Overview of Purification and Analysis of Synthetic Nucleic Acids

Chapter 10 presents information on evaluating, analyzing, and purifying nucleic acids, a broad class of biologically essential molecules, ranging from nucleobases, nucleosides, nucleotides, and oligonucleotides to native, high-molecular-weight, duplex DNA and RNA. Additionally, these methods will be useful for nucleic acid analogs, which possess many biological, diagnostic, and therapeutic effects. Emphasis is placed on the 5 to 10 million oligonucleotides synthesized each year currently for a large array of molecular biology applications. Steady improvements in the efficiency of synthesis chemistry and in the automated DNA synthesizers have made production of oligonucleotides routine. To produce an adequate level of quality and purity, rapid and convenient analytical methods are necessary for the dozens of oligonucleotides produced each day by a DNA synthesis laboratory. Basic principles are discussed to guide selection of the appropriate protocols to ensure functionality of oligonucleotides in the myriad of molecular biology applications.

Synthetic oligonucleotides should be evaluated for concentration and purity before use, and purified if necessary. At the current levels of synthesizer efficiency and reagent quality, the crude oligonucleotide may be in an appropriate form and sufficient purity level to function as a sequencing or PCR primer—two relatively low-stringency applications. However, careful purification, analysis, quantitation, and other preparations may be necessary for oligonucleotides to perform in more stringent experiments such as mutagenesis, gene construction, and those requiring the cloning and expression of synthetic oligonucleotide sequences. Post-synthesis protocols for preparing ready-to-use oligonucleotides are provided in this chapter.

Practical nucleic acid separations are based on four physical principles: velocity in an electric field (electrophoresis), sorption, size exclusion, and ultrafiltration. Useful methods of nucleic acid separation should identify samples accurately on the basis of molecular size (length) and possess near-single-nucleotide resolution. The analytical and purification methods detailed in this chapter are (1) polyacrylamide gel electrophoresis (PAGE; *UNIT 10.4*); (2) high performance liquid chromatography (HPLC; *UNITS 10.5 & 10.6*); (3) cartridge purification (*UNIT 10.7*); (4) mass spectroscopy (*UNITS 10.1 & 10.2*). Other useful methods, such as gel capillary electrophoresis, will be covered in future supplements.

SELECTING ANALYSIS/PURIFICATION METHODS

Analysis and purification methods differ in complexity, expense, and time requirements. It is advisable to become familiar with all available options. Some factors to consider when selecting strategies for analysis and purification are (1) quantity needed for a particular experiment(s); (2) purity level required for the experiment; (3) time constraints; (4) equipment required; (5) sample composition (charge, hydrophobicity, solubility, sequence, length, salt form, labels, and modifications).

Highly efficient syntheses yielding pure oligonucleotides lessen the demands on purification and facilitates analysis. Using the standard synthesis routine of automated, solid-phase, phosphoramidite chemistry (Beaucage and Iyer, 1992; Caruthers and Beaucage, 1983; Caruthers and Matteucci, 1984; *UNIT 3.3; APPENDIX 3C*), the oligonucleotide can be prepared with a final average yield per base addition of ~98%. At this level of efficiency, a 20-nt oligonucleotide will result in ~70% of the theoretical yield. Depending on the detection method and definition of purity, the crude sample will also be ~70% pure by mass. The remaining ~30% is a heterogeneous population of failure sequences, shorter oligonucleotides that failed to couple during synthesis and were effectively capped to prevent further extension.

Size (length) of the oligonucleotide is usually the discriminating factor which allows efficient separations. The sequence or base content of an oligonucleotide does not generally affect synthesis performance or the choices in analysis and purification methods. During synthesis, the four bases, as A,G,C,T phosphoramidite monomers, react at comparable rates and little compensation for their differences is necessary. Similarly, during analysis and purification, the chemical differences imparted by the bases are not usually significant. Exceptions are evident primarily under nondenaturing conditions, such as reversed-phase

Contributed by Alex Andrus and Robert G. Kuimelis
Current Protocols in Nucleic Acid Chemistry (2000) 10.3.1-10.3.6

HPLC. Occasionally, some sequences may show anomalous behavior, probably caused by hydrogen-bonding or secondary structure. For example, sequences containing four or more contiguous G bases are often problematic. Relative to other analytes, the narrow range of chromatographic and electrophoretic differences amongst oligonucleotides make this class of molecules relatively straightforward to analyze and purify.

POST-SYNTHESIS PROCEDURES FOR OLIGONUCLEOTIDES

Post-synthesis processing, completed before analysis and purification, includes dimethoxytrityl (DMTr) assay, cleavage/deprotection, UV yield quantitation, desalting, and any other manipulations required to prepare the oligonucleotide for the experiment. Some of these steps may be omitted after success has been demonstrated in a particular application. For example, in some repetitive, well controlled sequencing or PCR experiments, the primers are simply dried from the ammonium hydroxide solution, after cleavage and deprotection, and dissolved in buffer, assuming an approximate concentration that contains the presumed amount of crude oligonucleotide. Caution is required when using oligonucleotides without analysis or purification. At a minimum, the DMTr fractions from synthesis and UV quantitation should be checked visually. Solutions of crude oligonucleotides should be assumed to contain no more than 70% of the correct sequence when making experimental calculations.

Synthesis Yield

The flexibility of DNA/RNA synthesizers allows the choice of different synthesis scales. The smallest scale, ~40 nmol, provides more than enough oligonucleotide for common applications such as sequencing and PCR primers. At the opposite end of the scale spectrum, the 10-µmol scale is used primarily for physical studies, antisense experiments, or for commercially produced probes and primers. At the 40-nmol scale, synthesis of a 20-nt crude oligonucleotide typically yields 5 to 10 OD units (165 to 330 µg), enough for hundreds of PCR and sequencing experiments (Table 10.3.1). An optical density (OD) unit is the absorbance of a 1-mL solution, typically in water, measured at 260 nm (A_{260}), in a 1-cm path-length cuvette. The actual amount of pure oligonucleotide that can be attained after purification is dependent on the synthesis efficiency and can vary for numerous reasons. In general, the small synthesis scales are more efficient and generate purer oligonucleotide products. The approximate yields (Table 10.3.1) are applicable for phosphorothioate or normal, phosphodiester sequences. The conversion [33 µg oligonucleotide per OD unit (A_{260nm})] between absorption (OD units) and mass depends on an averaged extinction coefficient (ε) per base of 10,000. Table 10.3.1 gives the approximate crude product yield at each scale to be expected when all instrument and reagent parameters are optimized.

DMTr Assay

Determining the stepwise yield of coupling reactions by the DMTr (dimethoxytrityl, or trityl) assay during synthesis is a useful yet indirect indication of final product quality (Ellington, 1995; *APPENDIX 3C*). The DMTr group is the hydroxyl protecting group at the 5′ terminus of the oligonucleotide (*UNIT 2.3*). The DMTr assay is useful for immediate feedback on the performance of an automated DNA synthesizer as the DMTr cation is liberated at each detritylation step in the synthesis cycle and can be quantitated by absorbance measurement on a spectrophotometer. Most DNA synthesizers can deliver the DMTr cation–containing detritylation effluent from each cycle to a fraction collector, advanced by a signal from the synthesizer. The ISCO Cygnet fraction collector is convenient for one-column instruments, and the ISCO Retriever II, with a four-column adaptor, is suitable for multi-column instruments. Both can accommodate various test tube sizes. The DMTr fraction is diluted to 10.0 mL (5.0 mL for 40-nmol scale synthesis) using 0.1 M *p*-toluenesulfonic acid (Aldrich) in acetonitrile. The solution is mixed well and the absorbance read in a 1.0-cm path length cuvette at 490 nm, near the absorbance maxima for the bright orange dimethoxytrityl cation. The total yield is calculated by converting the final DMTr absorbance to a percent of the initial DMTr absorbance. The average stepwise yield is calcu-

Table 10.3.1 Approximate Yield of a Crude 20mer Oligonucleotide

Scale	Average overall yield	OD unit[a]	Amount
40 nmol	>75%	5-10	165-330 µg
0.2 µmol	50%-75%	20-30	660-1000 µg
1 µmol	50%	100	3.3 mg
10 µmol	40%	800	26 mg

[a] 33 µg single-stranded DNA = 1 OD unit (A_{260nm}).

lated by raising the total yield to the power of the inverse number of trityl fractions (Andrus, 1992b).

$$\text{Overall yield} = \text{lowest / highest } (OY = A_n/A_2)$$

$$\text{Stepwise yield} = \text{Overall yield}^{1/\text{couplings}} (SY = OY^{1/n})$$

Some synthesizers feature an automated conductivity measurement of the DMTr cation, which calculates real-time average stepwise yields (Andrus, 1992a; Kaufman et al., 1993) while other instrument designs measure the absorbance of the detritylation effluent.

An unusually low trityl absorption value corresponding to the detritylation of the first nucleoside on the column is sometimes noted. Typically, this first base absorption value is slightly less than the second due to spontaneous detritylation of the support-bound nucleoside during storage of the column. The DMTr cation cleaved in this manner is lost in the initial washes of the synthesis cycle and is not collected in the DMTr cation fraction. While this has no effect on the synthesis, it can affect the trityl assay. Therefore, the first trityl value is not used in calculations.

A coupling failure with efficient capping would be detected in the trityl assay as a large drop in absorbance at one specific fraction, with subsequent fractions showing equally low absorbance. The oligonucleotides that failed to couple in the low yield coupling reaction were completely capped, thus eliminating their ability to react in subsequent coupling reactions. A failure of this type can be confirmed by the appearance of a major peak or band in HPLC, gel capillary electrophoresis (Andrus, 1992c, 1994), or PAGE analysis. A coupling failure that occurs with the first coupling, noted by a large decrease in absorbance of the second trityl fraction, could be indicative of inadequate reagent delivery (e.g., base, tetrazole, or trichloroacetic acid), most likely stemming from a failure to purge the phosphoramidite and tetrazole lines prior to synthesis.

Unfortunately, the DMTr cation assay will not detect several conditions that may be detrimental to the final product quality. In some situations, the DMTr assay may give positive results while the product oligonucleotide is of poor quality, or there may be no product at all. For example, if oxidation is inefficient due to impaired flow, an empty bottle, or degraded reagent, then the internucleotide phosphite triester will be cleaved by the acidic detritylating reagent which is the next step in the synthesis cycle. The DMTr cation is liberated normally, giving a false indication of synthesis efficiency.

Stepwise yields from phosphoramidite synthesis on commercial synthesizers should be in the 98 ± 0.5% range. Yields below 97% may indicate synthesizer or reagent problems or some other unoptimized condition. Calculation and documentation of DMTr-assay yield data is very useful in the early diagnosis of instrument-related problems. Low DMTr assay average stepwise yields will inform the user of inferior product quality more quickly than analysis of the oligonucleotide. However, the DMTr assay should only be regarded as indirect evidence of purity and yield. Direct analysis of the oligonucleotide mixture is the only way to positively confirm product quality.

Cleavage/Deprotection

After the repetitive addition of phosphoramidite monomers to complete the extension of the growing oligonucleotide, the ester linkage to the solid support is cleaved, usually with concentrated ammonium hydroxide (Ellington, 1995; *APPENDIX 3C*). Cleavage is automated on most synthesizers. Typically, ~90% of the succinate ester linkages (*UNIT 3.2*) are cleaved and the protected oligonucleotide is liberated into solution within 30 min. To ensure complete cleavage, 1 hr at room temperature is required. Cleavage on a polystyrene support is slightly slower than on controlled-pore-glass (CPG; McCollum and Andrus, 1991; *UNITS 3.1 & 3.2*). On the other hand, CPG dissolves slightly in ammonium hydroxide during the typical 1-hr cleavage, leaving a white residue of glass with the oligonucleotide upon concentration. During the cleavage period, removal of the other oligonucleotide protecting groups begins. Complete deprotection of an oligonucleotide continues in the same concentrated ammonium hydroxide cleavage solution and entails removing the protecting groups from the phosphate groups and the exocyclic amino groups on the bases. No extra addition of ammonium hydroxide is required. Deprotection requires ~8 hr at 55°C when the standard set of nucleobase protection is used (A^{bz}, G^{ibu}, C^{bz}, T). The time period may be halved with each 10°C increase in temperature. For example, deprotection may alternatively be carried out for 4 hr at 65°C.

Incomplete deprotection of synthetic oligonucleotides will impede analysis and purification, as well as adversely affect biological activity. Alternative reagents and conditions for cleavage and deprotection are sometimes employed (Bellenson and Smith, 1992; Reynolds and Buck, 1992; Boal et al., 1996). Since the

isobutyryl group on G bases is by far the slowest protecting group on the oligonucleotide to be removed, substitution of it with dimethylformamide (DMF) and use of the proven benzoyl protecting group on A and C has the benefits of fast deprotection, high yield, and high purity (Theisen et al., 1993). It is important that only fresh ammonium hydroxide be used. Concentrated ammonium hydroxide should be purchased in small bottles (500 mL or less) and stored at 4°C. Each bottle should be dated when opened and kept refrigerated. Bottles that have been open for more than one month or that have warmed to room temperature should not be used for deprotection. Laboratory personnel should be warned that keeping concentrated ammonium hydroxide sealed at high temperatures causes very high pressure of a hazardous reagent. Eye protection, gloves, and shields should be used when handling samples.

Quantitation

Quantitation of oligonucleotides by UV absorbance is the most common and practical method to determine synthesis yield. The theoretical yield of any particular oligonucleotide can be calculated based on length and synthesis scale. The measurement of absorbance of light of a sample of DNA can be accurately used to convert to mass and thus molar amounts. The method is nondestructive and the sample is easily recovered. According to Beer's Law, $A = \varepsilon Cl$ (where A is the absorbance; ε is the molar extinction coefficient; C is the concentration in mol/L; and l is the path length in cm; typically, a 1-cm path-length cuvette is used). The conditions are defined at a certain wavelength and temperature and with a certain medium, all of which influence ε. The aromatic bases of DNA and RNA strongly absorb light with maxima near 260 nm. An average molar extinction coefficient of ε = 10,000 for each of the 4 bases is a useful approximation in a neutral pH range from 7.0 to 7.4. Using this, and other approximations, the absorbance can be translated to mass and concentration of oligonucleotides. One OD unit represents ~33 µg of single-stranded oligodeoxynucleotide (DNA). For example, 1 mg of an oligonucleotide is ~30 OD units. Conversely, for concentration purposes, 1 µmol of oligonucleotide will absorb 10 times the number of bases, in OD units. For example, 0.2 µmol of an 18-mer would be ~36 OD units.

There is a strong correlation between the measured yield (crude OD units), yield as a percentage of the theoretical yield (scale of synthesis), and oligonucleotide purity (quality). For example, an 18-nt oligonucleotide made at the 0.2 µmol scale, which gives 20 crude OD units is probably significantly more pure than one which yields 10 crude OD units. A low overall yield may indicate poor coupling reactions during synthesis, incomplete cleavage from the support, or inaccurate calculation from the absorbance reading.

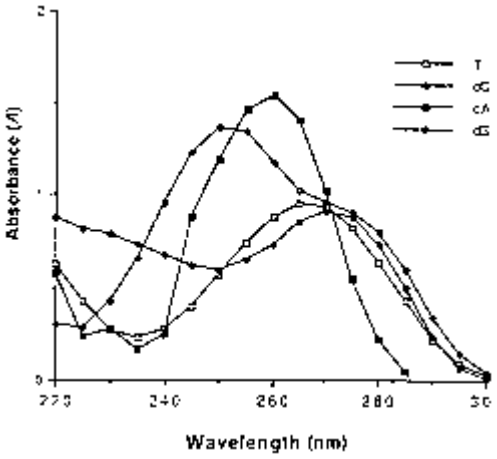

Figure 10.3.1 Absorbance spectra of four natural deoxynucleosides, dA, dC, dG, and T in aqueous solution at pH of 7.0.

The absorbance spectra of each of the four nucleosides are shown in Figure 10.3.1. A typical spectrum of a crude 18-nt oligonucleotide is shown in Figure 10.3.2. The average absorbance maximum of the four nucleosides is ~260 nm. Oligonucleotides that are very rich in either purines or pyrimidines could have absorbance maxima above or below 260 nm, depending upon the base composition.

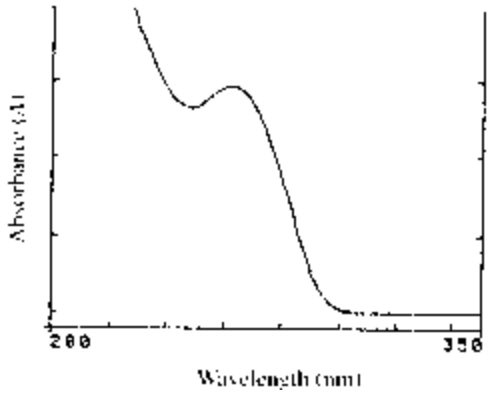

Figure 10.3.2 Absorbance spectra of an 18-nt oligonucleotide 5′ TCA CAG TCT GAT CTC GAT 3′ in 0.1 M TEAA, pH 7.0.

Desalting

The crude mixture, dissolved in ammonium hydroxide from the final deprotection, contains a variety of failure oligonucleotides along with some extraneous ammonium salts from the removed protecting groups. The crude mixture includes (1) the desired full-length oligonucleotide product; (2) shorter, failure-sequence oligonucleotides resulting from incomplete coupling and subsequent capping; (3) a collection of byproduct oligonucleotides resulting from other low-level side reactions, which were cleaved during ammoniolysis; (4) small molecule byproducts of synthesis including: benzamide, isobutyramide, ammonium acetate, ammonium trichloroacetate, pyridine, acetonitrile, and other solvents/reagents; and (5) trace amounts of other impurities.

Although desalting the crude oligonucleotide to remove these contaminants may be unnecessary for some applications, one may wish to consider desalting the crude sample before analysis and purification. At the very least, this provides for more accurate quantitation and subsequent concentration calculations. It also provides the opportunity to exchange the ammonium counterions for others, such as sodium, which may be better enzymatic substrates or confer different solubility properties.

There are several methods of desalting oligonucleotides, including precipitation, size-exclusion gel media (*UNIT 10.7*), and affinity cartridges (*UNIT 10.7*).

The preferred method depends on the quantity of oligonucleotide to be desalted, the materials, and the time available. Precipitation of oligonucleotides in an aqueous alcohol medium is a quick, efficient, and inexpensive desalting method, with the capacity to desalt large quantities. Only one precipitation is needed to remove virtually all soluble salts, byproducts, and short failure-sequence oligonucleotides. A convenient and efficient protocol for precipitation of oligonucleotides is as follows:

1. Dissolve the oligonucleotide in 30 µL water (20 µL at large scale) and 5 µL of 3 M sodium acetate per OD unit of oligonucleotide.
2. Add 100 µL of ethanol per OD unit of oligonucleotide and mix by vortexing. For very short oligonucleotides (<15mers), isopropanol may be substituted for ethanol to ensure complete precipitation.
3. Store at 4°C or –20°C for ~30 min, then centrifuge at high speed, i.e., ~10,000 × *g*, for 5 min.
4. Remove the supernatant with a pipet or micropipet, or decant, being careful not to disturb the pellet. Small quantities (<100 µg) may not be visible.
5. Add another 100 µL of ethanol, mix briefly, and centrifuge for 1 to 5 min.
6. Remove the supernatant and discard, being careful not to disturb the pellet. The oligonucleotide pellet can be dried by vacuum centrifugation, or air dried.
7. Resuspend the detritylated, desalted oligonucleotide in aqueous medium and quantitate by A_{260nm}.

MIXED-BASE OLIGONUCLEOTIDES

Oligonucleotides that have degenerate, or mixed-base sites, produce a more complex crude product mixture. The heterogeneity of mixed-base-containing oligonucleotides can present ambiguities and problems in analysis and purification. Equivalent incorporation of each base at a mixed-base site where more than one base has been delivered is usually desired. The four phosphoramidites, being different molecules, couple at slightly different rates. The order of reactivity, T>C>G>A, is largely compensated for by the order of the phosphoramidite positions on the valve block of some synthesizers, where A is positioned closest to the column and T the furthest. Virtually equal incorporation of the four bases in the oligonucleotide can result by concomitant delivery of the four equimolar and equivolume monomers. More precise and controlled incorporation of bases will result by premixing the desired phosphoramidites in a single bottle. Since mixed-base oligonucleotides are of the same length and contain the same number of phosphate anions, methods that separate on the basis of charge (electrophoresis and anion-exchange HPLC) will be more predictable than those that separate by hydrophobicity difference (reversed-phase HPLC).

MODIFIED OLIGONUCLEOTIDES

Although standard oligonucleotides make up the vast majority of chemically synthesized nucleic acids, modified oligonucleotides (containing one or many modifications) are becoming increasingly important (see Chapter 4). Modifications can include fluorescent moieties and other types of labels attached to the 3′ or 5′ terminus, or attached at internal positions. Other common modifications include amino or thiol linkers, which are used for subsequent attachment schemes. Nucleobase, carbohydrate, and phosphodiester modifications are commonly used to alter the physical or enzyme

substrate properties of an oligonucleotide. Many of these modifications also alter the chromatographic or electrophoretic behavior of the oligonucleotide and therefore influence the ultimate choice of a purification or analytical scheme. For example, fluorescent labels (and their associated linkers) usually impart substantial hydrophobicity to an oligonucleotide, and accordingly, affinity or reversed-phase separations would be employed. Certain special applications of oligonucleotides, especially modified oligonucleotides, may require very rigorous purification. In these situations it can be helpful to utilize an orthogonal purification scheme, whereby multiple purifications are performed in series on the same sample, but different modes of separation are employed at each step. For example, one might first use anion-exchange HPLC (charge-based separation) followed by reversed-phase HPLC (hydrophobicity-based separation). Performing the reversed-phase HPLC after ion exchange provides the added benefit of desalting the oligonucleotide. Each separation technique removes different types of impurities. Although time-consuming, orthogonal purification can be very effective when necessary.

LITERATURE CITED

Andrus, A. 1992a. AutoAnalysis: Trityl monitoring of DNA synthesizer operation by conductivity. Research News, Applied Biosystems, Foster City, Calif.

Andrus, A. 1992b. Evaluating and Isolating Synthetic Oligodeoxynucleotides. Applied Biosystems, Foster City, California.

Andrus, A. 1992c. Oligodeoxynucleotide analysis by gel capillary electrophoresis. *In* Methods: A Companion to Methods in Enzymology, Vol. 4 (J. Wiktorowicz, ed.) pp. 213-226. Academic Press, San Diego.

Andrus, A. 1994. Gel-capillary electrophoresis analysis of oligonucleotides. *In* Protocols for Oligonucleotide Conjugates (S. Agrawal, ed.) pp. 277-300. Humana Press, Totowa, N.J.

Beaucage, S.L. and Iyer, R.P. 1992. Advances in the synthesis of oligonucleotides by the phosphoramidite approach. *Tetrahedron* 48:2223-2311.

Bellenson, J. and Smith, A.J. 1992. Increasing DNA synthesizer throughput via off-instrument cleavage and deprotection. *BioTechniques* 12:219-220.

Boal, J.H., Wilk, A., Harindranath, N., Max, E., Kempe, T., and Beaucage, S.L. 1996. Cleavage of oligodeoxyribonucleotides from controlled-pore glass supports and their rapid deprotection by gaseous amines. *Nucleic. Acids Res.* 24:3115-17.

Caruthers, M.H. and Beaucage, S.L. 1983. Phosphoramidite compounds and processes. United States Patent 4,415,732.

Caruthers, M.H. and Matteucci, M.D. 1984. Process for preparing polynucleotides. United States Patent 4,458,066

Ellington, A. 1995. Synthesis and purification of oligonucleotides. *In* Current Protocols in Molecular Biology, Vol. 1, (F. Ausubel et al, eds.) Unit 2.11. John Wiley & Sons, New York.

Kaufman, J., Le, M., Ross, G., Hing, P., Budiansky, M., Yu, E., Campbell, E., Yoshimura, V., Fitzpatrick, V., Nadimi, K., and Andrus, A. 1993. Trityl monitoring of automated DNA synthesizer operation by conductivity: A new method of real-time analysis. *BioTechniques* 14:834-839.

McCollum, C. and Andrus, A. 1991. An optimized polystyrene support for rapid, efficient oligonucleotide synthesis. *Tetrahedron Lett.* 32:4069-4072.

Reynolds, T.R. and Buck, G.A. 1992. Rapid deprotection of synthetic oligonucleotides. *BioTechniques* 12:518-521.

Theisen, P., McCollum, C., and Andrus, A. 1993. N-6-Dialkylformamidine-2′-deoxyadenosine phosphoramidites in oligodeoxynucleotide synthesis. Rapid deprotection of oligodeoxynucleotides. *Nucleosides Nucleotides* 12:1033-1046.

Contributed by Alex Andrus
PE Applied Biosystems
Foster City, California

Robert G. Kuimelis
Phylos, Inc.
Lexington, Massachusetts