The Institute of Cancer Research

PHD STUDENTSHIP PROJECT PROPOSAL:

PROJECT DETAILS	
Project Title:	A fragment-based approach to identify protein-protein interaction inhibitors of Shelterin complex assembly
SUPERVISORY TEAM	
Primary Supervisor(s):	Professor Ian Collins and Dr Sebastian Guettler
Associate Supervisor(s):	Dr Mariola Zaleska
Backup Supervisor: (must have IRS status)	Dr Swen Hoelder
Lead contact person for the project:	Professor Ian Collins
DIVISIONAL AFFILIATION	
Primary Division:	Division of Cancer Therapeutics
Primary Team:	Medicinal Chemistry 2
Other Division (if applicable):	Division of Structural Biology
Other Team (if applicable)):	Structural Biology of Cell Signalling

PROJECT PROPOSAL

BACKGROUND TO THE PROJECT

The ends of our linear chromosomes, the telomeres, are specialised chromatin structures essential to genomic integrity and stability[1]. Telomeric DNA, which in humans spans 8-15 kb in length, consists of TTAGGG repeats that are decorated with telomere-binding proteins. In mammals, six subunits form the so-called Shelterin complex that plays important roles in *telomere protection* and *telomere extension*[2,3]. **Telomere protection** refers to the suppression of an illicit DNA damage response (DDR) at telomeres by distinguishing natural chromosome ends from damage-induced DNA double strand breaks. In other words, whilst an efficient DNA repair programme at sites of DNA damage is essential to ensure cell viability, an unwanted DDR at telomeres would lead to a loss of genomic integrity and needs to be averted. **Telomere extension** refers to the addition of new telomeric repeats onto telomeres by the enzyme telomerase, as telomeric repeats are gradually lost during semi-conservative DNA replication. The Shelterin complex contains six proteins[2] (Figure 1). These include the DNA double-strand binders TRF1 and TRF2 and the DNA single-strand binder POT1. The TIN2 subunit bridges TRF1 and TRF2 and interacts with TPP1, which in turn binds to POT1. RAP1 associates with TRF2. Many of the protein-protein contacts within Shelterin are mediated by domain-peptide interactions[4] (Figure 1).

The Shelterin subunits collaborate to suppress the DDR at telomeres[2]. Whilst all Shelterin components, with the exception of RAP1, are required for telomere protection, different subunits suppress different arms of the complex DDR. Among the major DDR pathways repressed are the ataxia telangiectasia mutated (ATM) kinase and classical

non-homologous end joining (NHEJ) pathways by TRF2, as well as the ataxia telangiectasia and Rad3-related (ATR) kinase signalling by POT1[2].

Telomere length is maintained by recruitment of telomerase to the TPP1 subunit of Shelterin, but TRF1 can control telomere length by a poorly characterised but telomerase-dependent mechanism[5].

There are currently no tool compounds available to inhibit the interactions between specific Shelterin subunits. Such compounds would be valuable research tools to study the biological functions of different Shelterin components and subcomplexes. Moreover, such compounds may induce acute telomere deprotection, which has been shown to selectively kill certain cancer cells (e.g., glioblastoma and lung cancer), limit tumour growth in $p53^{-/-}$ and $Ink4Arf^{-/-}$ mouse cancer models, and been proposed to provide a more promising strategy to target cancer cells than telomerase inhibition[6].

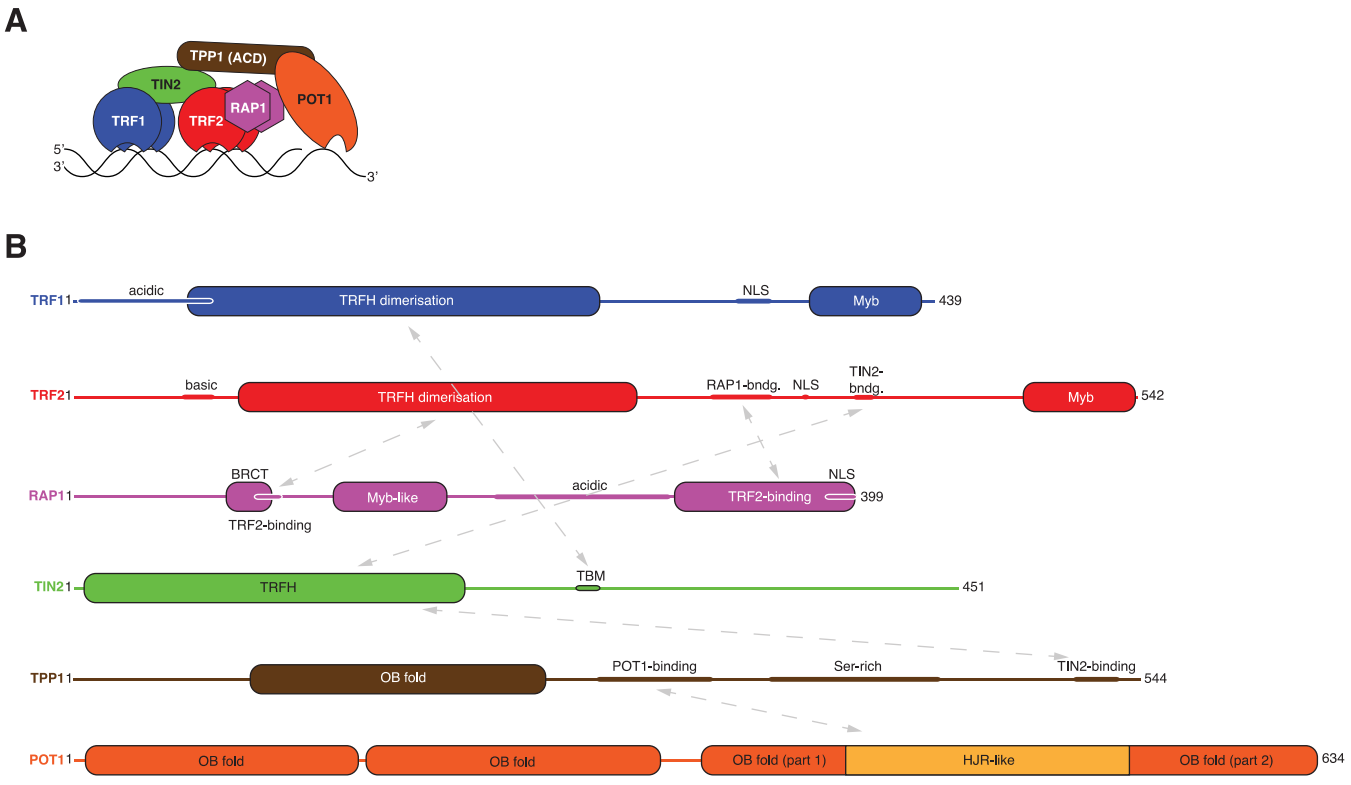

Figure 1: Components and domain organisation of the human Shelterin complex. (A) Schematic of the telomeric Shelterin complex. **(B)** Domain organisation of the human Shelterin components. Known protein-protein interactions are illustrated by arrows.

PROJECT AIMS (up to 5 bullet points)

- to identify small molecule fragments that bind to globular domains in Shelterin components, using differential scanning fluorimetry (DSF) / Thermofluor and ligand-observed NMR methods
- to validate identified domain binders using orthogonal biophysical techniques
- to characterise fragment binding sites using protein-observed NMR and, if possible, X-ray crystallography
- to characterise fragment-protein interaction affinities using protein-observed NMR and ITC

- to perform early structure-activity relationship (SAR) studies

RESEARCH PROPOSAL

In the first part of the project, the student will express and purify, to high quality, the full complement of domains present in the Shelterin complex, either in *E. coli* or using the insect cell/baculovirus system for those domains that cannot be obtained in soluble form from bacteria. In total, there will be 11 domains: the TRF1 and TRF2 TRF homology (TRFH) and DNA-binding MYB domains, the RAP1 BRCA1 C-terminal (BRCT), MYB-like and RAP1 C-terminal (RCT) domains, the TIN1 TRFH-like domain, the TPP1 oligonucleotide/oligosaccharide (OB) fold domain, and the POT1 OB fold domains, which encompass an N-terminal tandem OB domain pair and a C-terminal split OB fold domain[4]. For all domains, the final purification step will be size exclusion chromatography. The domains will be characterised by size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) to verify molecular weight and stoichiometry/homogeneity.

In the next and main step of the project, the student will screen the ICR fragment library for binders of these domains. We will employ two orthogonal screening approaches, as for our recent study to identify binders of tankyrase substrate recruitment modules[7]. First, we will exploit the ability of individual fragments to stabilise the domains, using differential scanning fluorimetry (DSF), also known as Thermofluor[8]. Hits will next be validated using newly purchased fragments, quality-controlled by HPLC-MS. The student will validate the fragments by two ligand-observed NMR approaches: relaxation-edited NMR (Carr-Purcell-Meiboom-Gill, CPMG) and WaterLOGSY[9]. In an orthogonal primary screen, the student will use relaxation-edited ligand-observed NMR to interrogate pools of fragments for domain binding. Fragment pools will next be deconvoluted by re-purchasing and quality-controlling individual fragments and analysing them by the two orthogonal ligand-observed NMR approaches.

In the final step of the project, we will select a subset of domains to perform more focussed biophysical analyses to study the interaction mode of selected fragments, their affinity as well as perform initial structure-activity relationship (SAR) studies. The choice of domain will depend on its function in Shelterin as well as the obtained hit matter. For example, targeting TRF1 may be attractive because of the demonstrated cancer-specific effect of TRF1 depletion on cell survival[6] and the reported role of TRF1 in controlling telomere length[10]. TIN2 will be an interesting target due to its role as a central and allosterically regulated Shelterin scaffolder[4]. The student will first validate fragment binding by protein-observed NMR using ^{15}N-labelled recombinant proteins. If significant chemical shift perturbations are observed, protein-observed NMR will be used to estimate fragment binding affinities. This will be complemented by isothermal titration calorimetry (ITC) studies. The student will next use relaxation-edited ligand-observed NMR methods to explore whether fragments can be competed from the protein domains using peptide motifs from interacting Shelterin complex components known to bind the particular domains[4]. To directly map the fragment binding sites on the proteins, the student will use X-ray crystallography to obtain co-crystal structures of relevant domain-fragment complexes, either by co-crystallisation or crystal soaking. If co-crystal structures cannot be obtained, we will, in collaboration with Dr Mark Pfuhl (King's College London), perform a backbone NMR assignment of the relevant domains, which will enable the direct identification of the fragment binding sites and *in-silico* docking studies. The student will further validate fragment binding through testing analogues of primary hits, either obtained commercially or made through straightforward synthetic routes in Professor Ian Collins' laboratory. Analysing these analogues will generate data on initial SAR and inform modelling of the binding modes.

The completed PhD project will provide a foundation for the development of potent protein-protein interaction inhibitors that interfere with the assembly of the Shelterin complex to induce acute telomere deprotection or impact telomere length homeostasis.

For general information on activities in the host laboratories, please visit:
www.icr.ac.uk/our-research/researchers-and-teams/professor-ian-collins
www.icr.ac.uk/our-research/researchers-and-teams/sebastian-guettler
www.sguettlerlab.org

LITERATURE REFERENCES

1. Shay, J. W. & Wright, W. E. Telomeres and telomerase: three decades of progress. *Nat. Rev. Genet.* **20,** 299–309 (2019).
2. de Lange, T. Shelterin-Mediated Telomere Protection. *Annu. Rev. Genet.* (2018). doi:10.1146/annurev-genet-032918-021921
3. Stewart, J. A., Chaiken, M. F., Wang, F. & Price, C. M. Maintaining the end: roles of telomere proteins in end-protection, telomere replication and length regulation. *Mutat. Res.* **730,** 12–19 (2012).
4. Chen, Y. The structural biology of the shelterin complex. *Biol. Chem.* **400,** 457–466 (2019).
5. Nandakumar, J. & Cech, T. R. Finding the end: recruitment of telomerase to telomeres. *Nat Rev Mol Cell Biol* **14,** 69–82 (2013).
6. Bejarano, L. *et al.* Safety of Whole-Body Abrogation of the TRF1 Shelterin Protein in Wild-Type and Cancer-Prone Mouse Models. *iScience* **19,** 572–585 (2019).
7. Pollock, K. *et al.* Fragment-based screening identifies molecules targeting the substrate-binding ankyrin repeat domains of tankyrase. *biorxiv.org* **20,** 737
8. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2,** 2212–2221 (2007).
9. Ma, R., Wang, P., Wu, J. & Ruan, K. Process of Fragment-Based Lead Discovery-A Perspective from NMR. *Molecules* **21,** 854 (2016).
10. van Steensel, B. & de Lange, T. Control of telomere length by the human telomeric protein TRF1. *Nature* **385,** 740–743 (1997).

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: e.g. BSc or equivalent in specific subject area(s)	BSc Hons in chemistry, biophysics, biochemistry or similar (some laboratory experience is required), scientific curiosity and a keen interest in research
Intended learning outcomes:	• Expert in protein expression, purification and protein biochemistry • Expert in the theory and practice of fragment-based ligand discovery • Experience and skills in the application of NMR and other biophysical techniques to measure ligand binding • Knowledge of protein crystallisation and X-ray crystallography • Experience and skills in the application of *in silico* methods to rationalise structure-activity relationships

ADVERTISING DETAILS

Project suitable for a student with a background in:	☒ Biological Sciences ☐ Physics or Engineering ☒ Chemistry ☐ Maths, Statistics or Epidemiology ☐ Computer Science ☐ Other (provide details)
Keywords:	**1. PhD chemical biology**
	2. protein biochemistry
	3. fragment screening
	4. protein-protein interactions
	5. biophysics
	6. telomeres