ANALYTICAL BIOCHEMISTRY **242,** 84–89 (1996)
ARTICLE NO. 0432

Real-Time DNA Sequencing Using Detection of Pyrophosphate Release

Mostafa Ronaghi, Samer Karamohamed, Bertil Pettersson, Mathias Uhlén, and Pål Nyrén
Department of Biochemistry, Royal Institute of Technology, S-100 44 Stockholm, Sweden

Received May 6, 1996

An approach for real-time DNA sequencing without the need for electrophoresis has been developed. The approach relies on the detection of DNA polymerase activity by an enzymatic luminometric inorganic pyrophosphate (PP_i) detection assay (ELIDA) (Nyrén, P. (1987) *Anal. Biochem.* 167, 235–238). The PP_i formed in the DNA polymerase reaction is converted to ATP by ATP sulfurylase and the ATP production is continuously monitored by the firefly luciferase. In the sequencing procedure, immobilized single-stranded template was used in a repeated cycle of deoxynucleotide extension. Real-time signals in the ELIDA, proportional to the amount of incorporated nucleotide, were observed when complementary bases were incorporated. An increased signal-to-noise ratio was obtained by substitution of deoxyadenosine α-thiotriphosphate (dATPαS) for the natural deoxyadenosine triphosphate, dATPαS is efficiently used by the DNA polymerase, but is not recognized by the luciferase. As a model, 15 bases of a single-stranded PCR product were sequenced. The possibility for parallel processing of many samples in an automated manner is discussed.

One of the largest challenges in the field of genomic research is to develop new methods for DNA sequencing. The need for simple, fast, and automated methods for sequence-based analysis is evident. This demand for robust and high-throughput protocols is also great for clinical applications in the field of genetic diseases, cancer diagnosis, and analysis of infectious diseases. Although the enzymatic chain termination method (1) has recently been improved greatly for high-throughput sequencing (2), the method still relies on gel electrophoresis to separate the formed DNA fragments according to their size. Alternative strategies have therefore been initiated, most notably scanning tunnel electron microscopy (3), sequencing by hybridization (4, 5), sequencing by mass spectrometry (6), and single molecule detection (7). Another alternative approach for sequence-based analysis is to detect the incorporation of each nucleotide during the extension reaction by a DNA polymerase. This procedure, often called sequencing by synthesis (8), offers the advantage of real-time detection.

Nyrén has earlier (9) described the use of an enzymatic system consisting of DNA polymerase, ATP sulfurylase, and luciferase to couple the release of pyrophosphate obtained when a nucleotide is incorporated by the polymerase with light emission, which can be easily detected by a luminometer. Nyrén *et al.* (10) showed that pyrophosphate release could be used for minisequencing where dideoxynucleotide incorporation was detected after adding all four nucleotides. However, using this method, only one base position could be analyzed at each time.

Here, we investigated if real-time sequencing by synthesis can be achieved by a proper choice of enzyme and substrate in a solid-phase format. We also show that substituting dATP with deoxyadenosine α-thiotriphosphate (dATPαS)[1] allows efficient incorporation by the polymerase with low background signal due to absence of interaction between dATPαS and luciferase.

MATERIALS AND METHODS

Synthesis and Purification of Oligonucleotides

The oligonucleotides E2PN (55-mer: 5′-CGACGATCTGAGGTCATAGCTGTTTCCTGTGTGAACTGGCCGTCGTTTTACAACG-3′), E3PN (35-mer: 5′-GCTGGAATTCGTCAGACTGGCCGTCGTTTTACAAC-3′), NUSPT (5′-GTAAAACGACGGCCAGT-3′), RIT 28, RIT 29, and USP (11) were synthesized by phosphoramidite chemistry on an automated DNA synthesis ap-

[1] Abbreviations used: dATPαS, deoxyadenosine α-thiotriphosphate; ELIDA, enzymatic luminometric inorganic pyrophosphate; APS, adenosine 5′-phosphosulfate.

0003-2697/96 $18.00

paratus (Gene Assembler Plus, Pharmacia Biotech, Uppsala, Sweden). Purification was performed on a fast protein liquid chromatography pepRPC 5/5 column (Pharmacia, Biotech, Uppsala, Sweden).

In Vitro Amplification and Template Preparation

PCR reactions were performed on the multilinker of plasmid pRIT 28 with 7.5 pmol of general primers, RIT 28 and RIT 29 (biotinylated), according to Hultman *et al.* (11). The biotinylated PCR products were immobilized onto streptavidin-coated super paramagnetic beads Dynabeads M280-Streptavidin or M450-Streptavidin (Dynal A.S., Oslo, Norway). Single-stranded DNA was obtained by removing the supernatant after incubation of the immobilized PCR product in 0.10 M NaOH for 5 min. Washing of the immobilized single-stranded DNA and hybridization to sequencing primers was carried out as described earlier (10).

Real-Time DNA Sequencing

The oligonucleotide E3PN and the above described PCR product were used as templates for real-time DNA sequencing. The oligonucleotide E3PN was immobilized onto streptavidin-coated super paramagnetic beads (Dynabeads M280-Streptavidin or M450-Streptavidin) as described above, and a primer was hybridized to the immobilized template. The immobilized DNA fragments were incubated with either a modified T7 DNA polymerase (Sequenase 2.0; U.S. Biochemical, Cleveland, OH), Klenow DNA polymerase (Pharmacia, Biotech, Uppsala, Sweden) or exonuclease-deficient (exo$^-$) Klenow DNA polymerase (Amersham, UK). The sequencing procedure was carried out by stepwise elongation of the primer strand upon sequential addition of the different deoxynucleoside triphosphates (Pharmacia, Biotech). Washing of the immobilized DNA fragments between each nucleotide addition was performed in two steps: first with a buffer containing 10 mM Tris–HCl (pH 7.5), 0.25 M NaCl, 0.1% Tween 20, and then with 10 mM Tris–acetate (pH 7.5). The PP_i released due to nucleotide incorporation was detected by the ELIDA (9). The luminescence was measured using an LKB 1250 luminometer connected to a potentiometric recorder. The luminometer was calibrated to give a response of 10 mV for the internal light standard. The luminescence output was calibrated by the addition of a known amount of ATP or PP_i. The standard assay volume was 0.2 ml and contained the following components: 0.1 M Tris–acetate (pH 7.75), 2 mM EDTA, 10 mM magnesium acetate, 0.1% bovine serum albumin, 1 mM dithiothreitol, 5 μM adenosine 5′-phosphosulfate (APS), 0.4 mg/ml polyvinylpyrrolidone (360 000), 100 μg/ml D-luciferin (BioOrbit, Finland), 4 μg/ml L-luciferin (BioOrbit, Finland), 0.3 U/ml ATP sulfurylase (ATP:sulfate adenylyltransferase; EC 2.7.7.4) (Sigma Chemical Co., St. Louis, MO), purified luciferase (Sigma Chemical Co.) in an amount giving a response of 200 mV for 0.1 μM ATP. One picomole of the immobilized DNA fragment and 3 pmol DNA polymerase were added to the solution described above. The sequencing reaction was started by adding 40 pmol of one of the nucleotides (Pharmacia, Biotech). The reaction was carried out at room temperature. When the effect of dATP and dATPαS on the luciferase reaction was studied, both APS and ATP sulfurylase were omitted from the assay.

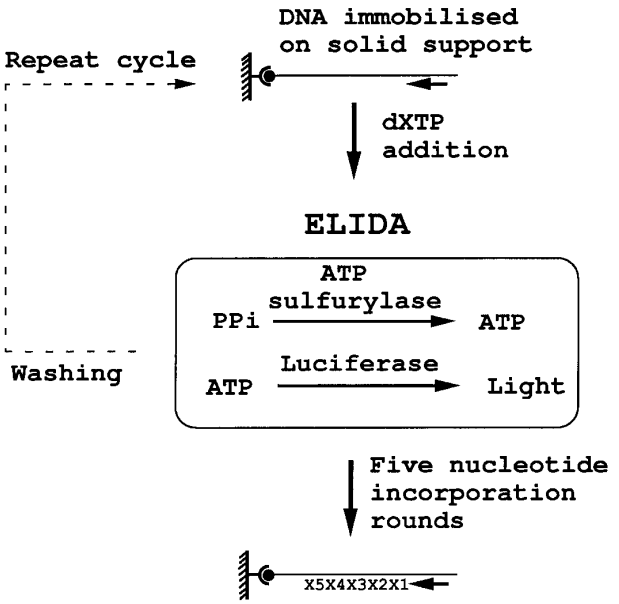

FIG. 1. A schematic representation of the real-time DNA sequencing method. The four different nucleotides are added stepwise to the immobilized template hybridized to a primer. The PP_i released in the DNA polymerase-catalyzed reaction is detected by the ATP sulfurylase- and luciferase-catalyzed reactions. The height of the signal is proportional to the number of bases which have been incorporated. After each base addition a washing step is performed. These steps are repeated in a cycle and the sequence of the template is deduced. See the text for further details.

Semiautomated Solid-Phase DNA Sequencing

The sequence data obtained from the real-time DNA sequencing were confirmed by semiautomated solid-phase sequencing (12).

RESULTS

Principle of the Sequencing Method

The principle of the sequencing method is illustrated in Fig. 1. A specific DNA fragment of interest is immobilized onto a solid support (e.g., by biotin/streptavidin coupling) and subsequently converted into single-stranded form. A sequencing primer is hybridized to

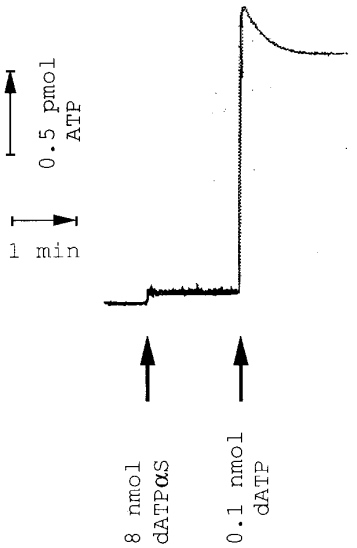

FIG. 2. The effect of dATP and dATPαS on the luciferase reaction. The 0.1 nmol dATP and 8 nmol dATPαS were added as indicated and the luminescence output was detected. The experimental conditions were as described under Materials and Methods.

the single-stranded DNA, and a repeated cycle of deoxynucleotide incubation and washing is performed. The synthesis of DNA is accompanied by release of PP_i equal in molarity to that of the incorporated nucleotide. Thereby, real-time signals are obtained by the ELIDA only when complementary bases are incorporated. In the ELIDA the produced PP_i is converted to ATP by ATP sulfurylase and the amount of ATP is then determined by the luciferase assay (Fig. 1). From the ELIDA results the sequence after the primer is deduced.

Effect of dATP and dATPαS on the Luciferase System

We have earlier shown that a luminescence system can be used for monitoring DNA polymerase activity (9). However, we also observed that dATP interfered with the detection system. This interference is a major problem when the method is used to detect a single-base incorporation event. Several approaches to decrease this background activity were tested (data not shown) and among those the largest positive effect was achieved by replacing the natural dATP with dATPαS. Figure 2 shows the results of using dATP and dATPαS during the luciferase assay. An addition of 0.1 nmol dATP induced an instantaneous increase in the light emission followed by a slow decrease until it reached a steady-state level. The steady-state level increase in light emission after adding dATP corresponds to 1–2% of the emission from an equivalent addition of ATP. This effect of dATP makes it impossible to start a sequencing reaction by adding dATP; the reaction must instead be started by addition of DNA polymerase. The signal-to-noise ratio will also become higher for dATP compared to the other nucleotides. On the other hand, addition of 8 nmol dATPαS (80-fold higher amount than dATP) had only a minor effect on the luciferase (Fig. 2). From Fig. 2 it can be deduced that dATPαS is less than 0.05% as effective as dATP as a substrate for luciferase. According to these results there is therefore a great advantage in using dATPαS instead of dATP if a DNA polymerase that can accept this nucleotide is used in the sequencing by synthesis protocol.

Solid-Phase Technique

Several different parameters important for the success of the real-time DNA sequencing approach were optimized in a model system using two different synthetic DNA templates. To simplify sequencing of several bases, the DNA was immobilized on a solid-phase. Here we have used two types of streptavidin-coated super paramagnetic beads from Dynal: M280 and M450. Both types of beads have a high binding capacity. We found that the larger beads (M450) allow a faster washing procedure due to their higher sedimentation rate (data not shown). To eliminate blunt-end DNA polymerase activity (13), sequence primers annealing at least one base inside from the 3′ end of the template were chosen. In Fig. 3, a single-base incorporation event is shown for two different concentrations of primer/template. The reactions were started by addition of the next correct base (dCTP) and the traces show the release of PP_i during the incorporation of the base. No release of PP_i was observed if a noncomplementary base was added (data not shown). In the subsequent experiments 1 pmol of primer/template was used and

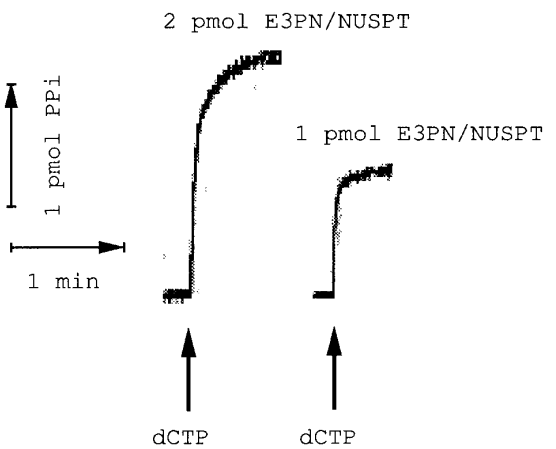

FIG. 3. The extent of PP_i synthesis as a function of template concentration. Three picomoles (exo$^-$) Klenow was incubated with 1 or 2 pmol E3PN/NUSPT as indicated. The reactions were started by the addition of 40 pmol dCTP. The PP_i released were detected by the ELIDA. The experimental conditions were as described under Materials and Methods.

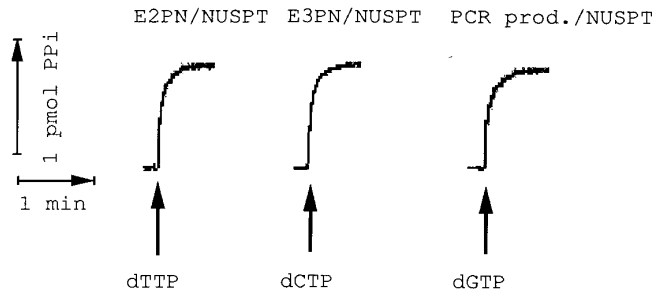

FIG. 4. Real-time detection of one base incorporation on three different templates. The 1.5 pmol of the indicated templates was incubated with 3 pmol of (exo$^-$) Klenow. The reactions were started by addition of 40 pmol of the indicated deoxynucleotide and the PP_i released was detected by the ELIDA. The experimental conditions were as described under Materials and Methods.

the relevant signal difference was recorded (Fig. 3). Both the initial rate and the extent of PP_i formed in the ELIDA are proportional to the DNA concentration within a broad interval (10). The upper limit for the assay is 200 pmol PP_i formed (1 μM) (14). The lower limit is mainly determined by the volume used, and by contamination of PP_i and ATP in the different solutions.

Effect of DNA Polymerase Concentration

In the next series of experiments the effect of polymerase concentration on the sequencing procedure was studied. We found that it was important to use an excess of polymerase over primer/template to be sure that all free 3′ ends were extended. At lower polymerase concentrations biphasic kinetics (a fast phase followed by a slower phase) was observed (data not shown). The amplitude of the fast phase is stoichiometric with the amount of enzyme present, and the slow phase is the same as the rate of steady-state incorporation. The rate-limiting steps for the slow phase are the dissociation of the polymerase from the extended primer/template and the subsequent binding to a nonextended primer/template. The incorporation rate as a function of nucleotide concentration was also studied. We observed a K_m, for one base incorporation, of 0.2 and 0.4 μM for Klenow and Sequenase 2.0, respectively (not shown). The latter results are in accordance with data from the literature (15).

Real-Time DNA Sequencing

Different synthetic templates as well as a PCR product were sequenced in order to investigate the feasibility of the new approach. Extension of one base on three different primer/templates is shown in Fig. 4. Both the rate and extent (slope and height of the signals) of nucleotide incorporation were similar for all three types of templates tested. In Fig. 5 real-time DNA sequencing of 15 bases of a 291-base-long single-stranded PCR product is shown. The sequencing procedure was started by addition of dATPαS. No PP_i release due to base incorporation was detected in the ELIDA. The small signal observed is due to PP_i contamination in the nucleotide solution. After a washing step (the polymerase reaction was allowed to proceed about 1 min before the washing was started), dGTP was added; a signal corresponding to incorporation of one residue was observed. The next base added was dCTP; a signal corresponding to incorporation of two identical residues was now detected. The subsequent addition of dTTP gave no signal. Thereafter, dATPαS was added again. This time the incorporation of two identical residues was noted. The latter detected incorporation confirmed earlier observations (16) that dATPαS is efficiently incorporated into the primer/template by Klenow polymerase. A signal corresponding to incorporation of one residue was obtained after the next addition which was dGTP. By continuing this cyclic procedure further information about the sequence was obtained. It is important to note that enough nucleotides must be added to allow longer extensions when there is a stretch of identical residues. The sequencing procedures were repeated several times on the same template with the same result. The decrease in signal due to loss and aggregation of beads during the washing procedure (measured by the decrease in optical density) has been compensated for in Fig. 5. The loss was lower for the M450 beads (about 2% of the beads were lost per wash) than for the M280. The obtained sequence was confirmed by semiautomated solid-phase Sanger sequencing (data not shown).

DISCUSSION

The results presented here describe a new approach for DNA sequencing with no need for electrophoresis or labeled primers. The sequencing reactions are continuously monitored in real-time. A procedure for rapid detection of pyrophosphate release can be envisioned. The ELIDA reactions have been estimated to take less than 2 s. (14). The rate-limiting step is the conversion of PP_i to ATP by ATP sulfurylase, while the luciferase reaction is fast and takes less than 0.2 s. The incorporation rate for the Klenow DNA polymerase has been estimated by different methods (17–19). Complete incorporation of one base takes less than 0.5 s; thus, the estimated total time for detection of the incorporation of one base with the ELIDA is approximately 3 s. In this study, the real-time sequencing procedure was performed at room temperature. However, by using a more thermostable luciferase (20) it should be possible to perform the reactions at higher temperatures.

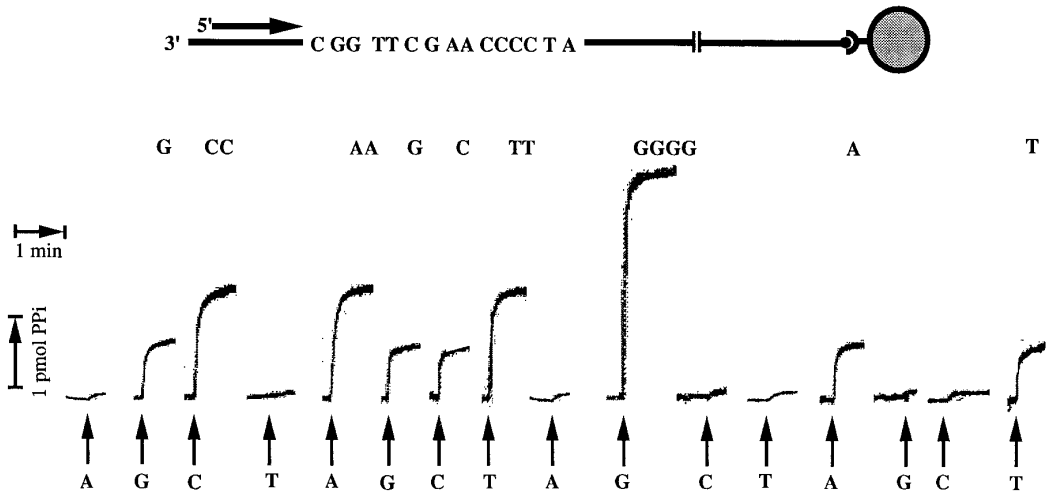

FIG. 5. Real-time DNA sequencing performed on a 291-base-long PCR-generated single-stranded template immobilized on streptavidin-coated paramagnetic beads. About 0.8 pmol of the template/primer (NUSPT) was incubated with 3 pmol (exo$^-$) Klenow. The reaction was started by the addition of 40 pmol of the indicated deoxynucleotide and the PP_i released was detected by the ELIDA. Between each nucleotide addition the beads were washed. The given ELIDA signals are compensated for the loss of beads during the washing procedures. The DNA sequence after the primer, as confirmed by semiautomated solid-phase DNA sequencing, is inserted in the figure. The experimental conditions were as described under Materials and Methods. See the text for further details.

In this way the reaction time might be further decreased.

The real-time pyrophosphate detection opens up the possibility for an automated approach for large-scale, nonelectrophoretic, solid-phase DNA sequencing. The procedure allows for continuous measurement of the progress of the polymerization reaction with time. Obviously, there is a need for a DNA polymerase with high efficiency in each extension step due to the rapid increase of background signal if templates which are not fully extended accumulate. A high fidelity in each step is also desired. A very high fidelity can be achieved by using polymerases with exonuclease activity. However, this has the disadvantage that primer degradation can be obtained. Although the exonuclease activity of the Klenow polymerase is low, we have found that the 3′ end of the primer was degraded with longer incubations in the absence of nucleotides. An induced-fit binding mechanism in the polymerization step (21) selects very efficiently for binding of the correct dNTP with a net contribution toward fidelity of 10^5–10^6. Exonuclease-deficient polymerases, such as (exo$^-$) Klenow or Sequenase 2.0, catalyzed incorporation of a nucleotide was only observed when the complementary dNTP was present, confirming a high fidelity of these enzymes even in the absence of proofreading exonuclease activity. The main advantage of using (exo$^-$) Klenow DNA polymerase over Sequenase 2.0 is its lower K_m for nucleotides, allowing a high rate of nucleotide incorporation even at low nucleotide concentrations.

In the new DNA sequencing approach one of the natural deoxynucleotides, dATP, was replaced by dATPαS. The reason for this replacement is that, in contrast to dATP, dATPαS does not interfere with the luciferase assay (Fig. 2). There are also other advantages to substituting dATPαS for dATP. For Klenow there is an elemental effect of 13 for the misincorporation of dATPαS into a primer/template and 2 for the correct incorporation (22). A strong elemental effect also exists for the other dNTPαSs (23, 24), and thus it is likely that the misincorporation rate could be decreased considerably by substituting all α-thiotriphosphate analogs for all natural nucleotides. For the shorter manually performed real-time sequencing presented in this paper no difference was observed when all dNTPs were replaced by dNTPαSs, but for longer sequencing projects the advantage might be more pronounced. In addition, by replacing all dNTPs by dNTPαSs it may be possible to use DNA polymerases with exonuclease activity (25).

Although the manually performed real-time sequencing works well for shorter sequencing projects, automation is necessary for larger projects. For a larger project both the stability of the template and the cost per base sequenced must be improved. For instance, by immobilization of the DNA template in a capillary the template loss observed for the paramagnetic beads could be avoided. In addition, in a flow system, with small volumes, high speed and low cost can be obtained.

In summary, we have successfully combined solid-phase technology (DNA bound to super paramagnetic

beads) with a sensitive luminometric assay system to obtain a sequencing by synthesis protocol. The simplicity of the method renders it suitable for routine analysis of PCR material. The new approach has several advantages compared to standard sequencing methods (i) the method is suitable for handling of multiple samples in parallel, (ii) relatively cost-effective instruments can be envisioned, and (iii) the method avoids the use of electrophoresis and thereby the loading of samples and casting of gels.

ACKNOWLEDGMENTS

This work was supported by grants from the Carl Tryggers stiftelse för Vetenskaplig Forskning, Magnus Bergvalls Stiftelse, Riksbankens Jubileumsfond, and the Swedish Board for Technical Development.

REFERENCES

1. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
2. Adams, M. D., Kerlavage, A. R., Kelley, J. M., Gocayne, J. D., Fields, C., Fraser, C. M., and Venter, J. C. (1994) *Nature* **368,** 474–475.
3. Driscoll, R. J., Youngquist, M. G., and Baldeschwieler, J. D. (1990) *Nature* **346,** 294–296.
4. Bains, W., and Smith, G. C. (1988) *J. Theor. Biol.* **135,** 303–307.
5. Drmanac, R., Labat, I., Brukner, I., and Crkvenjakov, R. (1989) *Genomics* **4,** 114–128.
6. Jacobson, K. B., Arlinghaus, H. F., Buchanan, M. V., Chen, C-H., Glish, G. L., Hettich, R. L., and McLuckey, S. A. (1991) *GATA* **8,** 223–229.
7. Jett, J. H., Keller, R. A., Martin, J. C., Marrone, B. L., Moyzis, R. K., Ratliff, R. L., Seitzinger, N. K., Shera, E. B., and Stewart, C. C. (1989) *Biomol. Struct. Dynamics* **7,** 301–306.
8. Hyman, E. D. (1988) *Anal. Biochem.* **174,** 423–436.
9. Nyrén, P. (1987) *Anal. Biochem.* **167,** 235–238.
10. Nyrén, P., Pettersson, B., and Uhlén, M. (1993) *Anal. Biochem.* **208,** 171–175.
11. Hultman, T., Murby, M., Ståhl, S., Hornes, E., and Uhlén, M. (1990) *Nucleic Acids Res.* **18,** 5107–5112.
12. Hultman, T., Bergh, S., Moks, T., and Uhlén, M. (1991) *BioTechniques* **10,** 84–93.
13. Clak, J. M. (1991) *Gene* **104,** 75–80.
14. Nyrén, P., and Lundin, A. (1985) *Anal. Biochem.* **151,** 504–509.
15. Van Draanen, N. A., Tucker, S. C., Boyd, F. L., Trotter, B. W., and Reardon, J. E. (1992) *J. Biol. Chem.* **267,** 25019–25024.
16. Vosberg, H. P., and Eckstein, F. (1977) *Biochemistry* **16,** 3633–3640.
17. Frey, M. W., Sowers, L. C., Millar, D. P., and Benkovic, S. J. (1995) *Biochemistry* **34,** 9185–9192.
18. Dahlberg, M. E., and Benkovic, S. J. (1991) *Biochemistry* **30,** 4835–4843.
19. Kuchta, R. D., Benkovic, P., and Benkovic, S. J. (1988) *Biochemistry* **27,** 6716–6725.
20. Kaliyama, N., and Nakano, E. (1994) *Biosci. Biotechnol. Biochem.* **58,** 1170–1171.
21. Wong, I., Patel, S. S., and Johnson, K. A. (1991) *Biochemistry* **30,** 526–537.
22. Benkovic, S., and Camron, C. E. (1995) *Methods Enzymol.* **262,** 257–269.
23. Patel, S. S., Wong, I., and Johnson, K. A. (1991) *Biochemistry* **30,** 511–525.
24. Wong, I., Patel, S. S., and Johnson, K. A. (1991) *Biochemistry* **30,** 526–537.
25. Gupta, A. P., Benkovic, P. A., and Benkovic, S. J. (1984) *Nucleic Acids Res.* **12,** 5897–5903.