Over the past few decades, advances in the computer science and informatics have made possible development of new strategies in molecular biology that enable to collect and analyze the data about complex interactions within biological systems. A new inter-disciplinary field of study, called systems biology, has been created. It analyzes and integrates information from genomics (DNA level), transcriptomics (RNA level), proteomics (protein level) and other related fields (e.g. metabolomics and other -omics). Specialized tools to view and analyze biological data and computerized databases to store, organize, and index the data are required. Thus, bioinformatics, the interface between the biological and computational sciences, is indispensable to obtain a clearer insight into the fundamental biology of organisms.

9.1 DNA Sequencing

DNA sequencing is the laboratory method used to determine the exact order of base pairs in a DNA fragment. The DNA sequence contains information necessary for organisms to survive and reproduce, thus reading the sequence is useful in fundamental research into why and how organisms live. With the sequencing of genomes, genomics was developed, which can be defined as the mapping of genes of various organisms by large-scale DNA-sequence analysis.

9.1.1 Chain Termination Sequencing

The first techniques of sequencing were established in the 1970-1980s. One of the major methods is known as chain termination sequencing, dideoxy sequencing, or Sanger sequencing after its inventor Frederick Sanger (Nobel Prize in Chemistry in 1980, shared with Walter Gilbert). The single stranded DNA fragment (which is the sequence to be determined) serves as a template
Fig. 9.1 Chain termination method of DNA sequencing. The synthesis of a new chain stops after incorporation of dideoxynucleotide (ddNTP). If ddNTPs are labeled differently, products of the reaction can be analyzed together, if not - samples with each ddNTP should be analyzed separately.