What I got wrong about shelterin

The ASBMB 2018 Bert and Natalie Vallee award in Biomedical Sciences honors our work on shelterin, a protein complex that helps cells distinguish the chromosome ends from sites of DNA damage. Shelterin protects telomeres from all aspects of the DNA damage response, including ATM and ATR serine/threonine kinase signaling and several forms of double-strand break repair. Today, this six-subunit protein complex could easily be identified in one single proteomics step. But, it took us more than 15 years to piece together the entire shelterin family be identified in one single proteomics step. But, it took us more than 15 years to piece together the entire shelterin complex, one protein at a time. Although we did a lot of things right, here I tell the story of shelterin's discovery with an emphasis on the things that I got wrong along the way.

TRF: Hesitation

Once Harold Varmus and I had cloned human telomeres and found they contained TTAGGG repeats (this was in early 1988, but we were scooped by Robert Moyzis’ paper later that year (1)), I wrote in my notebook that I should start looking for a protein that recognizes this sequence. At that time, there was a single precedent for such a protein. David Prescott had identified a protein complex that strongly bound to the ends of gene-sized DNA molecules that make up the macronuclear genome of the ciliate Oxytricha (2). Dan Gottschling (first in Tom Cech’s lab and then in Ginger Zakian’s lab) had published two papers on Prescott’s terminal complex, now called TEBPα/β, showing that it is specific for the 3’ single-stranded (ss) end of the telomeres (3, 4). So, although the bulk of the newly identified human telomeric repeats was likely to be in double-stranded form, I thought I should look for a protein that could bind ssTTAGGG repeats.

Lily Shue, a technician who was paid from my Lucille P. Markey Trust award, was tasked with looking for such an activity. She used HeLa nuclear extract provided by Grant Herzog, then a graduate student in Rick Myers’ transcription-focused lab and now Lily’s husband. Grant tutored Lily and me on gel-shift analysis, making probes, and doing specificity assays. Almost immediately, Lily observed a very abundant activity that bound to ssTTAGGG repeats. But within few months, Lily delivered devastating news to me at a New Year’s Eve party: the activity she was looking at had higher affinity for RNA than DNA. A few years later, first Howard Cooke and then Tom Cech published on the activity that Lily had detected, which turned out to be due to hnRNPA1 and other similar RNA-binding proteins (5, 6). So, our New Year’s resolution for 1989 was to look for proteins bound to double-stranded (ds) TTAGGG repeats instead. This search yielded a much less abundant binding activity that was specific for DNA, not RNA. We called it TRF (telomeric repeat binding factor).

In June 1989, the Rockefeller University offered me a position, which was a miracle in my view because my main paper on the shortening of human telomeres in somatic cells and cancers had just been rejected from Science (it came out in MCB in 1990 (7)). Upon moving to Rockefeller University in the mid-1990s, I set out to execute two experiments that I had initiated in Harold Varmus’ laboratory. Both were designed to look for in vivo evidence for a TRF-like factor at telomeres. I reasoned that if telomeres contained something like TRF, their nucleosomial structure might be altered. These experiments, completed with graduate student Henrik Tommerup, indeed showed an unusually blurred nucleosomal (MNase) pattern in the ds part of the telomeres, suggestive of nonhistone proteins associated with the telomeric DNA (8). The second experiment was designed to determine whether telomeres were bound to the nuclear matrix. It was known that certain sites in the genome were bound to the nuclear matrix due to the presence of nonhistone proteins, and I reasoned that if telomeres bound to the nuclear matrix, it would be evidence for the nonhistone proteins being present. I found that telomeres stuck to the nuclear matrix very strongly, again consistent with something like TRF existing in vivo (9).

While I was finishing the chromatin and nuclear matrix experiments, a technician, Shawn Kaplan, and a graduate student, Zhong Zhong, were following up on Lily Shue’s TRF activity. Zhong and Shawn determined the biochemical features of TRF, documenting a satisfying preference for TTAGGG repeats over telomeric sequences from other organisms (e.g. TTGGGG repeats from Tetrahymena) (10). But Zhong also showed that TRF had a very high off-rate, raising doubts that TRF could be functional at telomeres given its labile binding, and he left the lab in 1991 to work on the then new STAT pathway with James Darnell.

In a final attempt to reassure myself that telomeres contained a dsDNA-binding protein, I asked a postdoc, John Hanish, to set up an assay for formation of new telomeres in human cells. The idea was to test whether the sequence requirements for telomere formation were similar to those of TRF. Transfection of...
If I had not been side-tracked by all the experiments aimed to investigate the role of Rap1 in telomeric DNA, I would have already discovered a new DNA-binding factor in budding yeast by Kim Nasmyth and David Shore. This factor was shown to bind to yeast ds telomeric DNA in vitro (12, 13). In 1990, Art Lustig and David Shore as well as Ginger Zakian’s lab showed that Rap1 affects the length of telomeres (14, 15). Although this effect could have been indirect (resulting from a transcription effect of Rap1), the simpler interpretation was that Rap1 directly bound to the ds telomeric DNA in yeast.

Despite my findings and the work in yeast, I was unable to convince people in my lab to work on TRF after Zhong’s departure. I decided to do it myself, together with technician Laura Chong. I had very little experience in biochemistry and none in protein purification. It took us 2 years to figure out how to purify TRF. It involved six columns and delivered two detectable protein bands. Isolation and renaturation of the 60-kDa band showed that this was the TRF activity. Knowing that TRF was in this band was good news, but it also revealed that our yield was devastatingly low. Paul Tempst at Memorial Sloan Kettering Cancer Center had told me he needed 5 µg of a 60-kDa protein to derive peptide sequences by Edman degradation. This meant that we needed to purify TRF from close to >500 liters of HeLa cell lysate, which was more than I could grow in my laboratory. I forked over $17,500 to buy HeLa cells (shipped as pellets on dry ice), a sum close to the supply budget of my single National Institutes of Health (NIH) GM RO1 grant. It worked! Paul Tempst was able to derive TRF peptide sequences that allowed Dominique (Kiki) Broccoli and me to isolate the TRF cDNA from a phage library.

After several months of cDNA cloning and sequencing, Kiki, Laura, and I determined the TRF protein sequence in late 1993. To my disappointment, TRF was not homologous to Rap1. Worse, TRF looked like a transcription factor, with an Myb-type DNA-binding domain at the C terminus and an acidic domain at the N terminus. Had I just spent 2 years and most of my NIH grant on an irrelevant transcription factor with spurious TTAGGG repeat binding activity? Two pieces of information pulled me out of the funk that I was experiencing in early 1994. First, Bas van Steensel used indirect immunofluorescence to show that TRF was present at HeLa cell chromosome ends. Second, I heard from David Shore that Daniela Rhodes at the Laboratory of Molecular Biology (Cambridge, UK) had just determined the crystal structure of Rap1 and had found two Myb-like folds in its DNA-binding domain.

It took us a year to finalize the paper, which came out at the end of 1995, more than 7 years after we first observed TRF (16). If I had not been side-tracked by all the experiments aimed to validate a TRF-like protein in vivo and if I had had a stronger conviction and more courage, we would have isolated the TRF1 gene and probably the rest of shelterin several years earlier. Although these experiments kept the laboratory reasonably productive in its early years, they also extended the risky TRF gamble further in time.

TRF2, Rap1, TIN2: Missed opportunities

In early 1996, I found a short cDNA sequence in the EST database that looked similar to the Myb domain of TRF1. Could there be a second TRF? I considered this unlikely because I had done super-shift experiments with a TRF1 antibody; all the detectable HeLa dsTTAGGG-binding activity was shifted up by incubation with the TRF1 antibody, suggesting that TRF1 was the only TTAGGG repeat binding factor in the nuclear extract. I had advertised this view to my lab members who were now (nearly) all working on TRF1. So, once again, I could not find anybody to work on this mysterious Myb domain protein. Fortunately, an M.D. Ph.D. student, Agata Smogorzewska, joined the lab for a rotation in the summer of 1996, and she was unawake (or dismissive) of the lab lore. Agata cloned the cDNA of what now is called TRF2 (17). In addition to the Myb domain, TRF1 and TRF2 showed a large region of homology, now known to represent their highly conserved homodimerization domains (18). We called this the TRF-homology (TRFH) domain by analogy to the Src homology (SH) domains, a nod to my Varmus lab oncogene roots. At the N terminus, TRF2 carried a basic domain, rather than the acidic domain that makes TRF1 resemble a transcription factor. The basic N terminus was the reason for my failure to detect TRF2 by gel-shift assays, creating a complex that was hung up in the slot (17). So if Lily and I had used a different gel system, I could have isolated TRF1 and TRF2 at the same time. A missed opportunity . . .

By 1996, I had caught on to the idea that there might be more proteins at telomeres (additional factors besides Rap1 had by now been identified in yeast), and many lab members started looking for them. A two-hybrid screen done by postdoctoral fellow Susan Smith had delivered tankyrase, a poly(ADP-ribose) polymerase (PARP) that binds to TRF1 (19). A second two-hybrid screen by postdoctoral fellow Bibo Li revealed a TRF2-binding protein that we called HC1 until I realized, while writing the paper, that it was the mammalian ortholog of Rap1 (20).

We were pleased with our new proteins and stopped doing two-hybrid screens, turning to the newly developed MS approach instead. This was yet another mistake. Both of our two-hybrid screens had missed a key protein in shelterin, TIN2, which we now know binds to both TRF1 and TRF2. TIN2 (TRF1-interacting nuclear factor 1) was isolated by Judy Campisi’s lab in a two-hybrid screen with TRF1 (21). TIN2 did not come out of our TRF2 mass spectrometry efforts either because the TIN2–TRF2 link was unstable at the salt condition used for protein isolation (22, 23). It took several years before Jeff Ye, a postdoctoral fellow in my lab and also the Songyang lab, demonstrated that TIN2 could bind to both TRF1 and TRF2 (24, 25).
Rif1 and Tel2: Following the yeast trail to nowhere

Because Bibo Li had found the mammalian ortholog of yeast Rap1, I was convinced we should be able to find other budding yeast telomere factors in human cells. In early 2001, Celera allowed academic researchers to search Craig Venter’s genome sequence for the steep fee (at least for my lab) of $7,500 per year. As soon as we had access to his DNA sequence, we used budding yeast and fission yeast data to find the most conserved regions of candidate telomere–relevant genes and then BLASTed them through Venter’s DNA. Our gene list included the Oxytricha telomeric proteins (α and β), Tel2, and the two Rap1-interacting factors Rif1 and Rif2. The latter three were known to affect telomere length regulation in budding yeast (26–28).

Postdoctoral fellow Diego Loyaza rapidly found a human ortholog of Oxytricha TEBPα. He started working on this candidate, cloning the full-length cDNA and then looking at whether the protein localized to telomeres. Before he got very far, we attended the 2002 meeting on Telomeres and Telomerase at Cold Spring Harbor Laboratory (CSHL), where Peter Baumann from Tom Cech’s lab announced Diego’s gene as POT1 (Protection of Telomeres 1) and showed that the POT1 protein bound human telomeric DNA in vitro (29). After this painful episode, Diego recovered and studied the effect of POT1 on telomere length regulation (30).

Venter’s genome did not only deliver POT1, but also Rif1 and Tel2 (Rif2 is a budding yeast invention). Rif1 and Tel2 became the thesis projects of two M.D. Ph.D. students: Rich Wang and Josh Silverman. My expectation was that they would find that Rif1 and Tel2 affect telomere length, so that they could graduate in time to finish their medical training. However, we learned that neither protein had anything to do with telomeres. Rif1 is a DNA damage–response factor as Josh showed (31), and Tel2 is a chaperone that regulates the stability of PI3K-related kinases as postdoctoral fellow Hiro Takai figured out later (32). Rich had given up on Tel2 and graduated on his finding that TRF2 protects telomeres from t-loop cleavage (33). Ever since this episode, I have been hesitant to take budding yeast information too literally as a guide for our work on mammalian telomeres.

TPP1: The final link

In 2002, Jeff Ye had found that TIN2 interacts with both TRF1 and TRF2 and that this interaction stabilizes the two TRFs on telomeres in vivo (24), leading us to think that telomeres contained a complex composed of TRF1–TIN2–TRF2–Rap1 and a separate POT1 protein. But Diego Loayza had found that POT1 is somehow linked to TRF1 and TIN2 (30). Diego had also shown that POT1 does not have to bind telomeric DNA to accumulate at telomeres, consistent with a protein-based recruitment. Jeff ran HeLa nuclear extract over a sizing column and indeed found evidence of a large complex that contained TRF1–TIN2–TRF2–Rap1 as well as POT1 (24). So, most likely, POT1 was recruited to telomeres by an (indirect) interaction with the TRF1–TIN2–TRF2–Rap1 complex. For these reasons, Jeff collaborated with our colleague Brian Chait in a mass spectrometry analysis on TIN2-associated proteins and discovered TPP1 in 2003. We first called it POT1-interacting protein 1 (PIPI1) because Jeff had quickly found that TPP1 interacted with POT1, now linking all known telomeric proteins together by protein–protein interactions. Jeff presented his data at the 2004 CSHL meeting where two other labs (Smith and Songyang) announced the same result, and a session in the bar made us all settle on the name TPP1 (the two Ps stand for PIPI1 and Songyang’s PTOP (34) and the T reflects Smith’s TINT1 (35)).

Shelterin: Commitment

In June 2005, I arrived at the LXX CSHL Symposium on Quantitative Biology and realized that I was slated to give a special talk, the Reginald Harris lecture. Somehow this had dropped out of my memory. Panicked, I sat in my room for a full day to change my talk into something appropriate for this honor. I reviewed everything we knew about the telomeric protein complex and decided to declare victory. In my talk, I announced that telomeres contain a six-subunit complex, which I proposed to call shelterin. I discussed how shelterin regulates telomere length and proposed that shelterin blocks the DNA damage response by forming the telomeric t-loop structure (36). Coming back from the meeting, I learned that my lab hated the name shelterin. But there was no turning back after my CSHL talk and my promise to Terri Grodzicker that I would write a review on shelterin (37). In the review, I made the argument that shelterin was the main mechanism by which telomeres prevent activation of the DNA damage response (37). I relegated tankyrase, the Mre11 complex, and many other protein–protein interactions. Jeff presented his data at the 2004 CSHL meeting where two other labs (Smith and Songyang) announced the same result, and a session in the bar made us all settle on the name TPP1 (the two Ps stand for PIPI1 and Songyang’s PTOP (34) and the T reflects Smith’s TINT1 (35)).

Acknowledgments—I thank all current and past members of my lab for their remarkable dedication to the problems we have worked on and for making my lab life heavenly.

Addendum—John Hanish’s experiment also showed that telomeres from one organism (e.g. Tetrahymena) do not function in another (human cells). This was an important point because Jack Szostak and Liz Blackburn had done a similar experiment in Saccharomyces cerevisiae where Tetrahymena appeared to stabilize the ends of a linear DNA (38). Their experiment was generally interpreted to evidence trans-kingdom conservation of telomere function, which would be a strong argument against sequence specific proteins that recognize the ds telomeric DNA. However, on closer inspection, these Tetrahymena telomeres always had yeast telomeric DNA added to them. My interpretation was that the yeast telomeric DNA was needed for stable ends, and the Tetrahymena sequence had functioned as a telomerase primer. John’s telomere healing experiments also argued against the idea that telomeres were protected by the G4 structure. This had been a popular model after G4 structures were found to be a common feature of telomeric repeats. Our experiments showed that G4 structures were certainly not enough for telomere function because they could be formed by several of the sequences that did not seed new telomeres.
References

1. Moyzis, R. K., Buckingham, J. M., Cram, L. S., Dani, M., Deaven, L. L., Jones, M. D., Meyne, J., Ratliff, B. L., and Wu, J. R. (1988) A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6622–6626. CrossRef Medline

2. Lipps, H. J., Grüssem, W., and Prescott, D. M. (1982) Higher order DNA structure in macronuclear chromatin of the hypotrichous ciliate *Oxytricha nova*. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2495–2499. CrossRef Medline

3. Gottschling, D. E., and Cech, T. R. (1984) Chromatin structure of the molecular ends of *Oxytricha* macronuclear DNA: phased nucleosomes and a telomeric complex. *Cell* **38**, 501–510. CrossRef Medline

4. Gottschling, D. E., and Zakian, V. A. (1986) Telomere proteins: specific recognition and protection of the natural termini of *Oxytricha* macronuclear DNA. *Cell* **47**, 195–205. CrossRef Medline

5. McKay, S. J., and Cooke, H. (1992) A protein which specifically binds to single stranded TTAGGGn repeats. *Nucleic Acids Res.* **20**, 1387–1391. CrossRef Medline

6. Ishikawa, F., Matunis, M. I., Dreyfuss, G., and Cech, T. R. (1993) Nuclear proteins that bind the pre-mRNA 3’ splice site sequence r(UUAG/G) and the human telomeric DNA sequence d(TTAGGG)n. *Mol. Cell. Biol.* **13**, 4301–4310. CrossRef Medline

7. de Lange, T., Shiu, L., Myers, R. M., Cox, D. R., Naylor, S. L., Killery, A. M., and Varmus, H. E. (1990) Structure and variability of human chromosome ends. *Mol. Cell. Biol.* **10**, 518–527. CrossRef Medline

8. Tommerup, H., Dousmanis, A., and de Lange, T. (1994) Unusual chromatin in human telomeres. *Mol. Cell. Biol.* **14**, 5777–5785. CrossRef Medline

9. de Lange, T. (1992) Human telomeres are attached to the nuclear matrix. *EMBO J.* **11**, 717–724. Medline

10. Zhong, Z., Shiu, L., Kaplan, S., and de Lange, T. (1992) A mammalian factor that binds telomeric TTAGGG repeats in vitro. *Mol. Cell. Biol.* **12**, 4834–4843. CrossRef Medline

11. Hanish, J. P., Yanowitz, J. L., and de Lange, T. (1994) Stringent sequence requirements for the formation of human telomeres. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8861–8865. CrossRef Medline

12. Longtine, M. S., Wilson, N. M., Petracek, M. E., and Berman, J. (1989) A yeast telomere binding activity binds to two related telomere sequence motifs and is indistinguishable from RAP1. *Curr. Genet.* **16**, 225–239. CrossRef Medline

13. Shore, D., and Nasmyth, K. (1987) Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. *EMBO J.* **6**, 351–361. CrossRef Medline

14. Conrad, M. N., Wright, J. H., Wolf, A. J., and Zakian, V. A. (1990) RAP1 protein interacts with yeast telomeres in vivo: overproduction alters telomere length and decreases chromosome stability. *Cell* **63**, 739–750. CrossRef Medline

15. Luckst, A. J., Kurtz, S., and Shore, D. (1990) Involvement of the silencer and UAS binding protein RAP1 in regulation of telomere length. *Science* **250**, 549–553. CrossRef Medline

16. Chong, L., van Steensel, B., Broccoli, D., Erdjument-Bromage, H., Hanish, J., Tempst, P., and de Lange, T. (1995) A human telomeric protein. *Science* **270**, 1663–1667. CrossRef Medline

17. Broccoli, D., Smogorzewska, A., Chong, L., and de Lange, T. (1997) Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat. Genet.* **17**, 231–235. CrossRef Medline

18. Fairall, L., Chapman, L., Moss, H., de Lange, T., and Rhodes, D. (2001) Structure of the TRF1 dimerization domain of the human telomeric proteins TRF1 and TRF2. *Mol. Cell* **8**, 351–361. CrossRef Medline

19. Smith, S., Giriati, L., Schmitt, A., and de Lange, T. (1998) Tankyrase, a poly(ADP-ribose) polymerase at human telomeres. *Science* **282**, 1484–1487. CrossRef Medline

20. Li, B., Oestreicher, S., and de Lange, T. (2000) Identification of human Rapi: implications for telomere evolution. *Cell* **101**, 471–483. CrossRef Medline

21. Kim, S. H., Kaminker, P., and Campisi, J. (1999) TIN2, a new regulator of telomere length in human cells. *Nat. Genet.* **23**, 405–412. CrossRef Medline

22. Zhu, X. D., Küster, B., Mann, M., Petrini, J. H., and de Lange, T. (2000) Cell-cycle-regulated association of RAD50/MRE11/NBS1 with TRF2 and human telomeres. *Nat. Genet.* **25**, 347–352. CrossRef Medline

23. Zhu, X. D., Niedernhofer, L., Kuster, B., Mann, M., Hoeijmakers, J. H., and de Lange, T. (2003) ERCC1/XPF removes the 3’ overhang from uncapped telomeres and represses formation of telomeric DNA-containing single minute chromosomes. *Mol. Cell* **12**, 1489–1498. CrossRef Medline

24. Ye, J. Z., Donigian, J. R., van Overbeek, M., Loayza, D., Luo, Y., Krutchinsky, A. N., Chait, B. T., and de Lange, T. (2004) TIN2 binds TRF1 and TRF2 simultaneously and stabilizes the TRF2 complex on telomeres. *J. Biol. Chem.* **279**, 47264–47271. CrossRef Medline

25. Liu, D., O’Connor, M. S., Qin, J., and Songyang, Z. (2004) Telosome, a mammalian telomere-associated complex formed by multiple telomeric proteins. *J. Biol. Chem.* **279**, 51338–51342. CrossRef Medline

26. Lustig, A. J., and Petes, T. D. (1986) Identification of yeast mutants with altered telomere structure. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1398–1402. CrossRef Medline

27. Hardy, C. F., Sussel, L., and Shore, D. (1992) A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation. *Genes Dev.* **6**, 801–814. CrossRef Medline

28. Wotton, D., and Shore, D. (1997) A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to regulate telomere length in *Saccharomyces cerevisiae*. *Genes Dev.* **11**, 748–760. CrossRef Medline

29. Baumann, P., and Cech, T. R. (2001) Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* **292**, 1171–1175. CrossRef Medline

30. Loayza, D., and De Lange, T. (2003) POT1 as a terminal transducer of TRF1 telomere length control. *Nature* **423**, 1013–1018. CrossRef Medline

31. Silverman, J., Takai, H., Buonomo, S. B., Eisenhaber, F., and de Lange, T. (2004) Human Rif1, ortholog of a yeast telomeric protein, is regulated by ATM and 53BP1 and functions in the S-phase checkpoint. *Genes Dev.* **18**, 2108–2119. CrossRef Medline

32. Takai, H., Wang, R. C., Takai, K. K., Yang, H., and de Lange, T. (2007) Tel2 regulates the stability of PI3K-related protein kinases. *Cell* **131**, 1248–1259. CrossRef Medline

33. Wang, R. C., Smogorzewska, A., and de Lange, T. (2004) Homologous recombination generates T-loop-sized deletions at human telomeres. *Cell* **119**, 355–368. CrossRef Medline

34. Liu, D., Safari, A., O’Connor, M. S., Chan, D. W., Laegeler, A., Qin, J., and Songyang, Z. (2004) POT1 interacts with POT1 and regulates its localization to telomeres. *Nat. Cell Biol.* **6**, 673–680. CrossRef Medline

35. Houghtaling, B. R., Cuttonaro, L., Chang, W., and Smith, S. (2004) A dynamic molecular link between the telomere length regulator TRF1 and the chromosome end protector TRF2. *Curr. Biol.* **14**, 1621–1631. CrossRef Medline

36. Griffith, J. D., Comeau, L., Rosenfield, S., Stansel, R. M., Bianchi, A., Moss, H., and de Lange, T. (1999) Mammalian telomeres end in a large duplex loop. *Cell* **97**, 503–514. CrossRef Medline

37. de Lange, T. (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev.* **19**, 2100–2110. CrossRef Medline

38. Szostak, J. W., and Blackburn, E. H. (1982) Cloning yeast telomeres on linear plasmid vectors. *Cell* **29**, 245–255. CrossRef Medline