Hence, the bulk of the end sequences obtained during the first year will be derived from the chromosome 14 enriched set and should result in a sequence ready clone collection covering about 100 Mbp of the human genome. The purity of the chromosome 14 PAC collection will be characterized in a number of different ways, including testing with independent markers not used as probes and by FISH analysis of a representative set of PAC clones. To test the usefulness of the end sequence resource, the Sanger Centre will sequence chromosome 14 PACs from our collection and identify overlapping clones by virtual screening, using our end-sequence database.

If overlapping clones can not be found with the expected level of redundancy in the end-sequence database, we will screen the original PAC library with probes or STS markers derived from the sequenced PAC clones.

Subcontract under Glen Evans' DOE Grant No. DE-FC03-96ER62294.

Multiple-Column Capillary Gel Electrophoresis

Norman Dovichi
Department of Chemistry; University of Alberta;
Edmonton, Alberta, Canada T6G 2G2
403/492-2845, Fax: -8231, *norm.dovichi@ualberta.ca*
http://hobbes.chem.ualberta.ca

The objective of this project is to develop high-throughput DNA sequencing instrumentation. A two-dimensional arrayed capillary electrophoresis instrument is under development.

We have developed multiple capillary DNA sequencers. These instruments have several important attributes. First, by operation at electric fields greater than 100 V/cm, we are able to separate DNA sequencing fragments rapidly and efficiently. Second, the separation is performed with 3%T 0%C polyacrylamide. This low viscosity, non-crosslinked matrix can be pumped from the capillary and replaced with fresh material when required. Third, we operate the capillary at elevated temperature. High temperature operation eliminates compressions, speeds the separation, and increases the read length. Fourth, our fluorescence detection cuvette is manufactured locally by means of microlithography technology. These detection cuvettes provide robust and precise alignment of the optical system. Currently, 5, 16, and 90 capillary instruments are in operation in our lab; 32 and 576 capillary devices are under development. Fourth, we use both avalanche photodiode photodetectors and CCD cameras for high sensitivity detection. We have obtained detection limits of 120 fluorescein molecules injected onto the capillaries. High sensitivity is important in detecting the low concentration fragments generated in long sequencing reads. This combination of low concentration acrylamide, high temperature operation, and high sensitivity detection allows separation of fragments over 800 bases in length in 90 minutes.

DOE Grant No. DE-FG02-91ER61123.

DNA Sequencing with Primer Libraries

John J. Dunn, Laura-Li Butler-Loffredo, and **F. William Studier**
Biology Department; Brookhaven National Laboratory;
Upton, NY 11973
516/344-3012, Fax: -3407, *dunn@genome1.bio.bnl.gov*
http://genome5.bio.bnl.gov

Primer walking using oligonucleotides selected from a library is an attractive strategy for large-scale DNA sequencing. Strings of three adjacent hexamers can prime DNA sequencing reactions specifically and efficiently when the template is saturated with a single stranded DNA-binding protein (1), and a library of all 4,096 hexamers is manageable. We would like to be able to sequence directly on 35-kbp fesmid templates, but the signal from a single round of synthesis is relatively weak and triple-hexamer priming has not yet been adapted for cycle sequencing. We reasoned that a hexamer library might be used for cycle sequencing if combinations of hexamers could be selectively ligated by using other hexamers as the template for alignment. In this way, the longer primers needed for cycle sequencing could be generated easily and economically without the need for complex machines for de novo synthesis.

We found that ordered ligation of 3 hexamers to form an 18-mer occurs readily on a template of the 3 complementary hexamers (offset by three base pairs) that can base pair unambiguously to form a double-stranded complex of indefinite length (2). Each hexamer forms three complementary base pairs with two other hexamers, generating complementary chains of contiguous hexamers with strand breaks staggered by three bases. Two adjacent hexamers in the chain to be ligated contain 5' phosphate groups and the others are unphosphorylated. Both T4 and T7 DNA ligase can ligate the phosphorylated hexamers to their neighbors in such a complex at hexamer concentrations in the 50-100 M range, producing an 18-mer and leaving three unphosphorylated hexamers. The products of these ligation reactions can be used directly for fluorescent cycle sequencing of 35-kbp templates.

Unambiguous ligation requires that alternative complexes with perfect base pairing not be possible with the combination of hexamers used. Since the combination of hexamers is dictated by the sequence of the desired ligation product, some oligonucleotides cannot be produced unambiguously by this method. However, 82.5% of all possible 18-mers could potentially be generated starting with a library of all