ICR The Institute of Cancer Research

UKRI Medical Research Council

PhD Project Proposal

Funder details

Studentship funded by: MRC iCase

Project details

Project title: Investigating the Regulation and Replication of R-loops at Telomeres

Supervisory team

Primary Supervisor: Max Douglas

Associate Supervisor(s):Jana Wolf (Tesselate Bio)

Secondary Supervisor:Chris Lord

Divisional affiliation

Primary Division:Cancer Biology

Primary Team:Telomere Biology

Site:Chelsea

Project background

Linear eukaryotic chromosomes terminate in specialised structures called telomeres, which prevent the natural chromosome ends from being recognised as double strand breaks. In human cells, telomeres are composed of approximately 10 kilobases of the repeat sequence TTAGGG bound to a six-membered protein complex called Shelterin. Shelterin protects the chromosome end by inhibiting DNA repair enzymes such as the MRN complex, and enables recruitment of the enzyme telomerase, allowing stable linear chromosomes of the correct length to be maintained from one generation to the next.

While telomeres bear chromatin marks associated with transcriptionally silent regions of the genome, they are in fact transcribed into a repetitive class of non-coding RNAs called TERRA, which range from several hundred to several thousand base pairs in length. TERRA is thought to frequently remain bound to the template after transcription due to the G-rich nature of telomeric repeats. The resulting 3-stranded ‘R-loop’ structure (Fig. 1) is known to inhibit DNA replication and telomeres containing elevated R-loops levels, such as in ALT type cancer cells, show high levels of replication stress and increased rates of telomere loss, consistent with R-loops preventing proper replication of the chromosome end.

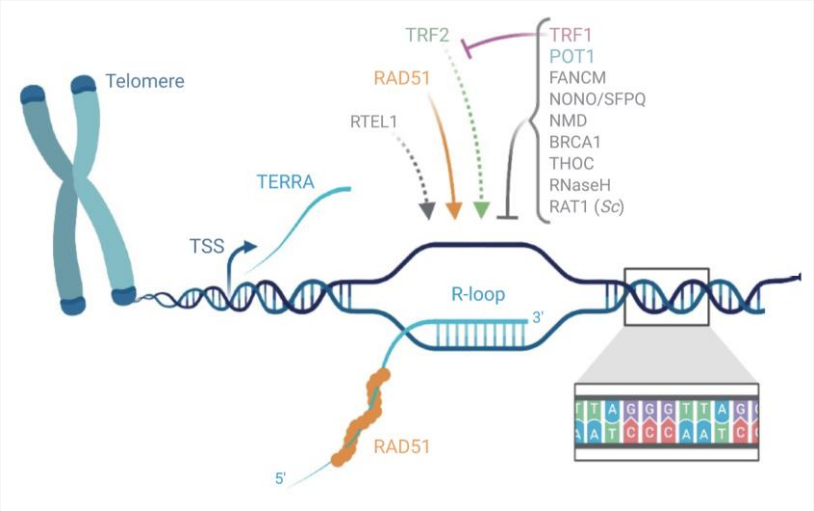

Figure 1. Graphic depiction of telomeric R-loops, and factors that have been proposed to regulate them. Adapted from Fernandez et al, Cell Cycle, 2021

A handful of factors have been proposed to regulate TERRA-mediated R-loops (Fig. 1). However, a comprehensive analysis of R-loop binding proteins at telomeres is yet to be performed, and our molecular understanding of how the factors that are known to affect R-loops is currently lacking. Furthermore, whilst R-loops are known to inhibit DNA replication, which aspects of the replication process are affected and how is largely unknown. Through the experimental strategy described below, we will address these questions using a combination of biochemical and genetic approaches.

Project aims

- Reconstitute telomeric R-loops in vitro
- Identify novel, telomere-specific R-loop binding proteins in extract-based pulldown assays
- Verify R-loop or telomere binding factors using cell-based assays
- Examine the effect of telomeric R-loops and binding proteins in reconstituted biochemical assays including an in vitro system for human DNA replication.

Research proposal

1. Identifying novel telomeric R-loop binding proteins.

Despite their crucial biological importance, a systematic analysis of proteins that bind telomeric R-loops is yet to be performed, likely due to the low abundance of these structures in most human cells. The starting point for the project is therefore to reconstitute telomeric and non-telomeric R-loops using T7 RNA polymerase and short DNA templates, and to use these structures to comprehensively identify telomeric and non-telomeric R-loop binding proteins. We will proceed in two steps:

i) We will use pulldown assays with human cell extracts and mass spectrometry to identify proteins that bind telomeric and non-telomeric R-loops. This work will be performed in collaboration with the ICR proteomics facility and builds on an established pulldown protocol we have developed to identify telomere end binding proteins. The outcome will be a list of factors that are enriched on R-loops per se, and factors that are enriched on telomeric R-loops specifically.

ii) To determine which factors from step i) are genuine R-loop or telomere binding factors, the student will spend a period of 6-months at Tesselate Bio to examine candidates using cell-based assays. this analysis will include a) colocalization experiments to identify candidate factors that can bind telomeres or R-loops within cells and

b) Crispr-mediated knock-out experiments to identify factors that affect R-loop associated events such as replication stress. Tesselate bio's expertise in studying R-loop-associated factors in vivo is central to the success of this part of the project. The outcome will be to verify which factors from part i) are genuine telomere or R-loop binding proteins, and to identify factors that are functionally important in vivo.

2. Mechanistic studies of R-loop regulating factors

Whilst part 1 will identify R-loop binding proteins, the second part of the project will harness our expertise in reconstitution biochemistry to examine how these factors affect R-loops and how R-loops influence DNA replication. This analysis will include proteins previously identified to regulate telomeric R-loops such as the branch point translocase FANCM and the helicase protein RTEL1. Our analysis will be divided into two parts:

i) We will use the baculovirus/insect cell system to purify factors of interest and combine them with reconstituted telomeric or non-telomeric R-loop structures. Our analysis will focus on whether candidate proteins can simply bind R-loops, or are able modify/unwind them. For factors specifically enriched on telomeric R-loops we will test whether they are specifically recruited to telomeric DNA/RNA sequences, or can bind purified telomeric factors such as Shelterin or the CST complex, which we already have in hand. The outcome will be to determine how novel and known R-loop regulating proteins affect R-loop assembly/disassembly, and to determine how some factors are specifically recruited to telomeric R-loops.

ii) We will examine the impact of R-loops and R-loop regulating proteins on a reconstituted human replisome. R-loops are known to act as a source of genome instability mainly due to an inhibitory effect on DNA replication. However, which aspects of DNA replication are affected and how is currently unclear. We have recently used purified proteins to reconstitute a human replisome that can replicate both leading and lagging strands at in vivo-like rates. Using the R-loop templates developed above, we will examine which aspects of DNA replication are affected by telomeric R-loops, and whether FANCM, RTEL1 or factors identified in part 1 can help to alleviate any inhibitory effects we observe. The outcome will depend on the success of part 1 and the time available for this aim, but could be a new mechanistic description of how telomeric R-loops are replicated efficiently.

In summary, telomeric R-loops are thought to be a major source of replication stress in ALT-type cancer cells, which have a particularly poor prognosis in the clinic. The work described above should lead to a new, mechanistic view of the factors that regulate this process.

Literature references

Candidate profile

Note: the ICR's standard minimum entry requirement is a relevant undergraduate Honours degree (First or 2:1).

Pre-requisite qualifications of applicants:

Intended learning outcomes:

- Ability to work with purified proteins in complex biochemical assays
- • Ability to perform and analyse pulldown proteomics experiments
- • Ability to perform cell-based assays relating to telomeres and replication stress
- • Ability to clearly present data and findings

Advertising details

Project suitable for a student with a background in:

- [x] Biological Sciences
- [ ] Physics or Engineering
- [ ] Chemistry
- [ ] Maths, Statistics or Epidemiology
- [ ] Computer Science