

DNA ENDS: JUST THE BEGINNING

Nobel Lecture, December 7, 2009

by

JACK W. SZOSTAK

Harvard Medical School; Howard Hughes Medical Institute; Massachusetts General Hospital, Boston, MA, U.S.A.

The contributions of my laboratory to our understanding of telomere function and maintenance by telomerase were made over a limited period of time early in the development of this story, from 1980 to 1989. What I would like to discuss here are some of the problems that we had to overcome, especially the preconceptions we had about models for telomere function and how hard it was to let go of those models. Fortunately the evidence we uncovered was strong enough to bring us to the right conclusions! Then, since I left the telomere field fairly early on, I would like to take this opportunity to briefly review some of the work that we've done since, primarily to show students who are just entering science that it is not only possible but really fun to address very different questions in different fields during one's career.

There were two well-known and long-standing puzzles associated with the nature of eukaryotic chromosome ends, or telomeres: the problem of the stability of the ends of chromosomes, and the problem of complete replication. My first introduction to these issues came when I was an undergraduate student at McGill University in Montreal. The first of those two problems, the reactivity of chromosome ends, had been a puzzle for many decades, ever since the pioneering work of Hermann Muller¹ and Barbara McClintock² in the 1930s. Muller used X-rays to create breaks in DNA, while McClintock used cytogenetic tricks to break chromosomes. But both came to the same conclusion, which is that the ends of broken chromosomes are very reactive and do things that normal chromosome ends never do. This is dramatically illustrated by the famous breakage-fusion-bridge cycle explored by McClintock (Figure 1). The basic observation is that the replication of a chromosome with a broken end results in two ends that can join together, generating a chromosome with two centromeres. When those centromeres are pulled towards opposite poles of the spindle during cell division, the chromosome is broken again, regenerating chromosomes with broken ends. This results in continuing cycles of fusion and breakage, a consequence of which is the formation of cells that have lost important parts of chromosomes. Not surprisingly many dead cells are generated in this process.

The contributions of my laboratory to our understanding of telomere function and maintenance by telomerase were made over a limited period of time early in the development of this story, from 1980 to 1989. What I would like to discuss here are some of the problems that we had to overcome, especially the preconceptions we had about models for telomere function and how hard it was to let go of those models. Fortunately the evidence we uncovered was strong enough to bring us to the right conclusions! Then, since I left the telomere field fairly early on, I would like to take this opportunity to briefly review some of the work that we've done since, primarily to show students who are just entering science that it is not only possible but really fun to address very different questions in different fields during one's career.

There were two well-known and long-standing puzzles associated with the nature of eukaryotic chromosome ends, or telomeres: the problem of the stability of the ends of chromosomes, and the problem of complete replication. My first introduction to these issues came when I was an undergraduate student at McGill University in Montreal. The first of those two problems, the reactivity of chromosome ends, had been a puzzle for many decades, ever since the pioneering work of Herman Müller¹ and Barbara McClintock² in the 1930s. Müller used X-rays to create breaks in DNA, while McClintock used cytogenetic tricks to break chromosomes. But both came to the same conclusion, which is that the ends of broken chromosomes are very reactive and do things that normal chromosome ends never do. This is dramatically illustrated by the famous breakage-fusion-bridge cycle explored by McClintock (Figure 1). The basic observation is that the replication of a chromosome with a broken end results in two ends that can join together, generating a chromosome with two centromeres. When those centromeres are pulled towards opposite poles of the spindle during cell division, the chromosome is broken again, regenerating chromosomes with broken ends. This results in continuing cycles of fusion and breakage, a consequence of which is the formation of cells that have lost important parts of chromosomes. Not surprisingly many dead cells are generated in this process. Normal chromosomes never do this, so it was clear that there was something very special and different going on at the ends of normal chromosomes that prevents end-to-end joining. But at the time of this work it wasn't even known that DNA was the genetic material, so they had no way to think in molecular terms about what was going on.

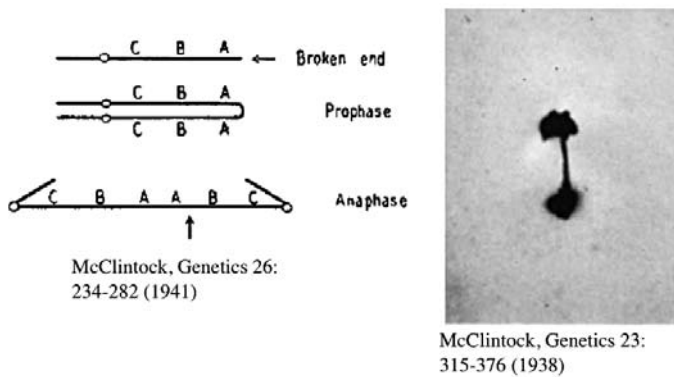


Figure 1. The chromosomal breakage-fusion-bridge cycle explored by Barbara McClintock. Left: After the replication of a broken chromosome, the two broken ends join together, creating a dicentric chromosome. When the two centromeres are pulled to opposite poles of the dividing cell, the chromosome breaks, and the new broken chromosomes continue the cycle. Right: micrograph of a dicentric chromosome bridging the two poles of a mitotic spindle.

Much later on, long after it was recognized that DNA was the genetic material in chromosomes, an additional problem was discussed by Watson³ and by Olovnikov⁴, who recognized that the replication of the very ends of DNA molecules posed a special problem (Figure 2). When a replication fork heads towards the end of the chromosome, the leading strand can go all the way to the end, but the lagging strand cannot since it is generated by the extension of an RNA primer by DNA polymerase. If this RNA primer is generated at an internal site, any distal DNA will remain unreplicated; even if the RNA primer was made at the very end, after the RNA primer is degraded, a short region of unreplicated DNA would remain. In the absence of some compensatory mechanism, the ends should get shorter and shorter, and since that doesn't happen, there must be some unknown process to counter-balance the necessarily incomplete replication.

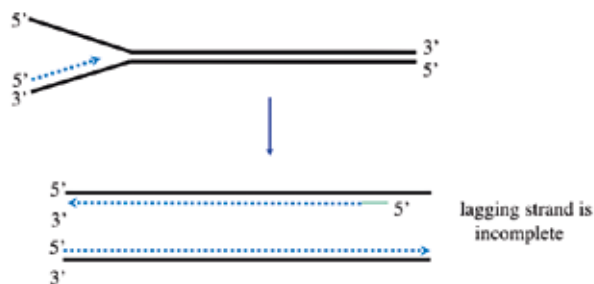


Figure 2. The end-replication problem as posed by Watson (³) and by Olovnikov (⁴). When a replication fork reaches the end of a chromosome, the lagging strand will necessarily be incomplete as a result of the removal and potentially internal location of the last primer generated by primase.

Although I learned about these problems as a student, I can't say that they made a very big impression on me and I didn't really think about them very much until years later, when I began working on the molecular reactions engaged in by broken pieces of DNA. This was work that I started as a post-doc at Cornell with Ray Wu, working in collaboration with my friend and colleague, Dr. Rodney Rothstein. Means for introducing DNA molecules into yeast cells, a process referred to as yeast transformation, had just been discovered down the road from our lab at Cornell in Gerry Fink's lab⁵. The ability to do this opened up a huge number of interesting experiments. Rod and I started to examine some variations on the initial procedure, such as cutting the circular DNA molecules before putting them into yeast. Shortly thereafter, when I moved to Boston and was setting up my lab at the Sidney Farber Cancer Institute, we continued this collaboration with the additional participation of my first graduate student, Terry Orr-Weaver^{6,7}.

In the course of our experiments on transformation and recombination, we observed a process that is analogous to the fusion events studied by McClintock in maize decades earlier (Figure 3). We began with a circular DNA molecule that was able to replicate as a circular DNA plasmid in yeast because it contained a yeast origin of DNA replication^{8,9}. Intact circular DNA of that plasmid yielded a high frequency of yeast transformants, because chromosomal integration was not required for plasmid maintenance. When we made a cut in the DNA with a restriction enzyme, in a region of the DNA that is not found in any yeast chromosome, we recovered many fewer transformants. When we analyzed the few transformants that we did recover, the cut DNA ends had been joined back together, presumably by the action of the enzyme DNA ligase⁷. In many cases some DNA was lost as the ends were chewed back by exonucleases before being joined together by ligase. As with McClintock's much earlier results, these DNA reactions are very different from anything that would happen at the ends of natural chromosomes.

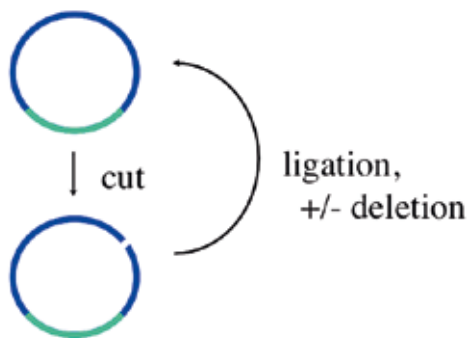


Figure 3. Non-homologous end-joining in yeast. A circular plasmid, cut with a restriction enzyme in a region of DNA that is not homologous to any yeast chromosomal DNA adjacent to the cut site may be degraded prior to ligation of the ends.

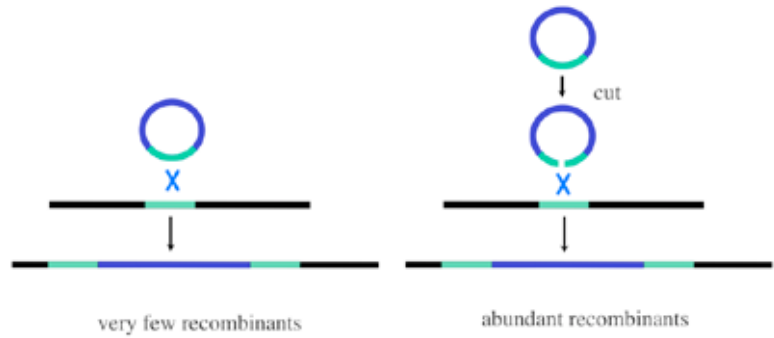


Figure 4. Double-strand breaks in DNA stimulate recombination. Intact circular DNA lacking a replication origin yields few transformants, because recombination events leading to chromosomal integration are rare. The same plasmid, when cut within a region of homology to a yeast chromosome, yields many more integrated transformants.

Terry, Rod and I actually spent most of our effort looking at what happened when we made cuts in regions of plasmid DNA that were homologous to a segment of a yeast chromosome (Figure 4). When a circular DNA molecule containing a region of homology with a chromosome is used to transform yeast, the occasional recombination event will occur, resulting in the plasmid becoming integrated into the yeast chromosome. This was the pathway found in the Fink lab in their early studies of transformation. What Terry, Rod and I found was that cutting the DNA in this region of homology led to a greatly increased frequency of such recombination events⁶. We continued to follow this up by studying the reactions that broken DNA ends undergo (Figure 5). If a DNA molecule is broken by cutting with a restriction enzyme, then in the cell the ends can be chewed back by nucleases, and exonucleases can generate single-stranded ends that can invade a homologous sequence. Strand invasion allows repair synthesis to begin using DNA polymerases, and Holliday junctions can be formed which can branch migrate. After repair synthesis, the Holliday junctions can be resolved by special enzymes called resolvases, to yield crossover or non-crossover configurations. This work eventually led us to propose, along with Frank Stahl, that cells entering meiosis engage in the programmed breakage of their chromosomal DNA as a means of initiating meiotic recombination by double-strand-break repair¹⁰. So broken DNA ends do a lot of things, but they are all things that don't happen with normal chromosome ends. I mention them here because these are the reactions I was thinking of before I entered the telomere field.

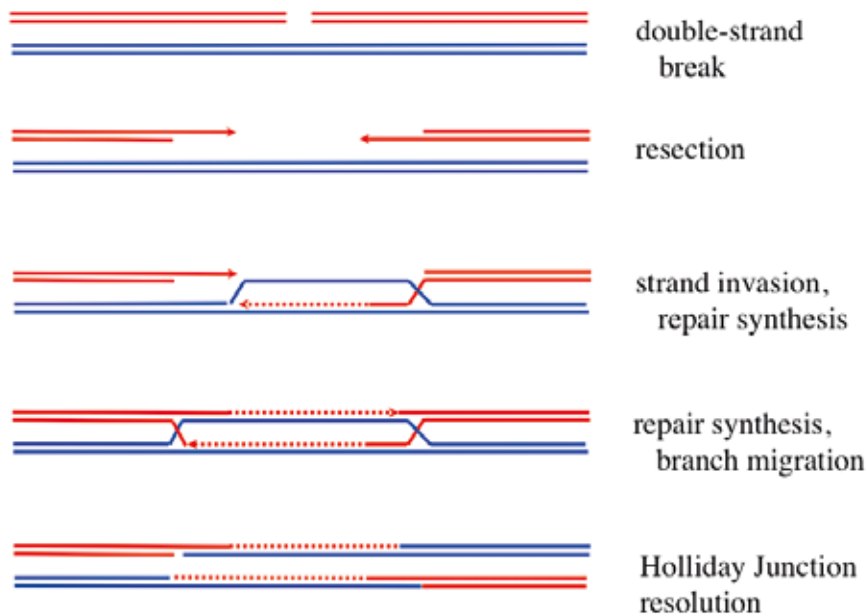


Figure 5. The double-strand break repair model for recombination. Two homologous chromosomes (red and blue) recombine when one is broken. The initial cut is further processed by nucleases, exposing single-stranded DNA, which invades the homologous duplex. Repair synthesis and branch migration generate Holliday junctions, the resolution of which generates recombinant DNA products.

In the summer of 1980, I attended the Nucleic Acids Gordon Conference and heard, for the first time, Elizabeth Blackburn talk about her amazing work on the stable DNA ends from *Tetrahymena thermophila*¹¹. This unicellular organism is very divergent from metazoans, and has an unusual cell biology characterized by the presence of both a micronucleus with normal chromosomes and a macronucleus in which the chromosomal DNA has been chopped into thousands of small fragments, many of which become highly amplified. Liz talked about the very simple repetitive sequences, just stretches of a GGGGTT repeats, that she had found at the ends of these very abundant short DNA molecules in the large macronucleus of *Tetrahymena* (Figure 6). It was incredibly striking that these little pieces of DNA were stable ends, and were apparently fully replicable, i.e. they seemed to behave just like normal chromosomal telomeres. They clearly behaved completely differently from the DNA ends that we were studying in my lab, in yeast cells. After Liz's talk I sought her out to discuss these experiments, and we realized that there was a really simple and potentially very interesting experiment that we could do to see if the telomeric ends from *Tetrahymena* would work as stable telomeric ends in yeast cells. Neither of us thought that the experiment was very likely to work, because *Tetrahymena* and yeast are so very distantly related. On the other hand, we had all the necessary bits and pieces and technically the experiment was quite trivial, so we decided to go ahead. Liz sent me some DNA that she had painstakingly purified from *Tetrahymena*,

and I took this little restriction fragment from the end of the ribosomal DNA of *Tetrahymena* and put it into yeast to see how it would function.

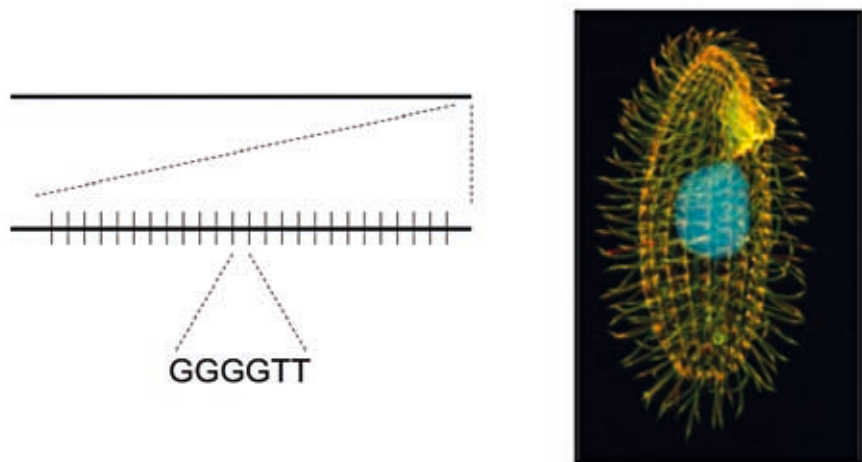


Figure 6. Telomeres from *Tetrahymena*. Left: DNA from the macronuclear fragments end in a series of tandem repeats of the hexanucleotide GGGGTT. These DNA ends are stable and fully replicated. Right: Image of *Tetrahymena*, showing the large macronucleus (blue).

There is an amazing aspect of this piece of DNA from *Tetrahymena* that I would like to comment on before describing the yeast experiment (Figure 7). Right next door to the telomere sequence, just a couple of kilobases in, is the primary ribosomal RNA transcript of *Tetrahymena*. In that transcript there is a little intron, just over 400 bases long, and that intron is the first self-splicing intron ever discovered¹², in the work for which Tom Cech was awarded the Nobel Prize in chemistry in 1989. A very nice piece of DNA indeed!

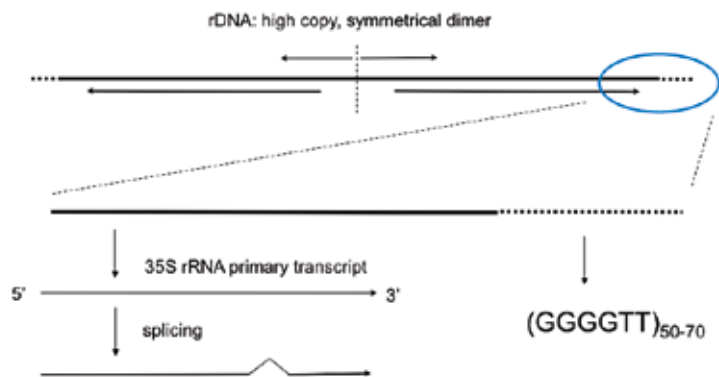


Figure 7. A very special piece of DNA. The *Tetrahymena* ribosomal DNA fragment from the macronucleus is a symmetrical dimer. The ends are telomeres and consist of GGGGTT repeats. Close to the ends is a region of the rRNA genes coding for a self-splicing intron.

Returning to the experimental test of *Tetrahymena* telomere function in yeast, what we really wanted to do was to test the idea that the biochemical machinery underlying telomere function might have been very highly conserved. If that turned out to be true, then the mechanisms that were being learned about in *Tetrahymena* might apply broadly to eukaryotic organisms, which would make the whole process much more significant. This was the motivation for the experiment that Liz Blackburn and I collaborated on. What we had available at that time, in my lab, were circular DNA plasmids containing yeast genes^{13,14} so that we could select for yeast transformants, i.e. cells that had taken up the DNA. These plasmids also contained origins of replication (known then as autonomous replication sequences or ARS elements^{8,9}) so that they could replicate independently of integration into the chromosome. When intact circular plasmid DNA of this type is used to transform yeast cells, many transformants are recovered and they almost all contain replicating circular DNA molecules. As I explained above, if the plasmid DNA is cut with a restriction enzyme (in a region that is not homologous to yeast genomic DNA) so as to generate linear DNA with 'broken ends', those ends do not function as stable telomeric ends and as a result very few transformants are recovered.

The critical experiment was to take the little pieces of telomeric *Tetrahymena* DNA ending in G₄T₂ repeats, and ligate them onto each end of the linearized plasmid DNA (Figure 8). I carefully purified the ligated DNA, put that into yeast, and recovered transformants. I was then able to ask whether the plasmid DNA was replicating as a linear molecule, which would mean the telomeres were working, or whether I had only recovered standard replicating circular plasmids. I distinguished between linear and circular DNA forms by preparing DNA from a dozen or so transformants, and analyzing the DNA by gel electrophoresis. When DNA molecules are separated by gel electrophoresis, circles generate a series of bands corresponding to monomers and multimers, and relaxed and supercoiled forms, leading to a complicated pattern. Linear DNA molecules don't have any of those alternative forms, so they migrate as a single band. The two possible results of the DNA analysis were therefore quite distinct. When I analyzed the DNA from the transformants that I had recovered, about half of them contained plasmid DNA that migrated as a single band on the gel. This was perhaps the most clear-cut experiment I have ever done. It was immediately obvious that the experiment had worked, and that the *Tetrahymena* ends were able to act as functional telomeres in yeast¹⁵. We therefore knew immediately that the underlying biochemical machinery must be very broadly conserved because these two organisms were so distantly related to each other. It also meant that we could now use all of the tools of yeast genetics and molecular biology to study telomeres in yeast.

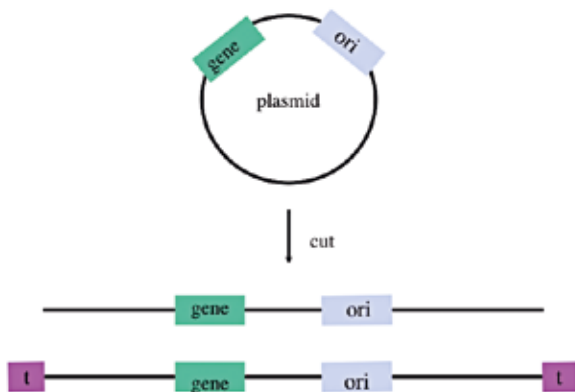


Figure 8. Moving *Tetrahymena* telomeres into yeast. A yeast plasmid vector containing selectable markers and an origin of replication was linearized by digestion with a restriction enzyme. *Tetrahymena* telomeres were ligated onto both ends, and the ligated DNA was purified and used to transform yeast cells. The resulting transformants contained replicating linear plasmids.

One of the first things that I wanted to do with the new linear plasmid with two *Tetrahymena* ends was to use it as a vector for cloning natural telomeres from the ends of yeast chromosomes. That experiment was extremely simple conceptually (Figure 9). I began with yeast chromosomal DNA, and cut it up with restriction enzymes into lots of pieces. Most of them were internal fragments, but the occasional fragment from the end of a chromosome would have one restriction cut end and one end derived from a yeast telomere. I then took our vector DNA, the linear plasmid with two *Tetrahymena* ends, cut off one end, and carefully purified the resulting DNA. This DNA molecule, which had one functional telomeric end and one non-functional ‘broken’ end, could not be maintained in yeast cells. The yeast telomere cloning experiment then simply involved joining the yeast DNA fragments and the purified vector DNA together using DNA ligase. Every now and then, this would result in a molecule with a *Tetrahymena* telomere at one end and a normal yeast telomere at the other end, and those rare molecules were expected to be able to replicate as linear molecules in yeast cells. I did recover some transformants with the expected linear structure¹⁵, and I was able to confirm through a variety of tests that one end was indeed a yeast telomeric DNA fragment. This allowed us to start looking at the structures found in normal yeast telomeres, including the DNA sequences characteristic of yeast telomeres. We didn’t expect the repeat sequences to be the same, since the *Tetrahymena* sequences didn’t cross-hybridize with yeast DNA. Other hybridization experiments, done in collaboration with Tom Petes¹⁶, showed that yeast telomeres contained stretches of alternating GT repeats. Still, when Janis Shampay, a graduate student in Liz’s lab, sequenced the yeast telomeres I had cloned, we were all a bit surprised to see a somewhat irregular sequence, summarized as G₁₋₃T repeats¹⁷. This was independently confirmed in the Tye and Petes laboratories based on the

cloning of telomeric ends by hybridization with (GT)_n probes¹⁸. While yeast did fit the general finding of a GT rich 3'-terminal strand, the absence of simple repeats was puzzling, and didn't seem to fit easily into the prevailing recombination-based models of telomere replication (19). It was the resolution of that puzzle that would eventually lead us to telomerase.

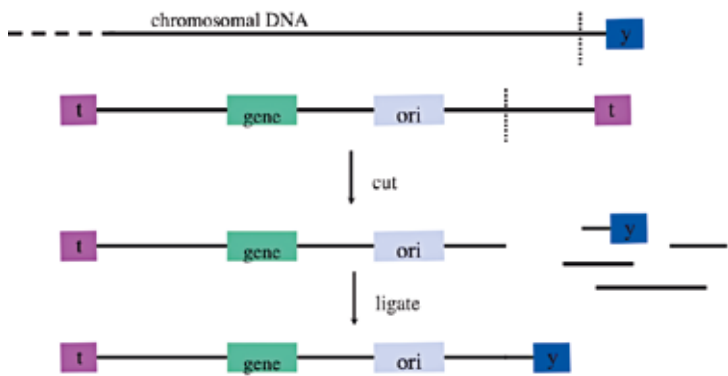


Figure 9. Cloning yeast telomeres. Yeast chromosomal DNA was digested with a restriction enzyme, as was the linear plasmid with two *Tetrahymena* telomeres. The purified vector fragment was ligated to the yeast DNA fragments, and the resulting mixture was used to transform yeast. A few linear plasmids were recovered, in which one end of the linear vector was replaced by a yeast telomere.

At this point, I would like to take a little digression to describe how we used these new telomeric DNA fragments as a tool to study the requirements for proper chromosome function in yeast. This work was done by Andrew Murray, my second graduate student. What we did was to take an engineering approach to seeing if we really understood the elements of chromosome structure. With telomeres in hand, we thought that we had all of the pieces that would be required to generate a fully functional chromosome. We had centromeric DNA, first cloned in John Carbon's lab²⁰; we had various genes such as LEU2 and HIS3^{13,14}, and we had origins of replication, first cloned by Kevin Struhl and Dan Stinchcomb in Ron Davis's lab^{8,9}. Those were all of the elements known at the time to be important in terms of chromosomal function. We thought that it would be interesting to put them all together and see if we could make something that behaved like a natural chromosome. To do this we constructed a circular plasmid that had all of the known chromosomal elements (Figure 10), linearized it so that it had two telomeric ends, and put it into yeast. Despite the fact that this DNA molecule had all the pieces (an origin of replication, a centromere, genes, and telomeres), when we put it into yeast it didn't behave at all like a proper chromosome. During mitosis it displayed a very high frequency of segregation errors, so that instead of being maintained over many cell cycles it was lost at a high frequency²¹. This was a very interesting result, because it said there was something going on that we didn't understand. What could

be missing? What were the potential problems that prevented accurate inheritance of this mini-chromosome? We tested many possible explanations. Eventually, Andrew figured out that what was missing was just more DNA^{21,22}. By simply adding enough non-yeast DNA from phage lambda to our small artificial chromosomes, he was able to make much bigger DNA molecules that now exhibited stable inheritance and behaved much more like natural yeast chromosomes (Figure 11). We considered various models for this, and based on the observation that the linear centromeric plasmid was much less mitotically stable than a similar circular centromeric plasmid, we proposed that the intertwining of DNA after the completion of DNA replication²³ played a role in holding sister chromatids together. This was long before the modern story of cohesin and separase²⁴ and the complex biochemistry that underlies the adherence and separation of sister-chromatids after replication. Our artificial chromosomes were also technically useful, at least for a little while, in the early days of genomic sequencing because it turns out that they are very nice vectors for cloning extremely large pieces of DNA, up to a megabase or two in length²⁵.

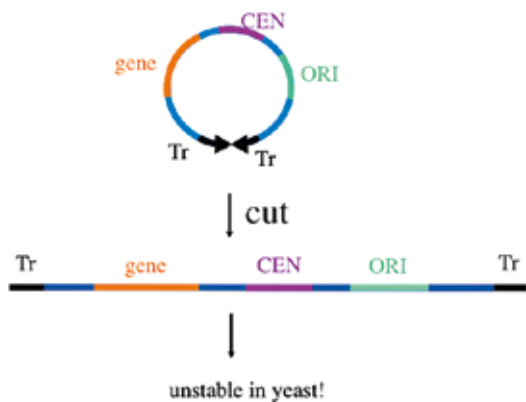


Figure 10. Our first attempt to make an artificial chromosome. We constructed a circular plasmid containing yeast genes, an origin of replication, a centromere, and telomeric DNA (Tr). This was linearized by cutting between the telomeric sequences, then introduced into yeast, where the DNA was maintained as a linear plasmid. Unexpectedly, this DNA molecule did not behave like a normal chromosome – it was mitotically unstable due to a high frequency of segregation errors.

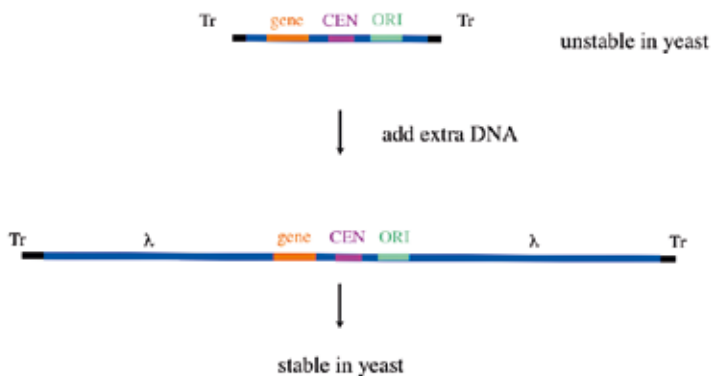


Figure 11. Successful construction of a yeast artificial chromosome. The addition of 50 to 150 kb of non-yeast DNA from phage λ greatly improved the mitotic stability of the DNA molecule, conferring improved chromosome-like behavior.

Returning once more to the story of telomeres and how they are fully replicated, all of our early models for thinking about this problem were based on recombination and the various kinds of reactions known to be engaged in by DNA ends. A very simple model that seemed quite attractive after Liz Blackburn's discovery of the short repetitive sequences of *Tetrahymena* telomeres was that recombination between different ends, perhaps biased in some way, could generate ends that were longer than either of the input DNA ends (Figure 12)²⁶. Alternatively, strand-invasion by the 3' end of one telomere into the repeats of another telomere could lead to repair synthesis which would result in elongation of that end (Figure 12). Another model that we considered invoked Holliday junction resolution. This model was based on idea that the very end of telomeric DNA was actually a hairpin, i.e. the strand loops around at the end. That was attractive because it meant that there was no actual DNA end, and a hairpin could act as a relatively inert DNA terminus. Replication would generate an inverted repeat structure, which could isomerize into a central Holliday junction, resolution of which by the corresponding recombination enzyme would generate two new hairpin terminated telomeres (Figure 13). A more complex variant of this model that originated in Piet Borst's lab²⁷ was that internal nicks within the repeats were sites of unpairing followed by gap-filling synthesis, leading to synthesis of new repeat units. These were the kinds of recombination based models that we discussed in the early years of thinking about telomere replication. How did we finally let go of these models and come to the correct explanation? Remarkably, we were driven to the answer by analyzing the sequences of *Tetrahymena* telomeres after their replication in yeast.

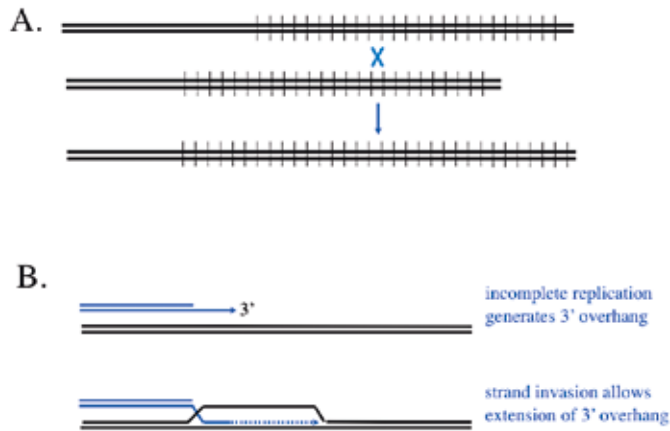


Figure 12. A. Telomere lengthening by recombination. B. Telomere lengthening by repair synthesis.

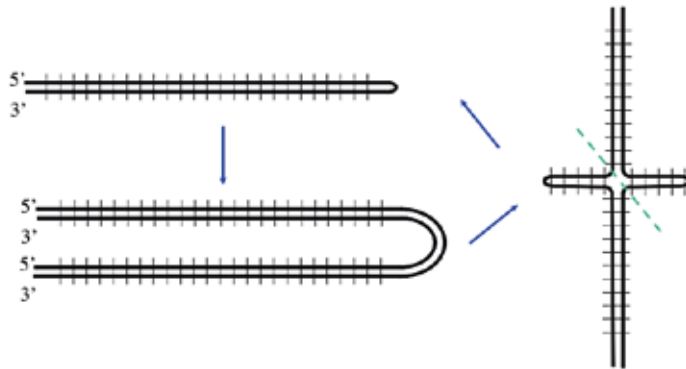


Figure 13. Telomere replication by Holliday junction resolution. Early models of telomeric DNA proposed a hairpin terminus. Replication would generate an inverted repeat, which could isomerize to form a Holliday junction, resolution of which would regenerate the original structures.

To understand why the replication of *Tetrahymena* telomeres in yeast was so important, consider again the linear plasmid with *Tetrahymena* ends. Those telomeric ends began as a restriction fragment of a certain size, but we noticed that after their maintenance in yeast that they had grown longer, by as much as a few hundred base-pairs, as well as becoming heterogeneous in size. We didn't know where this extra DNA had come from, but there were several possible explanations. It could have been, for example, a result of recombination between *Tetrahymena* ends on different molecules, or a result of strand-invasion and repair synthesis. Eventually, we cloned some of these lengthened *Tetrahymena* ends and, in a continuation of the collaboration with Liz, sent those DNA samples to Liz's lab where once again Janice Shampay

did the actual sequencing. To our complete shock, we found that the actual structure consisted of G_4T_2 repeats from the *Tetrahymena* ends joined directly to the irregular $G_{1-3}T$ repeats that were characteristic of yeast telomeres (Figure 14)¹⁷. Thus the reason the DNA had become longer was that the yeast-specific sequence had become appended to the *Tetrahymena* ends. This new DNA seemed to have just dropped out of the sky. Such a different and irregular sequence couldn't possibly have been generated by any recombinational process, so we immediately knew that all of our early models were wrong. The new sequencing data led directly to the idea that there must be a specific new enzyme that adds extra DNA to chromosomal ends. Shortly after these results and our prediction of this new enzyme, of course, Carol Greider went on to identify the predicted enzyme activity biochemically²⁹. Characterization of the purified enzyme, later named telomerase, showed that it is a ribonucleoprotein enzyme that contains an RNA template that specifies the telomeric repeat sequences, which are synthesized by a reverse transcriptase component of the enzyme³⁰. We now know that the different telomeric repeats found in different organisms are specified by the RNA templates of their particular telomerase enzymes. A great deal of work has been done to characterize telomerase in many organisms, including *Tetrahymena*, yeast and humans, by Elizabeth Blackburn, Carol Greider, Tom Cech and many other people.

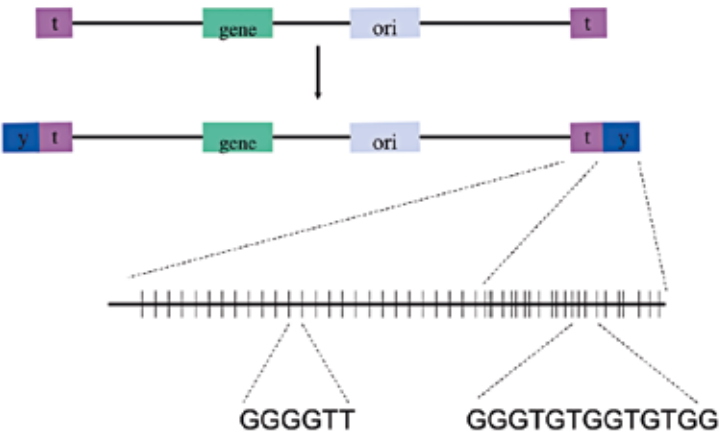


Figure 14. Yeast adds new DNA to *Tetrahymena* telomeres. Cloning and sequence analysis of *Tetrahymena* telomeres after replication in yeast (as the telomeres of a linear plasmid) revealed the addition of yeast telomeric sequences.

It is interesting to revisit the end replication problem in light of the activity of telomerase. As mentioned above, one of our early models was that the actual end was a hairpin structure. Of course that also turned out to be wrong, and the proper structure is a 3'-end overhang consisting of GT-rich repeats (Figure 15). This was originally worked out in a different ciliated protozoan, *Oxytricha*, in the lab of David Prescott²⁸, and then found to be

a universally conserved aspect of telomere structure. If we consider the replication of DNA with a 3' overhang, the end-replication problem is actually a little bit different from that noted earlier by Watson³ and by Olovnikov⁴. A replication fork heading towards this kind of end retains the previously noted problem of incomplete replication of the 3' end strand, but a much worse problem in that the leading strand can go to the end, but can't regenerate a 3' overhang. The 3' overhang will therefore be lost in every cycle of replication, unless there is a compensatory process. This, of course, is the role of the telomerase enzyme, which adds extra repeats to telomeric ends and thereby on average maintains the proper telomeric length and structure. The regulation of proper telomere length and structure has turned out to be quite elaborate, and the biochemistry of the corresponding protein-DNA interactions is remarkably complex and interesting³¹.

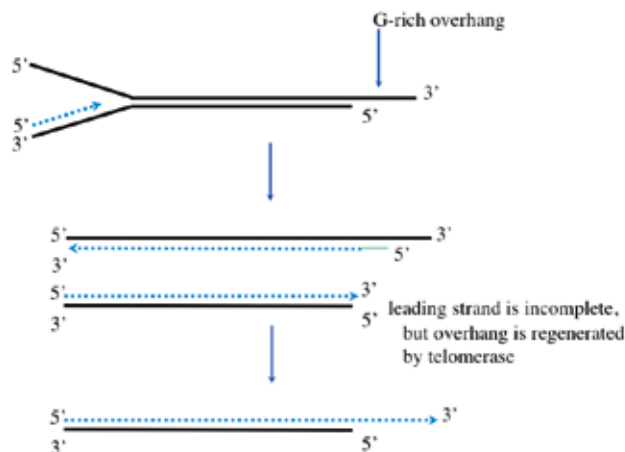


Figure 15. New model for telomere shortening, and the role of telomerase in telomere maintenance. When a replication fork reaches the end of a DNA duplex, the leading strand cannot regenerate the 3'-overhang. This is done by telomerase.

The activity of telomerase and its associated regulatory machinery in controlling telomere length turns out to have important biological consequences. Cells with high levels of telomerase activity can divide without limit, because they maintain functional telomeres. In contrast, cells with insufficient telomerase activity cannot maintain telomere length, and as a result have limited division potential. This prediction was initially verified by Vicki Lundblad, who came to my lab as a postdoc and decided to address this issue genetically in yeast³². What Vicki did was to set up a large and actually quite difficult screen for mutants that would be unable to maintain telomeres at their proper average length. She was able to recover mutants that had the property we were looking for, namely that telomeres would get shorter and shorter over an increasing number of cell divisions (Figure 16, part A). The first mutation with that property was named *est-1*, for 'ever shorter telomeres'. The most interesting property of this mutation (and similar mutations

recovered later) is that it confers a delayed senescence phenotype, just as predicted (Figure 16, part B). This phenotype is visually apparent in colonies of the mutant strain of yeast that have been grown for different numbers of generations. After 25 generations, the mutant colonies look just like wild-type colonies. After 46 generations the colonies are a little more irregular, and there are some small colonies; by 60 to 70 generations they are quite small and irregular, and after 80 to 90 generations the mutant strain can hardly grow at all. There are many dead cells in the small colonies, and there is a very high level of chromosome loss. Because the telomeres are getting shorter and shorter, eventually proper telomeric structure isn't maintained. As a consequence, ends are getting joined together leading to chromosome breakage and loss, so that cells are generated that are missing big chunks of their DNA. This was the first experimental demonstration that an inability to maintain normal telomere length would lead to a senescence phenotype, and therefore this inability to maintain telomeres might have an important role in problems of cellular senescence in higher organisms. At about the same time very similar experiments were done in Liz's lab, using *Tetrahymena*, and led to the same conclusion³³. We thought this was a potential explanation for the senescence seen during repeated passage of primary cells in tissue culture, and by extension perhaps to problems of aging related to a gradual decline in tissue renewal, perhaps due to limited cell division potential. The shortening of telomeres during passage of fibroblasts was soon demonstrated by Carol Greider³⁴, and the causal role of this shortening in cellular senescence was later proven³⁵. Of course, this has turned out to be a very important aspect of our growing understanding of ageing and age-related diseases³⁶. The complementary aspect of this has turned out to be very important for our understanding of cancer. In the vast majority of cancer cells, which have unlimited division potential, the telomerase gene has been up-regulated and functional telomeres are maintained indefinitely^{36,37}.

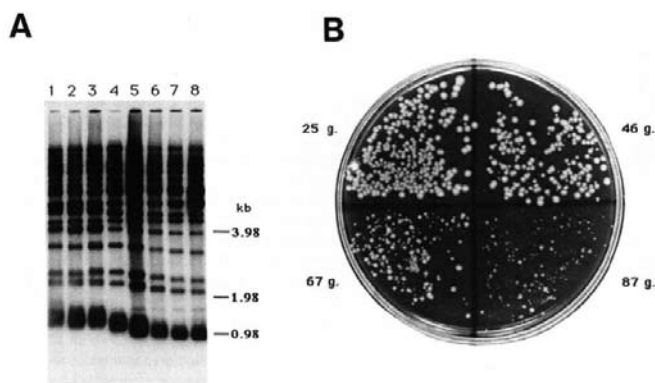


Figure 16. Senescence of yeast EST-1 cells. A: Telomeric yeast DNA fragments from an EST-1 mutant strain are visualized by Southern blotting. Lanes 1 through 8 represent increasing numbers of generations of growth. B: A mutant EST-1 strain streaked out on an agar plate after 25, 46, 67 and 87 generations of prior growth.

At that point in my career it became clear that many people would soon be exploring the roles of telomeres and telomerase in cancer and aging. I felt that the main questions were clear, and that they would be addressed whether or not I remained active in the field of telomere biology. I therefore began to look for other interesting questions that could be addressed experimentally, but where there were not too many people trying to look at the same issues.

Even as Vicki was doing her genetic work on telomere maintenance in yeast, I was already becoming interested in ribozymes, because Tom Cech's discovery of the self-splicing introns was very new and exciting¹². I thought there were many interesting questions, and I was surprised that more people weren't entering that field. In particular, I was attracted by the RNA world hypothesis³⁸ and the idea that RNA might be able to catalyze its own replication without protein enzymes. Since the experiments were largely molecular biology in nature, I thought that we might be able to make some contributions to that nascent field. For several years we studied the group I introns and tried to use various molecular techniques to force them to catalyze RNA replication reactions. Several of my students including Jennifer Doudna³⁹ and Rachel Green⁴⁰ worked on that problem, with some success. But eventually we came to the conclusion that the ribozymes available from nature were not good enough. Those ribozymes were doing jobs that they had evolved to do in modern organisms, and what we were primarily interested in were questions about what RNA could have done much earlier.

In the late 1980s we started to think about ways of evolving new RNA molecules that would do things that we were interested in. The basic idea was simple: prepare huge collections of random sequences, and then isolate the rare functional molecules that did what we wanted. The technology for doing this *in vitro* selection, or directed evolution, was worked out by Andy Ellington when he was a postdoc in my lab⁴¹, and independently by Craig Tuerk in Larry Gold's lab⁴². We spent most of the 90s applying this kind of selection technology to the laboratory evolution of RNA and DNA molecules that could do all kinds of interesting things. For example, an RNA molecule isolated by Mandana Sassanfar when she was a postdoc in the lab folds up into a three-dimensional shape that contains a binding site for ATP (Figure 17)⁴³. Subsequently, we and others were able to show that it is possible to evolve, in the laboratory, RNA and DNA sequences that will fold into defined shapes that can bind almost any target molecule of interest. Ongoing studies in several different labs and companies are aimed at exploring potential therapeutic uses of these target binding RNA molecules, known as aptamers, perhaps doing some of the things that we use antibodies to do today.

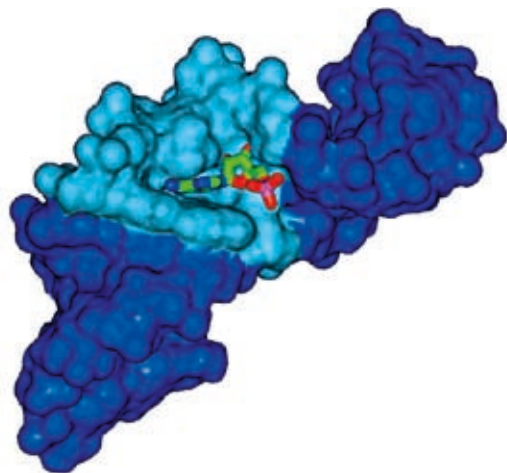


Figure 17. An ATP binding RNA molecule. This RNA was evolved from an initially random population of sequences. Dark blue: double-helical regions; light blue: folded recognition loop; stick figure presents bound AMP.

Once we were able to evolve aptamers routinely we turned our attention to evolving RNA molecules that could catalyze interesting reactions. Dave Bartel, when he was a graduate student in the lab, isolated a surprisingly intricate RNA molecule that catalyzes a joining reaction between two adjacent RNAs aligned on a template (Figure 18)⁴⁴. It uses the same chemistry that RNA and DNA polymerases use, i.e. the 3'-prime hydroxyl of one RNA substrate attacks the α -phosphate of the triphosphate of the other RNA substrate, generating a new phosphodiester bond. The ribozyme has an intricate folded secondary⁴⁵ and three-dimensional structure⁴⁶. This was a very exciting demonstration that RNA could catalyze the chemistry of RNA replication. Subsequently, in his own lab at the Whitehead Institute at MIT, Dave Bartel further evolved this ribozyme into an actual RNA polymerase that can copy RNA templates using nucleoside triphosphates as substrates⁴⁷. This is a marvelous 'proof-of-principle' of the plausibility of the RNA world hypothesis. Unfortunately the current versions of this RNA polymerase are not yet good enough to copy themselves and exhibit full cycles of replication, so there is plenty of scope for additional evolutionary optimization.

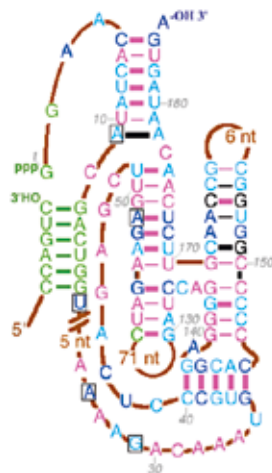


Figure 18. Secondary structure of the class I ribozyme ligase. This ribozyme catalyzes template-directed RNA-RNA ligation. It was evolved from an initially random population of RNA sequences.

More recently we have applied RNA *in vitro* selection to the analysis of human genomic sequences, in work done by Kouresh Salahi-Ashtiani and Andrej Luptak when they were postdocs in my lab (Figure 19)⁴⁸. Kourosh began this project by generating a large library of pieces of human DNA. He then transcribed them into RNA, and selected for molecules that could cut themselves at a unique site. He recovered four distinct self-cleaving RNAs or ribozymes. One of these is found in the CPEB3 gene, which has been implicated in memory⁴⁹, possibly through a role in controlling localized protein translation at synapses. There are two interesting things about this self-cleaving human genomic ribozyme. One is that it turns out to have exactly the same structure as a well known viral ribozyme, the HDV ribozyme of the hepatitis delta virus. The fact that there is a version of this ribozyme in the human genome suggests that the viral ribozyme may be derived from the genomic copy. Another potentially very interesting observation is that there is a polymorphism in the human population at a position within this ribozyme that affects its activity. A recent genetic study done by a group in Switzerland⁵⁰ has found an association between this polymorphism and performance on a word-recall memory test. A lot more work needs to be done on this, but the possibility that a self-cleaving catalytic RNA may play a role in human memory is fascinating.

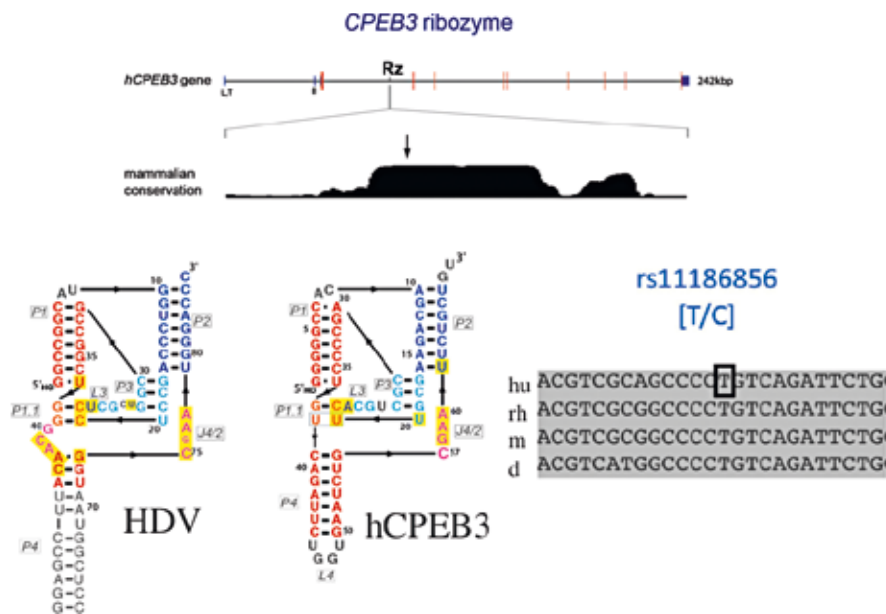


Figure 19. An HDV ribozyme in the human genome. Top: The self-cleaving ribozyme is located within the second intron of the CPEB3 gene. The ribozyme sequence is highly conserved relative to flanking intron sequences. Bottom left: The secondary structures of the human genomic ribozyme and the HDV ribozyme are virtually identical. Bottom right: A polymorphism with the ribozyme sequence affects ribozyme activity, and may affect human memory.

In the 1990s we extended our work on RNA and DNA directed evolution by developing methods for evolving proteins. Rich Roberts developed a clever means of tricking the ribosome into chemically linking a nascent peptide or protein chain to its own mRNA⁵¹, so that selection for a functional protein would also enrich the corresponding coding mRNA. This approach was used in later work done by Tony Keefe, who isolated a small ATP-binding protein from a library of completely random protein sequences⁵². This little protein domain looks indistinguishable from any natural biological protein domain. These kinds of laboratory evolution experiments showed that it is relatively easy to evolve functional RNAs, DNAs, and even proteins out of completely random collections of sequences.

The above experiments showed very directly that Darwinian evolution, applied to populations of molecules, is a powerful means of generating functional sequences. That led us to deeper questions: how did evolution get started? How did the transition from chemistry to Darwinian evolution first happen on the early earth? These are the central questions concerning the origin of life, and addressing these questions has become the main focus of my laboratory. The approach that we are taking is essentially a synthetic or engineering approach. We have a simple model for what we think early cells might have looked like (Figure 20)⁵³. This is not by any means a universally accepted model, but it is our view of what a very primitive cell might have looked like, and we are trying to construct such systems in order to define

possible pathways from chemistry to biology. We think that a primitive cell would have two critical components, the first of which is a cell membrane. In our experiments we make these membranes out of simple molecules that might have been around on the early earth, such as fatty acids. The cell membrane has to be able to grow spontaneously and divide to make daughter cells. The other important component of a primitive cell would be a polymer that could mediate the inheritance of genetic information. Here the big question is whether this could be RNA itself, or is it more likely to be some simpler progenitor material that was subsequently replaced by RNA? In either case, this material has to be able to replicate spontaneously without any of the highly sophisticated evolved machinery that is used by modern biology. The key question is therefore: how could both cell membranes and early genetic materials replicate prior to the evolution of complex biological machinery? The approach that we are taking is to try to divide this big problem up into simpler pieces that can be addressed separately. I will briefly describe a few of the experiments that we have done in the last six or seven years.

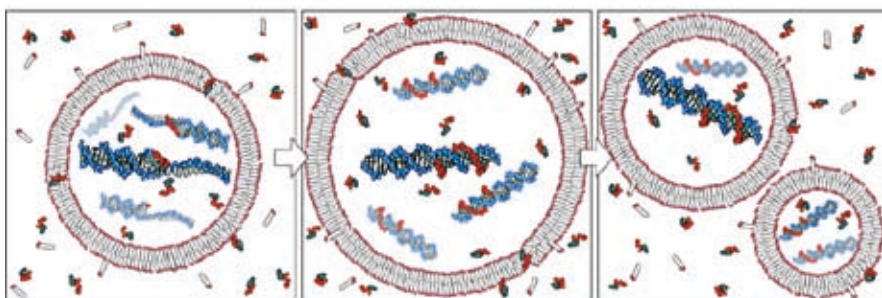


Figure 20. Schematic model of a protocell. A simple cell might be based on a replicating vesicle for compartmentalization, and a replicating genome to encode heritable information. A complex environment provides nucleotides, lipids and various sources of energy. Mechanical energy (for division), chemical energy (for nucleotide activation), phase transfer and osmotic gradient energy (for growth) may be used by the system.

About six years ago Marty Hanczyc, a postdoc, and Shelly Fujikawa, a graduate student in the lab, became interested in how protocell-like assemblies could be formed. They found that a common clay mineral, formed from volcanic ash and seawater, can facilitate this assembly process in a surprising way⁵⁴. This clay mineral is well known in the prebiotic chemistry community because it had been shown several years previously by Jim Ferris and Leslie Orgel to catalyze the assembly of RNA from activated nucleotides⁵⁵. Marty and Shelly showed that the same mineral could catalyze the assembly of membranes. Moreover, it can bring genetic polymers, such as RNA, into the vesicles it helps to assemble (Figure 21). Thus a common mineral can help to make genetic materials, help to assemble membranes, and bring them together⁵⁴, all of which is very attractive in terms of the assembly of early cellular structures.

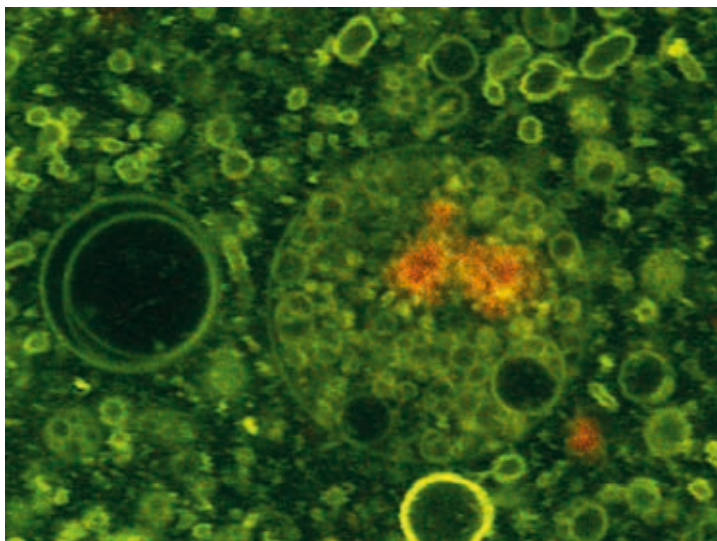


Figure 21. Montmorillonite can bring RNA into vesicles. Fluorescently labeled RNA (orange) on the surface of a clay particle is trapped inside a large vesicle (green) along with numerous small vesicles, all assembled as a result of the catalytic activity of the clay particle.

The replication of protocell-like structures is much more difficult than their assembly. However, the growth and division of the protocell membrane, which looked like an almost impossible problem just a few years ago, has actually turned out to be relatively simple. Our current model for what an early cell cycle might have looked like with respect to the cell membrane is based on the work of Ting Zhu, a graduate student in the lab⁵⁶. We prepare large multilamellar vesicles, and feed them with new fatty acids. Remarkably, they grow into long filaments, which are quite fragile; in response to gentle agitation, such as might result from waves on a pond, they break up into daughter cells (Figure 22). That generates a robust cycle that can be carried out indefinitely. Thus, the spontaneous growth and division of membrane compartments appears to be a relatively straightforward process.

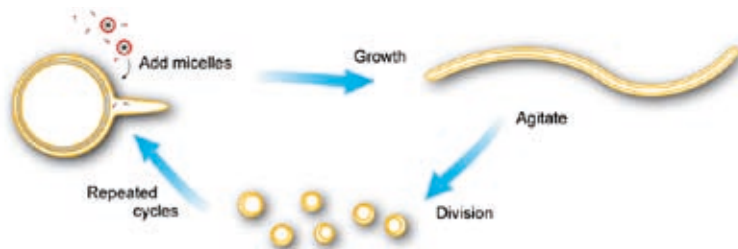


Figure 22. Cycles of growth and division of a model protocell membrane. Large multilamellar vesicles grow into long hollow vesicles following the addition of excess fatty acids. The filamentous vesicles are fragile and fragment in response to mild shear stress. The smaller daughter vesicles can grow and repeat the cycle.

What about the replication of genetic information? At the moment, this still seems to be difficult, because we don't understand how to accomplish this step. The RNA world hypothesis is based on the idea of RNA catalyzing its own replication³⁸, but that has turned out to be a harder problem than we thought. Could genetic replication have begun as a chemical, i.e. non-enzymatic, process? Almost twenty years ago, Leslie Orgel, one of the giants of prebiotic chemistry, proposed that chemical means of replicating genetic polymers should be found fairly easily by chemists, and that the solution to that problem would be relevant to the origin of life⁵⁷. That hasn't happened, perhaps because it's a harder problem than anybody thought, but also perhaps because there are not that many people working on this problem. I think that makes it a perfect problem to tackle because it is important, interesting and there are many reasonable experimental approaches. What we are doing is making synthetic nucleotides that are modified so as to become more reactive (Figure 23). For example, changing the hydroxyl nucleophile to an amine results in nucleotides that spontaneously extend a primer in a template-directed manner, without any enzyme⁵⁸. We do not yet have a robust and general replication system, but that is our goal.

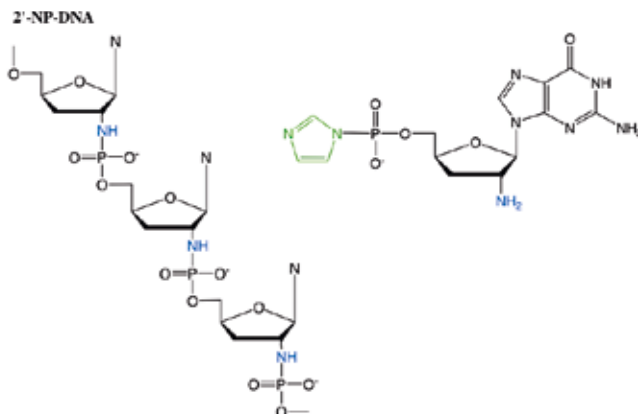


Figure 23. Typical monomer for spontaneous nucleic acid synthesis and the corresponding polymer. Left: 2'-5' linked phosphoramidate DNA. Right: The activated 2'-amino monomer. Note the 2'-amino nucleophile (blue) and the imidazole leaving group (green) on the 5'-phosphate. The combination of a good nucleophile with a good leaving group allows for rapid non-enzymatic polymerization of this class of monomer when aligned on an appropriate template.

There is an interesting aspect of the problem of chemical replication that we have just recently started to think about, namely, how can the very ends of our sequences be copied in the absence of telomerase? This turns out to be very interesting. Chemical replication results in spontaneous template-directed primer-extension, but once the end of the template is reached, the reaction slows down but often doesn't stop entirely. Depending on the conditions, we sometimes see chemical extension beyond the end of the template, generating a 3' overhang (Figure 24)⁵⁸. Thus complete replication of a template does not seem to be a problem, and in fact this process generates new sequences. It is interesting to speculate that this spontaneous chemical reaction might have something to do with eventual emergence of genetically encoded catalysts that would control and exploit this process, eventually leading to the telomerase enzyme that has been the main subject of my lecture.

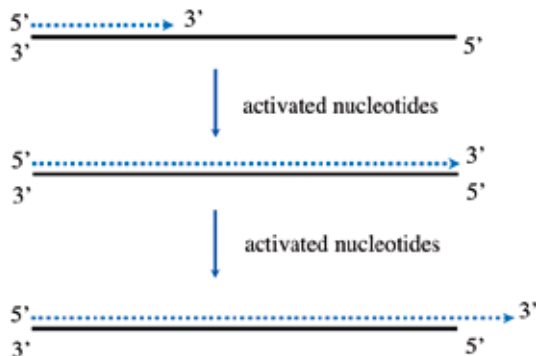


Figure 24. Origin of telomerase in spontaneous copying chemistry? Under certain conditions non-enzymatic primer-extension proceeds past the end of the template, generating a 3' overhang. Enzymatic control and elaboration of this chemical process could provide an evolutionary path towards telomerase.

ACKNOWLEDGEMENTS

I would like to thank all of the many brilliant students, postdocs, friends, and collaborators who contributed to this work.

REFERENCES

1. Muller, H.J., "The remaking of chromosomes," *Collecting Net* 13, 181–198 (1938).
2. McClintock, B., "Cytological observations of deficiencies involving known genes, translocations and an inversion in *Zea mays*," *Missouri Agr. Exp. Sta. Res. Bull.* 163, 4 (1931).
3. Watson, J.D., "Origin of concatameric T7 DNA," *Nature New Biol.* 239, 197–201 (1972).
4. Olovnikov, A.M., "A theory of marginotomy," *J. Theor. Biol.* 41, 181–190 (1973).
5. Hinnen, A., Hicks, J.B. and Fink, G.R., "Transformation of yeast," *Proc. Natl. Acad. Sci. USA* 75, 1929–1933 (1978).
6. Orr-Weaver, T.L., Szostak, J.W. and Rothstein, R.J., "Yeast transformation: A model system for the study of recombination," *Proc. Natl. Acad. Sci. USA* 78, 6354–6358 (1981).
7. Orr-Weaver, T.L. and Szostak, J.W., "Yeast recombination: The association between double-strand-gap repair and crossing-over," *Proc. Natl. Acad. Sci. USA* 80, 4417–4421 (1983).
8. Struhl, K., Stinchcomb, D.T., Scherer, S. and Davis, R.W., "High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules," *Proc. Natl. Acad. Sci. USA* 76, 1035–1039 (1979).
9. Stinchcomb, D.T., Struhl, K. and Davis, R.W., "Isolation and characterisation of a yeast chromosomal replicator," *Nature* 282, 39–43 (1979).
10. Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J. and Stahl, F., "The double-strand-break repair model for recombination," *Cell* 33, 25–35 (1983).
11. Blackburn, E.H. and Gall, J.G., "A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*," *J. Mol. Biol.* 120, 33–53 (1978).
12. Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. and Cech, T.R., "Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*," *Cell* 31, 147–157 (1982).

13. Ratzkin, B. and Carbon, J., "Functional expression of cloned yeast DNA in *Escherichia coli*," *Proc. Natl. Acad. Sci. USA* 74, 487–491 (1977).
14. Struhl, K. and Davis, R.W., "Production of a functional eukaryotic enzyme in *Escherichia coli*: cloning and expression of the yeast structural gene for imidazole-glycerolphosphate dehydratase (his3)," *Proc. Natl. Acad. Sci. USA* 74, 5255–5259 (1977).
15. Szostak, J.W. and Blackburn, E.H., "Cloning yeast telomeres on linear plasmid vectors," *Cell* 29, 245–255 (1982).
16. Walmsely, R., Petes, T.D. and Szostak, J.W., "Is there left-handed DNA at the ends of yeast chromosomes?" *Nature* 302, 84–86 (1983).
17. Shampay, J., Szostak, J.W. and Blackburn, E.H., "DNA sequences of telomeres maintained in yeast," *Nature* 310, 154–157 (1984).
18. Walmsley, R.W., Chan, C.S.M., Tye, B.-K. and Petes, T.D., "Unusual DNA sequences associated with the ends of yeast chromosomes," *Nature* 310, 157–160 (1984).
19. Szostak, J.W., "Replication and Resolution of Telomeres in Yeast," Cold Spring Harbor Symp. Quant. Biol. 47, 1187–1194 (1983).
20. Clarke, L. and Carbon, J., "Isolation of a yeast centromere and construction of functional small circular chromosomes," *Nature* 287, 504–509 (1980).
21. Murray, A.W. and Szostak, J.W., "Construction of artificial chromosomes in yeast," *Nature* 305, 189–193 (1983).
22. Murray, A.W., Schultes, N.P. and Szostak, J.W., "Chromosome length controls mitotic chromosome segregation in yeast," *Cell* 45, 529–536 (1986).
23. Sundin, O. and Varshavsky, A., "Arrest of segregation leads to accumulation of highly intertwined catenated dimers: dissection of the final stages of SV40 DNA replication," *Cell* 25, 659–669 (1981).
24. Nasmyth, K., "Segregating sister genomes: the molecular biology of chromosome separation," *Science* 297, 559–565 (2002).
25. Burke, D.T., Carle, G.F. and Olson, M.V., "Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors," *Science* 236, 806–812 (1987).
26. Blackburn, E.H. and Szostak, J.W., "The molecular structure of centromeres and telomeres," *Ann. Rev. Biochem.* 53, 163–194 (1984).
27. Bernards, A., Michels, P.A., Lincke, C.R. and Borst, P., "Growth of chromosome ends in multiplying trypanosomes," *Nature* 303, 592–597 (1983).
28. Klobutcher, L.A., Swanton, M.T., Donini, P. and Prescott, D.M., "All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus," *Proc. Natl. Acad. Sci. USA* 78, 3015–3019 (1981).
29. Greider, C.W. and Blackburn, E.H., "Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts," *Cell* 43, 405–413 (1985).
30. Greider, C.W. and Blackburn, E.H., "The telomere terminal transferase of *Tetrahymena* is a ribonucleoprotein enzyme with two kinds of primer specificity," *Cell* 51, 887–898 (1987).
31. de Lange, T., "How telomeres solve the end-protection problem," *Science* 326, 948–952 (2009).
32. Lundblad, V. and Szostak, J.W., "A mutant with a defect in telomere elongation leads to senescence in yeast," *Cell* 57, 633–643 (1989).
33. Yu, G.L., Bradley, J.D., Attardi, L.D. and Blackburn E.H., "In vivo alteration of telomere sequences and senescence caused by mutated *Tetrahymena* telomerase RNAs," *Nature* 344, 126–132 (1990).
34. Harley, C.B., Futcher, A.B. and Greider, C.W., "Telomeres shorten during ageing of human fibroblasts," *Nature* 345, 458–460 (1990).
35. Bodnar, A.G., Ouellette, M., Frolkis, M., Holt, S.E., Chiu, C.P., Morin, G.B., Harley, C.B., Shay, J.W., Lichtsteiner, S. and Wright, W.E., "Extension of life-span by introduction of telomerase into normal human cells," *Science* 279, 349–352 (1998).

36. Blackburn E.H., Greider C.W. and Szostak J.W., "Telomeres and telomerase: the path from maize, *Tetrahymena* and yeast to human cancer and aging," *Nat. Med.* 12, 1133–1138 (2006).
37. Greider, C.W., "Telomerase activation. One step on the road to cancer?" *Trends Genet.* 15, 109–112 (1999).
38. Gilbert, W., "The RNA World," *Nature* 319, 618 (1986).
39. Doudna, J.A. and Szostak, J.W., "RNA catalyzed synthesis of complementary strand RNA," *Nature* 339, 519–522 (1989).
40. Green, R. and Szostak, J.W., "Selection of a ribozyme that functions as a superior template in a self-copying reaction," *Science* 258, 1910–1915 (1992).
41. Ellington, A.E. and Szostak J.W., "In vitro selection of RNA molecules that bind specific ligands," *Nature* 346, 818–822 (1990).
42. Tuerk, C. and Gold, L., "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase," *Science* 249, 505–510 (1990).
43. Sassanfar, M. and Szostak, J.W., "An RNA motif that binds ATP," *Nature* 364, 550–553 (1993).
44. Bartel, D.P. and Szostak, J.W., "Isolation of new ribozymes from a large pool of random sequences," *Science* 261, 1411–1418 (1993).
45. Eklund, E.H., Szostak, J.W. and Bartel, D.P., "Structurally complex and highly active RNA ligases derived from random RNA sequences," *Science* 269, 364–370 (1995).
46. Shechner, D.M., Grant, R.A., Bagby, S.C., Koldobskaya, Y., Piccirilli, J.A. and Bartel D.P., "Crystal structure of the catalytic core of an RNA-polymerase ribozyme," *Science* 326, 1271–1275 (2009).
47. Eklund, E.H. and Bartel, D.P., "RNA-catalysed RNA polymerization using nucleoside triphosphates," *Nature* 383, 192–198 (1996).
48. Salehi-Ashtiani, K., Luptak, A., Litovchick, S. and Szostak, J.W., "A genome wide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene," *Science* 313, 1788–1792 (2006).
49. Theis, M., Si, K. and Kandel, E.R., "Two previously undescribed members of the mouse CPEB family of genes and their inducible expression in the principal cell layers of the hippocampus," *Proc. Natl. Acad. Sci. USA* 100, 9602–9607 (2003).
50. Vogler, C., Spalek, K., Aerni, A., Demougin, P., Müller, A., Huynh, K.-D., Pappasotiropoulos, A. and de Quervain, D.J.-F., "CPEB3 is associated with human episodic memory," *Front. Behav. Neurosci.* 3, 1–5 (2009).
51. Roberts, R.W. and Szostak, J.W., "RNA-peptide fusions for the in vitro selection of peptides and proteins," *Proc. Natl. Acad. Sci. USA* 94, 12297–12302 (1997).
52. Keefe, A.D. and Szostak, J.W., "Functional proteins from a random sequence library," *Nature* 410, 715–718 (2001).
53. Mansy, S.S., Schrum, J.P., Krishnamurthy, M., Tobé, S., Treco, D. and Szostak, J.W., "Template-directed synthesis of a genetic polymer in a model protocell," *Nature* 454, 122–125 (2008).
54. Hanczyc, M.M., Fujikawa, S.M. and Szostak, J.W., "Experimental models of primitive cellular compartments: Encapsulation, growth and division," *Science* 2003; 302, 618–622 (2003).
55. Ferris, J.P., Hill, A.R., Jr, Liu, R. and Orgel, L.E., "Synthesis of long prebiotic oligomers on mineral surfaces," *Nature* 381, 59–61 (1996).
56. Zhu, T.F. and Szostak, J.W., "Coupled growth and division of model protocell membranes," *J. Am. Chem. Soc.* 131, 5705–5713 (2009).
57. Orgel, L.E., "Molecular replication," *Nature* 358, 203–209 (1992).
58. Schrum, J., Ricardo, A., Krishnamurthy, K., Blain, J.C. and Szostak, J.W., "Efficient and rapid template-directed nucleic acid copying using 2'-amino-2', 3'-dideoxyribonucleoside-5'-phosphorimidazole monomers," *J. Am. Chem. Soc.* 131, 14560–14570 (2009).

Portrait photo of Professor Szostak by photographer Ulla Montan.