Quarterly Reviews of Biophysics **35**, 2 (2002), pp. 169–200. © 2002 Cambridge University Press
DOI: 10.1017/S0033583502003797 Printed in the United Kingdom

A review of DNA sequencing techniques

Lilian T. C. França[1], Emanuel Carrilho[2] and Tarso B. L. Kist[3]*

[1] Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil and Instituto de Biofísica Carlos Chagas Filho, CCS, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil (E-mail: lila@biof.ufrj.br)
[2] Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, SP, Brazil (E-mail: emanuel@iqsc.sc.usp.br)
[3] Departamento de Biofísica, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, 91501–970, Porto Alegre, RS, Brazil (E-mail: tarso@orion.ufrgs.br)

1. Summary

The four best known DNA sequencing techniques are reviewed. Important practical issues covered are read-length, speed, accuracy, throughput, cost, as well as the automation of sample handling and preparation. The methods reviewed are: (i) the Sanger method and its

* Author to whom correspondence should be addressed.
Tel.: 55 51 3316 7618; Fax: 55 51 3316 7003; E-mail: tarso@orion.ufrgs.br

most important variants (enzymic methods); (ii) the Maxam & Gilbert method and other chemical methods; (iii) the Pyrosequencing™ method – DNA sequencing in real time by the detection of released pyrophosphate (PPi); and (iv) single molecule sequencing with exonuclease (exonuclease digestion of a single molecule composed of a single strand of fluorescently labelled deoxynucleotides). Each method is briefly described, the current literature is covered, advantages, disadvantages, and the most suitable applications of each method are discussed.

2. Introduction

DNA sequencing techniques are key tools in many fields. A large number of different sciences are receiving the benefits of these techniques, ranging from archaeology, anthropology, genetics, biotechnology, molecular biology, forensic sciences, among others. A silent and remarkable revolution is under way in many disciplines; DNA sequencing is promoting new discoveries that are revolutionizing the conceptual foundations of many fields. At the same time new and very important issues are emerging with these developments, such as bioethical questions and questions related to public health and safety.

In this review we will follow the chronological development of the methods. We will start in Section 3 with the methods developed by Sanger and his collaborators in the 1970s. The Maxam & Gilbert method and other chemical methods are reviewed in Section 4. The PPi method – based on detection of PPi released on nucleotide incorporation during chain extension by polymerase – is reviewed in Section 5. The methods based on single molecule detection are reviewed in Section 6. Finally, the concluding remarks are given in Section 7.

3. Sanger's method and other enzymic methods

The first method described by Sanger and Coulson for DNA sequencing was called 'plus and minus' (Sanger & Coulson, 1975). This method used *Escherichia coli* DNA polymerase I and DNA polymerase from bacteriophage T4 (Englund, 1971, 1972) with different limiting nucleoside triphosphates. The products generated by the polymerases were resolved by ionophoresis on acrylamide gels. Due to the inefficacy of the 'plus and minus' method, 2 yr later, Sanger and his co-workers described a new breakthrough method for sequencing oligonucleotides via enzymic polymerization (Sanger *et al*. 1977). This method, which would revolutionize the field of genomics in the years to come, was initially known as the chain-termination method or the dideoxynucleotide method. It consisted of a catalysed enzymic reaction that polymerizes the DNA fragments complementary to the template DNA of interest (unknown DNA). Briefly, a ^{32}P-labelled primer (short oligonucleotide with a sequence complementary to the template DNA) was annealed to a specific known region on the template DNA, which provided a starting point for DNA synthesis. In the presence of DNA polymerases, catalytic polymerization of deoxynucleoside triphosphates (dNTP) onto the DNA occurred. The polymerization was extended until the enzyme incorporated a modified nucleoside [called a terminator or dideoxynucleoside triphosphate (ddNTP)] into the growing chain.

This method was performed in four different tubes, each containing the appropriate amount of one of the four terminators. All the generated fragments had the same 5′-end,

whereas the residue at the 3′-end was determined by the dideoxynucleotide used in the reaction. After all four reactions were completed, the mixture of different-sized DNA fragments was resolved by electrophoresis on a denaturing polyacrylamide gel, in four parallel lanes. The pattern of bands showed the distribution of the termination in the synthesized strand of DNA and the unknown sequence could be read by autoradiography. For a better understanding of the Sanger reaction, see Fig. 1. The enzymic method for DNA sequencing has been used for genomic research as the main tool to generate the fragments necessary for sequencing, regardless of the sequencing strategy. Two different approaches, shotgun and primer walking sequencing, are the most used (Griffin & Griffin, 1993). The main aspects of each strategy are described below in more detail.

3.1 Random approach

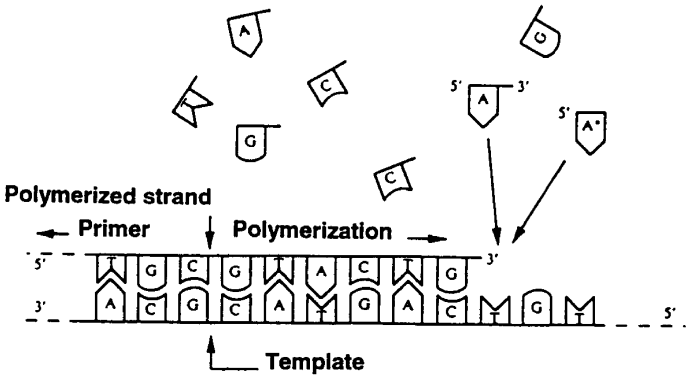
Also known as shotgun sequencing, this is a random process because there is no control of the region that is going to be sequenced, at least in the usual procedures (there are exceptions, for instance see the procedure described by Lander *et al*. 2001). Genomic DNA is randomly fragmented (by sonication, nebulization, or other scission methods) into smaller pieces, normally ranging from 2 to 3 kb. The fragments, inserted into a vector, are replicated in a bacterial culture. Several positive amplifications are selected, and the DNA is extensively sequenced. Due to the random nature of this process, the sequences generated overlap in many regions (Adams *et al*. 1996). The process of overlaying or alignment of the sequences is called sequence assembly. Shotgun sequencing normally produces a high level of redundancy (the same base is sequenced 6–10 times, in different reactions) which affects the total cost. A new variation of the method introduced by Venter *et al*. (1996) involved shotgunning a whole genome at once. This strategy depended enormously on computational resources to align all generated sequences. However, the efforts were rewarded with the sequencing of the *Haemophilus influenzae* genome in only 18 months (Fleischmann *et al*. 1995) and, more recently, the human genome (Venter *et al*. 2001).

Shotgun sequencing is well established, with ready availability of optimized cloning vectors, fluorescently labelled universal primers, and software for base calling and sequence assembly. The whole process has a high level of automation, from the cloning of the vectors and colony selection to the bases called. A simplified diagram of the shotgun process is summarized in Fig. 2. Although the random approach is fully compatible with automation, it can produce gaps in the sequence that can only be completed by direct sequencing of the region.

3.2 Direct approach

The other approach for genomic sequencing is the direct sequencing of unknown DNA within sites in which the sequence is known. For example, an unknown sequence of DNA is inserted into a vector and amplified. The first sequencing reaction is performed using the primers that hybridize to the vector sequence and polymerize the strand complementary to the template. A second priming site is then chosen inside the newly generated sequence, following the same direction as the first one. This approach is known as primer walking (Studier, 1989; Martin-Gallardo *et al*. 1992), and its major advantage is the reduced redundancy (Voss *et al*. 1993) because of the direct nature of the approach (opposite to

(*a*)



(*b*)

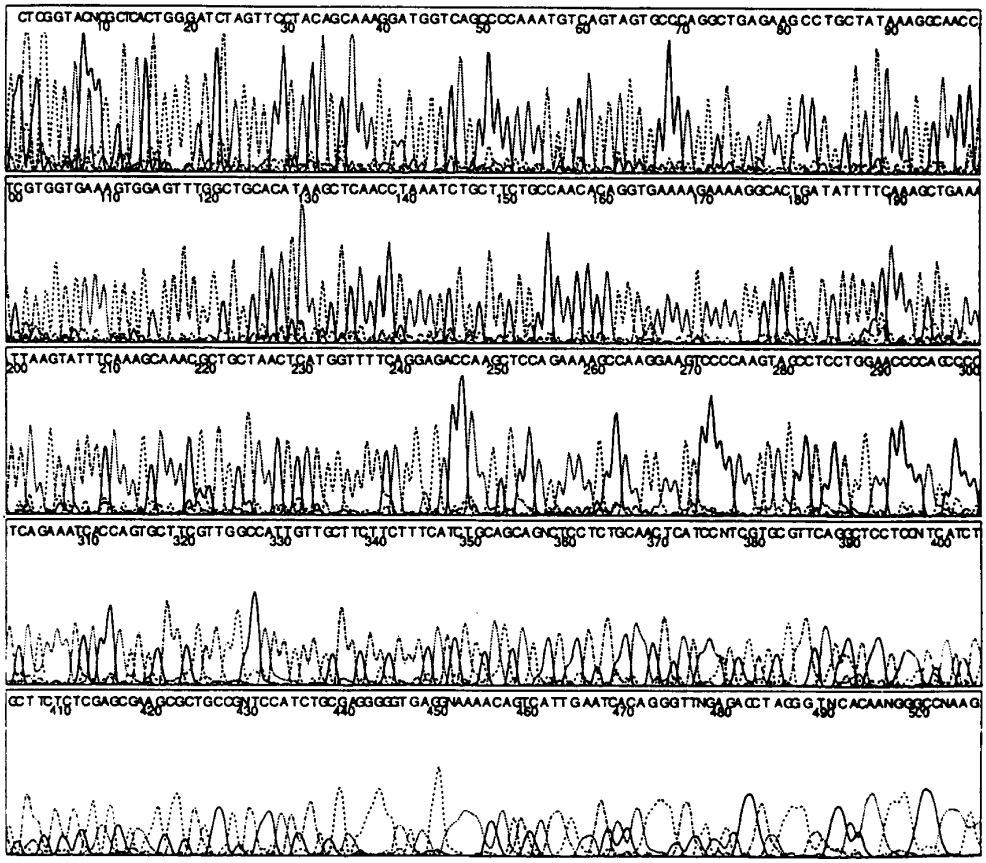


Fig. 1. Schematic representation of a sequencing process ('four-colour Sanger'): starting from many copies of the ssDNA to be sequenced, bearing a known 'marker' at the beginning of the unknown sequence, a short oligonucleotide 'primer' complementary to this marker is hybridized (i.e. paired) to the marker, in the presence of DNA polymerase and free nucleotides. This hybridization initiates reconstruction by the polymerase of a single strand complementary to the unknown sequence (*a*). Including in the nucleotide bath in which the polymerization takes place a small fraction of fluorescently labelled dideoxynucleotides (one different dye for each nucleotide type), which lack the OH group necessary for further extension of the strand, one is able to synthesize at random complementary strands with all possible stop points (i.e. all possible lengths with an integer number of nucleotides). These

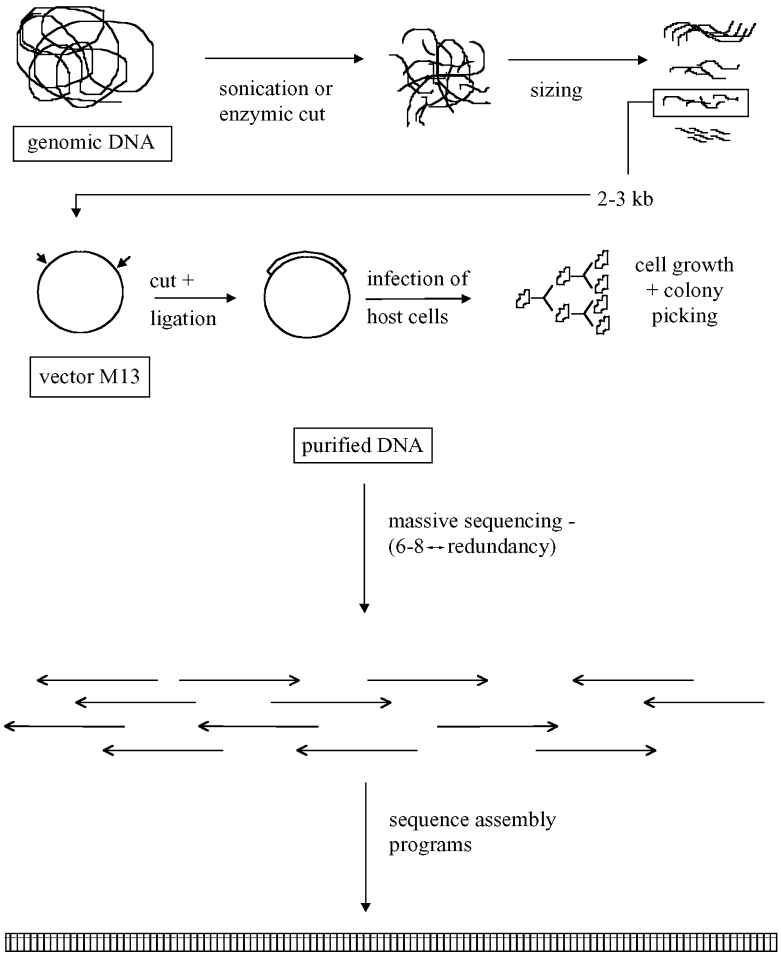


Fig. 2. Random sequencing approach or shotgun. The distinct processes involve first fragmentation of the DNA into 2–3 kbp range, fragments are then cloned into vectors and introduced into host cells for amplification. After purification, the DNA from individual colonies is sequenced, and the results are lined up with sequence-assembly programs.

random), as seen in Fig. 3. However, it requires the synthesis of each new primer, which, in the past, was time consuming and expensive, especially when dye-labelled primers were used.

Some alternatives were introduced to overcome the problems of time and cost (Ruiz-Martinez *et al*. 1996). Although slightly different, these approaches shared the same idea of using a short oligonucleotide library as a means to create a longer primer. The number of all sequences possible for an oligonucleotide with n bases is equal to 4^n. It was proposed by Kieleczawa *et al*. (1992) that a hexamer library containing 4096 oligonucleotides could be cost effective. While each new 18-mer primer is used only once for each new reaction site

---

newly synthesized ssDNAs are then separated by size electrophoretically [see electropherogram in (*b*)]: consecutive peaks correspond to DNA fragments differing by one base, and each line corresponds to one given nucleotide. Automated analysis of the data allows the determination of the sequence (symbols above the peaks). The symbol N indicates ambiguous determination. In the present case, the sequence was faultless up to 435 bases. (Reproduced from Viovy, 2000.)

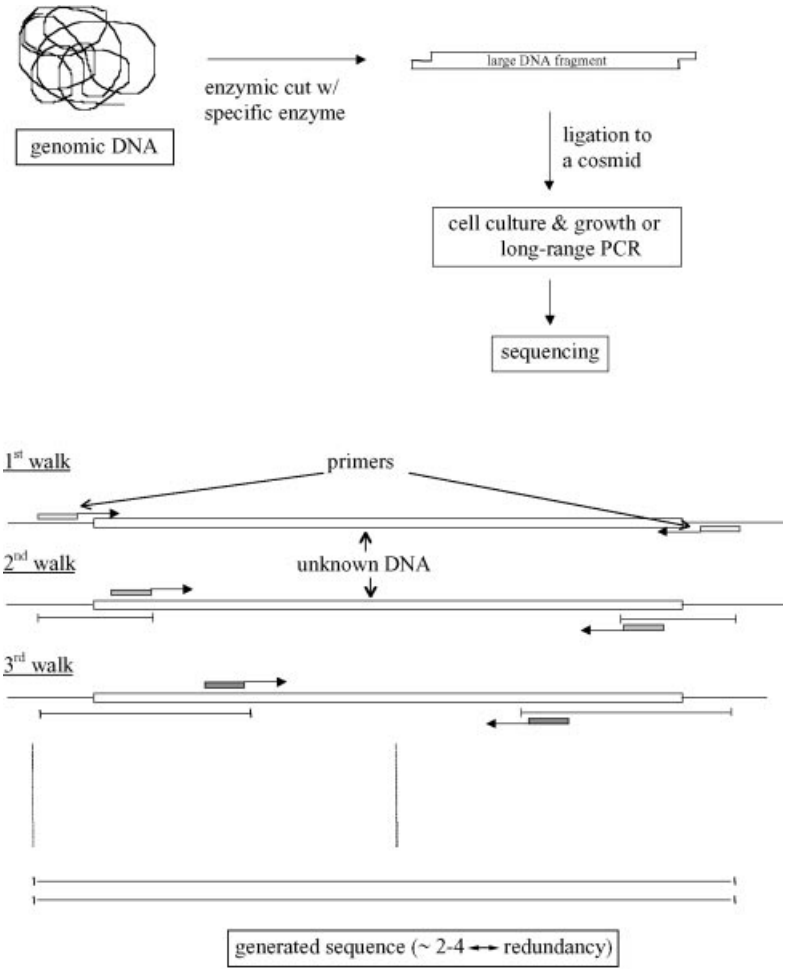


Fig. 3. DNA sequencing by the primer-walking strategy. In primer walking, the genomic DNA is cut into a large piece (~40 kbp) and inserted into a cosmid for growth. The sequencing is performed by walks, starting first from the known region of the cosmid. After the results from the first round are edited, a new priming site is located within the newly generated sequence. This procedure is repeated until the walks reach the opposite starting points.

(uniqueness is a requirement to avoid false priming), a 6-mer can be employed in many priming sites at different positions.

Using such short oligonucleotides leads to the possibility of mispriming since uniqueness is reduced with the reduced size of the oligonucleotide. For example, the use of three small oligonucleotides could result in several sites where one or two of them could hybridize to the template and initiate mispriming. To avoid this situation, a single-stranded DNA-binding protein (SSB) (Kieleczawa *et al*. 1992), or the stacking effects of selected modular primers (Kotler *et al*. 1993) were used.

Nowadays, the appeal of a cost-effective and time-saving method that uses small oligonucleotide libraries has disappeared with improvements in primer synthesis technology (Lashkari *et al*. 1995). However, the demand for a sequencing method that was able to provide long read-length (number of bases read per run), short analysis time, low cost, and high

accuracy has led to several modifications of the original Sanger method. In addition to several improvements in the procedures and in the reagents used in the sequencing reaction, further development in DNA separation technology was of paramount importance for the completion of the Human Genome Project. Several of the improvements that have been made in each step of enzymic DNA sequencing will be described.

3.3 Enzyme technology

Improvements in DNA polymerase enzymes have greatly contributed to the quality of the sequencing reactions and sequencing data. Initially, isothermal DNA polymerases were used in manual and automated DNA sequencing (Tabor & Richardson, 1987; Tabor *et al*. 1987). The reactions were performed at physiological temperatures (~37 °C) for a few minutes (~20 min). These enzymes (T4 or T7 DNA polymerases) evenly incorporated all four terminators, even the dye-labelled ones. The problem with these polymerases was that they were very sensitive to temperature and easily deactivated.

With the discovery of the polymerase chain reaction (PCR) and the use of a heat-stable DNA polymerase from *Thermus aquaticus* (Taq polymerase), the ability to perform sequencing reactions (cycle-sequencing) with reduced amounts of DNA template compared to isothermal enzymes became possible (Mullis *et al*. 1986; Mullis & Fallona, 1987). The major drawback of cycle-sequencing using Taq polymerase was the preference of the enzyme for ddNTPs rather than dNTPs. A single substitution of one amino acid in the primary sequence of the enzyme completely changed this effect and the rate of ddNTP incorporation was substantially equalized to that of dNTPs (Tabor & Richardson, 1995).

Many other enzymes are available for PCR and cycle sequencing. PCR enzymes require an extra feature, that is 3′- to 5′-exonuclease activity. This feature is called the proof-reading ability of the enzyme, i.e. its ability to correct mistakes made during incorporation of the nucleotides. For cycle-sequencing, this activity must be suppressed to avoid un-interpretable data.

Although largely improved, there was still significant variation in peak intensity for fluorescently labelled dye-terminators. The pattern of the termination was reproducible and predictable (Parker *et al*. 1996), but this variation made automatic base calling difficult. A few years later, one of the major suppliers of fluorescent sequencing kits introduced a modified set of fluorescent labels for ddNTPs. With this new dye-terminator kit, the signal was more even, and automated base calling improved significantly (see Section 3.5.3).

3.4 Sample preparation

The methodology for sample preparation often included the following steps: (i) DNA scission and cloning into a vector (e.g. M13 or M13mp18); (ii) vector amplification to produce a phage-infected culture; and (iii) purification from the cell culture to yield pure single-stranded (ss)DNA template (Martin & Davies, 1986), as illustrated in Figs 2 and 3. Among the strategies used to generate random fragments it is possible to mention: deletions generated by transposons (Ahmed, 1984), production of subclones by sonication of the DNA (Deininger, 1983), and restriction enzymes (Messing, 1983) such as DNAse (Anderson, 1981), exonuclease III (Henikoff, 1984) and T4 DNA polymerase resection clones (Dale *et al*. 1985).

An alternative strategy for sequencing projects on a large scale that involved procedures for amplification, purification, and selection of the M13 template was described by Beck &

Alderton (1993). The main innovation in the amplification step was the use of the PCR. For the purification step, a large number of systems that used agarose were commercially available. However, these systems were both expensive and time consuming, and used considerable quantities of PCR products. Several methodologies for purification of PCR products have been described; among them, a technique that uses exonuclease I and shrimp alkaline phosphatase to degrade the excess primers and non-incorporated nucleotides, the main factors interfering in the sequencing reactions (Werle *et al*. 1994). Another method for purification of the fragments generated in the PCR was based on precipitation by isopropyl alcohol (Hogdall *et al*. 1999). This method is inexpensive, fast, and efficient for PCR fragments of any length.

In another methodology for sequencing PCR products, a template generated by PCR using a biotinylated forward primer and a non-biotinylated reverse primer has been used (Van den Boom *et al*. 1997). The non-purified product was submitted to dye-terminator cycle-sequencing using the same primers as used for the PCR. They enhanced the probability for the extension reaction by employing a second DNA polymerase, which is insensitive to the ddNTP concentration needed for sequencing. This results in a combined amplification and sequencing reaction in a single reaction due to the two DNA polymerases with differential incorporation rates for dideoxynucleotides (Van den Boom *et al*. 1998).

Another method for directly sequencing from PCR products was suggested and is based on the substitution of the chain-terminator by chain-delimiters (Porter *et al*. 1997). In this case it was demonstrated that boranophosphates (dNT^bP: 2′-deoxynucleoside-5′-α-[P-borano]-triphosphate) were convenient for use as delimiters for direct PCR sequencing (Fig. 4). The boranophosphates were heat stable, therefore they could be incorporated into DNA by PCR and, once incorporated, they blocked the action of the exonucleases. After incorporation, the boranophosphate positions can be revealed by digestion with an exonuclease, thus generating a series of fragments with borane at the end. The resulting fragments were separated by gel electrophoresis in a standard sequencing reaction.

Finally, the widely used method of plasmid-based amplification in *E. coli* followed by alkaline lysis was originally described by Birnboim & Doly (1979). Actually, most of the column preparations currently being sold for DNA isolation, involve using a technique based on this work.

3.5 Labels and DNA labelling

3.5.1 Radioisotopes

The enzymic method, when it was first described, used ^{32}P as a label. Biggin *et al*. (1983) proposed the use of deoxyadenosine 5′-(α-[^{35}S]thio)triphosphate as the label incorporated into the DNA fragments. This strategy resulted in an increase in band sharpness on autoradiography as well as in the resolution of band separation.

3.5.2 Chemiluminescent detection

As an alternative to radioisotopes, a method based on chemiluminescence detection with the biotin–streptavidin system has been used (Beck *et al*. 1989; Gillevet, 1990; Olesen *et al*. 1993; Cherry *et al*. 1994). In this system, the 5′-end of an oligonucleotide linked to biotin was used as the primer in the sequencing reaction. The enzyme alkaline phosphatase is bound to the

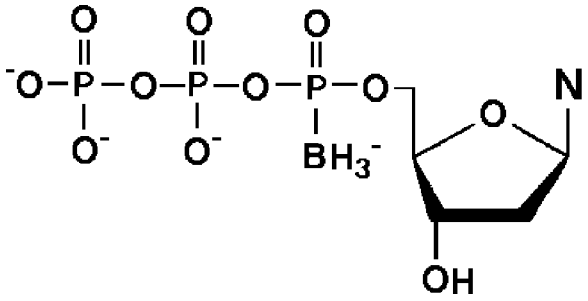


Fig. 4. Structure of 2′-deoxynucleoside-5′-α-[P-borano]-triphosphate (dNT^{b}P). N = adenine, cytosine, guanine or thymine. (Reproduced from Porter *et al.* 1997.)

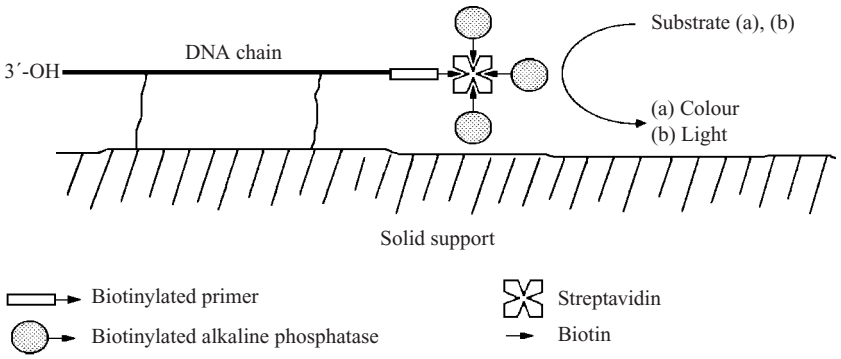


Fig. 5. Schematic diagram for the colorimetric (a) or chemiluminescent (b) detection of immobilized DNA using an enzyme-catalysed reaction. (Reproduced from Beck *et al.* 1989.)

5′-end of the oligonucleotide by a streptavidin conjugate. The enzyme catalysed a luminescent reaction (Fig. 5) and the emitted photons could be detected by a photographic film. There are at least three advantages to this method; first, the sequencing reactions were obtained directly from the PCR products; secondly, this method did not require cloning of the DNA before sequencing (Douglas *et al.* 1993; Debuire *et al.* 1993), and thirdly, it was possible to multiplex several reactions on the same gel and detect one at a time with appropriate enzyme-linked primers (Gillevet, 1990).

3.5.3 Fluorescent dyes

Although the Sanger method was fast and convenient, it still suffered from the use of radioisotopic detection, which was slow and potentially risky. Additionally, it required four lanes to run one sample because the label was the same for all reactions. To overcome such problems, Smith *et al.* (1986) developed a set of four different fluorescent dyes that allowed all four reactions to be separated in a single lane. The authors used the following fluorophore groups: fluorescein, 4-chloro-7-nitrobenzo-2-1-diazole (NBD), tetramethyl-rhodamine, and Texas Red (Smith *et al.* 1985, 1986), whose spectral properties are shown in Table 1.

Each of the four dyes was attached to the 5′-end of the primer and each labelled primer was associated with a particular ddNTP. For example, the fluorescein-labelled primer reaction was terminated with ddCTP (dideoxycytidine triphosphate), the tetramethyl-rhodamine-labelled primer reaction with ddATP (dideoxyadenosine triphosphate) and so on. All four

Table 1. *Spectral properties of some fluorophores used in automated DNA sequencing*

Dye	Absorption maximum (nm)	Emission maximum (nm)	Emission FWHM* (nm)
Fluorescein	493	516	60
4-Chloro-7-nitrobenzo-2-1-diazole (NBD)	475	540	79
Tetramethyl-rhodamine	556	582	52
Texas Red	599	612	42

* FWHM, full width at half maximum.

reactions were then combined and introduced onto a slab gel in a single lane. The bands were detected upon excitation of the fluorescent moiety attached to the DNA with a laser beam at the end of the gel. The fluorescent light was separated by means of four different coloured filters. After the 4-colour data was generated, the sequence read-out was straightforward, with the association of each colour to one base only.

DNA sequencing in slab gels with fixed-point fluorescence detection then became 'automated DNA sequencing' rather than 'manual DNA sequencing', which required exposure of the whole slab gel to a photographic plate for a fixed time and post-analysis detection (Griffin & Griffin, 1993; Adams *et al*. 1996). Automated DNA sequencing has been performed via two different labelling protocols. The first used a set of four fluorescent labels attached at the 5′-end of the primer, as described earlier. In the second method, the fluorescent moiety was linked to the ddNTP terminators, allowing the synthesis of all four ladders in a single vial. In the latter case, when the labelled ddNTP was incorporated, the enzyme terminated the extension at the same time as the ladder was labelled. Thus the C-terminated ladder contained one fluorescent dye, and the G-, A-, and T-terminated ladders had their own respective labels. The protocols are known as dye-labelled primer chemistry and dye-labelled terminator chemistry, respectively, and both labelling arrangements are shown in Fig. 6.

Alternative dyes were synthesized and linked to an M13 sequencing primer via a sulphydryl group and conjugated with tetramethyl-rhodamine iodoacetamide (Ansorge *et al*. 1986). This alternative dye used tetramethyl-rhodamine as the only fluorophore because of its high extinction coefficient, high quantum yield, and long wavelength of absorption ($\lambda_{exc} =$ 560 nm, $\lambda_{em} = 575$ nm, FWHM = 52 nm). One year later, the same group proposed a sulphydryl-containing M13 sequencing primer end-labelled with fluorescein iodoacetamine (Ansorge *et al*. 1987). Other dyes commonly linked to the primers includes carboxyfluorescein (FAM), carboxy-4′,5′-dichloro-2′,7′-dimetoxyfluorescein (JOE), carboxytetramethyl-rhodamine (TAMRA) and carboxy-X-rhodamine (ROX) (Swerdlow & Gesteland, 1990; Karger *et al*. 1991; Carson *et al*. 1993). These dyes have emission spectra with their maxima relatively well spaced, which facilitates colour/base discrimination. One drawback of this group of dyes was the need for two wavelengths for excitation; one at 488 nm for FAM and JOE dyes, and another at 543 nm for TAMRA and ROX dyes.

A different set of four base-specific succinylfluorescein dyes linked to chain-terminating dideoxynucleotides was described (Prober *et al*. 1987). These dyes were 9-(carboxyethyl)-3-hydroxy-6-oxo-6H-xanthenes or succinylfluoresceins (SF-xxx, where xxx represents the emission maximum in nanometres).

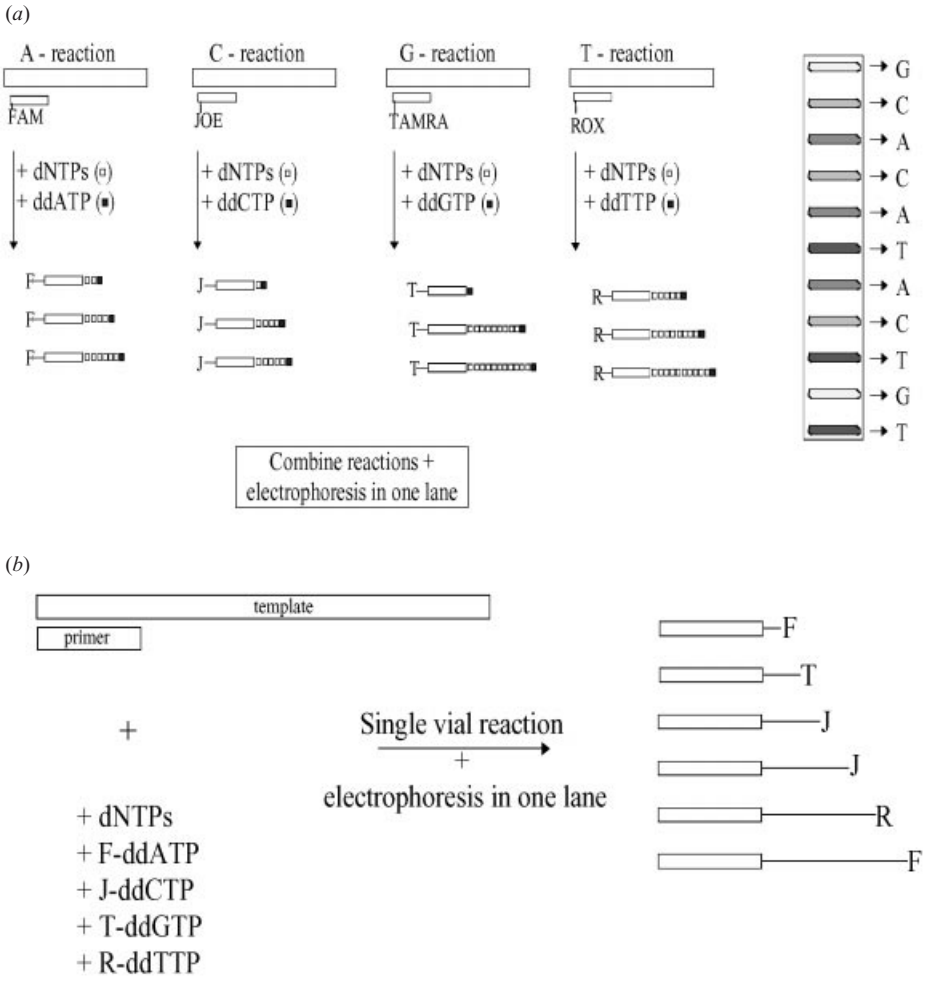


Fig. 6. Comparison of reactions for dye-labelled primer (*a*) and dye-labelled terminator (*b*) chemistries. Labelled primers require four separate reactions while labelled terminators only one. F, FAM; J, JOE; T, TAMRA; R, ROX.

Another modification in the original sequencing protocol used T7 DNA polymerase (or Sequenase™) with unlabelled primers but with a strategy of internal labelling. This helped to overcome ambiguous sequences that were occasionally observed (Wiemann *et al*. 1996). A new set of dyes, dipyrrometheneborondifluoride fluorophores (BODIPY) were shown to have better spectral characteristics than conventional rhodamine and fluorescein dyes. These dyes also showed uniform electrophoretic mobility, high fluorescence intensity, and consumed 30% less reagents per reaction than the conventional dyes (Metzker *et al*. 1996). A new dye set used for one-lane four-dye DNA sequencing with a set of fluorescent dyes with similar absorption and emission spectra, but different fluorescent lifetimes, has been described (Müller *et al*. 1997). A different strategy, based on a series of near-IR fluorescent dyes with an intramolecular heavy atom to alter the fluorescence lifetimes, was also suggested to produce a set of dyes for one-lane DNA sequencing (Flanagan *et al*. 1998).

A significant advance in dye-primer chemistry was the introduction of energy transfer (ET) dyes (Ju *et al*. 1995a, b). They consisted of two dyes per primer, one being a common donor

and the other an acceptor dye. The common donor can be either a fluorescein (FAM) or a cyanine (Cy5) derivative (Hung *et al*. 1996) at the 5′-end. The second dye, the discriminating one, is located about 10 bases along, with the separation between the dyes optimized for energy-transfer efficiency and minimum electrophoretic mobility shifts. The four acceptors are the commonly used ones in dye-primer chemistry; FAM, JOE, TAMRA and ROX (Ju *et al*. 1995a). The major advantages of ET dyes are that they can be almost evenly excited by a single wavelength (488 nm) and that the electrophoretic mobility shifts are minimal.[1] BODIPY dyes were used to produce similar ET primers offering narrower spectral bandwidth and better quantum efficiency (Metzker *et al*. 1996). Since their introduction, ET dyes have been widely used (Wang *et al*. 1995; Kheterpal *et al*. 1996, 1998). A new method of constructing ET primers using a universal cassette of ET was also developed. This cassette could be incorporated via conventional synthesis at the 5′-end of any primer sequence (Ju *et al*. 1996) allowing this technology to be used in primer-walking projects.

Any genome-sequencing project cannot be accomplished solely by the shotgun approach and, eventually, some part of the sequence has to be generated by primer walking. Because the synthesis of labelled primers is very expensive, dye-labelled terminator chemistry is the system of choice in such cases. Impressive advances have also been made in this field. As mentioned earlier, the first enzymes used in cycle-sequencing had severe problems in evenly incorporating the labelled terminators. To improve the sequencing performance, besides all modifications in the synthesis of the enzyme, significant changes in the dye structure were also made. Conventional dye-terminator chemistry used rhodamine and fluorescein derivatives. Depending on the enzyme used, these dyes showed a large variation in peak height, depending on the sequence. In addition, they required two different excitation wavelengths because the dyes that emitted fluorescence at longer wavelengths were poorly excited by the argon ion laser (488 nm); therefore, an additional laser had to be used. In order to improve the spectral features of such dye-terminators, dichlororhodamine derivatives were proposed and tested for peak pattern and enzyme discrimination. A further improvement was achieved with the concept of ET dyes, which was also successfully translated to dye-terminator protocol (Rosenblum *et al*. 1997; Lee *et al*. 1997). With this latest improvement, performing cycle-sequencing with energy-transfer terminators became routine and results were of high quality (Zakeri *et al*. 1998).

3.6 Fragment separation and analysis

Separation and analysis of DNA fragments generated by the Sanger method is a broad chapter and would be worthy a review on its own. However, it is impossible to discuss the Sanger method and DNA analysis without covering the important issues of electrophoresis and electrophoretic separation of DNA-sequencing samples.

3.6.1 Electrophoresis

The separation of labelled DNA fragments by polyacrylamide gel electrophoresis has been one of the greatest obstacles to complete automation of the enzymic DNA sequencing method. Among the main problems are gel preparation, sample loading, and post-

[1] Due to differences in charge and size, fluorescent dyes impart a differential migration pattern to the DNA. The effect is most pronounced for small fragments (< 200 bases).

electrophoresis gel treatment. However, a number of improvements in gel technology and electrophoresis have occurred, including the use of thinner gels (Garoff & Ansorge, 1981; Kostichka *et al*. 1992), gel gradient systems (Biggin *et al*. 1983), gel-to-plate binders, and the employment of devices to avoid temperature-induced band distortions (Garoff & Ansorge, 1981). Although significant progress in enzymic DNA sequencing was made, relying solely on slab gel technology was not enough to accomplish the challenges set by the Human Genome Project. In fact, in 1998 there was less than 6% of the genome published in the databases. The completion of the human genome was only possible due to several technological advances offered by capillary electrophoresis (CE) (Dovichi, 2000).

CE is a fast technique for separation and analysis of biopolymers (Jorgenson & Lukacs, 1983; Lauer & McManigill, 1986; Hjerten *et al*. 1987; Cohen & Karger, 1987). This technique uses narrow-bore fused silica capillaries (internal diameter less than 100 μm) and can resolve complex mixtures of biopolymers in a high electric field. The high surface-to-volume ratio of a small tube can efficiently dissipate the heat produced during electrophoresis and so the electric field can be higher than that used in slab gel electrophoresis. The higher the electric field, the faster the separation and, for this reason, CE is approximately 10 times faster than conventional slab gel electrophoresis.

The separation of oligonucleotides in DNA-sequencing samples is very challenging (for a review of the physical mechanisms of DNA electrophoresis, see Viovy, 2000). It is necessary to discriminate two fragments, which could be 100 or 1000 bases long, with only one base difference. Therefore, CE analysis must provide high separation efficiency and good selectivity. The use of CE with gel-filled capillaries for rapid separation and purification of DNA fragments has been proposed (Cohen *et al*. 1988). The first results of the use of gel-filled capillaries with laser-induced fluorescence for the separation of DNA fragments resulted in an excellent separation of more than 330 bases at single base resolution in approximately 1 h (Cohen *et al*. 1990). The method is very sensitive and has the advantage of allowing multiple injections on a single column. The applicability of capillary gel electrophoresis (CGE) to DNA-sequencing samples was demonstrated on two different instruments by Swerdlow and colleagues (Swerdlow & Gesteland, 1990; Swerdlow *et al*. 1990) and has been extensively investigated as a practical tool for DNA sequencing (Drossman *et al*. 1990; Guttman *et al*. 1990; Rocheleau & Dovichi, 1992; Luckey & Smith, 1993; Luckey *et al*. 1993; Lu *et al*. 1994). Although successful, CGE showed some features that were not compatible with high-throughput DNA sequencing, e.g. short column lifetime and injection-related problems (Swerdlow *et al*. 1992; Figeys *et al*. 1994). DNA sequencing using non-cross-linked polymer solutions was a major breakthrough introduced by Karger's group because it solved most of the CGE problems (Ruiz-Martinez *et al*. 1993). Replaceable linear polymer solutions made possible the reuse of the same capillary hundreds of times, with a fresh load of polymer solution for each sample (Salas-Solano *et al*. 1998a).

The first report on DNA sequencing by CE with replaceable linear polyacrylamide showed 350 bases in roughly 30 min (Ruiz-Martinez *et al*. 1993). Today, the sequencing rate with linear polyacrylamide is up to 1300 bases in 2 h (Zhou *et al*. 2000). However, scaling was not as straightforward as it may seem. Separation of DNA in sieving matrices is a very complex matter, and several issues had to be addressed in order to attain such results [for more details see reviews by Slater *et al*. (1998) and Quesada (1997)]. The major limitation in read-length is the onset of DNA stretching and alignment with the electric field, in which all DNA exhibits the same electrophoretic mobility, therefore losing size selectivity (Slater &

Noolandi, 1985). The 1000-base barrier to sequencing was broken after an extensive study on the separation matrix (linear polyacrylamide) composition, but the 2000 barrier seems to be extremely difficult to break, as predicted by theoretical considerations (Slater & Drouin, 1992). Experiments with polymer concentration and polymer molecular mass indicated that the larger the polyacrylamide the longer the read-length that can be obtained (Carrilho *et al*. 1996). The optimization of the separation conditions required a series of studies on temperature (Kleparnik *et al*. 1996), polyacrylamide polymerization (Goetzinger *et al*. 1998), base-calling software (Brady *et al*. 2000), sample purification (Ruiz-Martinez *et al*. 1998), and injection (Salas-Solano *et al*. 1998b). The knowledge obtained in each of these studies, when accumulated, allowed the sequencing read-length to reach 1300 bases in a single run by CE using entangled polymer solutions (Zhou *et al*. 2000).

Compared to slab gel electrophoresis, CE with polymer solutions was approximately 8–10 times faster per lane. Fortunately, this was not sufficient to compete in throughput owing to the parallel nature of slab gel instruments (which run 96 samples simultaneously) and this fact was the major driving force towards the development of a parallel CE instrument. The first instrument of capillary array electrophoresis (CAE) was introduced in 1992 by Mathies' group (Huang *et al*. 1992a, b). Over the years, several other groups developed instruments capable of fast, automated, sensitive and rugged operation (Kambara & Takahashi, 1993; Bay *et al*. 1994; Ueno & Yeung, 1994; Kheterpal *et al*. 1996; Quesada & Zhang, 1996; Madabhushi *et al*. 1997; Behr *et al*. 1999) and today four commercial companies produce seven different models of automated CAE instruments (Smith & Hinson-Smith, 2001).

CAE using polymer solutions was the technological breakthrough required for completion of the Human Genome Project many years ahead of time, and within the original budget. In fact, such technology allowed two different scientific groups to produce an initial draft of the complete sequence of the human genome early in 2001 (Venter *et al*. 2001; Lander *et al*. 2001). Obviously, the completion of the human genome does not mean that no further sequencing efforts are necessary. Indeed, the next technological development is intended to generate fast-sequencing information on microfabricated multichannel devices (microchips) in order to bring the power of sequencing analysis and diagnostics to hospitals and clinical laboratories (Carrilho, 2000). For example, an important drawback of the enzymic method is the amounts of reagents used. One of the solutions suggested to this problem was the development of a solid-phase nanoreactor directly coupled to CGE (Soper *et al*. 1998). This modification resulted in a reduction of approximately 300 times the amount of reagents used in the preparation of fragment sequences by conventional protocols. Such approaches demonstrated that the integration of sample preparation and analysis in a single microchip could decrease costs and increase speed.

3.6.2 Mass spectrometry – an alternative

Mass spectrometry (MS) has been viewed as the technique to allow the sequencing of hundreds of bases in a few seconds. Matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) MS (Karas & Hillenkamp, 1988), and electrospray ionization (ESI) MS (Fenn *et al*. 1990) are two of the most suitable MS techniques for sequencing DNA using the Sanger method. In the first, the sample is co-crystallized with an energy-absorbing compound, such as an aromatic amine or carboxylic acid. The sample-matrix mixture is hit with a pulse of laser light with the wavelength of the absorption maximum for the matrix.

The matrix vaporizes and expels the sample molecules. Through proton-exchange reactions, the matrix ionizes the sample with little or no fragmentation. Sample ions are then expelled and accelerated from the ionization chamber under the applied voltage and introduced into a field-free region (drift tube). In this tube, the sample ions fly through the evacuated tube and are separated according to the square-root of their mass-to-charge ratio. Nevertheless, even very large molecules take only few microseconds to reach the detector, making MALDI–TOF attractive for high-throughput DNA sequencing. Indeed, the first papers using MALDI–TOF for DNA sequencing were published as long ago as 1990 (Karas & Bahr, 1990; Spengler *et al*. 1990).

Electrospray of oligonucleotides was first demonstrated by Covey *et al*. (1988) with the detection of short oligomers by negative ion mode MS. Similar to MALDI, ESI was not as successful for the analysis of oligonucleotides as it was for peptides and proteins, mainly due to metal adduct formation and fragmentation. The intrinsic production of multiply charged ions by ESI creates an additional difficulty in the interpretation of the mass spectrum of mixtures.

MS is indeed a powerful tool for fast, accurate DNA sequencing, but the limitations in sensitivity and efficient ionization of large molecular sizes must be overcome before it becomes a high-throughput DNA-sequencing tool (Henry, 1997). The Human Genome Project has already been completed using electrophoretic methods, but certainly MS will be the technique of choice for probing small sequences and fragments generated by the Sanger method or mass determination of PCR fragments.

4. Maxam & Gilbert and other chemical methods

A sequencing method based on a chemical degradation was described by Maxam & Gilbert (1977). In this method, end-labelled DNA fragments are subjected to random cleavage at adenine, cytosine, guanine, or thymine positions using specific chemical agents (Table 2). The chemical attack is based on three steps: base modification, removal of the modified base from its sugar, and DNA strand breaking at that sugar position (Maxam & Gilbert, 1977). The products of these four reactions are then separated using polyacrylamide gel electrophoresis. The sequence can be easily read from the four parallel lanes in the sequencing gel (Fig. 7).

The template used in this sequencing method can be either double-stranded (ds)DNA or ssDNA from chromosomal DNA. In general, the fragments are first digested with an appropriate restriction enzyme (Maxam & Gilbert, 1980), but they can also be prepared from an inserted or rearranged DNA region (Maxam, 1980).

These DNA templates are then end-labelled on one of the strands. Originally, this labelling was done with [^{32}P]phosphate or with a nucleotide linked to ^{32}P and enzymically incorporated into the end fragment (Maxam & Gilbert, 1977). Alternatively, restriction fragments through [^{35}S]dideoxyadenosine 5′-(α-thio)triphosphate ([^{35}S]ddATPαS) and terminal deoxynucleotidyltransferase were used (Ornstein & Kashdan, 1985). These substitutions showed several advantages, including a longer lifetime, low-emission energy, increase in the autoradiograph resolution, and higher stability after labelling. Nevertheless, the use of radioactive labels is hazardous and a strategy based on a 21-mer fluorescein labelled M13 sequencing primer was therefore proposed. The fluorescent dye and its bound form to the oligonucleotide were shown to be stable during the chemical reactions used for the base-

Table 2. *Base-specific cleavage reactions*

Cleavage	Reagent
G > A[a]*	DMS followed by heating at pH 7/0·1 M alkali at 90 °C
A > G[a]*	DMS + acid/alkali
C + T[a]	Hydrazine at 20 °C
C[a]	Hydrazine + 2 M NaCl
G[b]	DMS
G + A[b]	Acid
C + T[b]	Hydrazine
C[b]	Hydrazine + salt
A > C[b]	Sodium hydroxide
G > A[b]	DMS heating at pH 7
G[c]	Methylene Blue
T[c]	Osmium tetroxide
T ⩾ G, C[d,e,f]	10^{-4} M $KMnO_4$ in H_2O
C[d]	N_2H_4–H_2O (3:1 v/v), 5 M N_2H_4.HOAc
C[d,e]	3 M NH_2OH–HCl in H_2O, pH 6·0
T > G ⩾ A, C[g]	1 M Cyclohexylamine in H_2O + UV irradiation
T[h]	1 M Spermine in H_2O + UV irradiation
G > T[h]	1 M Methylamine in H_2O + UV irradiation
T[i]	0·5 M $NaBH_4$ in H_2O, pH 8–10
T ⩾ C[i,j]	2–3 M H_2O_2 in carbonate buffer, pH 9·6
C[j]	2–3 M H_2O_2 in carbonate buffer, pH 8·3 or pH 7·4
G[c,k]	0·1% Methylene Blue + visible light
G[l,f]	4% DMS in formate buffer, pH 3·5
G ⩾ C[m]	0·3% Diethyl pyrocarbonate in cacodylate buffer, pH 8 at 90 °C
A + G[m]	0·1% Diethyl pyrocarbonate in acetate buffer, pH 5 at 90 °C
A + G[n,f]	60–80% Aqueous formic acid
A + G[e]	Citrate buffer, pH 4 at 80 °C
A + G[o]	2–3% Diphenylamine in 66% formic acid
G[p]	0·5% DMS in 50 mM cacodylate buffer, pH 8
A + G[p]	2% Diphenylamine in 66% formic acid
C + T[p]	N_2H_4–H_2O (7:4 v/v)
A[q]	K_2PdCl_4 at pH 2·0

Almost all the base-specific reactions (except *) were followed by treatment with hot aqueous piperidine. [a] Maxam & Gilbert (1977); [b] Maxam & Gilbert (1980); [c] Friedmann & Brown (1978); [d] Rubin & Schmid (1980); [e] Hudspeth *et al.* (1982); [f] Rosenthal *et al.* (1985); [g] Simoncsits & Török (1982); [h] Sugiyama *et al.* (1983); Saito *et al.* (1984); [i] Sverdlov & Kalinina (1984); [j] Sverdlov & Kalinina (1983); [k] Stalker *et al.* (1985); [l] Korobko *et al.* (1978); [m] Krayev (1981); [n] Ovchinnikov *et al.* (1979); [o] Korobko & Grachev (1977); [p] Banaszuk *et al.* (1983); [q] Iverson & Dervan (1987).

DMS, dimethyl sulphate.

specific degradations (Ansorge *et al.* 1988). For instance, fluorescein attached via a mercaptopropyl or aminopropyl linker arm to the 5′-phosphate of an oligonucleotide was described and shown to be stable during the reactions used in the chemical cleavage procedures (Rosenthal *et al.* 1990).

Another non-radioactive labelling strategy that was stable during the chemical reactions uses a biotin marker molecule chemically or enzymically attached to an oligonucleotide primer or enzymically attached to an end-filling reaction of restriction enzymes sites (Richterich, 1989). After fragment separation by direct blotting electrophoresis, the membrane-bound sequence pattern can be visualized by a streptavidin-bridged enzymic colour reaction.

An approach that made the automation of this labelling step possible was the use of PCR

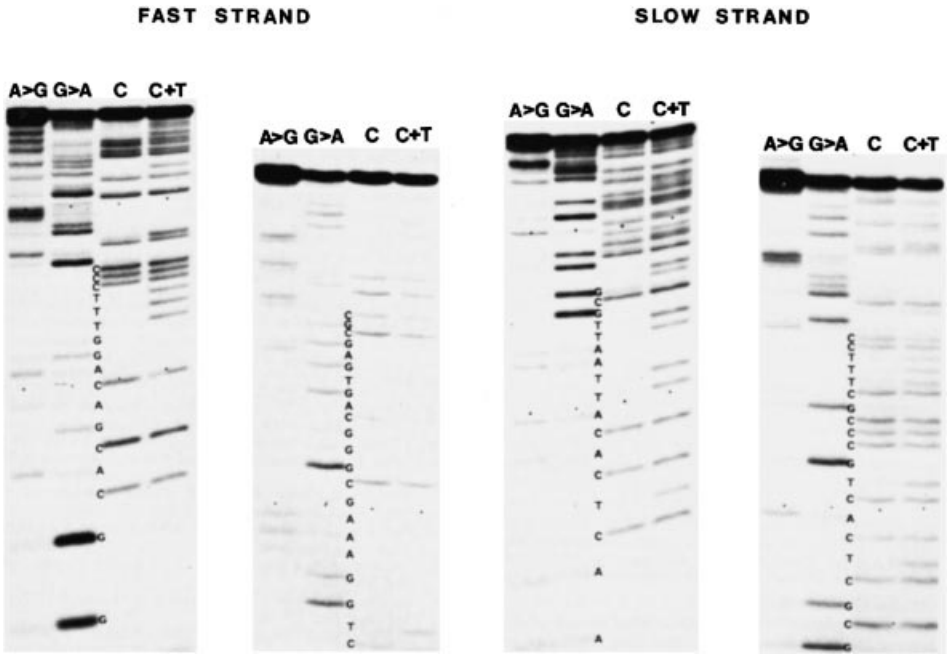


Fig. 7. Autoradiograph of a sequencing gel of the complementary strands of a 64-bp DNA fragment. Two panels, each with four reactions, are shown for each strand; cleavages proximal to the 5′-end are at the bottom left. A strong band in the first column with a weaker band in the second arises from an A; a strong band in the second column is a T. To derive the sequence of each strand, begin at the bottom of the left panel and read upwards until the bands are not resolved; then, pick up the pattern at the bottom of the right panel and continue upwards. The dimethyl sulphate treatment was 50 mM for 30 min to react with A and G; hydrazine treatment was 18 M for 30 min to react with C and T and 18 M with 2 M NaCl for 40 min to cleave C. After strand breakage, half of the products from the four reactions were layered on a 1·5 × 330 × 400 mm denaturing 20 % polyacrylamide slab gel, pre-electrophoresed at 1000 V for 2 h. Electrophoresis at 20 W (constant power), 800 V (average), and 25 mA (average) proceeded until the xylene cyanol dye had migrated halfway down the gel. Then the rest of the samples were layered and electrophoresis was continued until the new Bromphenol Blue dye moved halfway down. Autoradiography of the gel for 8 h produced the pattern shown. (Reproduced from Maxam & Gilbert, 1977.)

to amplify the products, where one of the primers was end-labelled (Nakamaye *et al*. 1988; Stamm & Longo, 1990; Tahara *et al*. 1990).

Among many dye- and fluorophore-labelling strategies, the chemiluminescent detection method showed competitive results. In this strategy, the chemically cleaved DNA fragment is transferred from a sequencing gel onto a nylon membrane. Specific sequences are then selected by hybridization to DNA oligonucleotides labelled with alkaline phosphatase or with biotin, leading directly or indirectly to the deposition of the enzyme. If a biotinylated probe is used, an incubation step with avidin-alkaline phosphatase conjugate follows. The membrane is soaked in the chemiluminescent substrate (AMPPD) and exposed to photographic film (Tizard *et al*. 1990).

Initially, all the steps of these chemical-sequencing methods were performed manually (Maxam & Gilbert, 1977, 1980). Years later, a system composed of a computer-controlled microchemical robot that carries out one of the four reactions (G, A + C, C + T, or C) in less than 2 h was described (Wada *et al*. 1983; Wada, 1984).

In order to eliminate DNA losses and to simplify the chemical reactions steps, DNA was immobilized by adsorption to DEAE paper (Whatman DE 81 paper). This method was called the simplified solid-phase technique for DNA sequencing and proved to be more efficient, much faster, and less laborious than the original method. Basically, in this solid-phase approach the end-labelled DNA fragments are adsorptively immobilized on DEAE paper, followed by specific chemical modifications and cleavage reactions (Chuvpilo & Kravchenko, 1984). However, the mechanical fragility of this support was an important drawback. This was overcome by using a new carrier medium, CCS anion-exchange paper (Whatman 540 paper activated with cyanuric chloride and then reacted with 2-bromo-ethylamine hydrobromide), which exhibited excellent stability during all operations (Rosenthal *et al*. 1985, 1986; Rosenthal, 1987). The solid-phase approach made possible the direct sequencing of fluorescently labelled amplified probes by chemical degradation, without the need for subcloning and purification steps (Voss *et al*. 1989).

This solid-phase approach is not applicable to very large DNA fragments. Thus a method based on reverse-phase chromatography (C_{18}-filled mini-columns), that works for both short and long DNA fragments, was proposed (Jagadeeswaran & Kaul, 1986). In this the DNA losses are minimized and the time-consuming steps of ethanol precipitation and lyophilization of piperidine are eliminated. Furthermore, by using solid-phase chromatography (with a modified Biomek 1000 automated workstation and glass-resin chromatography mini-columns), the authors also fully automated the Maxam–Gilbert chemical reactions (Boland *et al*. 1994).

Another solid-phase strategy was based on DNA immobilized on streptavidin-coated magnetic beads (Ohara & Ohara, 1995). An improvement was made by the use of a PCR-primer linked to biotin and fluorescein (in this order) at the 5′-end and replacement of the piperidine evaporation step with a magnetic-capture washing cycle (Ohara *et al*. 1997).

In another approach, the sequencing of phosphorothioate-linked oligonucleotides was carried out using 2-iodoethanol to cleave the sugar-phosphate backbone at thiolated sites (Polo *et al*. 1997). The fragments were then separated using MALDI–TOF MS instead of using polyacrilamide gel electrophoresis. MALDI–TOF MS was also used by other authors to separate the products of Maxam–Gilbert reactions (Isola *et al*. 1999). MALDI–TOF MS requires small sample amounts and short analysis times (< 90 s), which makes it an attractive alternative to gel electrophoresis if one is looking for short read-lengths (as discussed in Section 3.6).

The key points in the Maxam–Gilbert methods are the chemical reactions. They can be separated into two different groups: (i) four-lane methods, where four (or more) separate cleavage procedures are used (four base-specific modification protocols) and the information is displayed in four (or more) parallel gel lanes (the four original chemical reactions and some alternative reactions are shown in Table 2) and (ii) one-lane (or two-lane) method, where all reactions are based on only one chemical modification and electrophoresis is performed in a single (or two) lane(s) (see Ambrose & Pless, 1987, for a detailed comparison of one-lane methods with four-lane methods). The first report of a single-lane method was based on a chemical cleavage procedure that uses hot aqueous piperidine for several hours (Ambrose & Pless, 1985). Negri *et al*. (1991) described a two-lane method (which can become one-lane by mixing the products of the two reactions), where the labelled DNA fragment is heated in the presence of formamide. The result is an efficient cleavage of the phosphodiester bond at 3′ residues A, C and G, with relative efficiency $A = G > C$. The bias between A and G is

obtained through a pretreatment that consists of a photoreaction with Methylene Blue. In another method the DNA sequence is determined in a single electrophoretic lane by simply monitoring the intensities of the bands representing the products of cleavage at the four bases obtained by solvolysis in hot aqueous piperidine (10%) followed by treatment with hot formamide (Ferraboli *et al*. 1993). In this approach, the guanine sensitivity was increased by using inosine instead of guanosine residues (Di Mauro *et al*. 1994) and adenine sensitivity was decreased by substituting them with their diazo derivatives (Saladino *et al*. 1996). Alternatively, satisfactory results have been obtained with *N*-methylformamide in the presence of manganese (Negri *et al*. 1996).

In conclusion, the main advantages of the Maxam–Gilbert and other chemical methods compared with Sanger's chain termination reaction method are: (i) a fragment can be sequenced from the original DNA fragment, instead of from enzymic copies; (ii) no subcloning and no PCR reactions are required. Consequently, for the location of rare bases, the chemical cleavage analysis cannot be replaced by the dideoxynucleotide terminator method, as the latter analyses the DNA of interest via its complementary sequence, it can, thus, only give sequence information in terms of the four canonical bases; (iii) this method is less susceptible to mistakes with regard to sequencing of secondary structures or enzymic mistakes (Boland *et al*. 1994); (iv) some of the chemical protocols are recognized by different authors as being simple, easy to control, and the chemical distinctions between the different bases are clear (Negri *et al*. 1991).

Therefore, the chemical degradation methods have been used: (i) for genomic sequencing, where information about DNA methylation and chromatin structure could be obtained (Church & Gilbert, 1984); (ii) to confirm the accuracy of synthesized oligonucleotides or to verify the sequence of DNA regions with hairpin loops (Ornstein & Kashdan, 1985); (iii) to locate rare bases, such as Hoogsteen base-pairs (Sayers & Waring, 1993); (iv) to detect point-mutations (Ferraboli *et al*. 1993); (v) to resolve ambiguities that arise during dideoxy-sequencing (Goszczynski & McGhee, 1991); (vi) to analyse DNA–protein interactions (Isola *et al*. 1999); and (vii) to sequence short DNA fragments in general. This method, when described in 1977, had a read-length of approximately 100 nucleotides (Maxam & Gilbert, 1977). In 1980, it achieved 250 bases per assay (Maxam & Gilbert, 1980). Nowadays, with general improvements during the last few years, read-lengths close to 500 bp and automatic processing of multiple samples have been achieved (Dolan *et al*. 1995).

Despite all advantages, most of the protocols have some drawbacks. First, the chemical reactions of most protocols are slow and the use of hazardous chemicals requires special handling care. The worst problem, however, is the occurrence of incomplete reactions that decreases the read-lengths. The explanation for this is that incomplete reactions introduce electrophoretic mobility polidispersion (caused by chemical and physical inhomogeneities among the DNA chains within a given band); which enlarges the bandwidths and this in turn reduces the inter-band resolution.

5. Pyrosequencing – DNA sequencing in real time by the detection of released PPi

Pyrosequencing is a real-time DNA-sequencing method based on the detection of the PPi released during the DNA polymerization reaction (Nyrén & Lundin, 1985; Hyman, 1988;

Ronaghi *et al*. 1996). Initially, this approach was used for continuous monitoring of DNA polymerase activity (Nyrén, 1987). The cascade of enzymic reactions is shown in the diagram below:

$$(\mathrm{DNA})_n + \mathrm{dNTP} \xrightarrow{\text{DNA polymerase}} (\mathrm{DNA})_{n+1} + \mathrm{PPi}$$

$$\mathrm{PPi} + \mathrm{APS} \xrightarrow{\text{ATP sulphurylase}} \mathrm{ATP} + \mathrm{SO_4^{-2}}$$

$$\mathrm{ATP} + \text{luciferin} + \mathrm{O_2} \xrightarrow{\text{luciferase}} \mathrm{AMP} + \mathrm{PPi} + \text{oxyluciferin} + \mathrm{CO_2} + h\nu,$$

where APS stands for adenosine 5′-phosphosulphate and $h\nu$ represents a photon emitted by the bioluminescent reaction.

In the first step, each dNTP is tried in the nucleic acid polymerization reaction. PPi is released when one of the deoxynucleotides (dATP, dCTP, dGTP, or dTTP) is incorporated into the chain extension by DNA polymerase. This liberated PPi is then converted into ATP by ATP sulphurylase and light is emitted via the firefly luciferase that catalyses luciferin into oxyluciferin. The average number of emitted photons per template chain in a given step is proportional to the number of deoxynucleotides incorporated per chain at that step (this relation is linear only for a small number of incorporations). The sequence can then be determined by simply noting if incorporations occur and by counting the number of incorporations (by measuring the light intensity) in a given attempt (Fig. 8). The amount of light emitted can be measured by an avalanche photodiode, photomultiplier tube, or with a charged-coupled device camera (with or without a microchannel plate).

Currently, there are two different pyrosequencing approaches: solid-phase sequencing (Ronaghi *et al*. 1996) and liquid-phase sequencing (Ronaghi *et al*. 1998a). Solid-phase sequencing (three-enzyme mixture) requires a template-washing step between nucleotide additions to remove the non-incorporated deoxynucleotides and ATP resulting from sulphurylase action. Additionally, in this approach the template must be bound to a solid support (such as a magnetic bead) in order to avoid signal decrease. In liquid-phase sequencing (four-enzyme mixture) a nucleotide-degrading enzyme (such as apyrase) is added to eliminate the washing steps; the unreacted nucleotides and ATP produced are degraded by the enzyme.

Recently, a method using ssDNA-binding proteins was proposed. These proteins displace primers that bind non-specifically to the target DNA template, thus minimizing non-specific signals (Ronaghi, 2000). This strategy increased the efficiency of the enzymes, reduced mispriming, increased the signal intensity, yielded higher accuracy in reading the number of identical adjacent nucleotides in difficult templates, and gave read-lengths of more than 30 nucleotides.

Template preparation is a time-consuming step, because ssDNA, amplified by PCR, must be used. However, a simplified enzymic method for template preparation, which makes possible the use of dsDNA, was recently proposed (Nordstrom *et al*. 2000a, b). High-quality data have been obtained with several different enzyme combinations.

The main problem detected in all versions of pyrosequencing techniques was the interference of dATP in the detection of luminescence. This problem was solved by replacing dATP by dATPαS in the trial step. This substitution maintains efficient dATPαS

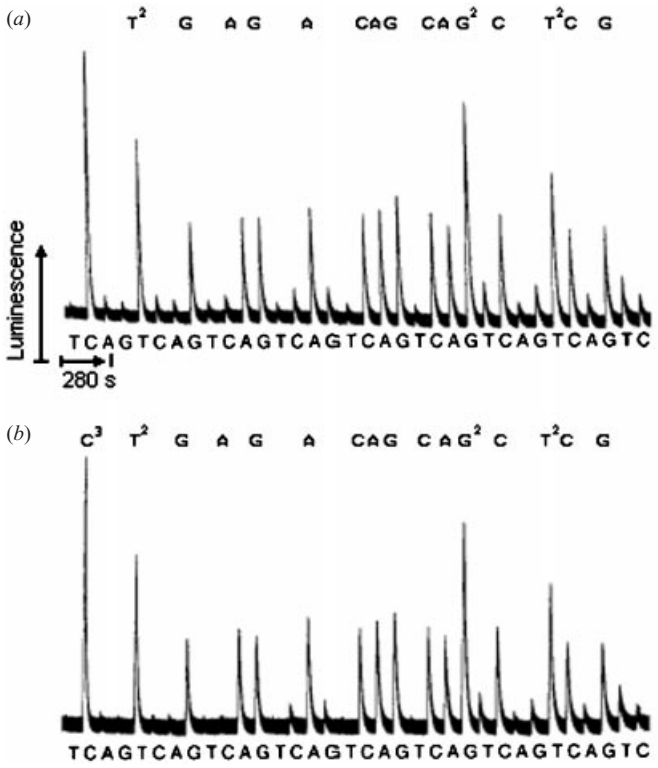


Fig. 8. Comparison between ssDNA and dsDNA as template for pyrosequencing. (*a*) Pyrosequencing was performed on 10 μl of PCR-generated template (320 bases long) enzymically treated with 1 unit shrimp alkaline phosphatase and 2 units exonuclease I for 20 min at 35 °C. (*b*) Pyrosequencing was performed on ssDNA template (320 bases long) produced using solid-phase template preparation. The order of nucleotide addition is indicated on the bottom of the traces. The correct sequence is indicated above the trace. (Reproduced from Nordström *et al*. 2000b.)

incorporation by DNA polymerase and, at the same time, reduces the background signal because dATPαS does not function as a substrate for luciferase (Ronaghi *et al*. 1996).

Pyrosequencing has shown several advantages: (i) this sequencing technique dispenses with the need for labelled primers, labelled nucleotides, and gel electrophoresis; (ii) detection is in real time with a cycle time of approximately 2 min (solid-phase); (iii) the sequencing reactions occur at room temperature and physiological pH; (iv) this method is cost-effective when compared with the traditional methods (Ronaghi, 2001); (v) the method is easily adapted for multiplexed sample processing; (vi) short chains can be satisfactorily sequenced: the signal-to-noise ratio remains relatively high even after 40 cycles (Ronaghi *et al*. 1998a).

On the other hand, this method has presented some disadvantages such as: (i) in the solid-phase approach the template must be washed completely after each nucleotide addition, resulting in a decreased signal due to loss of templates (Ronaghi *et al*. 1996); (ii) in the liquid-phase approach the apyrase activity is decreased in later cycles due to accumulation of intermediate products (Ronaghi *et al*. 1998a). Non-specific interaction between apyrase and DNA was observed; this results in loss of nucleotide-degrading activity (Ronaghi, 2000); (iii) it is difficult to determine the correct number of nucleotides incorporated into homopolymeric

regions due to a nonlinear light response following incorporation of more than five identical adjacent nucleotides. It has also been observed that this effect is less pronounced for G and C homopolymeric regions (Ronaghi *et al*. 1999). Using SSB this number increased by about 10 nucleotides (Ronaghi, 2000); (iv) contamination with PPi decreases the signal-to-noise ratio significantly, due to increased background signal. Incomplete incorporations (incomplete extensions by DNA polymerase in each nucleotide incorporation) also increase the background signal significantly and constitute the main reason for the short read-lengths obtained in this technique (Ronaghi *et al*. 1998a); (v) the fidelity of incorporation by the DNA polymerase reaction is not high enough due to the use of an exonuclease-deficient DNA polymerase. DNA polymerase with a relatively high strand-displacement activity is required to achieve a fast polymerization during the limited exposure time of nucleotide in the solution (Ronaghi *et al*. 1999). Faster polymerization also enables more efficient nucleotide incorporation, which simplifies reading the correct nucleotide number in homopolymeric regions – mainly before apyrase degradation (Ronaghi, 2000); (vi) for long sequences (and certain templates, such as GC-rich) the stability of the template and the cost/base must be improved; (vii) mispriming decreases intensity by the loss of DNA template in the reaction mixture due to fragmentation or enzymic degradation and might be eliminated by SSB addition. In order to minimize mispriming, the primer hybridization step has been eliminated using a stem-loop structure generated by PCR (Ronaghi *et al*. 1998b).

The main applications of this method includes the analyses of secondary structure, such as hairpin structures (Ronaghi *et al*. 1999), analysis of single-nucleotide polymorphisms (Ahmadian *et al*. 2000; Alderborn *et al*. 2000; Nordstrom *et al*. 2000a), mutation detection (Garcia *et al*. 2000), and *de novo* DNA sequencing for short- and medium-length DNA (Nordstrom *et al*. 2000b).

6. Single-molecule sequencing with exonuclease

Single-molecule sequencing was initially conceived as a laser-based technique that allows the fast sequencing of DNA fragments of 40 kb or more at a rate of 100–1000 bases per second (Jett *et al*. 1989). This technique is based on the detection of individual fluorescent nucleotides in a flowing sample stream (Shera *et al*. 1990; Harding & Keller, 1992). The method is divided into the following steps: fluorescent labelling of the bases in a single fragment of DNA (see Fig. 9*a*), attachment of this labelled DNA fragment onto a microsphere (Fig. 9*b*), movement of the supported DNA fragment into a flowing buffer stream, digestion of the labelled DNA with an exonuclease that sequentially cleaves the 3′-end nucleotides, and detection and identification of individual fluorescently labelled bases as they cross a focused laser beam (Fig. 9*c*) (Davis *et al*. 1991; Goodwin *et al*. 1997). Although a substantial single molecule sequencing experiment has not been performed yet, a combination of all the experimental procedures has been demonstrated (Stephan *et al*. 2001).

Since natural bases in DNA have intrinsic fluorescence quantum yields of less than 10^{-3} at room temperature, the single-molecule sequencing method requires the complete labelling of every base in one strand. Each nucleotide type must be labelled with a characteristic dye, with large fluorescence quantum yields and distinguishable spectral properties (Dörre *et al*. 1997).

After replacing all nucleotides by their fluorescent analogues, a single DNA strand is selected. A 5′-biotinylated DNA could be attached to a streptavidin-coated microsphere

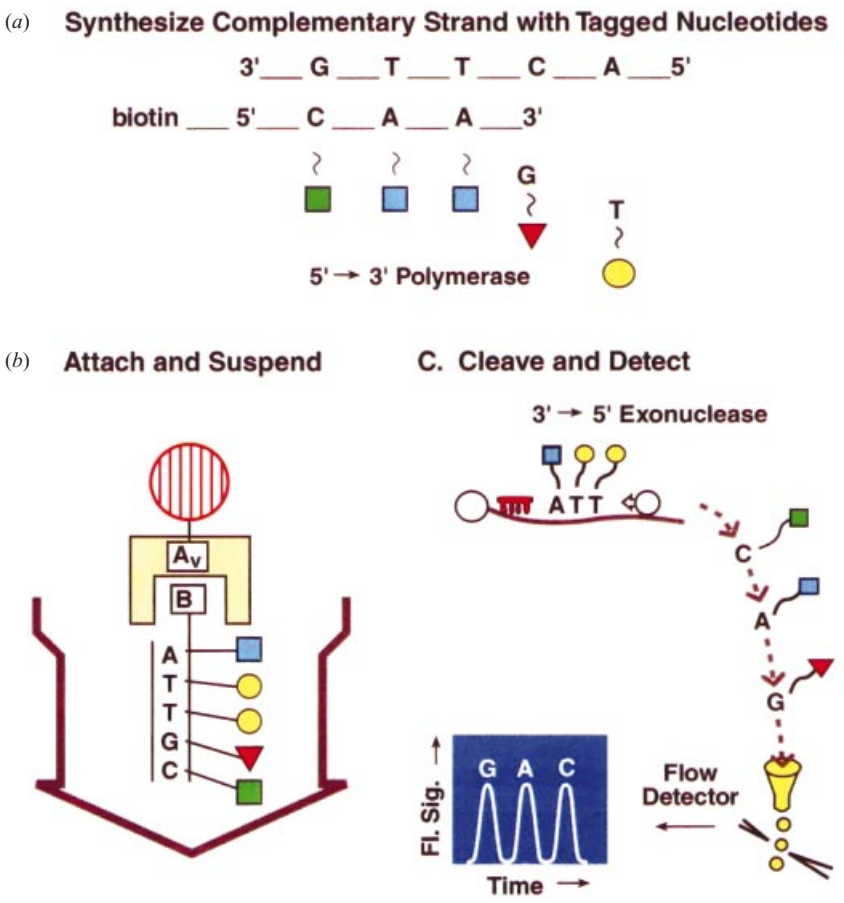


Fig. 9. Schematic representation of the steps involved in a single molecule sequencing experiment with a detection setup based on laser induced fluorescence. In (*a*) a biotinylated primer is subjected to the polymerase chain extension reaction using modified dNTPs (a different fluorophore for each N), the resulting ensemble of identical strands have one fluorophore attached to each nucleoside. In (*b*) on single strand is picked-up from the ensemble, immobilized on a microsphere or tip of a fibre, and then suspended in a flowing buffer stream. In (*c*) the resulting labelled dNMP from the strand digestion by exonuclease are sequentially detected and identified. (Reproduced from Davis *et al*. 1991.)

(Stephan *et al*. 2001). Once the fluorescently tagged DNA fragment is attached to the bead, it may either be sequenced in its double-stranded form (since only the fluorescently modified nucleotides will be detected) or it may be denatured prior to sequencing (since the fluorescent strand is attached to the bead through the biotin–streptavidin complex). Only one DNA fragment must be attached to the microsphere (Davis *et al*. 1991; Goodwin *et al*. 1997).

The addition of an exonuclease to the 3′-end of the labelled DNA fragment in the flow stream will start the sequential cleavage of the bases from the 3′- to the 5′-end of the DNA. The rate of cleavage can be adjusted by varying the exonuclease concentration, the cofactor concentration, the temperature, or by the use of inhibitors (Davis *et al*. 1991).

The detection limit is determined by the ability to distinguish each fluorescent molecule from the background. First, single-molecule sequencing was based on photon-burst

detection, where the photon bursts are correlated in time, which enables them to be distinguished from the background (Greenless *et al*. 1977). Later, time-correlated single-photon counting was used in combination with mode-locked picosecond-pulsed excitation, to allow the detection of single fluorescent molecules in the presence of significant solvent Raman and Rayleigh backgrounds (Wilkerson *et al*. 1993). The application of surface-enhanced Raman scattering has also been proposed for the detection of single DNA base molecules (Kneipp *et al*. 1998). Alternatively, a method was described that provides for detection and identification of single molecules in solution using a confocal set-up (Eigen & Rigler, 1994). A multiplex technique for identification of single fluorescent molecules in a flowing sample stream by correlated measurement of single-molecule fluorescence burst size and intraburst fluorescence decay rate has also been described (Van Orden *et al*. 1998). Weiss described how the use of single fluorescent-dye molecules attached covalently to macromolecules at specific sites can offer insight into molecular interactions (Weiss, 1999).

Since the ability to sequence large fragments of DNA is as important as speed, this approach will significantly reduce the amount of subcloning and the number of overlapping sequences required to assemble megabase segments of sequence information (Davis *et al*. 1991). The expected rate of sequencing is approximately 100–1000 bases per second, which is faster than all techniques so far described. Furthermore, this method is a powerful alternative for *de novo* sequencing of individual genomes (Stephan *et al*. 2001).

However, there are still many problems that remain to be solved. The buffer quality must be improved. A selection step must be integrated into the sequencing process. The biochemistry has to be developed to label complementary DNA strands with four different nucleotides. Finally, new polymerases as well as new exonucleases are required for rapid and efficient sequencing (Stephan *et al*. 2001).

7. Conclusion

The four best known techniques of DNA sequencing are reviewed and close to 200 references cited. These techniques are the Sanger method, the Maxam & Gilbert method, the Pyrosequencing™ method, and the method of single-molecule sequencing with exonuclease. There are good prospects for the emergence of new and non-conventional methods of DNA sequencing, which may one day revolutionize the field of DNA sequencing. Some of these candidates are methods based on atomic force microscopy, on the use of nanopores or ion channels, on quantum optics, DNA microarrays and TOF MS of aligned ssDNA fibres, among others (some of them are reviewed by Marziali & Akeson, 2001). All these new possibilities deserve a special review paper with a deep and critical analysis. Finally, we would like to apologize to the authors who have made significant contributions and that are not cited in this review.

8. Acknowledgements

The authors acknowledge the financial support from the Brazilian agencies FAPERGS (Fundação de Amparo a Pesquisa do Estado do Rio Grande do Sul, Porto Alegre, RS), FAPESP (Fundação de Amparo a Pesquisa do Estado de São Paulo, São Paulo, SP), CAPES

(Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasília, DF), and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brasília, DF). One of the authors (T. B. L. K.) acknowledges Professor Chistine Gaylarde for helpful suggestions and comments.

9. References

Adams, M. D., Fields, C. & Venter, J. C. (1996). *Automatic DNA Sequencing and Analysis*. San Diego: Academic Press.

Ahmadian, A., Gharizadeh, B., Gustafsson, A. C., Sterky, F., Nyrén, P., Uhlén, M. & Lundeberg, J. (2000). Single-nucleotide polymorphism analysis by pyrosequencing. *Analyt. Biochem.* **280**, 103–110.

Ahmed, A. (1984). Use of transposon-promoted deletions in DNA sequence analysis. *J. molec. Biol.* **178**, 941–948.

Alderborn, A., Kristoffersson, A. & Hammerling, U. (2000). Determination of single nucleotide polymorphisms by real-time pyrophosphate DNA sequencing. *Genome Res.* **10**, 1249–1258.

Ambrose, B. J. B. & Pless, R. C. (1985). Analysis of DNA sequences using a single chemical cleavage procedure. *Biochemistry* **24**, 6194–6200.

Ambrose, B. J. B. & Pless, R. C. (1987). DNA sequencing: chemical methods. *Meth. Enzym.* **152**, 522–539.

Anderson, S. (1981). Shotgun DNA sequencing using cloned DNaseI-generated fragments. *Nucleic Acids Res.* **9**, 3015–3027.

Ansorge, W., Rosenthal, A., Sproat, B., Schwager, C., Stegemann, J. & Voss, H. (1988). Non-radioactive automated sequencing of oligonucleotides by chemical degradation. *Nucleic Acids Res.* **16**, 2203–2206.

Ansorge, W., Sproat, B. S., Stegemann, J. & Schwager, C. (1986). A non-radioactive automated method for DNA sequence determination. *J. biochem. biophys. Meth.* **13**, 315–323.

Ansorge, W., Sproat, B., Stegemann, J., Schwager, C. & Zenke, M. (1987). Automated DNA sequencing: ultrasensitive detection of fluorescent bands during electrophoresis. *Nucleic Acids Res.* **15**, 4593–4602.

Banaszuk, A. M., Deugau, K. V., Sherwood, J., Michalak, M. & Glick, B. R. (1983). An efficient method for the sequence-analysis of oligodeoxyribonucleotides. *Analyt. Biochem.* **128**, 281–286.

Bay, S., Starke, H., Zhang, J. Z., Elliot, J. F., Coulson, L. D. & Dovichi, N. J. (1994). Capillary gel electrophoresis for DNA sequencing of a template from the malaria genome. *J. Capillary Electroph.* **1**, 121–126.

Beck, S. & Alderton, R. P. (1993). A strategy for amplification, purification, and selection of M13 templates for large-scale DNA sequencing. *Analyt. Biochem.* **212**, 498–505.

Beck, S., O'Keeffe, T., Coull, J. M. & Köster, H. (1989). Chemiluminescent detection of DNA: application for DNA sequencing and hybridization. *Nucleic Acids Res.* **17**, 5115–5123.

Behr, S., Matzig, M., Levin, A., Eickhoff, H. & Heller, C. (1999). A fully automated multicapillary electrophoresis device for DNA analysis. *Electrophoresis* **20**, 1492–1507.

Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983). Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. *Proc. natn. Acad. Sci. USA* **80**, 3963–3965.

Birnboim, H. C. & Doly, J. (1979). Rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**, 1513–1523.

Boland, E. J., Pillai, A., Odom, M. W. & Jagadeeswaran, P. (1994). Automation of the Maxam–Gilbert chemical sequencing reactions. *BioTechniques* **16**, 1088–1095.

Brady, D., Kocic, M., Miller, A. W. & Karger, B. L. (2000). A maximum-likelihood base caller for DNA sequencing. *IEEE Trans. biomed. Engng* **47**, 1271–1280.

Carrilho, E. (2000). DNA sequencing by capillary array electrophoresis and microfabricated array systems. *Electrophoresis* **21**, 55–65.

Carrilho, E., Ruiz-Martinez, M. C., Berka, J., Smirnov, I., Goetzinger, W., Miller, A. W., Brady, D. & Karger, B. L. (1996). Rapid DNA sequencing of more than 1000 bases per run by capillary electrophoreses using replaceable linear polyacrylamide solutions. *Analyt. Chem.* **68**, 3305–3313.

Carson, S., Cohen, A. S., Belenkii, A., Ruiz-Martinez, M. C., Berka, J. & Karger, B. L. (1993). DNA sequencing by capillary electrophoresis: use of a two-laser-two-window intensified diode array detection system. *Analyt. Chem.* **65**, 3219–3226.

Cherry, J. L., Young, H. H., Disera, L. J., Ferguson, F. M., Kimball, A. W., Dunn, D. M., Gesteland, R. F. & Weiss, R. B. (1994). Enzyme-linked fluorescent detection for automated multiplex DNA-sequencing. *Genomics* **20**, 68–74.

Church, G. M. & Gilbert, W. (1984). Genomic sequencing. *Proc. natn. Acad. Sci. USA* **81**, 1991–1995.

Chuvpilo, S. A. & Kravchenko, V. V. (1984). A simple and rapid method for sequencing DNA. *FEBS Lett.* **179**, 34–36.

Cohen, A. S. & Karger, B. L. (1987). High-performance sodium dodecyl-sulfate polyacrylamide-gel capillary electrophoresis of peptides and proteins. *J. Chromatogr.* **397**, 409–417.

Cohen, A. S., Najarian, D. R., Paulus, A., Guttman, A., Smith, J. A. & Karger, B. L. (1988). Rapid separation and purification of oligonucleotides by high-performance capillary gel electrophoresis. *Proc. natn. Acad. Sci. USA* **85**, 9660–9663.

Cohen, A. S., Najarian, D. R. & Karger, B. L. (1990). Separation and analysis of DNA sequence reaction products by capillary gel electrophoresis. *J. Chromatogr.* **516**, 49–60.

Covey, T. R., Bonner, R. F., Shushan, B. I. & Henion, J. (1988). The determination of protein, oligonucleotide and peptide molecular weights by ion-spray mass spectrometry. *Rapid Commun. Mass Spectrom.* **2**, 249–256.

Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985). A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: Application to sequencing the corn mitochondrial 18 S rDNA. *Plasmid* **13**, 31–40.

Davis, L. M., Fairfield, F. R., Harger, C. A., Jett, J. H., Keller, R. A., Hahn, J. H., Krakowski, L. A., Marrone, B. L., Martin, J. C., Nutter, H. L., Ratliff, R. L., Shera, E. B., Simpson, D. J. & Soper, S. A. (1991). Rapid DNA sequencing based upon single molecule detection. *Genetic Analysis – Biomolec. Engng* **8**, 1–7.

Debuire, B., Chabli, A. & Frenoy, N. (1993). Fast, manual, nonradioactive method for DNA sequencing. *Clin. Chem.* **39**, 1682–1685.

Deininger, P. L. (1983). Random subcloning of sonicated DNA: application to shotgun DNA sequence analysis. *Analyt. Biochem.* **129**, 216–223.

Di Mauro, E., Costanzo, G. & Negri, R. (1994). One-lane chemical sequencing of PCR amplified DNA: the use of terminal transferase and of the base analogue inosine. *Nucleic Acids Res.* **22**, 3811–3812.

Dolan, M., Ally, A., Purzycki, M. S., Gilbert, W. & Gillevet, P. M. (1995). Large-scale genomic sequencing: optimization of genomic chemical sequencing reactions. *BioTechniques* **19**, 264–237.

Dörre, K., Brakmann, S., Brinkmeier, M., Han, K. T., Riebeseel, K., Schwille, P., Stephan, J., Wetzel, T., Lapezyna, M., Stuke, M., Bader, R., Hinz, M., Seliger, H., Holm, J., Eigen, M. & Rigler, R. (1997). Techniques for single molecule sequencing. *Bioimaging* **5**, 139–152.

Douglas, A. M., Georgalis, A. M. & Atchinson, B. A. (1993). Direct sequencing of double-stranded PCR products incorporating a chemiluminescent detection procedure. *BioTechniques* **14**, 824–828.

Dovichi, N. J. (2000). Chemists' contribution. *Chem. Engng News: Letters* **78**, 10–10.

Drossman, H., Luckey, J. A., Kostichka, A. J., D'Cunha, J. & Smith, L. M. (1990). High-speed separations of DNA sequencing reactions by capillary electrophoresis. *Analyt. Chem.* **62**, 900–903.

Eigen, M. & Rigler, R. (1994). Sorting single molecules: application to diagnostics and evolutionary biotechnology. *Proc. natn. Acad. Sci. USA* **91**, 5740–5747.

Englund, P. T. (1971). Analysis of nucleotide sequences at 3′ termini of duplex deoxyribonucleic acid with the use of the T4 deoxyribonucleic acid polymerase. *J. biol. Chem.* **246**, 3269–3276.

Englund, P. T. (1972). The 3′-terminal nucleotide sequences of T7 DNA. *J. molec. Biol.* **66**, 209–224.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. (1990). Electrospray ionization-principles and practice. *Mass Spectrom. Rev.* **9**, 37–70.

Ferraboli, S., Negri, R., Di Mauro, E. & Barlati, S. (1993). One-lane chemical sequencing of 3′-fluorescent-labeled DNA. *Analyt. Biochem.* **214**, 566–570.

Figeys, D., Renborg, A. & Dovichi, N. J. (1994). Spatial and temporal depletion of ions from non-crosslinked denaturing polyacrylamide in capillary electrophoresis. *Electrophoresis* **15**, 1512–1517.

Flanagan, J. J. H., Owens, C. V., Romero, S. E., Waddell, E., Kahn, S. H., Hammer, R. P. & Soper, S. A. (1998). Near-infrared heavy-atom-modified fluorescent dyes for base-calling in DNA-sequencing applications using temporal discrimination. *Analyt. Chem.* **70**, 2676–2684.

Fleischmann, R. D., Adams, M. D., White, O., *et al.* (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–498.

Friedmann, T. & Brown, D. M. (1978). Base-specific reactions useful for DNA sequencing – methylene-blue – sensitized photo-oxidation of guanine and osmium tetraoxide modification of thymine. *Nucleic Acids Res.* **5**, 615–622.

Garcia, A. C., Ahmadian, A., Gharizadeh, B., Lundeberg, J., Ronaghi, M. & Nyrén, P. (2000). Mutation detection by pyrosequencing: sequencing of exons 5 to 8 of the p53 tumour supressor gene. *Gene* **253**, 249–257.

Garoff, H. & Ansorge, W. (1981). Improvements of DNA sequencing gels. *Analyt. Biochem.* **115**, 450–457.

Gillevet, P. M. (1990). Chemiluminescent multipex DNA sequencing. *Nature* **348**, 657–658.

Goetzinger, W., Kotler, L., Carrilho, E., Ruiz-

Martinez, M. C., Salas-Solano, O. & Karger, B. L. (1998). Characterization of high molecular mass linear polyacrylamide powder prepared by emulsion polymerization as a replaceable polymer matrix for DNA sequencing by capillary electrophoresis. *Electrophoresis* **19**, 242–248.

Goodwin, P. M., Cai, H., Jett, J. H., Ishaug-Riley, S. L., Machara, N. P., Semin, D. J., Orden, A. V. & Keller, R. A. (1997). Application of single molecule detection to DNA sequencing. *Nucleos. Nucleot.* **16**, 543–550.

Goszczynski, B. & McGhee, J. D. (1991). Resolution of sequencing ambiguities: a universal *Fok*I adapter permits Maxam–Gilbert re-sequencing of single-stranded phagemid DNA. *Gene* **104**, 71–74.

Greenless, G. W., Clark, D. L., Kaufman, S. L., Lewis, D. A., Tonn, J. F. & Broadhurst, J. H. (1977). High-resolution laser spectroscopy with minute samples. *Opt. Commun.* **23**, 236–239.

Griffin, H. G. & Griffin, A. M. (1993). DNA sequencing – recent innovations and future trends. *Appl. Biochem. Biotechnol.* **38**, 147–159.

Guttman, A., Cohen, A. S., Heiger, D. N. & Karger, B. L. (1990). Analytical and micropreparative ultra-high resolution of oligonucleotides by polyacrylamide-gel high-performance capillary electrophoresis. *Analyt. Chem.* **62**, 137–141.

Harding, J. D. & Keller, R. A. (1992). Single-molecule detection as an approach to rapid DNA sequencing. *Trends Biotechnol.* **10**, 55–57.

Henikoff, S. (1984). Unidirectional digestion with exonuclease-III creates targeted breakpoints for DNA sequencing. *Gene* **28**, 351–359.

Henry, C. (1997). Can MS really compete in the DNA world? *Analyt. Chem.* **69**, 243A–246A.

Hjerten, S., Elenbring, K., Kilar, F., Liao, J. L., Chen, A. J. C., Siebert, C. J. & Zhu, M. D. (1987). Carrier-free zone electrophoresis, displacement electrophoresis and isoelectric-focusing in a high-performance electrophoresis apparatus. *J. Chromatogr.* **403**, 47–61.

Hogdall, E., Boye, K. & Vuust, J. (1999). Simple preparation method of PCR fragments for automated DNA sequencing. *J. Cell Biochem.* **73**, 433–436.

Huang, X. C., Quesada, M. A. & Mathies, R. A. (1992a). Capillary array electrophoresis using laser-excited confocal fluorescence detection. *Analyt. Chem.* **64**, 967–972.

Huang, X. C., Quesada, M. A. & Mathies, R. A. (1992b). DNA sequencing using capillary array electrophoresis. *Analyt. Chem.* **64**, 2149–2154.

Hudspeth, M. E. S., Ainley, W. M., Shumard, D. S., Butow, R. A. & Grossman, L. I. (1982). Location and structure of the var1 gene on yeast mitochondrial-DNA – nucleotide-sequence of the 40.0 allele. *Cell* **30**, 617–626.

Hung, S. C., Ju, J., Mathies, R. A. & Glazer, A. N. (1996). Cyanine dyes with high absorption cross section as donor chromophores in energy transfer primers. *Analyt. Biochem.* **243**, 15–27.

Hyman, E. D. (1988). A new method of sequencing DNA. *Analyt. Biochem.* **174**, 423–436.

Isola, N. R., Allman, S. L., Golovlov, V. V. & Chen, C. H. (1999). Chemical cleavage sequencing of DNA using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Analyt. Chem.* **71**, 2266–2269.

Iverson, B. L. & Dervan, P. B. (1987). Adenine specific DNA chemical sequencing reaction. *Nucleic Acids Res.* **15**, 7823–7830.

Jagadeeswaran, P. & Kaul, R. K. (1986). Use of reverse-phase chromatography in the Maxam–Gilbert method of DNA sequencing. *Genet. Anal.: Tech. Appl.* **3**, 79–85.

Jett, J. H., Keller, R. A., Martin, J. C., Marrone, B. L., Moyzis, R. K., Ratliff, R. L., Seitzinger, N. K., Shera, E. B. & Stewart, C. C. (1989). High-speed DNA sequencing – an approach based upon fluorescence detection of single molecules. *J. biomolec. struct. Dyn.* **7**, 301–309.

Jorgenson, J. W. & Lukacs, K. D. (1983). Capillary zone electrophoresis. *Science* **222**, 266–272.

Ju, J., Kheterpal, I., Scherer, J. R., Ruan, C., Fuller, C. W., Glazer, A. N. & Mathies, R. A. (1995a). Design and synthesis of fluorescence energy transfer dye-labeled primers and their application for DNA sequencing and analysis. *Analyt. Biochem.* **231**, 131–140.

Ju, J., Ruan, C., Fuller, C. W., Glazer, A. N. & Mathies, R. A. (1995b). Fluorescence energy transfer dye-labeled primers for DNA sequencing and analysis. *Proc. natn. Acad. Sci. USA* **92**, 4347–4351.

Ju, J., Glazer, A. N. & Mathies, R. A. (1996). Cassette labeling for facile construction of energy transfer fluorescent primers. *Nucleic Acids Res.* **24**, 1144–1148.

Kambara, H. & Takahashi, S. (1993). Multiple-sheath flow capillary array DNA analyzer. *Nature* **361**, 565–566.

Karas, M. & Bahr, U. (1990). Laser desorption ionization mass-spectrometry of large biomolecules. *Trends Analyt. Chem.* **9**, 321–325.

Karas, M. & Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. *Analyt. Chem.* **60**, 2301–2303.

Karger, A. E., Harris, J. M. & Gesteland, R. F. (1991). Multiwavelength fluorescence detection for DNA sequencing using capillary electrophoresis. *Nucleic Acids Res.* **19**, 4955–4962.

Kheterpal, I., Li, L., Speed, T. P. & Mathies, R. A. (1998). A three-wavelength labeling approach for DNA sequencing using energy transfer primers and

capillary electrophoresis. *Electrophoresis* **19**, 1403–1414.

Kheterpal, I., Scherer, J. R., Clark, S. M., Radhakrishnan, A., Ju, J., Ginther, C. L., Sensabaugh, G. F. & Mathies, R. A. (1996). DNA sequencing using a four-color confocal fluorescence capillary array scanner. *Electrophoresis* **17**, 1852–1859.

Kieleczawa, J., Dunn, J. J. & Studier, F. W. (1992). DNA sequencing by primer walking with strings of contiguous hexamers. *Science* **258**, 1787–1791.

Klepárnik, K., Foret, F., Berka, J., Goetzinger, W., Miller, A. W. & Karger, B. L. (1996). The use of elevated column temperature to extend DNA sequencing read lengths in capillary electrophoresis with replaceable polymer matrices. *Electrophoresis* **17**, 1860–1866.

Kneipp, K., Kneipp, H., Kartha, V. B., Manoharan, R., Deinum, G., Itzkan, I., Dasari, R. R. & Feld, M. S. (1998). Detection and identification of a single DNA base molecule using surface-enhanced Raman scattering (SERS). *Phys. Rev.* (*E*) **57**, R6281–R6284.

Korobko, V. G. & Grachev, S. A. (1977). Sequence determination in DNA by a modified chemical method. *Bioorg. Khim*. **3**, 1420–1422.

Korobko, V. G., Grachev, S. A. & Kolosov, M. N. (1978). G-specific degradation of single stranded DNA – sequence determination in HAEIII restriction fragments of phage M13 DNA. *Bioorg. Khim*. **4**, 1281–1283.

Kostichka, A. J., Marchbanks, M. L., Brumley, J. R. L., Drossman, H. & Smith, L. M. (1992). High speed automated DNA sequencing in ultrathin slab gels. *Biotechnology* **10**, 78–81.

Kotler, L. E., Zevinsonkin, D., Sobolev, I. A., Beskin, A. D. & Ulanovsky, L. E. (1993). DNA sequencing – modular primers assembled from a library of hexamers or pentamers. *Proc. natn. Acad. Sci. USA* **90**, 4241–4245.

Krayev, A. S. (1981). The use of diethylpyrocarbonate for sequencing adenines and guanines in DNA. *FEBS Lett*. **130**, 19–22.

Lander, E. S., Linton, L. M., Birren, B., *et al*. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.

Lashkari, D. A., Hunicke-Smith, S. P., Norgren, R. M., Davis, R. W. & Brennan, T. (1995). An automated multiplex oligonucleotide synthesizer – development of high-throughput, low-cost DNA-synthesis. *Proc. natn. Acad. Sci. USA* **32**, 7912–7915.

Lauer, H. H. & McManigill, D. (1986). Capillary zone electrophoresis of proteins in untreated fused-silica tubing. *Analyt. Chem*. **58**, 166–170.

Lee, L. G., Spurgeon, S. L., Heiner, C. R., Benson, S. C., Rosenblum, B. B., Menchen, S. M., Graham, R. J., Constantinescu, A., Upadhya, K. G. & Cassel, J. M. (1997). New energy transfer dyes for DNA sequencing. *Nucleic Acids Res*. **25**, 2816–2822.

Lu, H., Arriaga, E., Chen, D. Y. & Dovichi, N. J. (1994). High-speed and high-accuracy DNA-sequencing by capillary gel-electrophoresis in a simple, low-cost instrument 2-color peak-height encoded sequencing at 40 °C. *J. Chromatogr*. (*A*) **680**, 497–501.

Luckey, J. A., Norris, T. B. & Smith, L. M. (1993). Analysis of resolution in DNA sequencing by capillary gel-electrophoresis. *J. phys. Chem*. **97**, 3067–3075.

Luckey, J. A. & Smith, L. M. (1993). Optimization of electric-field strength for DNA-sequencing in capillary gel-electrophoresis. *Analyt. Chem*. **65**, 2841–2850.

Madabhushi, R. S., Vainer, M., Dolnik, V., Enad, S., Barker, D. L., Harris, D. W. & Mansfield, E. S. (1997). Versatile low-viscosity sieving matrices for nondenaturing DNA separations using capillary array. *Electrophoresis* **18**, 104–111.

Martin, W. J. & Davies, W. R. (1986). Automated DNA sequencing: progress and prospects. *Biotechnology* **4**, 890–895.

Martin-Gallardo, A., McCombie, W. R. & Gocayne, J. D. (1992). Automated DNA sequencing and analysis of 106 kilobases from human. *Nat. Genet*. **1**, 34–39.

Marziali, A. & Akeson, M. (2001). New DNA sequencing methods. *Annu. Rev. biomed. Engng* **3**, 195–223.

Maxam, A. M. (1980). Sequencing the DNA of recombinant chromosomes. *Fed. Proc*. **39**, 2830–2836.

Maxam, A. M. & Gilbert, W. (1977). A new method for sequencing DNA. *Proc. natn. Acad. Sci. USA* **74**, 560–564.

Maxam, A. M. & Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Meth. Enzym*. **65**, 499–560.

Messing, J. (1983). New M13 vectors for cloning. *Meth. Enzym*. **101**, 20–78.

Metzker, M. L., Lu, J. & Gibbs, R. A. (1996). Electrophoretically uniform fluorescent dyes for automated DNA sequencing. *Science* **271**, 1420–1422.

Müller, R., Herten, D. P., Lieberwirth, U., Neumann, M., Sauer, M., Schulz, A., Siebert, S., Drexhage, K. H. & Wolfrum, J. (1997). Efficient DNA sequencing with a pulsed semicondutor laser and a new fluorescent dye set. *Chem. Phys. Lett*. **279**, 282–288.

Mullis, K. & Fallona, F. A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Meth. Enzym*. **155**, 335–350.

Mullis, K., Fallona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb. Symp*. **51**, 263–273.

Nakamaye, K. L., Gish, G., Eckstein, F. & Vosberg, H. P. (1988). Direct sequencing of polymerase chain reaction amplified DNA fragments through the incorporation of deoxynucleoside alpha-thiotriphosphates. *Nucleic Acids Res.* **16**, 9947–9959.

Negri, R., Costanzo, G. & Mauro, E. D. (1991). A single-reaction method for DNA sequence determination. *Analyt. Biochem.* **197**, 389–395.

Negri, R., Costanzo, G., Saladino, R. & Di Mauro, E. (1996). One-step, one-lane chemical DNA sequencing by N-methylformamide in the presence of metal ions. *BioTechniques* **21**, 910–917.

Nordstrom, T., Ronaghi, M., Forsberg, L., De Faire, U., Morgenstern, R. & Nyrén, P. (2000a). Direct analysis of single-nucleotide polymorphism on double-stranded DNA by pyrosequencing. *Biotechnol. Appl. Biochem.* **31**, 107–112.

Nordstrom, T., Nourizad, K., Ronaghi, M. & Nyrén, P. (2000b). Method enabling pyrosequencing on double-stranded DNA. *Analyt. Biochem.* **282**, 186–193.

Nyrén, P. (1987). Enzymatic method for continuous monitoring of DNA polymerase activity. *Analyt. Biochem.* **167**, 235–238.

Nyrén, P. & Lundin, A. (1985). Enzymatic method for continuos monitoring of inorganic pyrophosphate synthesis. *Analyt. Biochem.* **151**, 504–509.

Ohara, R. & Ohara, O. (1995). A new solid-phase chemical DNA sequencing method which uses streptavidin-coated magnetic beads. *DNA Res.* **2**, 123–128.

Ohara, R., Tanaka, A. & Ohara, O. (1997). Automated fluorescent DNA sequencing by a simplified solid-phase chemical sequencing method. *BioTechniques* **22**, 653–656.

Olesen, C. E. M., Martin, C. S. & Bronstein, I. (1993). Chemiluminescent DNA sequencing with multiplex labeling. *BioTechniques* **15**, 480–485.

Ornstein, D. L. & Kashdan, M. A. (1985). Sequencing DNA using ^{35}S-labeling: a troubleshooting guide. *BioTechniques* **3**, 476–484.

Ovchinnikov, Y. A., Guryev, S. O., Krayev, A. S., Monastyrskaya, G. S., Skryabin, K. G., Sverdlov, E. D., Zakharyev, V. M. & Bayev, A. A. (1979). Primary structure of an ecoR1 fragment of gamma-IMM434 DNA containing regions CI-CRO of phage 434 and CII-0 of phage lambda. *Gene* **6**, 235–249.

Parker, L. T., Zakeri, H., Deng, Q., Spurgeon, S., Kwok, P. Y. & Nickerson, D. A. (1996). AmpliTaq DNA polymerase, FS dye-terminator sequencing: analysis of peak height patterns. *BioTechniques* **21**, 694–699.

Polo, L. M., McCarley, T. D. & Limbach, P. A. (1997). Chemical sequencing of phosphorothioate oligonucleotides using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Analyt. Chem.* **69**, 1107–1112.

Porter, K. W., Briley, J. D. & Shaw, B. R. (1997). Direct PCR sequencing with boronated nucleotides. *Nucleic Acids Res.* **25**, 1611–1617.

Prober, J. M., Trainor, G. L., Dam, R. J., Hobbs, F. W., Robertson, C. W., Zagursky, R. J., Cocuzza, A. J., Jensen, M. A. & Baumeister, K. (1987). A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* **238**, 336–341.

Quesada, M. A. (1997). Replaceable polymers in DNA sequencing by capillary electrophoresis. *Curr. Opin. Biotechnol.* **8**, 82–93.

Quesada, M. A. & Zhang, S. (1996). Multiple capillary DNA sequencer that uses fiber-optic illumination and detection. *Electrophoresis* **17**, 1841–1851.

Richterich, P. (1989). Non-radioactive chemical sequencing of biotin labeled DNA. *Nucleic Acids Res.* **17**, 2181–2186.

Rocheleau, M. J. & Dovichi, N. J. (1992). Separation of DNA sequencing fragments at 53 bases minute by capillary gel-electrophoresis. *J. Microcol. Sep.* **4**, 449–453.

Ronaghi, M. (2000). Improved performance of pyrosequencing using single-stranded DNA-binding protein. *Analyt. Biochem.* **286**, 282–288.

Ronaghi, M. (2001). Pyrosequencing sheds light on DNA sequencing. *Genome Res.* **11**, 3–11.

Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlén, M. & Nyrén, P. (1996). Real-time DNA sequencing using detection of pyrophosphate release. *Analyt. Biochem.* **242**, 84–89.

Ronaghi, M., Uhlén, M. & Nyrén, P. (1998a). A sequencing method based on real-time pyrophosphate. *Science* **281**, 363–365.

Ronaghi, M., Petterson, B., Uhlén, M. & Nyrén, P. (1998b). PCR-introduced loop structure as primer in DNA sequencing. *BioTechniques* **25**, 876–883.

Ronaghi, M., Nygren, M., Lundeberg, J. & Nyrén, P. (1999). Analyses of secondary structures in DNA by pyrosequencing. *Analyt. Biochem.* **267**, 65–71.

Rosenblum, B. B., Lee, L. G., Spurgeon, S. L., Khan, S. H., Menchen, S. M., Heiner, C. R. & Chen, S. M. (1997). New dye-labeled terminators for improved DNA sequencing patterns. *Nucleic Acids Res.* **25**, 4500–4504.

Rosenthal, A. (1987). Sequencing of synthetic DNA fragments containing various 5-substituted pyrimidines by solid-phase chemical degradation using CCS paper. *Nucleos. Nucleot.* **6**, 419–420.

Rosenthal, A., Jung, R. & Hunger, H. D. (1986). Optimized conditions for solid-phase sequencing: simultaneous chemical cleavage of a series of long DNA fragments immobilized on CCS anion-exchange paper. *Gene* **42**, 1–9.

Rosenthal, A., Schwertner, S., Hahn, V. & Hunger, H. D. (1985). Solid-phase methods for

sequencing of nucleic acids I. Simultaneous sequencing of different oligodeoxyribonucleotides using a new, mechanically stable anion-exchange paper. *Nucleic Acids Res.* **13**, 1173–1184.

Rosenthal, A., Sproat, B., Voss, H., Stegemann, J., Schwager, C., Erfle, H., Zimmermann, J., Coutelle, C. & Ansorge, W. (1990). Automated sequencing of fluorescently labeled DNA by chemical degradation. *DNA Sequence* **1**, 63–71.

Rubin, C. M. & Schmid, C. W. (1980). Pyrimidine-specific chemical reactions useful for DNA sequencing. *Nucleic Acids Res.* **8**, 4613–4619.

Ruiz-Martinez, M. C., Berka, J., Belenkii, A., Foret, F., Miller, A. W. & Karger, B. L. (1993). DNA sequencing by capillary electrophoresis with replaceable linear polyacrylamide and laser-induced fluorescence detection. *Analyt. Chem.* **65**, 2851–2858.

Ruiz-Martinez, M. C., Carrilho, E., Berka, J., Kieleczawa, J., Miller, A. W., Foret, F., Carson, S. & Karger, B. L. (1996). DNA sequencing by capillary electrophoresis using short oligonucleotide primer libraries. *BioTechniques* **20**, 1058–1069.

Ruiz-Martinez, M. C., Salas-Solano, O., Carrilho, E., Kotler, L. & Karger, B. L. (1998). A sample purification method for rugged and high-performance DNA sequencing by capillary electrophoresis using replaceable polymer solutions. A. Development of the cleanup protocol. *Analyt. Chem.* **70**, 1516–1527.

Saito, I., Sugiyama, H., Matsuura, T., Ueda, K. & Komano, T. (1984). A new procedure for determining thymine residues in DNA sequencing – photoinduced cleavage of DNA fragments in the presence of spermine. *Nucleic Acids Res.* **12**, 2879–2885.

Saladino, R., Mincione, E., Crestini, C., Negri, R., Di Mauro, E. & Costanzo, G. (1996). Mechanism of degradation of purine nucleosides by formamide. Implications for chemical DNA sequencing procedures. *J. Am. Chem. Soc.* **118**, 5615–5619.

Salas-Solano, O., Carrilho, E., Kotler, L., Miller, A. W., Goetzinger, W., Sosic, Z. & Karger, B. L. (1998a). Routine DNA sequencing of 1000 bases in less than one hour by capillary electrophoresis with replaceable linear polyacrylamide solutions. *Analyt. Chem.* **70**, 3996–4003.

Salas-Solano, O., Ruiz-Martinez, M. C., Carrilho, E., Kotler, L. & Karger, B. L. (1998b). A sample purification method for rugged and high-performance DNA sequencing by capillary electrophoresis using replaceable polymer solutions. B. Quantitative determination of the role of sample matrix components on sequencing analysis. *Analyt. Chem.* **70**, 1528–1535.

Sanger, F. & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. molec. Biol.* **94**, 441–448.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. natn. Acad. Sci. USA* **74**, 5463–5467.

Sayers, E. W. & Waring, M. J. (1993). Footprinting titration studies on the binding of echinomycin to DNA incapable of forming Hoogsteen base-pairs. *Biochemistry* **32**, 9094–9107.

Simoncsits, A. & Török, I. (1982). A photoinduced cleavage of DNA useful for determining T-residues. *Nucleic Acids Res.* **10**, 7959–7964.

Shera, E. B., Seitzinger, N. K., Davis, L. M., Keller, R. A. & Soper, S. A. (1990). Detection of single fluorescent molecules. *Chem. Phys. Lett.* **174**, 553–557.

Slater, G. W. & Drouin, G. (1992). Why can we not sequence thousands of DNA bases on a polyacrylamide-gel. *Electrophoresis* **13**, 574–582.

Slater, G. W., Kist, T. B. L., Ren, H. J. & Drouin, G. (1998). Recent developments in DNA electrophoretic separations. *Electrophoresis* **19**, 1525–1541.

Slater, G. W. & Noolandi, J. (1985). New biased-reptation model for charged polymers. *Phys. Rev. Lett.* **55**, 1579–1582.

Smith, J. P. & Hinson-Smith, V. (2001). DNA sequencers rely on CE. *Analyt. Chem.* **73**, 327A–331A.

Smith, L. M., Fung, S., Hunkapiller, M. W., Hunkapiller, T. J. & Hood, L. E. (1985). The synthesis of oligonucleotides containing an aliphatic amino group at the 5′ terminus: synthesis of fluorescent DNA primers for use in DNA sequence analysis. *Nucleic Acids Res.* **13**, 2399–2412.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H. & Hood, L. E. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature* **321**, 674–679.

Soper, S. A., Williams, D. C., Xu, Y., Lassiter, S. J., Zhang, Y., Ford, S. M. & Bruch, R. C. (1998). Sanger DNA-sequencing reactions performed in a solid-phase nanoreactor directly coupled to capillary gel electrophoresis. *Analyt. Chem.* **70**, 4036–4043.

Spengler, B., Pan, Y., Cotter, R. J. & Kan, L. S. (1990). Molecular weight determination of underivatized oligodeoxyribonucleotides by positive-ion matrix-assisted ultraviolet laser-desorption mass spectrometry. *Rapid Commun. Mass Spectrom.* **4**, 99–102.

Stalker, D. M., Hiatt, W. R. & Comai, L. (1985). A single amino-acid substitution in the enzyme 5-enolpyruvylshikimate-3-phosphate synthase confers resistance to the herbicide glyphosate. *J. biol. Chem.* **260**, 4724–4728.

Stamm, S. & Longo, F. M. (1990). Direct sequencing of PCR products using the Maxam–Gilbert method. *Genet. Anal.: Tech. Appl.* **7**, 142–143.

Stephan, J., Dörre, K., Brakmann, S., Winkler, T., Wetzel, T., Lapezyna, M., Stuke, M., Angerer, B.,

Ankenbauer, W., Földes-Papp, Z., Rigler, R. & Eigen, M. (2001). Towards a general procedure for sequencing single DNA molecules. *J. Biotechnol.* **86**, 255–267.

Studier, F. W. (1989). A strategy for high-volume sequencing of cosmid DNAs – random and directed priming with a library of oligonucleotides. *Proc. natn. Acad. Sci. USA* **86**, 6917–6921.

Sugiyama, H., Saito, I., Matsuura, T., Ueda, K. & Komano, T. (1983). A new, convenient method for determining T residues in chemical DNA sequencing by using photoreaction with spermine. *Nucleic Acids Symp. Ser.* **12**, 103–106.

Sverdlov, E. D. & Kalinina, N. F. (1983). DNA interaction with hydrogen-peroxide – a method for determining pyrimidine-bases in DNA. *Bioorg. Khim.* **9**, 1696–1698.

Sverdlov, E. D. & Kalinina, N. F. (1984). Chemical modifications of double DNA as a method of detection of exposed individual base-pairs. *Dokl. Akad. Nauk SSSR* **274**, 1508.

Swerdlow, H., Dew-Jager, K. E., Brady, K., Grey, R., Dovichi, N. J. & Gesteland, R. (1992). Stability of capillary gels for automated sequencing of DNA. *Electrophoresis* **13**, 475–483.

Swerdlow, H. & Gesteland, R. (1990). Capillary gel electrophoresis for rapid, high resolution DNA sequencing. *Nucleic Acids Res.* **18**, 1415–1419.

Swerdlow, H., Wu, S., Harke, H. & Dovichi, N. J. (1990). Capillary gel electrophoresis for DNA sequencing. *J. Chromatogr.* **516**, 61–67.

Tabor, S., Huber, H. E. & Richardson, C. C. (1987). *Escherichia coli* thioredoxin confers processivity on the DNA-polymerase activity of the gene-5 protein of bacteriophage-T7. *J. biol. Chem.* **262**, 16212–16223.

Tabor, S. & Richardson, C. C. (1987). DNA-sequence analysis with a modified bacteriophage-T7 DNA-polymerase. *Proc. natn. Acad. Sci. USA* **84**, 4767–4771.

Tabor, S. & Richardson, C. C. (1995). A single residue in DNA-polymerases of the *Escherichia coli* DNA-polymerase-I family is critical for distinguishing between deoxyribonucleotides and dideoxyribonucleotides. *Proc. natn. Acad. Sci. USA* **92**, 6339–6343.

Tahara, T., Kraus, J. P. & Rosenberg, L. E. (1990). Direct DNA sequencing of PCR amplified genomic DNA by the Maxam–Gilbert method. *BioTechniques* **8**, 366–368.

Tizard, R., Cate, R. L., Ramachandran, K. L., Wysk, M., Voyta, J. C., Murphy, O. J. & Bronstein, I. (1990). Imaging of DNA sequences with chemiluminescence. *Proc. natn. Acad. Sci. USA* **87**, 4514–4518.

Ueno, K. & Yeung, E. S. (1994). Simultaneous monitoring of DNA fragments separated by electrophoresis in a multiplexed array of 100 capillaries. *Analyt. Chem.* **66**, 1424–1431.

Van den Boom, D., Ruppert, A., Jurinke, C. & Köster, H. (1997). Combined amplification and sequencing in a single reaction using two DNA polymerases with differential incorporation rates for dideoxynucleotides. *J. biochem. biophys. Meth.* **35**, 69–79.

Van den Boom, D., Jurinke, C., Ruppert, A. & Köster, H. (1998). Forward and reverse DNA sequencing in a single reaction. *Analyt. Biochem.* **256**, 127–129.

Van Orden, A., Machara, N. P., Goodwin, P. M. & Keller, R. A. (1998). Single-molecule identification in flowing sample streams by fluorescence burst size and intraburst fluorescence decay rate. *Analyt. Chem.* **70**, 1444–1451.

Venter, J. C., Adams, M. D., Myers, E. W., *et al.* (2001). The sequence of the Human Genome. *Science* **291**, 1304–1351.

Venter, J. C., Smith, H. O. & Hood, L. (1996). A new strategy for genome sequencing. *Nature* **381**, 364–366.

Viovy, J.-L. (2000). Electrophoresis of DNA and other polyelectrolytes: physical mechanisms. *Rev. Mod. Phys.* **72**, 813–872.

Voss, H., Schwager, C., Wirkner, U., Sproat, B., Zimmermann, J., Rosenthal, A., Erfle, H., Stegemann, J. & Ansorge, W. (1989). Direct genomic fluorescent on-line sequencing and analysis using in vitro amplification of DNA. *Nucleic Acids Res.* **17**, 2517–2527.

Voss, H., Wiemann, S., Grothues, D., Sensen, C., Zimmermann, J., Schwager, C., Stegemann, J., Erfle, H., Rupp, T. & Ansorge, W. (1993). Automated low-redundancy large-scale DNA-sequencing by primer walking. *BioTechniques* **15**, 714–721.

Wada, A. (1984). Automatic DNA sequencing. *Nature* **307**, 193.

Wada, A., Yamamoto, M. & Soeda, E. (1983). Automatic DNA sequencer: computer-programmed microchemical manipulator for the Maxam–Gilbert sequencing method. *Rev. Sci. Instrum.* **54**, 1569–1572.

Wang, Y., Ju, J., Carpenter, B. A., Atherton, J. M., Sensabaugh, G. F. & Mathies, R. A. (1995). Rapid sizing of short tandem repeat alleles using capillary array electrophoresis and energy-transfer fluorescent primers. *Analyt. Chem.* **67**, 1197–1203.

Weiss, S. (1999). Fluorescence spectroscopy of single biomolecules. *Science* **283**, 1676–1683.

Werle, E., Schneider, C., Renner, M., Völker, M. & Fiehn, W. (1994). Convenient single-step, one tube purification of PCR products for direct sequencing. *Nucleic Acids Res.* **22**, 4354–4355.

Wiemann, S., Schilke, A., Rechmann, S., Zimmermann, J., Voss, H. & Ansorge, W. (1996). Reducing 'double sequences' in automated DNA

sequencing with T7 DNA poymerase and internal labeling. *BioTechniques* **20**, 791–792.

Wilkerson, C. W., Goodwin, P. M., Ambrose, W. P., Martin, J. C. & Keller, R. A. (1993). Detection and lifetime measurement of single molecules in flowing sample streams by laser-induced fluorescence. *Appl. Phys. Lett.* **62**, 2030–2032.

Zakeri, H., Amparo, G., Chen, S. M., Spurgeon, S. & Kwok, P. Y. (1998). Peak height pattern in dichlororhodamine and energy transfer dye terminator sequencing. *BioTechniques* **25**, 406–414.

Zhou, H., Miller, A. W., Sosic, Z., Buchholz, B., Barron, A. E., Kotler, L. & Karger, B. L. (2000). DNA sequencing up to 1300 bases in two hours by capillary electrophoresis with mixed replaceable linear polyacrylamide solutions. *Analyt. Chem.* **72**, 1045–1052.