*Science* 17 July 1998:
Vol. 281. no. 5375, pp. 363 – 365
DOI: 10.1126/science.281.5375.363

TECH.SIGHT

DNA SEQUENCING:

# A Sequencing Method Based on Real-Time Pyrophosphate

**Mostafa Ronaghi, Mathias Uhlén, and Pål Nyrén***

DNA sequencing is one of the most important technologies in bioscience today. Whole-genome approaches (1) and human expressed sequence tag (EST) sequencing (2) have started to exert profound in□uence on biology and medicine. The need for robust, high-throughput methods to replace the elegant Sanger method, described more than 20 years ago (3), has led to the development of several new principles, such as array methods based on sequencing by hybridization (4). New applications, such as population-based biodiversity projects and brute-force genotyping using single-nucleotide polymorphism, make such efforts even more urgent, in particular, for simple and robust methods for sequencing short "tags" (1 to 20 bases) such as ESTs or biallelic markers and methods suitable for routine diagnostic applications.

Sequencing-by-synthesis is based on the detection of nucleotide incorporation, using a primer-directed polymerase extension. The sequence can be deduced iteratively (5). During the last decade, many researchers, including the groups of Rosenthal (6), Gibbs (7), and Jones (8), described various protocols based on □uorescently labeled nucleotides. The level of incorporation of these □uorescent nucleotides is low, however, as shown by Metzker *et al*. (7), and therefore, the protocols only permit detection of a few bases.

Recently, Ronaghi *et al*. (9) showed that natural nucleotides can be used to obtain efficient incorporation during a sequencing-by-synthesis protocol. The detection was based on the pyrophosphate (PPi) released during the DNA polymerase reaction, the quantitative conversion of pyrophosphate to ATP by sulfurylase, and the subsequent production of visible light by □re□y luciferase. However, this PPi-based sequencing method is not without drawbacks: The template must be washed

thoroughly between nucleotide additions to remove unincorporated nucleotides. Also, templates not bound to a solid support are difficult to sequence, and the addition of new enzymes to each cycle of deoxynucleotide (dATP, dTTP, dGTP, and dCTP) is required.

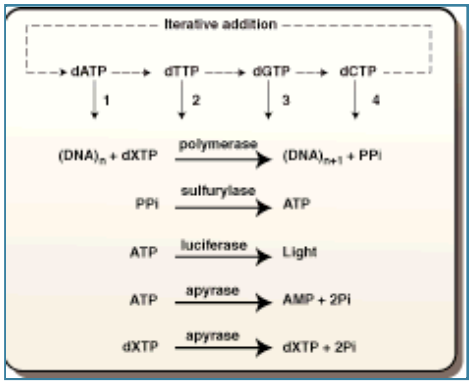


**Fig. 1.** In the new DNA sequencing method, four nucleotides are added stepwise to the template hybridized to a primer. The PPi released in the DNA polymerase–catalyzed reaction is detected by the ATP sulfurylase and luciferase in a coupled reaction. The added nucleotides are continuously degraded by a nucleotide-degrading enzyme. After the □rst added nucleotide has been degraded, the next nucleotide can be added. As this procedure is repeated, longer stretches of the template sequence are deduced. dXTP, one of the four nucleotides.

Here, we address these problems by a modi□cation in which the sequencing cycles can be performed without intermediate washing steps. This is achieved by the addition of a nucleotide-degrading enzyme to obtain a four-enzyme mixture. The principle of pyrosequencing is outlined in Fig. 1. The DNA fragment of interest (sequencing primer hybridized to a single-stranded DNA template) is incubated with DNA polymerase, ATP sulfurylase, □re□y luciferase, and a nucleotide-degrading enzyme (such as apyrase). Repeated cycles of deoxynucleotide addition are performed. A deoxynucleotide will only be incorporated into the growing DNA strand if it is complementary to the base in the template strand. The synthesis of DNA is accompanied by release of PPi equal in molarity to that of the incorporated deoxynucleotide. Thereby, real-time signals are obtained by the enzymatic inorganic pyrophosphate detection assay (10). In this assay the released PPi is converted to ATP by ATP sulfurylase and the concentration of ATP is then sensed by the luciferase. The amount of light produced in the luciferase-catalyzed reaction can readily be estimated by a suitable light-sensitive device such as a luminometer or a CCD (charge-coupled device) camera. Unincorporated deoxynucleotides and the produced ATP are degraded between each cycle by the nucleotide-degrading enzyme. The nucleotide-degrading enzyme must possess the following properties: First, the enzyme must hydrolyze all deoxynucleotide triphosphate at approximately the same rate. This includes the α-thio-dATP, which is used instead of dATP to improve the background in sequencing reactions (9). Second, it should also hydrolyze ATP to prevent the accumulation of ATP between cycles. Third, the time for nucleotide degradation by the nucleotide-degrading enzyme must be slower than nucleotide incorporation by the polymerase. Obviously, these two enzymes compete for the same substrate, and it is important that the yield of primer-directed incorporation is as close to 100% as possible before the nucleotide-degrading enzyme can degrade the nucleotide to a concentration below the $K_M$ for the

faster than the rate of ATP hydrolysis to obtain ATP concentrations and light production in proportion to the amount of PPi released.

To optimize the assay, several parameters were tested by using synthetic oligonucleotides as template (11). The assay solution consisted of the template–primer plus varying amounts of the four enzymes (12). The protocol consisted of simply adding a new nucleotide every other minute in an iterative manner and detecting the visible light produced. In Fig. 2 (Left), an example of the results is shown. Clear specific signals can be observed with low background. Note the higher signals obtained at cycles 21, 23, and 24 when two nucleotides were incorporated because of the presence of two identical adjacent bases in the template. Similar high-quality sequencing results were obtained for other oligonucleotide templates (13).

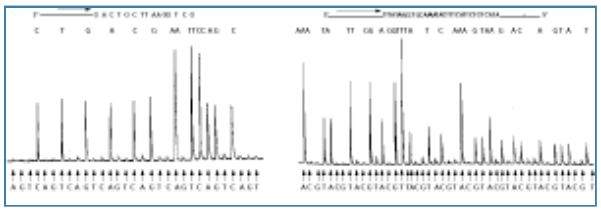

**Fig. 2.** Pyrosequencing was performed on a 35-base-long oligonucleotide template (**Left**) and a 130-base-long PCR product (**Right**). About 2 pmol of the template–primer was used in the assay. The reaction was started by the addition of the indicated deoxynucleotide, and the PPi released was detected by the described method. The DNA sequence after the primer is indicated.

For direct sequencing of polymerase chain reaction (PCR) products, single-stranded template was obtained by biotin capture on magnetic beads (14). In Fig. 2 (Right), the amount of light produced during approximately 40 subsequent cycles is shown. Sequence data covering 34 bases is obtained (15), and the signal-to-noise ratio remains relatively high even at 40 cycles. The low signals in the early cycles (2, 3, 6, and 7) show that the misincorporation of bases, which would produce PPi and consequently light, is not a major problem, despite the presence of "wrong" noncomplementary nucleotides. The background does increase during the later cycles, which is not surprising because relatively crude enzyme preparations were used in this work. Low amounts of contaminating enzymes such as exonucleases or kinases will give rise to nonsynchronized extensions on some templates, causing increased background and lower specific signal. Obviously, more work is needed to purify these enzymes from present contaminants. In addition, apyrase activity is decreased in later cycles, which is because of accumulation of intermediate products (such as deoxynucleoside diphosphate, or dNDP) and eventually undegraded dNTP. Removal of these nucleotides by enzymes such as nonspecific nucleoside diphosphatases (16) will increase the efficiency of apyrase in degradation of nucleoside triphophates and thereby allow longer reads. However, it is reassuring that the nonoptimal enzyme mixture used here allows accurate determination of more than 20 bases for many different PCR products tested (13).

An inherent problem with the described method is the difficulty in determining the number of incorporated nucleotides in homopolymeric regions due to the nonlinear light response following incorporation of more than three or four identical nucleotides. This can be demonstrated by the relatively low signal at cycle 16 (Fig. 2, Right) when four T bases are incorporated. However, this nonlinear response can

tag-sequencing applications, such as brute-force EST-sequencing, biallelic marker analysis, and confirmatory sequencing, this problem is not a major concern, because the number of bases, if present, will be known.

With this method, parallel processing of large numbers of samples can easily be envisioned with the use of high-density microtiter plates and microinjector technology. An automated instrument has recently been developed based on the precise delivery of submicroliter volumes of the four nucleotides by "ink-jet" technology into a microtiter plate coupled with simultaneous detection of all samples by a single CCD unit (17). Together with a robot (17) performing single-strand template preparation (from double-stranded PCR products), ready for pyrosequencing, it would be possible to analyze thousands of samples daily with little manual intervention.

**References and Notes**


1. R. D. Fleischmann *et al.*, *Science* **269**, 496 (1995); C. M. Fraser *et al.*, *ibid.* **270**, 397 (1995); C. J. Bult *et al.*, *ibid.* **273**, 1058 (1996); J. Tomb *et al.*, *Nature* **388**, 539 (1997); H. W. Mewes *et al.*, *ibid.* **387**, 7 (1997).
2. V. E. Velculescu, L. Zhang, B. Vogelstein, K. W. Kinzler, *Science* **270**, 484 (1995).
3. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
4. R. Drmanac, I. Labat, I. Brukner, R. Crkvenjakov, *Genomics* **4**, 114 (1989); H. Köster *et al.*, *Nature Biotechnol.* **14**, 1123 (1996).
5. E. D. Hyman, *Anal. Biochem.* **174**, 423 (1988).
6. A. Rosenthal, *International Patent Application Publication* 761107 (1989).
7. M. L. Metzker *et al.*, *Nucleic Acids Res.* **22**, 4259 (1994).
8. D. H. Jones, *Biotechniques* **22**, 938 (1997).
9. M. Ronaghi, S. Karamohamed, B. Pettersson, M. Uhlén, P. Nyrén, *Anal. Biochem.* **242**, 84 (1996).
10. P. Nyrén and A. Lundin, *ibid.* **151**, 504 (1985).
11. The oligonucleotides E3PN (5¢-GCTGGAATTCGTCAGACTGGCCGTCGTTTTACAAC-3¢), NUSPT (5¢-GTAAAACGACGGCCAGT-3¢), and JA80 (5¢-GATGGAAACCAAAAATGATAGG-3¢) were synthesized by phosphoramidite chemistry (Interactiva).
12. The oligonucleotide E3PN and the PCR product generated from cloned HIV-V3 were used as templates for DNA sequencing. The oligonucleotides and single-stranded PCR product were hybridized to the primers NUSPT and JA80, respectively. The hybridized DNA fragments were incubated with exo$^-$ Klenow or exo$^-$ T7 DNA polymerase (Sequenase 2.0), respectively (Amersham). The sequencing procedure was carried out by stepwise elongation of the primer-strand upon sequential addition of the different deoxynucleoside triphosphates and simultaneous degradation of nucleotides by apyrase (nucleoside 5¢-triphosphatase and nucleoside 5¢-diphosphatase; EC 3.6.1.5) (Sigma). The sequencing reaction was performed at room temperature and was started by adding a specific amount of one of the deoxynucleotides. The PPi released due to nucleotide incorporation was detected as described (**9**).
13. M. Ronaghi and P. Nyrén, data not shown.
14. The biotinylated PCR products were immobilized onto streptavidin-coated super paramagnetic beads [Dynabeads M280-Streptavidin (Dynal)]. Elution of single-

described (9).

15. The sequencing data obtained from the pyrosequencing method was con□rmed by semiautomated solid-phase Sanger sequencing (18).
16. H. D. Doremus and D. G. Belvins, *Plant Physiol.* **87**, 36 (1988).
17. The pyrosequencing instrument was based on a cassette containing the four separate nucleotides on an x-ray robotic arm (B. Ekström, M. Ronaghi, T. Nordström, P. Nyrén, M. Uhlén, unpublished data). The sample preparation robot was based on streptavidin-coated magnetic particles for PCR-capture and handling (A. Holmberg and M. Uhlén, unpublished data).
18. T. Hultman, S. Ståhl, E. Hornes, M. Uhlén, *Nucleic Acids Res.* **17**, 4937 (1989).
19. We thank K. Nourizad for very helpful technical assistance and A. Scott for critical review. Supported by grants from PyroSequencing AB and the Swedish Research Council for Engineering Science (TFR).



---

*The authors are in the Department of Biochemistry and Biotechnology, The Royal Institute of Technology, SE-10044 Stockholm, Sweden. E-mail: paaln@biochem.kth.se*

**To whom correspondence should be addressed.*