Extending Virtual PCR to Multiple Thermocycles

Gary Hon, UCSD

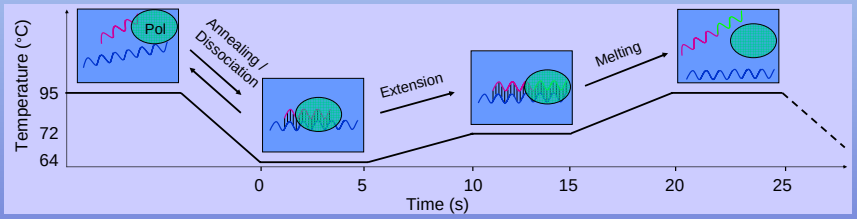
EEBI Division of Computations Directorate , Shea Gardner
Lawrence Livermore National Laboratory

PCR stands among the keystone technologies for analysis of biological sequence data. But despite its ubiquity, we lack the precise predictive capability that would enable detailed optimization of PCR reactions dynamics. Electronic PCR (ePCR) is an *in silico* method to model the kinetic, thermodynamic, and biological processes of PCR reactions. Here, we introduce speed improvements and approximate methods to make simulation of >40 thermocycles computationally feasible.

Introduction

Question: Can we simulate, in reasonable time, the dynamics of PCR, taking into account experimental conditions including time-temperature schedule, ion concentrations, and complex sequence mixtures?

The underlying biochemistry including DNA-hybridization kinetics and polymerase behavior has been extensively studied. Stochastic simulation methods have been perfected years before. The goal here is to bring this all together.

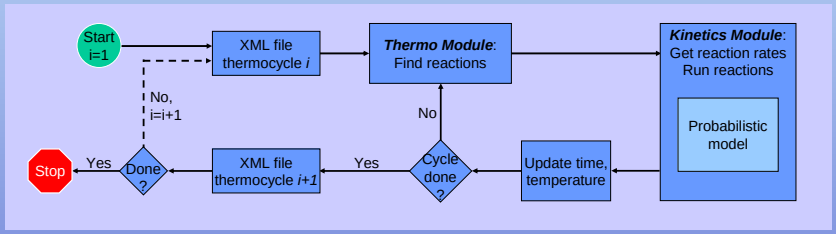


The model of a single thermocycle. We model 4 types of reactions: annealing, dissociation, polymerase extension, and melting. Although each of these predominantly occurs in a specific temperature regime, they are in constant competition throughout the thermocycle.

Methods

The simulation algorithm consists of a mixture of Gillespie's stochastic chemical kinetics algorithms: exact SSA, improved τ-leap SSA, and trapezoidal τ-leap SSA (with KINSOL).

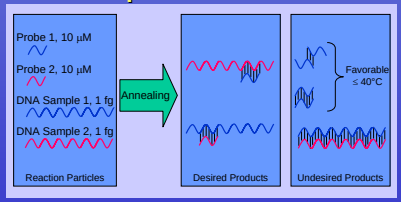
Hybridization kinetics come from Wetmur's theory. Polymerase extension kinetics are consistent with Taq polymerase data.



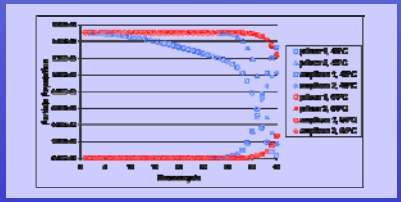
Simplified code overview.

Results

The updated code can simulate 40 thermocycles, giving results that make intuitive sense. The example below averages 10 seconds per thermocycle for a total of ~7 minutes, simulating an experiment that would have taken ~41 minutes to run. Future modifications aim to improve this time.



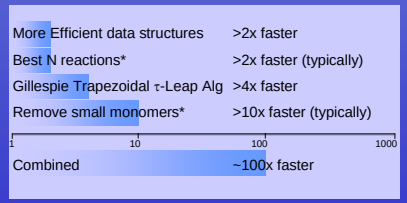
Example competing annealing reactions. At lower soak temperatures, annealing reactions are generally more favorable, opening pathways to undesired products.



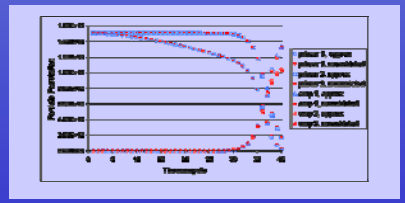
The system to the left is simulated through 40 thermocycles, at soak temperatures of 40°C and 64 °C. Note that more desired amplicons are output at 40 °C, where pathways to undesired products exist. Lowering the soak temperature increases the rates of all annealing reactions, which in this case outweighs the addition of undesired reactions.

Discussion

To improve running time, we introduce more efficient data structures, a speedier simulation algorithm, and reasonable approximations. These changes improve running time, and the approximations maintain accuracy. Simulation of a typical thermocycle on a moderately difficult problem takes < 20 seconds, making simulation of >40 thermocycles feasible.



When combined, the speed improvements give running times that are 2 orders of magnitude faster. *Approximations



The approximations maintain accuracy. The sample system is simulated for 40 thermocycles at 40 °C, with or without approximations.

This work was performed under the auspices of the U.S. Department of Energy by University of California Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

UCRLxx-xxxxxxx