Molecular Computation
A DNA-Based Model for Solving Mathematical Problems

Salvatore F. Priore
Dr. Wendy K. Pogozelski
Dr. Anthony Macula
Topics

- What is DNA computing?
- What are the prerequisites for creating a DNA computing model?
- The two phases of the project
  - Phase 1: Testing DNA strand binding
  - The Geneseo problem
  - Phase 2: Modeling the problem
- Future work
What is DNA Computing?

- An application of nanotechnology
- First developed in 1994 by Leonard Adleman
- Uses single stranded DNA molecules to represent the variables in a mathematical problem
- Selective hybridization then allows for the isolation of the correct answer
Why DNA?

- DNA has the ability to compute in parallel
- Standard computers compute in series
- This gives DNA the ability to sort through massive sets of data much quicker than a standard computer
- The key is simultaneous hybridization of DNA in solution

Example: Adding Numbers

Standard Computer:
- 5+3=8
- 4+2=6
- 3+6=9

Parallel:
- 5+3=8
- 4+2=6
- 3+6=9
Prerequisites

- Method for generating sequences of DNA to represent all the variables of a given problem

- Must also generate “well-behaved” strands

- Procedure for validating this method

- Strands that mis-pair when hybridizing will create errors in a DNA problem
Phase I

- Generated sequences using a computer algorithm developed by Dr. Anthony Macula
- Sequences were then synthetically prepared by Invitrogen
- Validated efficacy of the algorithm using SYBR Green I molecular dye
- This Dye binds to the minor groove of double stranded DNA (dsDNA) only
SYBR Green

- Place the generated strands in solution together
- Measure fluorescence to see if there is dsDNA present
- If there is notable fluorescence when it is not expected a DNA strand will be discarded and not used in the final experiment
Conclusion: SYBR Green was effective in screening for appropriate binding. The algorithm used to generate the sequences was efficient at producing “well-behaved” strands.
Other SYBR Experiments

Combined Data for SYBR Green as a function of Base Pairs

Fluorescence vs. Number of Base Pairs (PolyC - PolyG)

- 65.3°C
- 71°C
- 75.1°C
- 80.1°C
- 85.1°C
- 73.3°C
- 78.1°C
- 68.1°C
The SAT Problem

- Short for *Satisfiability*
- A series of logic clauses connected by “and” statements
- Each clause must be true for the entire series to be true
- Each Variable is represented by two strands: One for TRUE and one for FALSE

In a SAT problem all clauses must be true!

Example:
1) X1 or X2
2) ~X1 or X5
3) ~AX or X3 or X4
4) X3 or ~X5
5) X4 or X5
6) ~X3 or ~X4
7) ~X2 or X3

It is easy to see that these problems can get incredibly complex as more clauses are added.

Answer: A1=T, A2=T, A3=T, A4=F, A5=T
Phase II
Modeling the Problem

- The first step is to create all the possible answers to the problem.
- The individual DNA strands must be linked together to produce all possible combinations.
- Ex:

10 strands to code for T or F for 5 Variables

Each strand is 16 nucleotides long there are $2^5 = 32$ possible “solution” strands of length 80 nucleotides (16 nucleotides * 5 variables)
DNA Strands

- Library Sequences (5’—3’)
  - CCTTTTTTTTTTTTTTTTTTT, T1
  - TTTCCAAATTTTCTTTAT, F1
  - AACAACCACAAACCAAAA, T2
  - AATCCTCTTTTCTAAA, F2
  - ACACACACTTTTACTTT, T3
  - ACTATCACAAATATTAA, F3
  - CTTACTATTTCATTTCAT, T4
  - TACCTACCTCTCCTCTT, F4
  - CTAATACCATCATATTAT, T5
  - TAATTTCTCAACATTTA, F5
Building the Answers

- DNA polymerase extends a DNA molecule by adding complementary bases to a primer.
- We exploit this capability to create all combinations of the variable strands.
Connecting the Pieces

- Polymerase Chain Reaction (PCR): Repetitive heating and cooling of DNA sample to allow DNA polymerase to extend all the molecules.
- Using complements to each new variable position all possible combinations of strands are created one PCR reaction at a time.
- What is this slide talking about?
Stepwise PCR

\[X_1\]

\[\text{PCR}\]

\[X_1\]

\[X_2\]

\[X_1\]

\[X_2\]

\[\text{PCR}\]

\[X_1\]

\[X_2\]

\[\text{PCR}\]

\[X_1\]

\[X_2\]

\[X_3\]
Future Work

- The PCR reactions produce all 32 combinations, but it also creates their complements.
- We need single-stranded DNA.
- Use Biotin and Streptavidin to separate.
Once all strands are single stranded the problem can be solved.

Bind complement sequences that make the first clause True to Streptavidin, which is also associated with metallic beads on a column.

Strands with the right combination will adhere to the column.

Pass the captured strands through successive columns representing TRUE for each clause in the problem.

Anything remaining at the end codes for the answer.

Sequence the strand.
Special Thanks

- Supporters of this work:
  - The Air Force Research Laboratories, Rome NY
  - The National Science Foundation
    Biomathematics grant
  - The Geneseo Foundation undergraduate
    research grant
  - Dr. Wendy Pogozelski and Dr. Anthony Macula
    for their time and mentorship throughout this project