chen. 2017. "Structural and Functional Analyses of the Mammalian TIN2-TPP1-TRF2 Telomeric Complex." *Cell Research* 27(12):1485–1502.
- Huang, F. W., E. Hodis, M. J. Xu, G. V. Kryukov, L. Chin, and L. A. Garraway. 2013. "Highly Recurrent TERT Promoter Mutations in Human Melanoma." *Science* 339(6122):957–59.
- Huang, Ying Wen, Chung Chi Hu, Ming Ru Liou, Ban Yang Chang, Ching Hsiu Tsai, Menghsiao Meng, Na Sheng Lin, and Yau Heiu Hsu. 2012. "Hsp90 Interacts Specifically with Viral RNA and Differentially Regulates Replication Initiation of Bamboo Mosaic Virus and Associated Satellite RNA" edited by P. D. Nagy. *PLoS Pathogens* 8(5):e1002726.
- Israelachvili, Jacob N., and Jacob N. Israelachvili. 2011. "Electrostatic Forces between Surfaces in Liquids." *Intermolecular and Surface Forces* 291–340.
- Jafri, Mohammad A., Shakeel A. Ansari, Mohammed H. Alqahtani, and Jerry W. Shay. 2016. "Roles of Telomeres and Telomerase in Cancer, and Advances in Telomerase-Targeted Therapies." *Genome Medicine* 8(1).
- Katzen, Alisa L., Gordon M. Cann, and Elizabeth H. Blackburn. 1981. "Sequence-Specific Fragmentation of Macronuclear DNA in a Holotrichous Ciliate." *Cell* 24(2):313–20.
- Kemmerling, Simon, Jörg Ziegler, Gabriel Schweighauser, Stefan A. Arnold, Dominic Giss, Shirley

- A. Müller, Philippe Ringler, Kenneth N. Goldie, Nils Goedecke, Andreas Hierlemann, Henning Stahlberg, Andreas Engel, and Thomas Braun. 2012. “Connecting μ -Fluidics to Electron Microscopy.” *Journal of Structural Biology* 177(1):128–34.
- Khan, Shagufta H., and Raj Kumar. 2009. “An Overview of the Importance of Conformational Flexibility in Gene Regulation by the Transcription Factors.” *Journal of Biophysics* 2009:1–9.
- Killela, Patrick J., Zachary J. Reitman, Yuchen Jiao, Chetan Bettegowda, Nishant Agrawal, Luis A. Diaz, Allan H. Friedman, Henry Friedman, Gary L. Gallia, Beppino C. Giovanella, Arthur P. Grollman, Tong Chuan He, Yiping He, Ralph H. Hruban, George I. Jallo, Nils Mandahl, Alan K. Meeker, Fredrik Mertens, George J. Netto, B. Ahmed Rasheed, Gregory J. Riggins, Thomas A. Rosenquist, Mark Schiffman, Ie Ming Shih, Dan Theodorescu, Michael S. Torbenson, Victor E. Velculescu, Tian Li Wang, Nicolas Wentzensen, Laura D. Wood, Ming Zhang, Roger E. McLendon, Darell D. Bigner, Kenneth W. Kinzler, Bert Vogelstein, Nickolas Papadopoulos, and Hai Yan. 2013. “TERT Promoter Mutations Occur Frequently in Gliomas and a Subset of Tumors Derived from Cells with Low Rates of Self-Renewal.” *Proceedings of the National Academy of Sciences of the United States of America* 110(15):6021–26.
- Kim, S., P. Kaminker, and J. Campisi. 2000. “TIN2, a New Regulator of Telomere Length in Human Cells.” *Chemtracts* 13(10):700–704.
- Klobutcher, L. A., M. T. Swanton, P. Donini, and D. M. Prescott. 1981. “All Gene-Sized DNA Molecules in Four Species of Hypotrichs Have the Same Terminal Sequence and an Unusual 3' Terminus.” *Proceedings of the National Academy of Sciences* 78(5):3015–19.
- Kremer, James R., David N. Mastronarde, and J. Richar. McIntosh. 1996. “Computer Visualization of Three-Dimensional Image Data Using IMOD.” *Journal of Structural Biology* 116(1):71–76.
- Lagadari, Mariana, Nadia R. Zgajnar, Luciana I. Gallo, and Mario D. Galigniana. 2016. “Hsp90-Binding Immunophilin FKBP51 Forms Complexes with HTERT Enhancing Telomerase Activity.” *Molecular Oncology* 10(7):1086–98.

- Lai, Serene, R. 2005. "Epigenetic Control of Telomerase and Modes of Telomere Maintenance in Aging and Abnormal Systems." *Frontiers in Bioscience* 10(1–3):1779.
- de Lange, Titia. 2005. "Shelterin: The Protein Complex That Shapes and Safeguards Human Telomeres." *Genes & Development* 19(18):2100–2110.
- De Lange, Titia. 2009. "How Telomeres Solve the End-Protection Problem." *Science* 326(5955):948–52.
- Lehman, I. R., M. J. Bessman, E. S. Simms, and A. Kornberg. 1958. "Enzymatic Synthesis of Deoxyribonucleic Acid. I. Preparation of Substrates and Partial Purification of an Enzyme from *Escherichia Coli*." *The Journal of Biological Chemistry* 233(1):163–70.
- Levy, Michael Z., Richard C. Allsopp, A. Bruce Futcher, Carol W. Greider, and Calvin B. Harley. 1992. "Telomere End-Replication Problem and Cell Aging." *Journal of Molecular Biology* 225(4):951–60.
- Lewis, Karen A., and Deborah S. Wuttke. 2012. "Telomerase and Telomere-Associated Proteins: Structural Insights into Mechanism and Evolution." *Structure* 20(1):28–39.
- Li, Bibo, Stephanie Oestreich, and Titia De Lange. 2000. "Identification of Human Rap1: Implications for Telomere Evolution." *Cell* 101(5):471–83.
- Li, Xueming, Paul Mooney, Shawn Zheng, Christopher R. Booth, Michael B. Braunfeld, Sander Gubbens, David A. Agard, and Yifan Cheng. 2013. "Electron Counting and Beam-Induced Motion Correction Enable near-Atomic-Resolution Single-Particle Cryo-EM." *Nature Methods* 10(6):584–90.
- Li, Yinghui, Qi Ling Zhou, Wenjie Sun, Prashant Chandrasekharan, Hui Shan Cheng, Zhe Ying, Manikandan Lakshmanan, Anandhkumar Raju, Daniel G. Tenen, Shi Yuan Cheng, Kai Hsiang Chuang, Jun Li, Shyam Prabhakar, Mengfeng Li, and Vinay Tergaonkar. 2015. "Non-Canonical NF- κ B Signalling and ETS1/2 Cooperatively Drive C250T Mutant TERT Promoter Activation." *Nature Cell Biology* 17(10):1327–38.

- Liao, Hstau Y., and Joachim Frank. 2010. "Definition and Estimation of Resolution in Single-Particle Reconstructions." *Structure* 18(7):768–75.
- Lingner, Joachim, Timothy R. Hughes, Andrej Shevchenko, Matthias Mann, Victoria Lundblad, and Thomas R. Cech. 1997. "Reverse Transcriptase Motifs in the Catalytic Subunit of Telomerase." *Science* 276(5312):561–67.
- Lipps, Hans J., W. Gruissem, and D. M. Prescott. 1982. "Higher Order DNA Structure in Macronuclear Chromatin of the Hypotrichous Ciliate *Oxytricha Nova*." *Proceedings of the National Academy of Sciences* 79(8):2495–99.
- Lipps, Hans J., and Daniela Rhodes. 2009. "G-Quadruplex Structures: In Vivo Evidence and Function." *Trends in Cell Biology* 19(8):414–22.
- Ly, Hinh. 2003. "A Role for a Novel 'trans-Pseudoknot' RNA-RNA Interaction in the Functional Dimerization of Human Telomerase." *Genes & Development* 17(9):1078–83.
- Makarov, Vladimir L., Yoko Hirose, and John P. Langmore. 1997. "Long G Tails at Both Ends of Human Chromosomes Suggest a C Strand Degradation Mechanism for Telomere Shortening." *Cell* 88(5):657–66.
- Mao, Yu-Qian, and Walid A. Houry. 2017. "The Role of Pontin and Reptin in Cellular Physiology and Cancer Etiology." *Frontiers in Molecular Biosciences* 4(AUG).
- McMullan, G., A. R. R. Faruqi, D. Clare, and R. Henderson. 2014. "Comparison of Optimal Performance at 300keV of Three Direct Electron Detectors for Use in Low Dose Electron Microscopy." *Ultramicroscopy* 147:156–63.
- Medalia, O. 2002. "Macromolecular Architecture in Eukaryotic Cells Visualized by Cryoelectron Tomography." *Science* 298(5596):1209–13.
- Mendez-Bermudez, Aaron, Liudmyla Lototska, Serge Bauwens, Marie-Josèphe Giraud-Panis, Olivier Croce, Karine Jamet, Agurtzane Irizar, Macarena Mowinckel, Stephane Koundrioukoff, Nicolas Nottet, Genevieve Almouzni, Mare-Paule Teulade-Fichou, Michael Schertzer, Mylène

- Perderiset, Arturo Londoño-Vallejo, Michelle Debatisse, Eric Gilson, and Jing Ye. 2018. “Genome-Wide Control of Heterochromatin Replication by the Telomere Capping Protein TRF2.” *Molecular Cell* 70(3):449-461.e5.
- Merk, Alan, Alberto Bartesaghi, Soojay Banerjee, Veronica Falconieri, Prashant Rao, Mindy I. Davis, Rajan Pragani, Matthew B. Boxer, Lesley A. Earl, Jacqueline L. S. Milne, and Sriram Subramaniam. 2016. “Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery.” *Cell* 165(7):1698–1707.
- Meselson, M., and F. W. Stahl. 1958. “The Replication of DNA in Escherichia Coli.” *Proceedings of the National Academy of Sciences* 44(7):671–82.
- Milne, Jacqueline L. S., Mario J. Borgnia, Alberto Bartesaghi, Erin E. H. Tran, Lesley A. Earl, David M. Schauder, Jeffrey Lengyel, Jason Pierson, Ardan Patwardhan, and Sriram Subramaniam. 2013. “Cryo-Electron Microscopy - a Primer for the Non-Microscopist.” *FEBS Journal* 280(1):28–45.
- Mitchell, Meghan, Andrew Gillis, Mizuko Futahashi, Haruhiko Fujiwara, and Emmanuel Skordalakes. 2010. “Structural Basis for Telomerase Catalytic Subunit TERT Binding to RNA Template and Telomeric DNA.” *Nature Structural and Molecular Biology* 17(4):513–18.
- Morris, Danna K., and Victoria Lundblad. 1997. “Programmed Translational Frameshifting in a Gene Required for Yeast Telomere Replication.” *Current Biology* 7(12):969–76.
- Moye, Aaron L., Karina C. Porter, Scott B. Cohen, Tram Phan, Katherine G. Zyner, Natsuki Sasaki, George O. Lovrecz, Jennifer L. Beck, and Tracy M. Bryan. 2015. “Telomeric G-Quadruplexes Are a Substrate and Site of Localization for Human Telomerase.” *Nature Communications* 6(1):7643.
- Nakane, Takanori, Dari Kimanius, Erik Lindahl, and Sjors HW Scheres. 2018. “Characterisation of Molecular Motions in Cryo-EM Single-Particle Data by Multi-Body Refinement in RELION.” *ELife* 7.

- Nandakumar, Jayakrishnan, and Thomas R. Cech. 2013. "Finding the End: Recruitment of Telomerase to Telomeres." *Nature Reviews Molecular Cell Biology* 14(2):69–82.
- Naydenova, Katerina, Mathew J. Peet, and Christopher J. Russo. 2019. "Multifunctional Graphene Supports for Electron Cryomicroscopy." *Proceedings of the National Academy of Sciences of the United States of America* 116(24):11718–24.
- Nguyen, Thi Hoang Duong, Wojciech P. Galej, Xiao-chen Bai, Chris Oubridge, Andrew J. Newman, Sjors H. W. Scheres, and Kiyoshi Nagai. 2016. "Cryo-EM Structure of the Yeast U4/U6.U5 Tri-SnRNP at 3.7 Å Resolution." *Nature* 530(7590):298–302.
- Nguyen, Thi Hoang Duong, Wojciech P. Galej, Xiao-chen Bai, Christos G. Savva, Andrew J. Newman, Sjors H. W. Scheres, and Kiyoshi Nagai. 2015. "The Architecture of the Spliceosomal U4/U6.U5 Tri-SnRNP." *Nature* 523(7558):47–52.
- Nguyen, Thi Hoang Duong, Jane Tam, Robert A. Wu, Basil J. Greber, Daniel Toso, Eva Nogales, and Kathleen Collins. 2018. "Cryo-EM Structure of Substrate-Bound Human Telomerase Holoenzyme." *Nature* 557(7704):190–95.
- Nickell, Stephan, Oana Mihalache, Florian Beck, Reiner Hegerl, Andreas Korinek, and Wolfgang Baumeister. 2007. "Structural Analysis of the 26S Proteasome by Cryoelectron Tomography." *Biochemical and Biophysical Research Communications* 353(1):115–20.
- Noble, Alex J., Venkata P. Dandey, Hui Wei, Julia Brasch, Jillian Chase, Priyamvada Acharya, Yong Zi Tan, Zhening Zhang, Laura Y. Kim, Giovanna Scapin, Micah Rapp, Edward T. Eng, William J. Rice, Anchi Cheng, Carl J. Negro, Lawrence Shapiro, Peter D. Kwong, David Jeruzalmi, Amedee des Georges, Clinton S. Potter, and Bridget Carragher. 2018. "Routine Single Particle CryoEM Sample and Grid Characterization by Tomography." *ELife* 7:1–47.
- Nudelman, Fabio, Gijsbertus de With, and Nico A. J. M. Sommerdijk. 2011. "Cryo-Electron Tomography: 3-Dimensional Imaging of Soft Matter." *Soft Matter* 7(1):17–24.
- Nwanochie, Emeka, and Vladimir N. Uversky. 2019. "Structure Determination by Single-Particle

- Cryo-Electron Microscopy: Only the Sky (and Intrinsic Disorder) Is the Limit.” *International Journal of Molecular Sciences* 20(17):4186.
- Okamoto, Keiji, and Hiroyuki Seimiya. 2019. “Revisiting Telomere Shortening in Cancer.” *Cells* 8(2):107.
- Ortiz, Julio O., Friedrich Förster, Julia Kürner, Alexandros A. Linaroudis, and Wolfgang Baumeister. 2006. “Mapping 70S Ribosomes in Intact Cells by Cryoelectron Tomography and Pattern Recognition.” *Journal of Structural Biology* 156(2):334–41.
- Ouellette, Michel M., Martha Liao, Brittney-Shea Shea Herbert, Mari Johnson, Shawn E. Holt, Heidi S. Liss, Jerry W. Shay, and Woodring E. Wright. 2000. “Subsenescent Telomere Lengths in Fibroblasts Immortalized by Limiting Amounts of Telomerase.” *Journal of Biological Chemistry* 275(14):10072–76.
- Paavolainen, Lassi, Erman Acar, Uygur Tuna, Sari Peltonen, Toshio Moriya, Pan Soonsawad, Varpu Marjomäki, R. Holland Cheng, and Ulla Ruotsalainen. 2014. “Compensation of Missing Wedge Effects with Sequential Statistical Reconstruction in Electron Tomography” edited by E.-I. Pecheur. *PLoS ONE* 9(10):e108978.
- Paeschke, Katrin, Stefan Juranek, Tomas Simonsson, Anne Hempel, Daniela Rhodes, and Hans Joachim Lipps. 2008. “Telomerase Recruitment by the Telomere End Binding Protein- β Facilitates G-Quadruplex DNA Unfolding in Ciliates.” *Nature Structural & Molecular Biology* 15(6):598–604.
- Paeschke, Katrin, Tomas Simonsson, Jan Postberg, Daniela Rhodes, and Hans Joachim Lipps. 2005. “Telomere End-Binding Proteins Control the Formation of G-Quadruplex DNA Structures in Vivo.” *Nature Structural & Molecular Biology* 12(10):847–54.
- Palm, Wilhelm, and Titia de Lange. 2008. “How Shelterin Protects Mammalian Telomeres.” *Annual Review of Genetics* 42(1):301–34.
- Palovcak, Eugene, Feng Wang, Shawn Q. Zheng, Zanlin Yu, Sam Li, Miguel Betegon, David

- Bulkley, David A. Agard, and Yifan Cheng. 2018. “A Simple and Robust Procedure for Preparing Graphene-Oxide Cryo-EM Grids.” *Journal of Structural Biology* 204(1):80–84.
- Pantelic, Radosav S., Jannik C. Meyer, Ute Kaiser, Wolfgang Baumeister, and Jürgen M. Plitzko. 2010. “Graphene Oxide: A Substrate for Optimizing Preparations of Frozen-Hydrated Samples.” *Journal of Structural Biology* 170(1):152–56.
- Patrick, Eric M., Joseph Slivka, Bramyn Payne, Matthew J. Comstock, and Jens C. Schmidt. 2019. “Mechanism of Processive Telomerase Catalysis Revealed by High-Resolution Optical Tweezers.” 1–18.
- Pettersen, Eric F., Thomas D. Goddard, Conrad C. Huang, Gregory S. Couch, Daniel M. Greenblatt, Elaine C. Meng, and Thomas E. Ferrin. 2004. “UCSF Chimera?A Visualization System for Exploratory Research and Analysis.” *Journal of Computational Chemistry* 25(13):1605–12.
- Pike, Alexandra M., Margaret A. Strong, John Paul T. Ouyang, and Carol W. Greider. 2019. “TIN2 Functions with TPP1/POT1 To Stimulate Telomerase Processivity.” *Molecular and Cellular Biology* 39(21):1–15.
- Podlevsky, Joshua D., and Julian J. L. Chen. 2016. “Evolutionary Perspectives of Telomerase RNA Structure and Function.” *RNA Biology* 13(8):720–32.
- Porter, K. R. 1945. “A STUDY OF TISSUE CULTURE CELLS BY ELECTRON MICROSCOPY: METHODS AND PRELIMINARY OBSERVATIONS.” *Journal of Experimental Medicine* 81(3):233–46.
- Rai, Rekha, Hong Zheng, Hua He, Ying Luo, Asha Multani, Phillip B. Carpenter, and Sandy Chang. 2010. “The Function of Classical and Alternative Non-Homologous End-Joining Pathways in the Fusion of Dysfunctional Telomeres.” *The EMBO Journal* 29(15):2598–2610.
- Ravelli, Raimond B. G., Frank J. T. Nijpels, Rene J. M. Henderikx, Giulia Weissenberger, Sanne Thewessem, Abril Gijsbers, Bart W. A. M. M. Beulen, Carmen López-Iglesias, and Peter J. Peters. 2019. “Automated Cryo-EM Sample Preparation by Pin-Printing and Jet Vitrification.”

BioRxiv 651208.

- Rawson, S., M. G. Iadanza, N. A. Ranson, and S. P. Muench. 2016. “Methods to Account for Movement and Flexibility in Cryo-EM Data Processing.” *Methods* 100:35–41.
- Reboul, Cyril F., Michael Eager, Dominika Elmlund, and Hans Elmlund. 2018. “Single-Particle Cryo-EM-Improved Ab Initio 3D Reconstruction with SIMPLE/PRIME.” *Protein Science* 27(1):51–61.
- Ren, Xiaojun, Gérald Gavory, Haitao Li, Liming Ying, David Klenerman, and Balasubramanian Balasubramanian. 2003. “Identification of a New RNA·RNA Interaction Site for Human Telomerase RNA (HTR): Structural Implications for HTR Accumulation and a Dyskeratosis Congenita Point Mutation.” *Nucleic Acids Research* 31(22):6509–15.
- Rhodes, Daniela, and Hans J. Lipps. 2015. “G-Quadruplexes and Their Regulatory Roles in Biology.” *Nucleic Acids Research* 43(18):8627–37.
- Rigort, Alexander, and Jürgen M. Plitzko. 2015. “Cryo-Focused-Ion-Beam Applications in Structural Biology.” *Archives of Biochemistry and Biophysics* 581:122–30.
- Rohou, Alexis, and Nikolaus Grigorieff. 2015. “CTFFIND4: Fast and Accurate Defocus Estimation from Electron Micrographs.” *Journal of Structural Biology* 192(2):216–21.
- De Rosier, D. J., and A. Klug. 1968. “Reconstruction of Three Dimensional Structures from Electron Micrographs.” *Nature* 217(5124):130–34.
- Ruska, Ernst. 1987. “The Development of the Electron Microscope and of Electron Microscopy(Nobel Lecture).” *Angewandte Chemie International Edition in English* 26(7):595–605.
- Russo, Christopher J., and Lori A. Passmore. 2016. “Progress towards an Optimal Specimen Support for Electron Cryomicroscopy.” *Current Opinion in Structural Biology* 37:81–89.
- Russo, Christopher J., Steve Scotcher, and Martin Kyte. 2016. “A Precision Cryostat Design for

- Manual and Semi-Automated Cryo-Plunge Instruments.” *Review of Scientific Instruments* 87(11):114302.
- Sambrook, J., E. .. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*.
- Sandin, Sara, and Daniela Rhodes. 2014. “Telomerase Structure.” *Current Opinion in Structural Biology* 25:104–10.
- Sartori, Anna, Rudolf Gatz, Florian Beck, Alexander Rigort, Wolfgang Baumeister, and Juergen M. Plitzko. 2007. “Correlative Microscopy: Bridging the Gap between Fluorescence Light Microscopy and Cryo-Electron Tomography.” *Journal of Structural Biology* 160(2):135–45.
- Sauerwald, Anselm, Sara Sandin, Gaël Cristofari, Sjors H. W. Scheres, Joachim Lingner, and Daniela Rhodes. 2013. “Structure of Active Dimeric Human Telomerase.” *Nature Structural and Molecular Biology* 20(4):454–60.
- Schaffer, Miroslava, Stefan Pfeffer, Julia Mahamid, Stephan Kleindiek, Tim Laugks, Sahraddha Albert, Benjamin D. Engel, Andreas Rummel, Andrew J. Smith, Wolfgang Baumeister, and Juergen M. Plitzko. 2019. “A Cryo-FIB Lift-out Technique Enables Molecular-Resolution Cryo-ET within Native *Caenorhabditis Elegans* Tissue.” *Nature Methods* 16(8):757–62.
- Schaffitzel, C., I. Berger, J. Postberg, J. Hanes, H. J. Lipps, and A. Pluckthun. 2001. “In Vitro Generated Antibodies Specific for Telomeric Guanine-Quadruplex DNA React with Stylonychia Lemnae Macronuclei.” *Proceedings of the National Academy of Sciences* 98(15):8572–77.
- Scheres, Sjors H. W. 2012. “RELION: Implementation of a Bayesian Approach to Cryo-EM Structure Determination.” *Journal of Structural Biology* 180(3):519–30.
- Scheres, Sjors H. W., and Shaoxia Chen. 2012. “Prevention of Overfitting in Cryo-EM Structure Determination.” *Nature Methods* 9(9):853–54.
- Schmidt, Jens C., and Thomas R. Cech. 2015. “Human Telomerase: Biogenesis, Trafficking, Recruitment, and Activation.” *Genes and Development* 29(11):1095–1105.

- Schmidt, Jens C., Andrew B. Dalby, and Thomas R. Cech. 2014. "Identification of Human TERT Elements Necessary for Telomerase Recruitment to Telomeres." *ELife* 3:1–20.
- Schnapp, Gisela, Hans Peter Rodi, Wolfgang J. Rettig, Andreas Schnapp, and Klaus Damm. 1998. "One-Step Affinity Purification Protocol for Human Telomerase." *Nucleic Acids Research* 26(13):3311–13.
- Sexton, Alec N., Samuel G. Regalado, Christine S. Lai, Gregory J. Cost, Colleen M. O’Neil, Fyodor D. Urnov, Philip D. Gregory, Rudolf Jaenisch, Kathleen Collins, and Dirk Hockemeyer. 2014. "Genetic and Molecular Identification of Three Human TPP1 Functions in Telomerase Action: Recruitment, Activation, and Homeostasis Set Point Regulation." *Genes & Development* 28(17):1885–99.
- Sfeir, A., and T. de Lange. 2012. "Removal of Shelterin Reveals the Telomere End-Protection Problem." *Science* 336(6081):593–97.
- Shakeel, Shabih, Eeson Rajendra, Pablo Alcón, Francis O’Reilly, Dror S. Chorev, Sarah Maslen, Gianluca Degliesposti, Christopher J. Russo, Shaoda He, Chris H. Hill, J. Mark Skehel, Sjors H. W. Scheres, Ketan J. Patel, Juri Rappsilber, Carol V. Robinson, and Lori A. Passmore. 2019. "Structure of the Fanconi Anaemia Monoubiquitin Ligase Complex." *Nature* 575(7781):234–37.
- Shampay, Janis, Jack W. Szostak, and Elizabeth H. Blackburn. 1984. "DNA Sequences of Telomeres Maintained in Yeast." *Nature* 310(5973):154–57.
- Shcherbakova, D. M., K. A. Sokolov, M. I. Zvereva, and O. A. Dontsova. 2009. "Telomerase from Yeast *Saccharomyces Cerevisiae* Is Active in Vitro as a Monomer." *Biochemistry (Moscow)* 74(7):749–55.
- Slatko, Barton E., and Lisa M. Albright. 2001. "Denaturing Gel Electrophoresis for Sequencing." in *Current Protocols in Molecular Biology*. Hoboken, NJ, USA: John Wiley & Sons, Inc.
- Stark, Holger. 2010. "GraFix: Stabilization of Fragile Macromolecular Complexes for Single Particle Cryo-EM." Pp. 109–26 in.

- Steitz, T. A., J. Wang, S. H. Eom, J. Jaeger, T. Pestle, and D. Jamzalmi. 1996. "DNA and RNA Polymerases: Structural Diversity and Common Mechanisms." *FASEB Journal* 10(6):17395–99.
- Stern, J. L., K. G. Zyner, H. A. Pickett, S. B. Cohen, and T. M. Bryan. 2012. "Telomerase Recruitment Requires Both TCAB1 and Cajal Bodies Independently." *Molecular and Cellular Biology* 32(13):2384–95.
- Sundquist, Wesley I., and Aaron Klug. 1989. "Telomeric DNA Dimerizes by Formation of Guanine Tetrads between Hairpin Loops." *Nature* 342(6251):825–29.
- Tan, Yong Zi, Philip R. Baldwin, Joseph H. Davis, James R. Williamson, Clinton S. Potter, Bridget Carragher, and Dmitry Lyumkis. 2017. "Addressing Preferred Specimen Orientation in Single-Particle Cryo-EM through Tilting." *Nature Methods* 14(8):793–96.
- Theimer, Carla A., Jing Zhou, Robert D. Peterson, Nak-Kyoon Kim, Qi Zhang, Juli Feigon, Jing Zhou, Carla A. Theimer, Robert D. Peterson, Juli Feigon, Jing Zhou, Robert D. Peterson, Nak-Kyoon Kim, Qi Zhang, and Juli Feigon. 2008. "Solution Structure and Dynamics of the Wild-Type Pseudoknot of Human Telomerase RNA." *Journal of Molecular Biology* 384(5):1249–61.
- Thompson, Rebecca F., Matt Walker, C. Alistair Siebert, Stephen P. Muench, and Neil A. Ranson. 2016. "An Introduction to Sample Preparation and Imaging by Cryo-Electron Microscopy for Structural Biology." *Methods* 100:3–15.
- Tomlinson, Rebecca L., Tania D. Ziegler, Teerawit Supakorndej, Rebecca M. Terns, and Michael P. Terns. 2006. "Cell Cycle-Regulated Trafficking of Human Telomerase to Telomeres." *Molecular Biology of the Cell* 17(2):955–65.
- Truman, Caroline M., Christopher Phillips, Christopher M. Johnson, Domagoj Baretic, Alan R. Fersht, Hannah K. Pollard, David I. Fisher, Roger L. Williams, Balaji Santhanam, and Tomas Kouba. 2017. "Structures of Closed and Open Conformations of Dimeric Human ATM." *Science Advances* 3(5):e1700933.
- Tsurimoto, Toshiki, Thomas Melendy, and Bruce Stillman. 1990. "Sequential Initiation of Lagging

- and Leading Strand Synthesis by Two Different Polymerase Complexes at the SV40 DNA Replication Origin.” *Nature* 346(6284):534–39.
- Turchi, John J., Lin Huang, Richard S. Murante, Yong Kim, and Robert A. Bambara. 1994. “Enzymatic Completion of Mammalian Lagging-Strand DNA Replication.” *Proceedings of the National Academy of Sciences of the United States of America* 91(21):9803–7.
- Turoňová, Beata, Florian K. M. Schur, William Wan, and John A. G. Briggs. 2017. “Efficient 3D-CTF Correction for Cryo-Electron Tomography Using NovaCTF Improves Subtomogram Averaging Resolution to 3.4 Å.” *Journal of Structural Biology* 199(3):187–95.
- Venteicher, Andrew S., and Steven E. Artandi. 2009. “TCAB1: Driving Telomerase to Cajal Bodies.” *Cell Cycle* 8(9):1329–31.
- Viviescas, Maria Alejandra, Maria Isabel Nogueira Cano, and Marcela Segatto. 2018. “Chaperones and Their Role in Telomerase Ribonucleoprotein Biogenesis and Telomere Maintenance.” *Current Proteomics* 16(1):31–43.
- Vogan, Jacob M., Xiaozhu Zhang, Daniel T. Youmans, Samuel G. Regalado, Joshua Z. Johnson, Dirk Hockemeyer, and Kathleen Collins. 2016. “Minimized Human Telomerase Maintains Telomeres and Resolves Endogenous Roles of H/ACA Proteins, TCAB1, and Cajal Bodies” edited by C. Greider. *ELife* 5:e18221.
- Wagner, Thorsten, Felipe Merino, Markus Stabrin, Toshio Moriya, Claudia Antoni, Amir Apelbaum, Philine Hagel, Oleg Sitsel, Tobias Raisch, Daniel Prumbaum, Dennis Quentin, Daniel Roderer, Sebastian Tacke, Birte Siebolds, Evelyn Schubert, Tanvir R. Shaikh, Pascal Lill, Christos Gatsogiannis, and Stefan Raunser. 2019. “SPHIRE-CrYOLO Is a Fast and Accurate Fully Automated Particle Picker for Cryo-EM.” *Communications Biology* 2(1):218.
- Wang, Feng, Zanlin Yu, Miguel Betegon, Melody Campbell, Tural Aksel, Jianhua Zhao, Sam Li, Shawn M. Douglas, Yifan Cheng, and David A. Agard. 2019. “Amino and PEG-Amino Graphene Oxide Grids Enrich and Protect Samples for High-Resolution Single Particle Cryo-

- Electron Microscopy.” *BioRxiv* 813972.
- Wang, Hong Wei, and Xiao Fan. 2019. “Challenges and Opportunities in Cryo-EM with Phase Plate.” *Current Opinion in Structural Biology* 58:175–82.
- Wang, Quan, Jia Quan Liu, Zhao Chen, Ke Wei Zheng, Chang Yue Chen, Yu Hua Hao, and Zheng Tan. 2011. “G-Quadruplex Formation at the 3’ End of Telomere DNA Inhibits Its Extension by Telomerase, Polymerase and Unwinding by Helicase.” *Nucleic Acids Research* 39(14):6229–37.
- Watson, J. D., and Francis Crick. 1953. “Molecular Structures of Nucleic Acids.” *Nature* 171(1):737–38.
- Wayne, Natalie, and Daniel N. Bolon. 2007. “Dimerization of Hsp90 Is Required for in Vivo Function.” *Journal of Biological Chemistry* 282(48):35386–95.
- Weaver, David T. 1998. “Telomeres: Moonlighting by DNA Repair Proteins.” *Current Biology* 8(14):492–94.
- Webb, Christopher J., and Virginia A. Zakian. 2016. “Telomerase RNA Is More than a DNA Template.” *RNA Biology* 13(8):683–89.
- Weinrich, S. L., R. Pruzan, L. Ma, M. Ouellette, V. M. Tesmer, S. E. Holt, A. G. Bodnar, S. Lichtsteiner, N. W. Kim, J. B. Trager, R. D. Taylor, R. Carlos, W. H. Andrews, W. E. Wright, J. W. Shay, C. B. Harley, and G. B. Morin. 1997. “Reconstitution of Human Telomerase with the Template RNA Component HTR and the Catalytic Protein Subunit HTRT.” *Nature Genetics* 17(4):498–502.
- Wenz, C. 2001. “Human Telomerase Contains Two Cooperating Telomerase RNA Molecules.” *The EMBO Journal* 20(13):3526–34.
- Wong, J. M. Y., and K. Collins. 2006. “Telomerase RNA Level Limits Telomere Maintenance in X-Linked Dyskeratosis Congenita.” *Genes & Development* 20(20):2848–58.
- Wong, Mandy S., Woodring E. Wright, and Jerry W. Shay. 2014. “Alternative Splicing Regulation of

- Telomerase: A New Paradigm?" *Trends in Genetics* 30(10):430–38.
- Wotton, David, and David Shore. 1997. "A Novel Rap1p-Interacting Factor, Rif2p, Cooperates with Rif1p to Regulate Telomere Length in *Saccharomyces Cerevisiae*." *Genes and Development* 11(6):748–60.
- Wright, Woodring E., Mieczyslaw A. Piatyszek, William E. Rainey, William Byrd, and Jerry W. Shay. 1996. "Telomerase Activity in Human Germline and Embryonic Tissues and Cells." *Developmental Genetics* 18(2):173–79.
- Wright, Woodring E., Valerie M. Tesmer, Martha L. Liao, and Jerry W. Shay. 1999. "Normal Human Telomeres Are Not Late Replicating." *Experimental Cell Research* 251(2):492–99.
- Wu, Robert Alexander, Yavuz S. Dagdas, S. Tunc Yilmaz, Ahmet Yildiz, Kathleen Collins, and United States. 2015. "Single-Molecule Imaging of Telomerase Reverse Transcriptase in Human Telomerase Holoenzyme and Minimal RNP Complexes." 1–26.
- Wu, Robert Alexander, Jane Tam, and Kathleen Collins. 2017. "DNA -binding Determinants and Cellular Thresholds for Human Telomerase Repeat Addition Processivity ." *The EMBO Journal* 36(13):1908–27.
- Wynford-Thomas, David, and David Kipling. 1997. "Cancer and the Knockout Mouse." *Nature* 389(6651):551–52.
- Xi, Linghe, and Thomas R. Cech. 2014. "Inventory of Telomerase Components in Human Cells Reveals Multiple Subpopulations of HTR and HTERT." *Nucleic Acids Research* 42(13):8565–77.
- Yan, Chuangye, Jing Hang, Ruixue Wan, Min Huang, Catherine C. L. Wong, and Yigong Shi. 2015. "Structure of a Yeast Spliceosome at 3.6-Angstrom Resolution." *Science* 349(6253):1182–91.
- Yang, Ting-Lin B., Shufei Song, and F. Brad Johnson. 2016. "Contributions of Telomere Biology to Human Age-Related Disease." Pp. 205–39 in *Handbook of the Biology of Aging*. Elsevier.

- Ye, Jeffrey Zheng-Sheng, Jill R. Donigian, Megan van Overbeek, Diego Loayza, Yan Luo, Andrew N. Krutchinsky, Brian T. Chait, and Titia de Lange. 2004. "TIN2 Binds TRF1 and TRF2 Simultaneously and Stabilizes the TRF2 Complex on Telomeres." *Journal of Biological Chemistry* 279(45):47264–71.
- Ye, Jeffrey Zheng Sheng, and Titia De Lange. 2004. "TIN2 Is a Tankyrase 1 PARP Modulator in the TRF1 Telomere Length Control Complex." *Nature Genetics* 36(6):618–23.
- Zaug, A. J., E. R. Podell, and T. R. Cech. 2005. "Human POT1 Disrupts Telomeric G-Quadruplexes Allowing Telomerase Extension in Vitro." *Proceedings of the National Academy of Sciences* 102(31):10864–69.
- Zaug, A. J., E. R. Podell, J. Nandakumar, and T. R. Cech. 2010. "Functional Interaction between Telomere Protein TPP1 and Telomerase." *Genes & Development* 24(6):613–22.
- Zaug, Arthur J., Elaine R. Podell, and Thomas R. Cech. 2008. "Mutation in TERT Separates Processivity from Anchor-Site Function." *Nature Structural and Molecular Biology* 15(8):870–72.
- Zeng, Zhixiong, Wei Wang, Yuting Yang, Yong Chen, Xiaomei Yang, J. Alan Diehl, Xuedong Liu, and Ming Lei. 2010. "Structural Basis of Selective Ubiquitination of TRF1 by SCFFbx4." *Developmental Cell* 18(2):214–25.
- Zhang, Kai. 2018. "Gautomatch." Retrieved (www.mrc-lmb.cam.ac.uk/kzhang/).
- Zhang, Q., N. K. Kim, and J. Feigon. 2011. "Architecture of Human Telomerase RNA." *Proceedings of the National Academy of Sciences* 108(51):20325–32.
- Zhang, Xiaofeng, Chuangye Yan, Xiechao Zhan, Lijia Li, Jianlin Lei, and Yigong Shi. 2018. "Structure of the Human Activated Spliceosome in Three Conformational States." *Cell Research* 28(3):307–22.
- Zhang, Yi, Liuh Yow Chen, Xin Han, Wei Xie, Hyeung Kim, Dong Yang, Dan Liu, and Zhou Songyang. 2013. "Phosphorylation of TPP1 Regulates Cell Cycle-Dependent Telomerase

- Recruitment.” *Proceedings of the National Academy of Sciences of the United States of America* 110(14):5457–62.
- Zhao, Yong, Eladio Abreu, Jinyong Kim, Guido Stadler, Ugur Eskiocak, Michael P. Terns, Rebecca M. Terns, Jerry W. Shay, and Woodring E. Wright. 2011. “Processive and Distributive Extension of Human Telomeres by Telomerase under Homeostatic and Nonequilibrium Conditions.” *Molecular Cell* 42(3):297–307.
- Zhao, Zhizhen, and Amit Singer. 2013. “Rotationally Invariant Image Representation for Viewing Direction Classification in Cryo-EM.”
- Zheng, Shawn Q., Eugene Palovcak, Jean-Paul Armache, Kliment A. Verba, Yifan Cheng, and David A. Agard. 2017. “MotionCor2: Anisotropic Correction of Beam-Induced Motion for Improved Cryo-Electron Microscopy.” *Nature Methods* 14(4):331–32.
- Zhong, F., S. A. Savage, M. Shkreli, N. Giri, L. Jessop, T. Myers, R. Chen, B. P. Alter, and S. E. Artandi. 2011. “Disruption of Telomerase Trafficking by TCAB1 Mutation Causes Dyskeratosis Congenita.” *Genes & Development* 25(1):11–16.
- Zhong, Franklin L., Luis F. Z. Batista, Adam Freund, Matthew F. Pech, Andrew S. Venteicher, and Steven E. Artandi. 2012. “TPP1 OB-Fold Domain Controls Telomere Maintenance by Recruiting Telomerase to Chromosome Ends.” *Cell* 150(3):481–94.
- Zhu, Lingxiang, Karen S. Hathcock, Prakash Hande, Peter M. Lansdorp, Michael F. Seldin, and Richard J. Hodes. 1998. “Telomere Length Regulation in Mice Is Linked to a Novel Chromosome Locus.” *Proceedings of the National Academy of Sciences of the United States of America* 95(15):8648–53.
- Zivanov, Jasenko, Takanori Nakane, and Sjors H. W. W. Scheres. 2019. “A Bayesian Approach to Beam-Induced Motion Correction in Cryo-EM Single-Particle Analysis.” *IUCrJ* 6(1):5–17.
- Zvereva, M. I., D. M. Shcherbakova, and O. A. Dontsova. 2011. “Telomerase: Structure, Functions, and Activity Regulation.” *Biochemistry (Moscow)* 75(13):1563–83.