DNA sequencing with the Sanger method

The interdependence of information stored in DNA and its physical representation

Before looking at the details of the experimental methods used to read the information stored in DNA, it seems worthwhile to step back and to consider the relationship between DNA as a physical object and the information stored in this object.

We all know that DNA stores information in the sequence of the bases along its backbone. Remarkably, this base sequence has almost no influence on the physico-chemical properties of a DNA molecule[1] or on the speed, efficiency, and accuracy, with which DNA enzymes such as DNA polymerase process it.

This uncoupling of information content from the physical properties of the storage medium has two important consequences: 1) The same DNA-sequencing methods can be applied to virtually any piece of DNA, regardless of the specific information stored in that piece of DNA. A DNA strand encoding the amino-acid sequence of an enzyme can be read in the same way as a DNA strand encoding a promoter region of a gene or even a completely random sequence.

2) Determining the base sequence of a piece of DNA captures the most relevant biological information about this piece of DNA. This somewhat subtle point becomes clear when we contrast DNA with proteins. In the latter case, the sequence of a protein molecule is, of course, also of interest, but what really matters for understanding the biological function of a protein is not the amino-acid sequence, but the physical and chemical properties of its folded three-dimensional structure. By contrast, it can be argued that, in the case of DNA, the information stored in the base sequence is of primary interest while the physical properties and structure of this piece of DNA convey little additional information[2].

Sanger sequencing was the first reliable and universal method for determining the base sequence of a strand of DNA

Sanger sequencing is named after its inventor Frederick Sanger who not only invented the first reliable und universal method for DNA sequencing, but prior to that also developed a method to determine the amino-acid sequence of proteins. Each of these discoveries earned him a Nobel prize!

Sanger himself called his DNA-sequencing technique the "dideoxynucleotide chain-termination method" - a name, which nicely captures the central idea of his method.

For his sequencing method, Sanger assembled all the components that are needed for a DNA synthesis reaction in a test tube: a single-stranded template DNA, a primer, a polymerase and triphosphate deoxynucleotides dATP, dGTP, dTTP, and dCTP[3] (figure 1). When mixed together, the primer anneals to its complementary sequence on the template, the polymerase binds to the formed duplex and begins to extend the primer by incorporating complementary nucleotides according to the sequence of the template strand. In the synthesis reaction, the α-phosphate group of the incoming nucleotide attaches to the 3'-hydroxyl group of the previously incorporated nucleotide.

If only these components were present in this reaction mix, DNA synthesis would proceed in much

[1] Exceptions to this general statement do, of course, exist. The most obvious example being extended repeats of the GCGC sequence, which will induce a Z-conformation in the DNA.

[2] Where DNA structure does play an important role is in the way the DNA double helix is packed into chromatin. There the quaternary DNA structure controls the extend, to which sections of the cell's DNA are accessible to the transcriptional machinery. This in turn determines how accessible the information stored in this section of DNA is to the cell at a particular point in time.

[3] The "d" indicates that we are dealing with the DNA precursors and not RNA precursors (ATP, GTP etc.), which contain an additional hydroxy group at 2' position of their sugar moiety.

the same way as it does during DNA replication and would end only when the end of the template is reached, or the reaction runs out of deoxynucleotides.

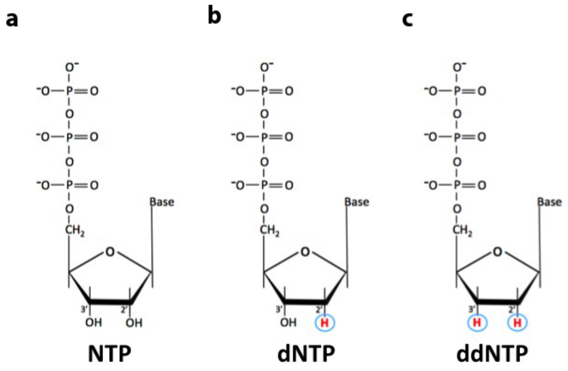

Figure 1: Comparison of the chemical structures of NTPs, dNTPs, and ddNTPs. (a) NTPs, such as ATP, serve as the energy currency of the cell and form the building blocks of RNA. (b) dNTPs are the building blocks of DNA. They lack the hydroxyl group onthe 2' carbon of the sugar ring. This gives DNA an increased chemical stability against hydrolysis compared to RNA. (c) ddNTPs are synthetically generated compounds that are used as chain terminators in Sanger sequencing. In ddNTPs, the 3'-hydroxyl group of the sugar ring, which is needed for the attachment of the next nucleotide to the growing DNA strand, has been replaced by a hydrogen atom.

But Sanger added one additional ingredient to the synthesis reaction: a dideoxynucleotide (ddNTP, Figure 1c). This is a chemically synthesized nucleotide analog, in which the 3'-hydroxyl group is replaced by a hydrogen atom. This nucleotide is incorporated by the polymerase just as any other nucleotide. But, the lack of a 3'-hydroxyl group prevents the addition of any further nucleotides to the growing DNA strand. Hence, as the original name of the method indicates, the dideoxynucleotide terminates the synthesis of the growing DNA chain.

Let us now consider what happens if we add a small amount of just one type of dideoxynucleotide (e.g., ddATP) to the DNA-synthesis reaction. The synthesis proceeds normally until the polymerase encounters a thymine (T) on the template strand. At that time the polymerase can either incorporate a natural dATP and synthesis will continue with the next nucleotide; or it incorporates an artificial ddATP, in which case the synthesis of this particular strand is terminated. Which of the two happens is determined by chance and by the relative concentration of dATP and ddATP. Those strands where a "normal" dATP was incorporated will continue to grow until the polymerase encounters another T, in which case there is again a chance that synthesis is terminated by addition of a ddATP etc.

The result of this synthesis reaction is a population of DNA strands of different lengths. The reason the Sanger methods works is that these lengths contain information about the positions, in which the template strand contained a base that is complementary to the dideoxynucleotide used in the reaction. If, for example (see figure 2), the template contained a T at positions 2, 4, and 9, then a ddATP nucleotide may have been incorporated at these positions in one of the growing strands and would have terminated the synthesis of this strand. As a result, the products of the reaction will contain newly synthesized strands that are 2, 4, and 9 bases long. At positions where the leading strand contained any of the other three bases, synthesis will have simply continued. As a result, none of the newly synthesized strands will be 1, 3, 5, 7, or 8 bases long.

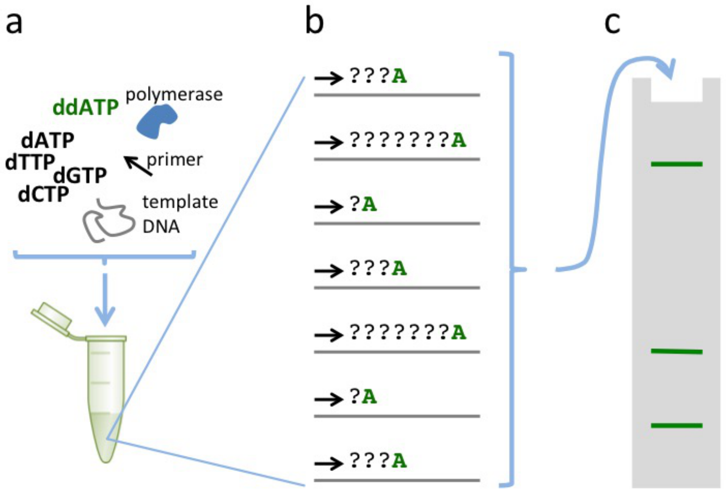

Figure 2: Sanger DNA sequencing uses base-specific random termination of DNA synthesis to determine the locations of a specific base type in a DNA sequence. During the synthesis reaction (a) a specific dideoxynuclotide (e.g. ddATP) included in the synthesis reaction may be inserted in the growing DNA chain wherever a complementary base is present in the template strand, thus causing termination of the chain at this position. The result is a population of DNA strands (b), each terminating with an adenine. Separating these DNA strands by gel electrophoresis (c) reveals the relative position of adenosine nucleotides in the newly synthesized DNA strands.

By performing DNA synthesis in four separate reactions, each containing a different dideoxynucleotide, one obtains four populations of newly synthesized strands. Together the lengths of these strands encode the complete base sequence of the DNA strand.

The second step of Sanger sequencing is to read out this information by separating the newly synthesized strands by gel electrophoresis. Shorter DNA strands will move through the gel more quickly than longer strands. By using the right gels and running conditions, it is possible to distinguish DNA molecules that differ by a single base in length.

To visualize the positions of the DNA bands on the gel, early versions of the Sanger method used radioactively labeled ddNTP molecules and separated the products of the four sequencing reactions on adjacent lanes of the same electrophoresis gel. Exposing the gel to a photographic film then revealed a staircase pattern of bands corresponding to the different lengths of newly synthesized DNA strands (figure 3). By stepping along this staircase from the bottom of the gel (shortest strand) to the top, the sequence of the newly synthesized strand can be read off in the 5'-to-3' direction. The sequence that is read out from the gel is complementary to the originally sequence DNA strand.

This radioactivity-based version of the Sanger method was soon replaced by a fluorescence-based version (figure 4). In this fluorescence-based approach, each of the four dideoxynucleotides is labeled with a different fluorophore (for an example, see figure 4b). In this way, the type of nucleotide (A, T, C, or G) that had terminated the synthesis of a strand is encoded in the color of the fluorophore attached to this strand. This makes it possible to combine the four previously separate synthesis reactions into a single reaction and to separate all of the newly synthesized DNA strands on a single lane of a gel. Most importantly, reading out the base sequence could now be automated by placing a color-sensitive fluorescence detector at the bottom of the gel that detects each of the fluorescing bands as they pass the detector. The resulting chromatograms can then be interpreted by computerized "base-calling" algorithms to determine the DNA sequence automatically.

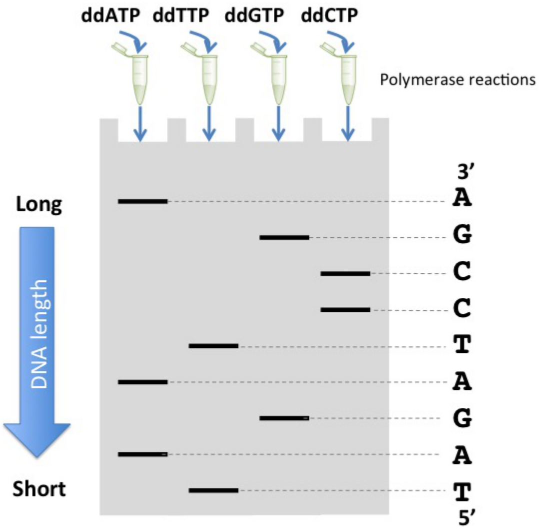

Figure 3: Reading out the DNA sequence from an autoradiograph generated by the traditional Sanger method. The products of the four separate sequencing reactions are separated on a gel. The gel is then exposed to a photographic film and the blackened bands on the film reveal the relative length of the DNA strands generated in the four reactions. By "stepping" from the smallest to the largest strand the sequence of the entire strand can be read off in the 5'-to-3' direction.

Performance parameters of Sanger sequencing

In its initial implementation, the Sanger method developed in the 1970's was labor intensive, involved the use of radioactive compounds, and the length of usable sequences (called the read length) was less then 100 bases. At that time, a dedicated and skillful scientist would be able to determine the sequence of a few hundred bases per day. Through steady development and automation of all aspects of the technique, the efficiency of the Sanger method was improved continuously over the next 25 years to the point where a technical staff member with a state-of-the-art instrument could sequence about 1 million bases per day at a cost of 2'000 - 3'000 US dollars. And in fact, the Sanger-sequencing method using fluorescent chain terminators was used in a massive community effort to generate a draft sequence of the human genome, a landmark achievement in the life sciences.

Yet, the fundamental principle of sequencing - terminating the growth of DNA strands by random incorporation of modified nucleotides and then separating the resulting DNA molecules by gel electrophoresis - has remained unchanged.

In its most optimized form, the Sanger method is able to reach a read length of 1000 bases. This maximal read length is fundamentally limited by the relative size resolution obtainable by gel electrophoresis. The challenge can be understood by considering that to read the 100th base in a sequence, the electrophoresis step has to separate two strands of DNA that are 99 and100 bases long. This corresponds to a relative size difference of 1/100. But, to read the 1000th base the electrophoresis needs to separate two strands that are 999 and 1000 bases long. This brings the required relative resolution to 1/1000, which approaches the current technical limit of gel electrophoresis.

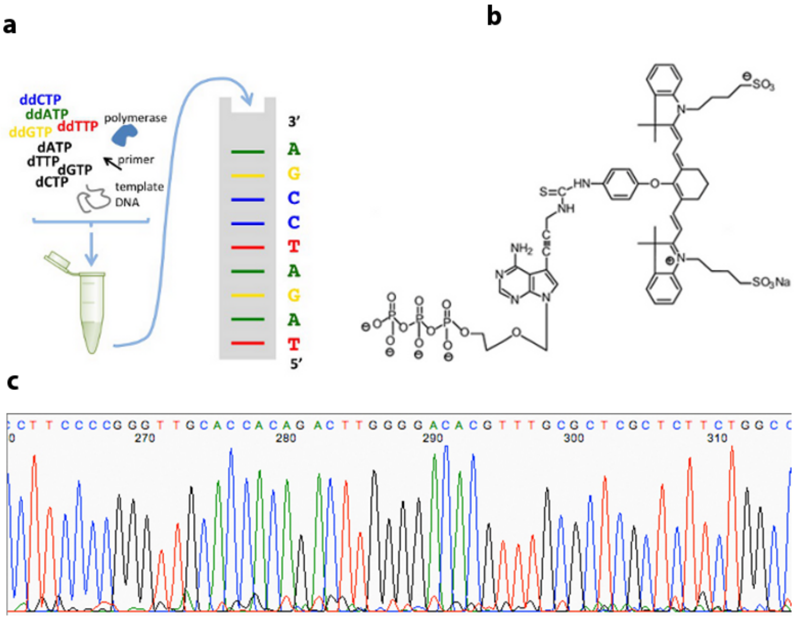

Figure 4: In fluorescence-based Sanger sequencing, (a) fluorescently labeled ddNTP analogs, each fluorescing in a different color, are used to distinguish the different chain-termination products of the polymerase reaction. (b) Chemical structure of a fluorescent chain-terminating dATP analog. The structure incorporates the chain-terminating and fluorescent functionalities, but is still recognized and incorporated into the growing DNA strand by the DNA polymerase. (c) Chromatogram recorded by a fluorescence detector placed at the bottom of a capillary gel. Each peak corresponds to a band on the gel. The sequence in 5'-to-3' order is read from left to right.

The error rate of Sanger sequencing typically lies in the range of one error per 10'000 - 100'000 base pairs, which is extremely low. Additionally, problematic areas of the sequence likely to contain errors can usually be spotted in the chromatograms and can then be removed during data processing.

Sanger sequencing is still popular for many applications

Most large-scale DNA-sequencing projects (e.g., whole-genome analysis) now employ the second and third generation methods that will be discussed in an upcoming lesson. Still, Sanger sequencing remains popular for many day-to-day applications in molecular-biology laboratories and for medical diagnostics. The reason for this continued popularity is two-fold: First, the Sanger method employs sequencing primers that can be chosen by the experimenter. This allows targeted sequencing of a specific portion of a larger DNA molecule while the 2nd and 3rd generation workflows will sequence the entire DNA in the sample. The second strength of the Sanger method is that it can read longer stretches of DNA in a single reaction with very low error rates. In medical diagnostics, it is therefore still common practice to verify candidate mutations found by 2nd and 3rd generation-sequencing methods via the Sanger method.