PROGRESS IN BIOMEDICAL OPTICS AND IMAGING
Vol. 2, No. 21
ISSN 1605-7422

Genomics and Proteomics Technologies

Ramesh Raghavachari
Weihong Tan
Chairs/Editors

22–23 January 2001
San Jose, USA

Proceedings of SPIE
Volume 4264

PROGRESS IN BIOMEDICAL OPTICS AND IMAGING
Vol. 2, No. 21

Genomics and Proteomics Technologies

Ramesh Raghavachari
Weihong Tan
Chairs/Editors

22–23 January 2001
San Jose, USA

Sponsored and Published by
SPIE—The International Society for Optical Engineering

Proceedings of SPIE
Volume 4264

The papers appearing in this book compose the proceedings of the technical conference cited on the cover and title page of this volume. They reflect the authors' opinions and are published as presented, in the interests of timely dissemination. Their inclusion in this publication does not necessarily constitute endorsement by the editors or by SPIE. Papers were selected by the conference program committee to be presented in oral or poster format, and were subject to review by volume editors or program committees.

Please use the following format to cite material from this book:

Author(s), "Title of paper," in *Genomics and Proteomics Technologies*, Ramesh Raghavachari, Weihong Tan, Editors, Proceedings of SPIE Vol. 4264, page numbers (2001).

ISSN 1605-7422
ISBN 0-8194-3942-8

Published by
SPIE—The International Society for Optical Engineering
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone 1 360/676-3290 (Pacific Time) • Fax 1 360/647-1445
http://www.spie.org/

Printed in the United States of America.

Conference Committee

Conference Chairs

Ramesh Raghavachari, Corning Inc. (USA)
Weihong Tan, University of Florida (USA)

Program Committee

John Babiak, Pharmacopeia (USA)
Rex M. Bitner, Promega Corporation (USA)
Xiaohong Fang, University of Florida (USA)
Robin A. Felder, University of Virginia (USA)
Brij M. Khorana, Rose-Hulman Institute of Technology (USA)
Patrick A. Limbach, Louisiana State University (USA)
Shuming Nie, Indiana University (USA)
Daniel J. Simpson, Promega Corporation (USA)
Eiichi Tamiya, Japan Advanced Institute of Science and Technology

Session Chairs

Advances in Nuleic Acid Chemistries and Relevant Technologies
Ramesh Raghavachari, Corning Inc. (USA)

Proteomics
Rex M. Bitner, Promega Corporation (USA)

Mass Spectroscopy and Imaging
Patrick A. Limbach, Louisiana State University (USA)

Fluorescence Applications and Probe Development in Genomics and Proteomics
Daniel J. Simpson, Promega Corporation (USA)

Advanced Assay Technologies for Drug Discovery
Robin A. Felder, University of Virginia (USA)
John Babiak, Pharmacopeia (USA)

Molecular Beacon DNA Probe: Development and Applications
Xiaohong Fang, University of Florida (USA)

Contents

Novel fluorescent nonnucleoside triphosphates as terminators of enzyme-directed DNA synthesis

Stephen C. Roemer*[a], Craig M. Johnson[b], Vince R. Boveia[b], Philip R. Buzby[c], James J. DiMeo[c], Dan Draney[b], Narasimhachari Narayanan[b], D. Michael Olive[b]
[a]Orchid BioSciences, Inc., Princeton, NJ; [b]LI-COR, Inc., Lincoln, NE;
[c]PerkinElmer Life Sciences, Boston, MA

ABSTRACT

Near infrared (NIR) fluorescent acycloterminators were tested as substrates in the Sanger enzymatic method of DNA sequencing. The acyclic triphosphates of adenosine, uridine, guanosine, and cytidine (AcyNTP) were labeled with a heptamethine carbocyanine dye via a propargylamino linker to the purine or pyrimidine base. Dye-labeled AcyNTPs which are lacking in the sugar moiety positions equivalent to the C-2 and C-3 of the ribose functioned similarly to chain-terminating dideoxynucleotides (ddNTPs). These fluorescent nonnucleotide analogs were incorporated by a mutant, thermostable Taq DNA polymerase with the same efficacy and fidelity as traditional ddNTPs. Sequence read length and base-calling accuracy were comparable for both dye-acycloterminator and dye-primer sequencing methods. In two primer walking projects, cycle sequencing with fluorescent AcyNTPs achieved a mean sequence read length of 1,090 bases with 99.1% accuracy at one kilobase read length. The cyanine dye-labeled acycloterminators produced electropherograms in which weak T peaks follow G peaks. In cases of polymorphism, such peak height variability may make it difficult to distinguish the presence or absence of a heterozygote at a specific site.

Keywords: acycloterminators, DNA sequencing, near infrared fluorescence, nonnucleoside triphosphates

1. INTRODUCTION

Terminators of enzyme-directed DNA synthesis are pivotal to the Sanger method of DNA sequencing[1]. Specifically, a DNA polymerase extends an oligonucleotide-primed DNA template by incorporating 2'-deoxyribonucleoside-5'-triphosphates (dNTPs) complementary to the template. Synthesis of new DNA continues until the reaction is halted by the occasional inclusion of a chain terminating nucleotide. The most familiar terminators used in Sanger sequencing are the 2',3'-dideoxyribonucleoside-5'-triphosphates (ddNTPs). These nucleotide analogs are incapable of further chain elongation since the ribose moiety of the ddNTP lacks the 3'-hydroxyl necessary for forming a phosphodiester bond with the next incoming dNTP. The resulting mixture of partially extended DNA sequencing fragments is analyzed by gel electrophoresis using either radioactive or fluorescent labels to reveal the underlying sequence[2].

Early reports of automated DNA sequence analysis using dideoxynucleotides labeled with fluorescein and rhodamine dye derivatives revealed that T7 DNA polymerase, AMV reverse transcriptase, and Taq DNA polymerase incorporate these fluorescent analogs with concomitant termination of DNA synthesis following proper Watson-Crick base pairing[3,4]. One disadvantage of dye-terminator chemistry is that for any given enzyme dye-labeled ddNTPs are incorporated less efficiently than dNTPs, and their incorporation is dependent on the local sequence context leading to variable band intensities or uneven peak heights in sequence chromatograms[4-6]. However, wild-type Taq DNA polymerase was engineered to remove discrimination against ddNTPs[7,8], and thus improve the rate of incorporation of the dye-labeled analogs[9]. Recently, three new dye classes (cyanine, dichlororhodamine, and energy-transfer BigDye™) were used to tag dideoxynucleotides in order to improve both the accuracy of base-calling and the ability to identify mutations based on peak height differences[10-12].

Other terminators of enzyme-directed DNA synthesis have been proposed for Sanger-based DNA

*Corresponding author; Orchid BioSciences, 303 College Road East, Princeton, NJ 08540; Tel: 609-750-2337; Fax: 609-750-2250; Email: sroemer@orchid.com

Genomics and Proteomics Technologies, Ramesh Raghavachari, Weihong Tan, Editors,
Proceedings of SPIE Vol. 4264 (2001) © 2001 SPIE · 1605-7422/01/$15.00

sequencing. These include modified dNTPs[13], modified ddNTPs[14] and nonnucleotide (acyclic) terminators[15-17]. None of these alternative substrates have been used to generate long accurate sequence reads with any automated DNA sequencing system.

We have evaluated the use of near infrared (NIR) dye-labeled nonnucleoside triphosphates for automated DNA sequencing. These fluorescent acyclic analogs (AcyNTPs) were labeled with a heptamethine carbocyanine dye through a propargylamino linker to the purine or pyrimidine base. Dye-labeled AcyNTPs which are lacking in the sugar moiety positions equivalent to the C-2 and C-3 of ribose functioned similarly to the chain-terminating ddNTPs in Sanger-based sequencing protocols. A mutant, thermostable DNA polymerase incorporates the fluorescent acycloterminators in cycle sequencing reactions with the efficacy and fidelity of traditional dideoxynucleotides. Sequence readlength and base-calling accuracy is similar for both dye-labeled acycloterminator and dye-primer sequencing methods.

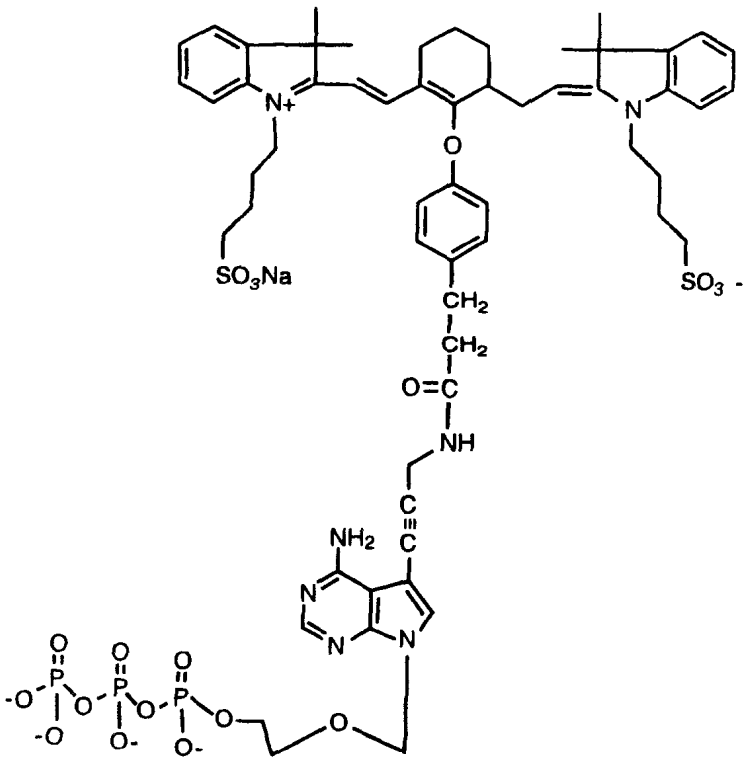


Figure 1. Structure of the cyanine dye-labeled acyclic ATP terminator

2. METHODS

The heptamethine carbocyanine dye (IRD800) bearing a single carboxylic acid functional group was synthesized as previously described[18] by the LI-COR Biotechnology Division (Lincoln, NE). Acycloterminators labeled with IRD800 were obtained from PerkinElmer Life Sciences (Boston, MA). Absorbance measurements of the IRD800-AcyNTPs were made with an HP8453 spectrometer equipped with a diode array detector. Fluorescence emission was measured using a LI-COR near infrared spectrofluorometer[19,20].

Plasmid DNA was purified using QIAprep® Spin Miniprep Kit 27104 (QIAGEN Inc., Valencia, CA). Oligonucleotide sequencing primers were designed with GeneRunner Ver. 3.0 (Hastings Software Inc., Hastings, NY). Primers were 18 to 22 bases in length, with the GC content ranging between 25% to 58% and the T_M between 26°C to 60°C (thermodynamic method).

Table 1. Final concentrations of dye-labeled acycloterminators and dNTPs in each termination mix.

Acyclo-terminator	Dye	Linker	Conc. (μM) Dye-terminator	Conc. (μM) dNTPs	Mix Ratio
AcyATP	IRD800	PA[+]	3	150	1:50
AcyUTP	IRD800	PA	20	60	1:3
AcyGTP	IRD800	PA	2	150	1:75
AcyCTP	IRD800	PA	20	60	1:3

[+] Linker is propargylamino (PA).

Termination mixes were prepared from 0.25 mM stock solutions of dye-labeled acycloterminators, sequencing grade deoxynucleotides (dATP, dTTP, c^7dGTP, dCTP) from Amersham Pharmacia Biotech (Piscataway, NJ), and water that had been purified using a Milli-Q® system (Millipore Corp., Bedford, MA). Each mix was optimized by adjusting the ratio of dye-terminator to deoxynucleotides in extension-termination reactions until a balanced signal was obtained through 1200 bases of the sequence ladder. The final composition of the four termination mixes is provided in Table 1.

The fluorescent AcyNTPs were used in cycle sequencing protocols to test their efficacy as substitutes for ddNTPs. Briefly, a master mix was prepared by combining 0.5 pmole of purified DNA template, 3 pmole of sequencing primer, 8 U of Thermo Sequenase™ DNA polymerase (USB Corp., Cleveland, OH), and 1.5 μl of reaction buffer (260 mM Tris-HCl, pH 9.5, 65 mM $MgCl_2$) and adjusted to a final volume of 27 μl with deionized water. Two microliters of the appropriate termination mix were combined with 6.5 μl of the master mix in a MicroAmp® reaction tube (PE Biosystems, Foster City, CA). The reaction mixtures were denatured by heating at 95°C for 2 min, followed by 30 thermal cycles of 95°C for 30 sec, 50°C for 30 sec, and 72°C for 1 min in either a DNA Engine Model PTC-200 or MiniCycler Model PTC-150 (MJ Research, Watertown, MA).

Following completion of the sequencing reactions, unincorporated dye terminators were removed from the extension products by either ethanol precipitation with a mixture of 3 volumes of 100% ethanol and 3M sodium acetate pH 5.2, precipitation with Vivid Violet linear polyacrylamide (CHIMERIX, Madison, WI), or by Sephadex™ G50 chromatography (Amersham Pharmacia Biotech). Purified reaction products were dried under reduced pressure in a Speed Vac® centrifuge (Savant Instruments, Holbrook, NY). Dried samples were suspended in 3-4 μl of IR^2 stop solution (LI-COR Inc.) and denatured by heating at 95°C for 3 min.

Aliquots (1.5 μl) of the denatured sequencing reactions were loaded on a 0.2 mm, 3.7% KBPlus slab gel (LI-COR Inc.) with a 56 cm well-to-read length. Electrophoresis was carried out with the IR^2 DNA Sequencer System (PerkinElmer Life Sciences) operated at 3000 V for 10 hours with a gel temperature of 45°C. Sequence data were analyzed using the LI-COR version 4.0 automatic base caller.

3. RESULTS AND DISCUSSION

3.1 Spectroscopy

The structure of acycloadenosine triphosphate is shown in Fig. 1. The NIR fluorescent dye IRD800 is covalently attached to the AcyNTP at the 5 position in the pyrimidines and the 7 position in the 7-deazapurines via a 7-atom propargylamino linker (Fig. 1). The fluorophore has a strong NIR absorbance (extinction coefficient = 250,000 M^{-1} cm^{-1}) compared to the ultraviolet (UV) absorbance of the acyclic nucleoside triphosphate. IRD800-AcyATP has a NIR absorption maximum of 769 nm and fluorescence emission maximum of 801 nm allowing detection with previously described infrared DNA analysis systems[19,21].

3.2 DNA sequencing

In order to evaluate the efficacy of the new NIR fluorescent acycloterminators as substitutes for ddNTPs in cycle sequencing, we selected a thermostable DNA polymerase known to have a relatively broad spectrum of substrate tolerance. Thermo Sequenase DNA polymerase has been engineered to remove discrimination against ddNTPs by a phenylalanine-to-tyrosine mutation in the polymerase domain (the F667Y variant of Taq DNA polymerase)[7,8]. This amino acid change may also render Thermo Sequenase more effective in accepting dye-labeled terminators compared to other DNA polymerases[9].

Optimal concentrations of the fluorescent acyclic triphosphates of adenosine, uridine, cytidine, and guanosine in Thermo Sequenase cycle sequencing reactions were determined by titrating each acyclo-terminator and selecting the concentration which gave approximately equivalent signal at base positions 100, 500, and 1000 in the sequence ladder. For example, uniform signal strength to 1000 bases was evident in G- and A-ladders using 2 and 3 μM concentrations of the respective dye-terminator along with 150 μM dNTPs. In contrast, the fluorescent AcyUTP and AcyCTP were incorporated less efficiently by Thermo

Table 2. Comparison of read length[1] and accuracy to 1000 bases for two primer walking projects using either dye-primer or dye-acycloterminator sequencing.

	Dye-Terminator		Dye-Primer	
Custom Primer	Read Length	Accuracy (%)	Read Length	Accuracy (%)
Template p4CG9				
T3	1157	99.1	1067	99.4
T7	1090	99.4	1099	99.8
967T7F	1126	98.9	1163	100.0
218T7R-1	1192	98.4	1120	99.9
914T3F	1059	99.4	1129	99.8
284T3R	1093	99.4	1120	100.0
1872T7F	1090	99.8	1174	100.0
1874T3F	1162	99.5	1063	99.8
310T7R	1020	98.4	1154	99.6
221T3R	1120	99.5	1138	99.9
2742T3F	1118	97.6	1039	99.9
251T3R	1178	99.2	1140	99.9
2844T7F	1083	98.8	1085	99.8
218T7R-2	1010	98.2	1083	99.6
3624T3F	1177	99.4	1147	99.4
211T3R	1066	99.6	1224	99.9
3701T7F	1121	99.4	1112	99.6
228T7R	1036	98.0	1053	99.4
Template pUC19				
M13F	1005	98.9	1092	99.5
508F	1073	99.9	1198	99.8
928F	1062	98.9	1164	99.7
1243F	1093	99.1	1108	100.0
1585F	1040	99.2	1184	99.6
1991F	1040	99.4	1141	99.5
2286F	1037	99.4	1152	99.6
2482F	1116	99.7	1067	98.1
MEAN	**1090**	**99.1**	**1123**	**99.7**

[1] **Read length determined by the automatic termination of the base calling process when the predicted sequence accuracy drops below 99%.**

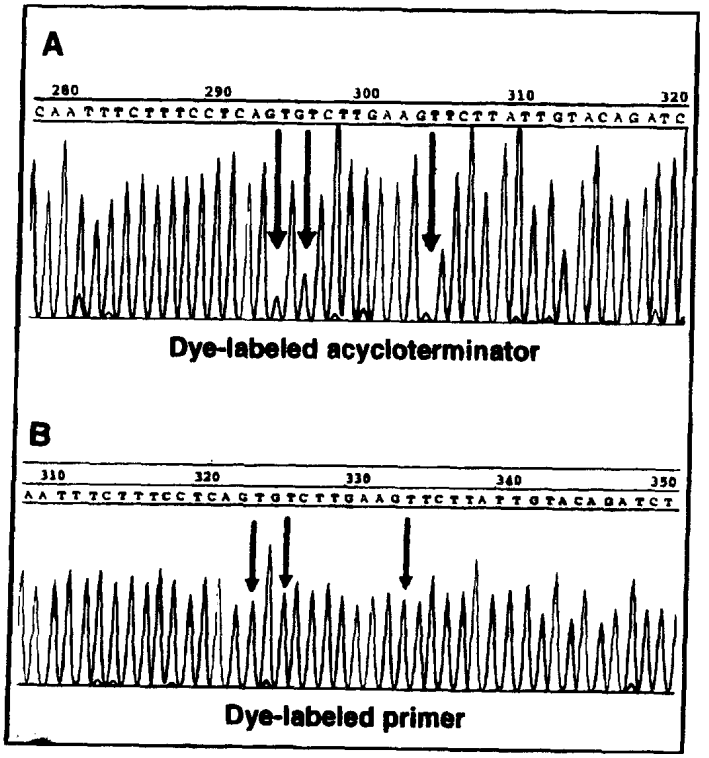
Sequenase DNA polymerase, requiring up to 10-fold more terminator in the reaction mixture (Table 1).

Other acyclic NTP analogs have been used to terminate sequencing reactions, but the concentrations needed for efficient termination were high (0.1 to 1 mM) compared to the corresponding ddNTPs[16]. There is one report of an allylic adenosine triphosphate which required a concentration of only 1 μM for complete termination of a short polymerization reaction using Ampli Taq® DNA polymerase FS (Perkin Elmer)[17]. Additionally, dideoxynucleotides have been tagged with the cyanine dye Cy5.5 which has a maximum fluorescence emission of 694 nm (Amersham Life Science, specifications for FluoroLink™ Cy5.5™ Monofunctional Dye 5-Pack). When used with Thermo Sequenase DNA polymerase, Cy5.5 dye-labeled ddNTPs generate sequence read lengths of 350-400 bases[11] with termination mixes containing 3 μM Cy5.5 dye terminator and 150 μM dNTPs (Amersham Pharmacia Biotech, instruction manual for product US79840).

Identical cycle sequencing reactions using either IRD800-acycloterminators or IRD800-dideoxynucleotides were directly compared (data not shown). Both types of fluorescent terminator produced high quality sequence ladders indicating that Thermo Sequenase recognized the new acyclic terminators as ddNTP substitutes.

The sequencing utility of NIR dye-labeled acycloterminators was tested by primer walking two plasmids, an 11.5 kilobase pBluescript® (Stratagene, LaJolla, CA) construct containing rat genomic DNA and the 2.6 kilobase pUC19 vector without insert. In both cases, long accurate sequence reads were obtained using

Figure 2. Sequence context effect with IRD800-acycloterminators. **(A)** Dye-terminator reaction performed with Thermo Sequenase DNA polymerase, DNA template p4CG9, and custom primer 914T3F using thermal cycling conditions described in text. Arrow denotes weak T peak following G peak. **(B)** Dye-primer reaction showing the same section of electropherogram as in Fig. 2A. Arrow shows no suppression of T peak after G peak



unlabeled walking primers with the fluorescent AcyNTPs. Moreover, the quality of the dye-terminator data

resembled the dye-primer sequencing data (Table 2). For 26 walking primers, cycle sequencing with acycloterminators produced a mean read length of 1,090 bases with a base-calling accuracy of 99.1 % at one kilobase.

3.3 Sequence context effects

A major disadvantage of cycle sequencing with dye-labeled ddNTPs is the generation of sequence traces with uneven peak heights or intensities[6,8,10,12]. Peak height variations represent the disproportionate rate of enzymatic incorporation of dNTPs versus dye-labeled ddNTPs, a fact which is highly dependent on the local sequence context as well as the DNA polymerase used[6,8-10,22]. In order to identify specific sequence context effects, we have analyzed a series of sequence ladders generated from different DNA templates using fluorescent AcyNTPs. The trace patterns and base calls from a section of plasmid template p4CG9 using either dye-terminator or dye-primer sequencing methods is shown in Fig. 2. Sequencing with IRD800-AcyNTPs results in weak T peaks following G peaks (Fig. 2A), a phenomenon not observed in dye-primer sequencing when unlabeled ddNTPs are used (Fig. 2B). These diminuitive T peaks are still larger than noise peaks and are accurately identified by the automatic base caller. Similar sequence-dependent patterns of nucleotide incorporation have been observed with other dye-labeled ddNTPs. For example, BigDye terminators show suppression of T peaks after G peaks[10]. With rhodamine terminators, weak G peaks follow A peaks[10] although this effect can be minimized with the addition of manganese citrate to the cycle sequencing reaction[23]. In cases of polymorphism, such peak height variability may make it difficult to distinguish the presence or absence of a heterozygote at a specific site[6,12,24].

Table 3. Coefficient of peak height variability (CV)[1] for dye-primer and dye-acycloterminator sequencing in the 100 – 400 bp region of pGEM

	A	T	G	C	All Peaks
Dye-primer[2]	15%	15%	18%	18%	17%
IRD800 - acycloterminators	28%	46%	31%	37%	37%

[1] CV = 100% • (Standard Deviation/Mean).
[2] Universal M13 Forward Primer 5' end-labeled with IRD800 (20)

Peak pattern uniformity in DNA sequence traces was evaluated either by visual inspection or by measuring the peak heights in a certain sequence and calculating the mean and standard deviation. The data were normalized by defining the coefficient of variability (CV) as the standard deviation expressed as a percentage of the mean peak height[25]. A completely uniform series of peaks would produce a CV value of 0%. Table 3 shows this relative measure of peak variation in the same region of pGEM® vector DNA (Promega Corp., Madison, WI) for dye-primer and dye-terminator sequencing. The coefficient of peak height variability for dye-labeled acycloterminators is 2-fold higher than for dye-primer sequencing. Moreover, the peak patterns generated by fluorescent acyclic pyrimidines have greater variability than their purine counterparts. However, the overall peak height variation observed with cyanine dye-labeled acycloterminators is similar to sequencing with 4,7-dichlororodamine terminators or energy transfer BigDye terminators[10].

4. CONCLUSIONS

Previous studies have shown that certain acyclic NTP analogs are incorporated by commercially available DNA polymerases with concomitant termination of the polymerization reaction[16,17]. However, these compounds are not practical as alternate substrates for Sanger-based DNA sequencing since the concentrations needed for efficient termination are high compared to the corresponding dideoxy-nucleotides. We report for the first time a set of cyanine dye-labeled nonnucleoside triphosphates which are efficiently incorporated by Thermo Sequenase DNA polymerase and provide a chain-terminating function to yield sequence data comparable to standard dye-primer sequencing methods.

Future work in cyanine dye synthesis, enzyme engineering, and optimization of linker arm (structure and location) will be necessary to further advance the performance of NIR dye-labeled acycloterminators. If sequence context effects could be eliminated with improved attributes of infrared dye terminators, the accuracy of automatic base calling would be enhanced coupled with the ability to identify the presence or absence of a heterozygote at a specific site based on peak height variation.

REFERENCES

1. F. Sanger, S. Nicklen, and A.R. Coulson, "DNA sequencing with chain terminating inhibitors," *Proc Natl Acad Sci USA* **74**, pp. 5463-5467, 1977.
2. L.M. Smith, "DNA sequence analysis: past, present, and future," *Intl Biotech. Lab* **7**, pp. 8-19, 1989.
3. J.M. Prober, G.L. Trainor, R.J. Dam, F.W. Hobbs, C.W. Robertson, R.J. Zagursky, A.J. Cocuzza, M.A. Jensen, and K. Baumeister, "A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides," *Science* **238**, pp.336-341, 1987.
4. L.G. Lee, C.R. Connell, S.L. Woo, R.D. Cheng, B.F. McArdle, C.W. Fuller, N.D. Halloran, and R.K. Wilson, "DNA sequencing with dye-labeled terminators and T7 DNA polymerase: effect of dyes and dNTPs on incorporation of dye-terminators and probability analysis of termination fragments," *Nucleic Acids Res* **20**, pp. 2471-2483, 1992.
5. T.E. Tracy and L.S. Mulcahy, "A simple method for direct automated sequencing of PCR fragments," *BioTechniques* **11**, pp. 68-75, 1991.
6. L.T. Parker, Q. Deng, H. Zakeri, C. Carlson, D.A. Nickerson, and P.Y. Kwok, "Peak height variations in automated sequencing of PCR products using Taq dye-terminator chemistry," *BioTechniques* **19**, pp. 116-121, 1995.
7. S. Tabor and C.C. Richardson, "A single residue in DNA polymerases of the Escherichia coli DNA polymerase I family is critical for distinguishing between deoxy- and dideoxyribonucleotides," *Proc Natl Acad Sci. USA* ***92***, pp. 6339-6343, 1995.
8. P.B. Vander Horn, M.C. Davis, J.J. Cunniff, C. Ruan, B.F. McArdle, S.B. Samols, J. Szasz, G. Hu, K.M. Hujer, S.T. Domke, S.R. Brummet, R.B. Moffet, and C.W. Fuller, "Thermo Sequenase™ DNA polymerase and T. acidophilum pyrophosphatase: new thermostable enzymes for DNA sequencing," *BioTechniques* **22**, pp. 758-765, 1997.
9. M.A. Reeve and C.W. Fuller, "A novel thermostable polymerase for DNA sequencing," *Nature* **376**, pp. 796-797, 1995.
10. B.B. Rosenblum, L.G. Lee, S.L. Spurgeon, S.H. Khan, S.M. Menchen, C.R. Heiner, and S.M. Chen, "New dye-labeled terminators for improved DNA sequencing patterns," **25**, pp. 4500-4504, 1997.
11. S. Kumar, C.W. Fuller, S. Nampalli, M. Khot, I. Livshin, L. Sun, S. Hamilton, S.B. Samols, J.A. Mamone, K.M. Hujer, B.F. McArdle, J.R. Nelson, and S. Duthie, "Uniform band intensities in fluorescent dye terminator sequencing," *Nucleosides & Nucleotides* **18**, pp. 1101-1103, 1999.
12. H. Zakeri, G. Amparo, S.-M. Chen, S. Spurgeon, and P.-Y. Kwok, "Peak height pattern in dichloro-rhodamine and energy transfer dye terminator sequencing," *BioTechniques* **25**, pp. 406-414.
13. M.L. Metzker, R. Raghavachari, S. Richards, S.E. Jacutin, A. Civitello, K. Burgess, and R.A. Gibbs, "Termination of DNA synthesis by novel 3'-modified-deoxyribonucleoside 5'-triphosphates," *Nucleic Acids Res.* **22**, pp. 4259-4267, 1994.
14. Z.G. Chidgeavadze, A.V. Scamrov, R.S. Beabealashvilli, E.I. Kvasyuk, G.V. Zaitseva, I.A. Mikhailopulo, G. Kowollik, and P. Langen, 3'-Fluoro-2',3'-dideoxyribonucleoside 5'-triphosphates: terminators of DNA synthesis," *FEBS* **183**, pp. 275-278, 1985.
15. J.M. Prober, R.J. Dam, C.W. Robertson, F.W. Hobbs, and G.L. Trainor, "Method, system, and reagents for DNA sequencing," *United States Patent* **5,332,666**.
16. C.I. Martinez, M.A. Ansari, R. Gibbs, and K. Burgess, "Acyclic nucleoside triphosphate analogs as terminators in biocatalytic DNA replication," *Bioorg Med Chem Lett.* **7**, pp. 3013-3016, 1997.
17. C.I. Martinez, L.H. Thoresen, R.A. Gibbs, and K. Burgess, "An allylic/acyclic adenosine nucleoside triphosphate for termination of DNA synthesis by DNA template-dependent polymerases," *Nucleic Acids Res.* **27**, 1271-1274, 1999.
18. N. Narayanan, and G. Patonay, "A new method for the synthesis of heptamethine cyanine dyes: synthesis of new near-infrared fluorescent labels," *J. Org Chem* **60**, pp. 2391-2395, 1995.
19. L. Middendorf, J. Amen, R. Bruce, D. Draney, D. DeGraff, J. Gewecke, D. Grone, P. Humphrey, G. Little, A. Lugade, N. Narayanan, A. Oommen, H. Osterman, R. Peterson, J. Rada, R. Raghavachari, and

S. Roemer, "Near-infrared fluorescence instrumentation for DNA analysis," In *Near-Infrared Dyes for High Technology Applications*, S. Daehne et al. (eds.), pp. 21-54, Kluwer Academic Publishers, The Netherlands, 1998.
20. N. Narayanan, G. Little, R. Raghavachari, J. Gibson, A. Lugade, C. Prescott, K. Reiman, S. Roemer, D. Steffens, S. Sutter, and D. Draney, "New NIR dyes: synthesis, spectral properties, and applications in DNA analyses," In *Near-Infrared Dyes for High Technology Applications*, S. Daehne et al. (eds.), pp. 141-158, Kluwer Academic Publishers, The Netherlands, 1998.
21. L.R. Middendorf, J.C. Bruce, R.C. Bruce, R.D. Eckles, D.L. Grone, S.C. Roemer, G.D. Sloniker, D.L. Steffens, S.L. Sutter, J.A. Brumbaugh, and G. Patonay, "Continuous, on-line DNA sequencing using a versatile infrared laser scanner/electrophoresis apparatus," *Electrophoresis* **13**, pp. 487-494, 1992.
22. L.T. Parker, H. Zakeri, Q. Deng, S. Spurgeon, P.-Y. Kwok, and D.A. Nickerson, "AmpliTaq® DNA polymerase, FS dye-terminator sequencing: analysis of peak height patterns," *BioTechniques* **21**, pp. 694-699, 1996.
23. C. Korch, and H. Drabkin, "Manganese citrate improves base-calling accuracy in DNA sequencing reactions using rhodamine-based fluorescent dye-terminators," *Nucleic Acids Res.* **27**, pp. 1405-1407, 1999.
24. P.-Y. Kwok, C. Carlson, T.D. Yager, W. Ankener, and D.A. Nickerson, "Comparative analysis of human DNA variations by fluorescence-based sequencing of PCR products," *Genomics* **23**, pp. 138-144, 1994.
25. R.G.D. Steel, and J.H. Torrie, *Principles And Procedures Of Statistics*, McGraw-Hill Book Company, New York, 1980.

Automation of DNA Extraction from Food and Plants using MagneSil™ Paramagnetic Particles

Rex Bitner and Susan Koller Promega Corporation, Madison, WI 53711-5399

ABSTRACT

MagneSil™ paramagnetic particles allow the flexibility of automating the isolation of DNA from as little as 20mg of plant material to as much as 500 grams of vegetable oil for use in testing for DNA sequences from genetically modified organisms (GMO), or plant breeding applications such as random amplification polymorphism detection (RAPD) or polymerase chain reaction (PCR). Given the wide variety of plant materials, foods and highly processed food ingredients that require testing, the purification system must be both scalable and flexible in its ability to purify DNA from such a wide array of sample types.

The procedures used in these purification systems are similar to other methods used for the walkaway automation of plasmid purification and DNA sequencing reaction cleanup used in genomics applications, as well as DNA purification of DNA from PCR reactions used for genetic interogations or DNA immobilizations. These purification systems can be used with a variety of robotic workstations in 96 well formats.
keywords: automation, paramagnetic particles, GMO testing

INTRODUCTION

Many applications involving the automation of nucleic acid purification and analysis can be achieved through the use of paramagnetic particles. In addition to the lower cost of goods generally demonstrated over multi-column systems, paramagnetic particles offer additional flexibility in both scaling down and scaling up sample volumes. Additional utilities of paramagnetic particles are based on more rapid solution kinetics (because the particles can bind nucleic acids while in suspension) and the particles' high binding capacities for nucleic acids, coupled with the ability to limit non-specific binding of contaminants through optimized binding chemistries and washes. The rapid response of MagneSil™ particles to magnetic fields, as well as the robust physical integrity of the particles in a wide variety of chemical environments are key elements in creating a robust, automated purification system.

It is very important that the particles are paramagnetic, rather than magnetic. If the particles possess even a small dipole moment, they will remain attracted to one another after the magnetic field has been removed, thus complicating the process of washing and eluting the nucleic acids. This generally places an upper limit on the size of the magnetic cores that can be employed in making paramagnetic particles. The use of paramagnetic cores such as neodymium enclosed in silica surfaces offers many of the traditional advantages of silica based nucleic acid purification, along with well characterized chemistries that can be used for surface modifications of the paramagnetic particles.

To expand the purification capabilities of MagneSil™ paramagnetic particles, we have modified their surface chemistries through the coupling of pH dependent ion exchange ligands to the silica surfaces, as previously described (1). Traditional anion exchange chromatography utilizes matrices that remain positively charged from pH 3 to pH 10. Negatively charged DNA or RNA molecules are electrostatically attracted to the positively charged surfaces, and this property can be used for their purification. However, using traditional anion exchange methods, after purification the nucleic acids must be eluted from the positively charged surface by displacing the negatively charged DNA using a relatively high concentration of negatively charged counter-ion, such as chloride. As a result, the DNA is eluted along with a high salt concentration, thus requiring that the nucleic acids be desalted prior to many downstream applications.

Genomics and Proteomics Technologies, Ramesh Raghavachari, Weihong Tan, Editors,
Proceedings of SPIE Vol. 4264 (2001) © 2001 SPIE · 1605-7422/01/$15.00

A key advantage of using pH dependent ion exchange purification is that the matrices are positively charged at acidic pH (such as pH 4-6), but are either neutrally charged or negatively charged at more basic pH (such as pH 8-10). Therefore, the anion exchange purification can be performed at pH 5 for example, and when elution of the DNA or RNA is desired, the matrix can be washed with water to remove salts, and the pH can then be raised so the matrix becomes neutrally or negatively charged. Thus the DNA can be eluted in a very weak buffer, such as 10mM Tris HCl at a pH such as 8.0, that is compatible with the desired downstream applications.

By using a combination of positively charged ligands, such as histidine, and negatively charged ligands such as propionic acid, the ratio of positively charged to negatively charged ligands determines the pH at which the matrix is neutral in overall charge. The amount of positively charged ligand attached to the matrix determines the amount of negative counterion required to elute the nucleic acid at a given pH where the matrix is positively charged (1). This combination of positively charged and negatively charged ligands is the basis for being able to produce pH dependent matrices with a desired overall positive charge at a given acidic pH, and a desired pH of neutrality, effectively allowing the "tuning" of the pH dependent ion exchange matrix to the desired purification system.

The ability to create such an array of pH dependent ion exchangers enables the use of mixed beds of particles consisting of two (or more) pH dependent ion exchange matrices. While mixed beds are traditionally column based mixtures of anion exchange and reverse phase materials (3) where the nucleic acid simultaneously interacts with both surfaces, the combination of two pH dependent ion exchangers allows for binding of the nucleic acid molecule to one type of particle under one set of conditions, and to the other type of particle under a different set of conditions. After the purification, the nucleic acid can be eluted from both sets of particles by washing with water and raising the pH through the use of an elution solution such as 10mM Tris HCl pH 8.0 (4).

The use of MagneSil™ paramagnetic particles in the walkaway automation of DNA sequencing using PE Biosystems' Big-Dye™ has been previously described (1). A key element in achieving the walkaway automation of this system is the elimination of the centrifugation step in the clearing of the bacterial lysate. This can be achieved through the use of a filtration step, but that also requires the additional cost of a filtration column in a 96-well configuration. The clearing of the bacterial lysate can also be achieved through the addition of paramagnetic particles to the bacterial debris in solution, followed by mixing which helps bind the cellular debris to the paramagnetic particles, and then the application of a magnetic field that removes the debris from solution. This type of magnetic clearing of cellular debris has also been demonstrated with lysates from blood and tissue (2).

While it is possible to use either filtration or magnetic clearing to remove debris from a wide variety of lysed materials, in other cases the use of centrifugation is a preferred method. Two such systems are the purification of nucleic acids from many food materials, and many plant materials also produce debris that are difficult to filter out or magnetically clear. In the case of food materials, there is often very little DNA present in a large volume of the food item (such as chocolate, lecithin or vegetable oil). In the case of plant materials, there are many examples of crusty, starchy or fibrous materials present in large quantities in the lysed materials that are more easily removed by centrifugation.

The automation of genomic DNA purification from food ingredients and plant materials requires the ability to handle both small scale (e.g. 20 mg) and large scale (e.g. 500 gm) samples. The downstream uses for the genomic DNA vary significantly from PCR, RAPD (Random Amplified Polymorphic DNA), AFLP ( ), or quantitative PCR applications such as PE BioSystem's TaqMan® or Roche's LightCycler™. Given this large range of sample types and sample sizes, and the diverse requirements of downstream

applications, any common solution to these purification problems must provide considerable flexibility, versatility and robustness.

RESULTS

Figure 1 shows one method of making pH dependent ion exchange matrices, using histidine and propionic acid as examples. In the absence of propionic acid, the pH where the matrix is neutrally charged is about pH 8.0 to 8.5. With incremental addition of propionic acid, the negative charges on the matrix surface increases and the pH where the matrix surface is neutrally charged is lowered.

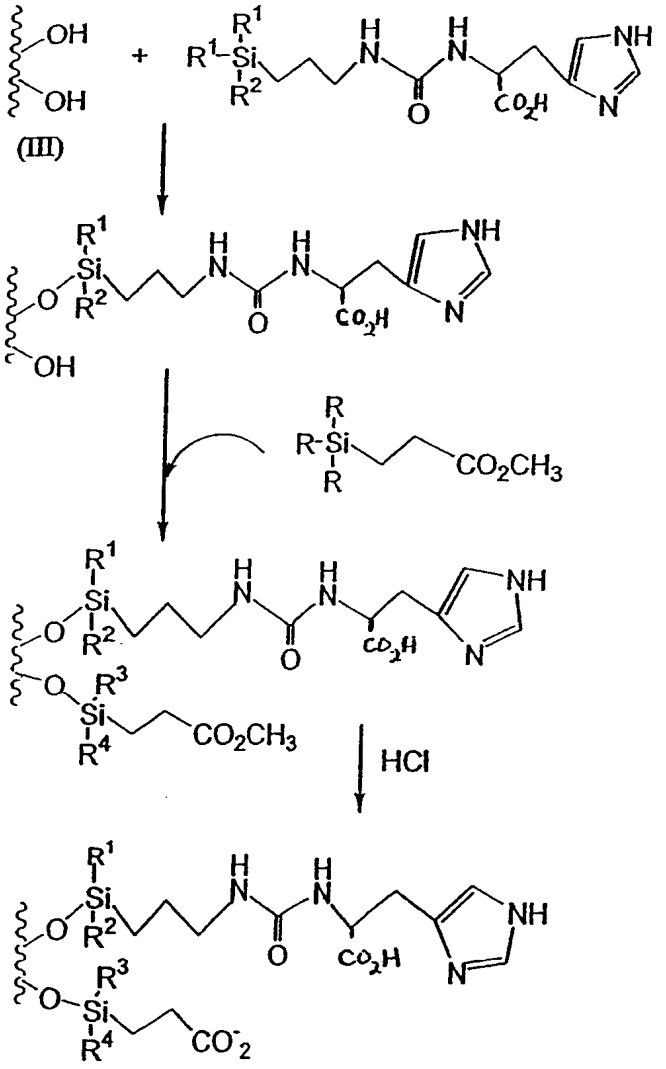


Figure 1 The synthesis of one type of pH dependent ion exchange matrix is based on the attachment of a positively charged ligand (histidine) and a negatively charged ligand (propionate) to the support material. The ratio of positively charged ligand to negatively charged ligand determines the pH at which the matrix is neutral in overall charge. The amount of positive ligand attached determines the amount of counterion required for nucleic acid elution at a given pH in which the matrix is positively charged. $R_1 - R_4$ in this example could be $-OH$, $-OCH_3$ or $-OCH_2CH_3$.