Micro Flow Bio Molecular Computation (MF-BMC)

Ashish Gehani and John Reif

Duke University

- Introduction
- MEMS
- Architecture
- Chamber Operations
- Routing
- Applications
- Lower Bounds

Introduction

- Background
- Motivation
- Approach
- Relation To Previous Work
- Assumptions

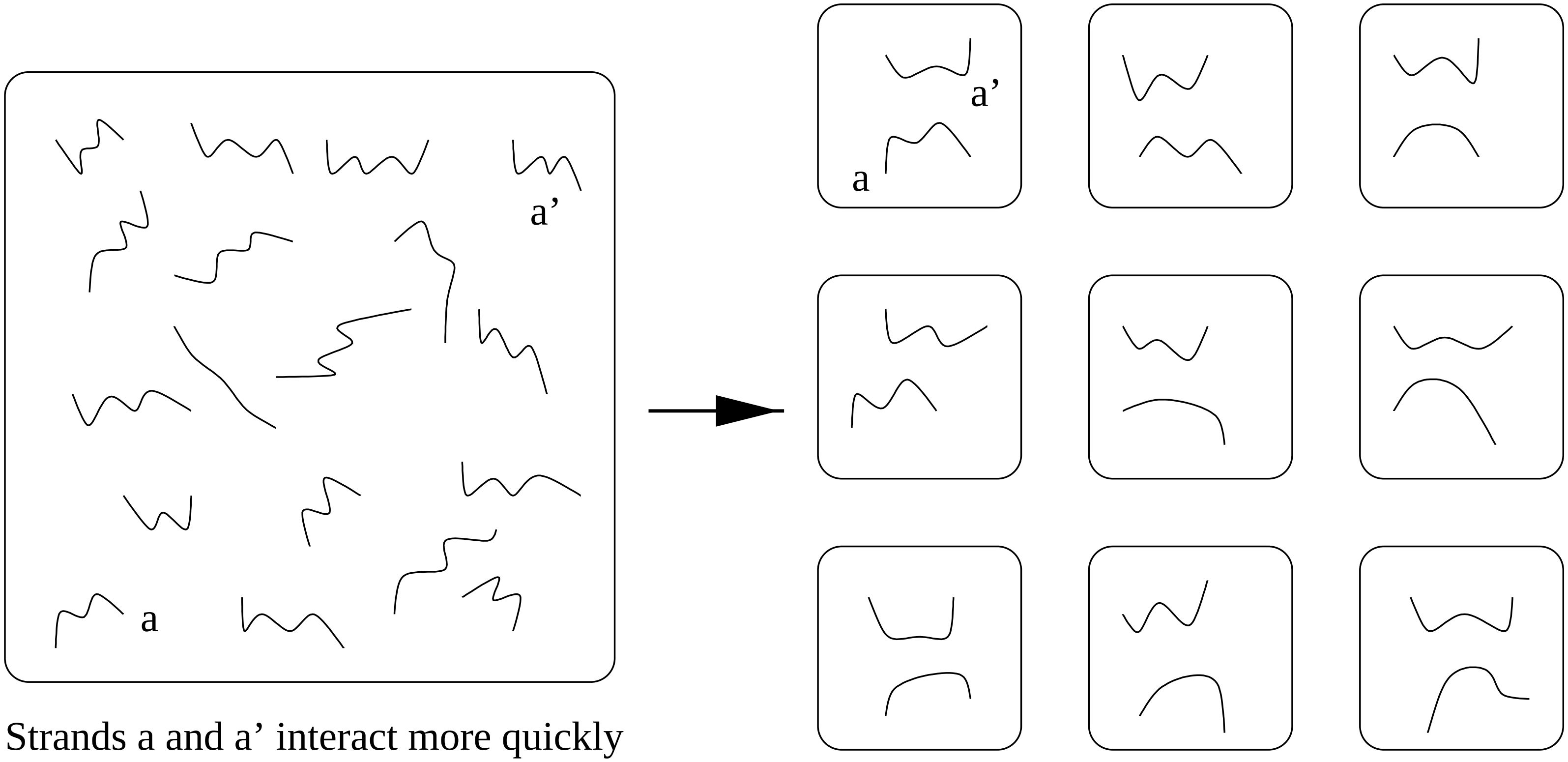
Background

Bio Molecular Computation (BMC):

- Can conceivably encode much data in limited space
- Can exploit massive parallelism of biochemical reactions
- Can control reactions using available RDNA technology

Motivation

- Kinetic models describe reactions statistically
- Predicting biochemical interactions at molecular level is hard
- Can predict result of interaction of two constituents
- Can not predict exactly when interaction will occur
- In smaller volumes specific reactants take less time to “meet”
- Operating independently on subsets of data is useful



Strands a and a' interact more quickly in a smaller volume.

Figure 1: Division of volume into multiple chambers

Approach

- Living systems process raw material using:

  (i) Chemical reactions, and (ii) Spatial control

- Spatial control is effected by:

  (a) Division of total volume into cells, and

  (b) Circulatory system routing material to cells

- Models of Adleman, Lipton, Reif focussed on (i)
- We propose a framework that addresses (ii)
- Divide total volume into many small chambers
- Embed devices in chambers to allow RDNA operations within
- Connect chambers with channels

Relation to Previous Work

- Number of laboratory steps limited due to:

  (i) Time for individual operations, and

  (ii) Human intervention required

- Small chamber volume alleviates (i)

- Three approaches have been taken to address (ii):

  (a) Molecular self-assembly by suitable encoding

  (b) Assume single technician, reformulate model

  (c) Introduce automation

- Micro Flow Bio Molecular Computation takes approach (c)

- Differs from previous approaches in two ways -
  (1) Highly parallel mechanical control, and
  (2) Efficient strand routing

Assumptions

Due to limits of fabrication technology:

- Small number of layers (in which chambers and channels exist)
- Size of mechanical components in system is lower bounded

For computation to be useful:

- Chamber volume is lower bounded
- Channel area of cross-section is lower bounded
- Rate of fluid transport must be sufficient

For reducing cumulative volume:

- Strand interaction is essentially random

Micro Electro Mechanical Systems (MEMS)

- MEMS used to effect automation
- MEMS are miniaturized mechanical devices
- Photofabrication using lithography allows low cost production
- MEMS used for controlling fluids are called micro-flow devices
- Examples are micro-actuators, micro-valves, micro-pumps, micro-sensors
- Each kind of device has many available implementations
- Choice must be based on factors such as cost and expected use
- E.g. micro-pump type depends on flow rate, channel gradient

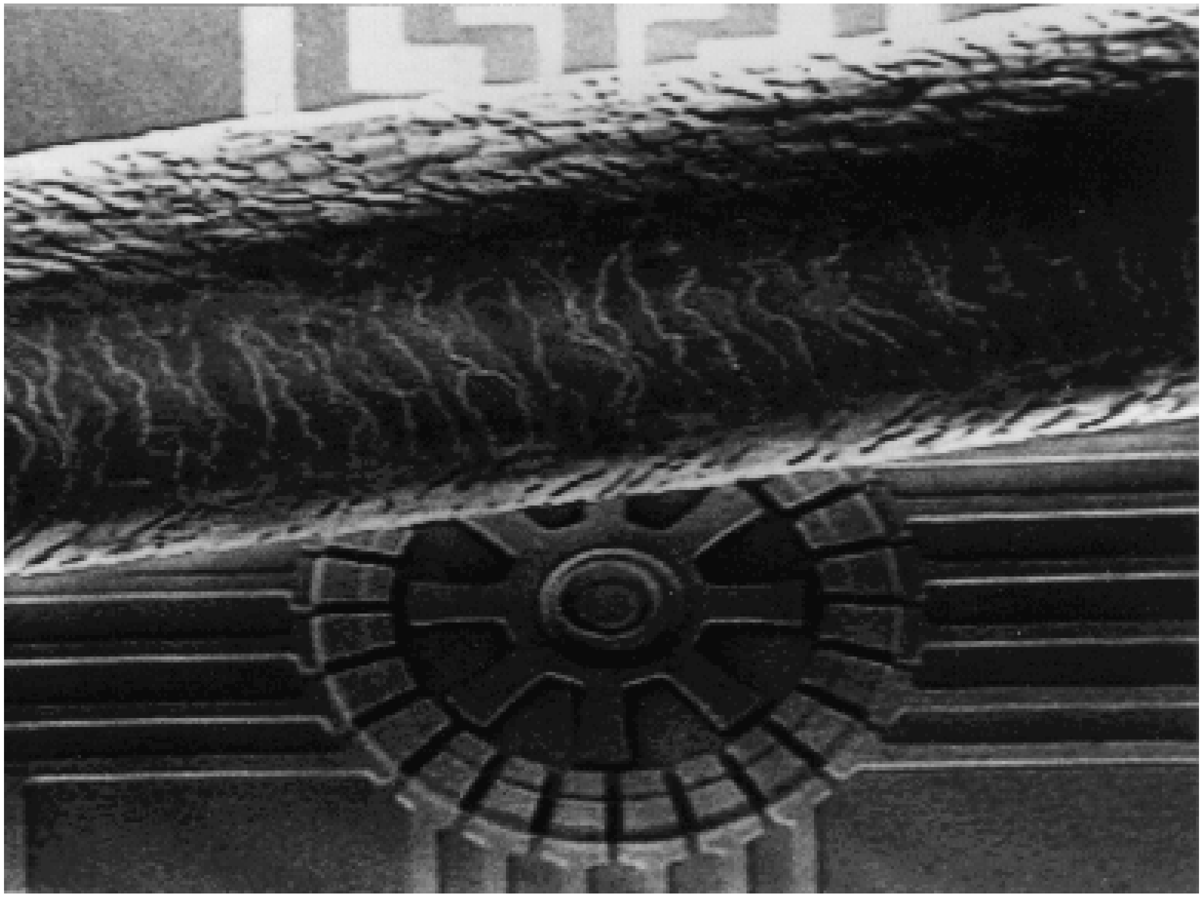
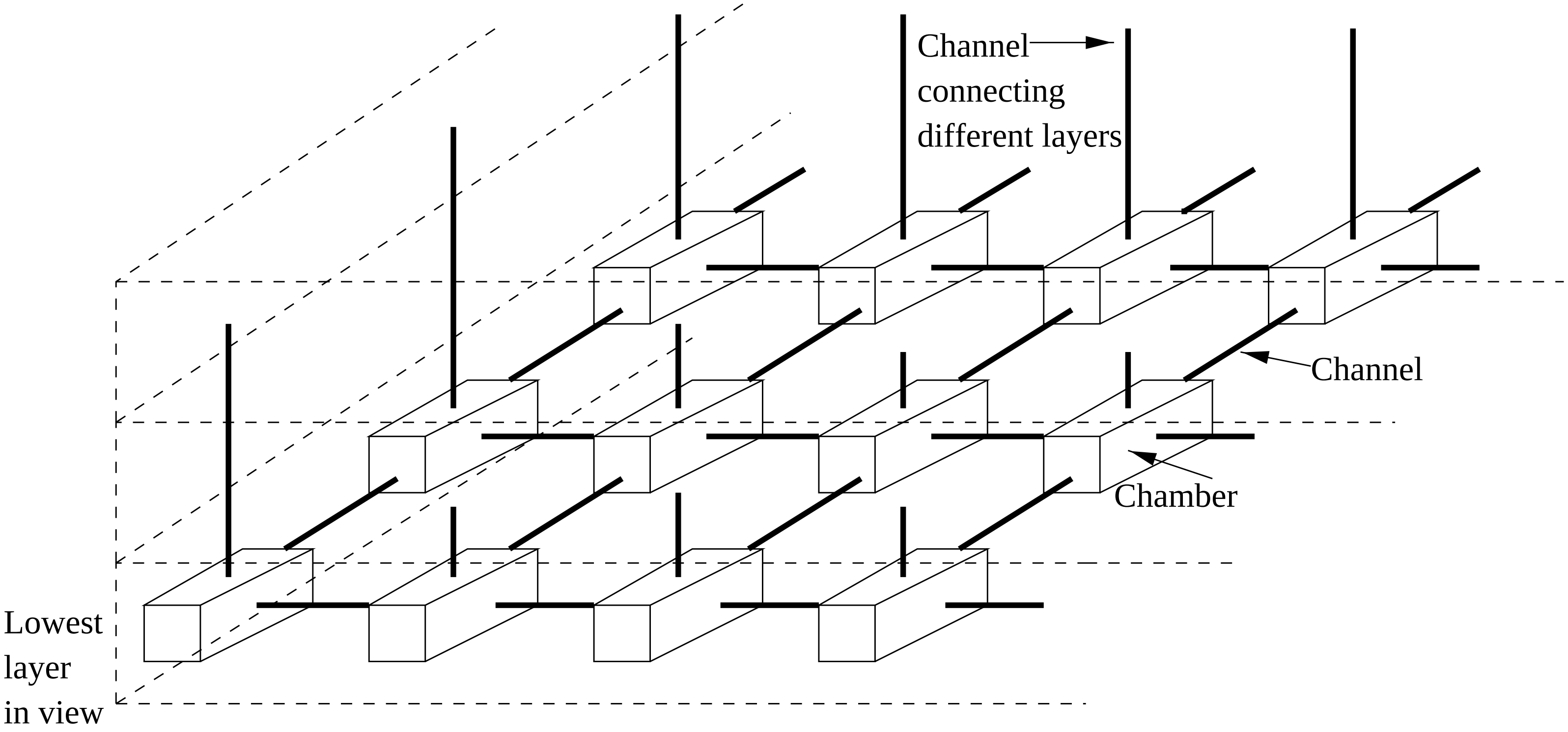


Figure 2: MEMS motor juxtaposed with hair follicle

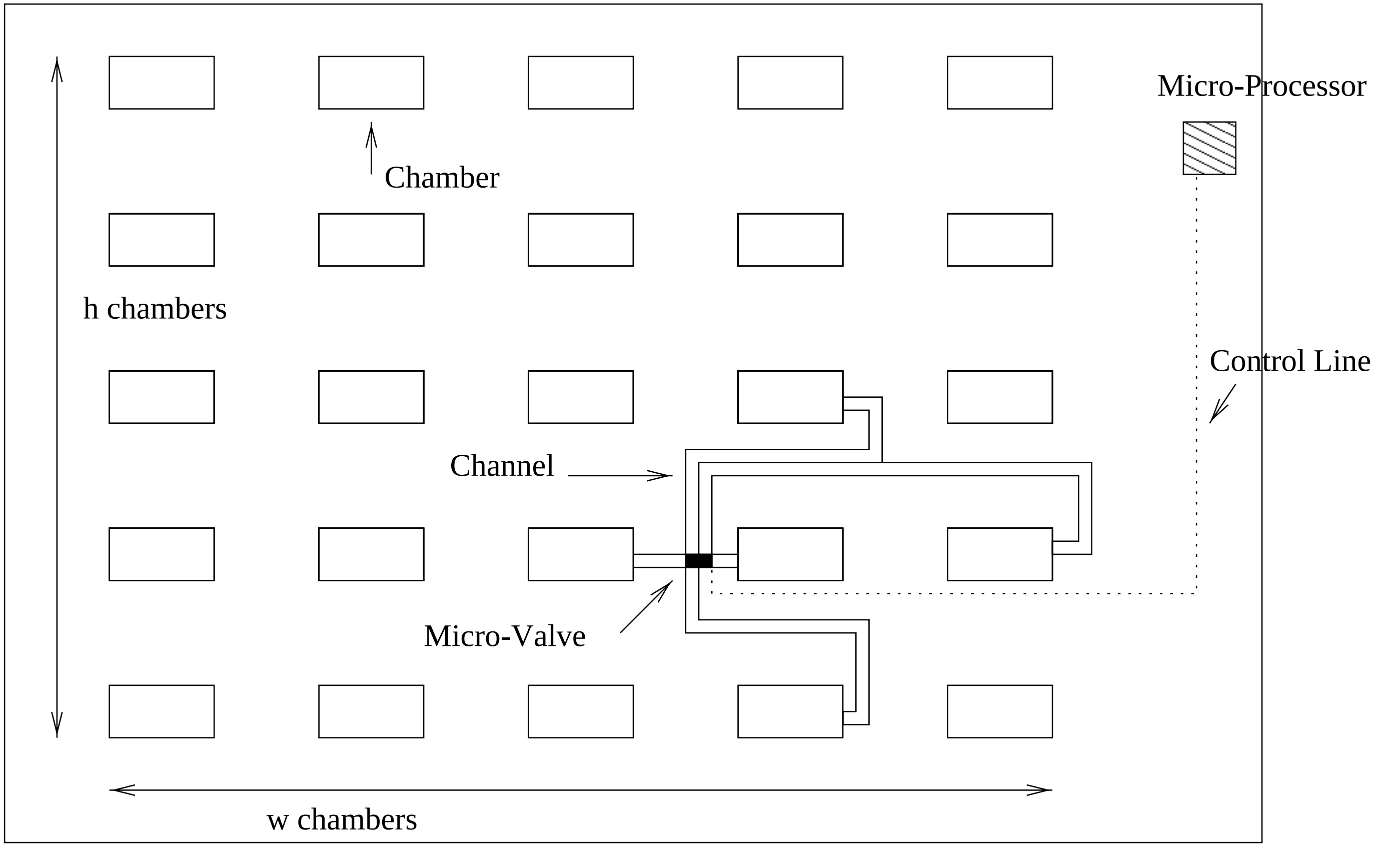
MF-BMC Architecture

- Fixed number of planar layers, due to fabrication constraints
- Each layer consists of a grid of n chambers
- Adjacent chambers are connected with channels
- Single microchip used to control all embedded devices
- Control lines from microchip to all chambers and channels



Proposed architecture for Micro-Flow Bio-Molecular Computation consists of several planar layers, in which an array of chambers connected by channels exists. Channels from layers above bring in data strands, reagents needed for reactions.

Figure 3: MF-BMC Architecture



In this example, we see a micro-valve that allows control over which of the four adjacent chambers' outlets are connected to a chamber's inlet. The micro-processor controls the micro-valve through a control line. (We use the convention that the outlets are located on the right and the inlets are on the left of the chambers.)

Figure 4: Controlling fluid movement with micro-valves

Chamber Description

- Contents from adjacent chambers introduced through inlet
- Contents moved to adjacent chambers through outlet
- Contents travel between layers through routing channels
- Separate valves and pumps control access to each portal
- Embedded components which enable RDNA operations:

  (i) Anchor strands with fixed sequence unique to the chamber,
  (ii) Micro-heater, (iii) Micro-thermometer,
  (iv) Optical emitter and (v) Optical sensor on opposing walls,
  (vi) Micro-solenoids along walls,
  (vii) Micro-cathode near inlet, (viii) Micro-anode near outlet

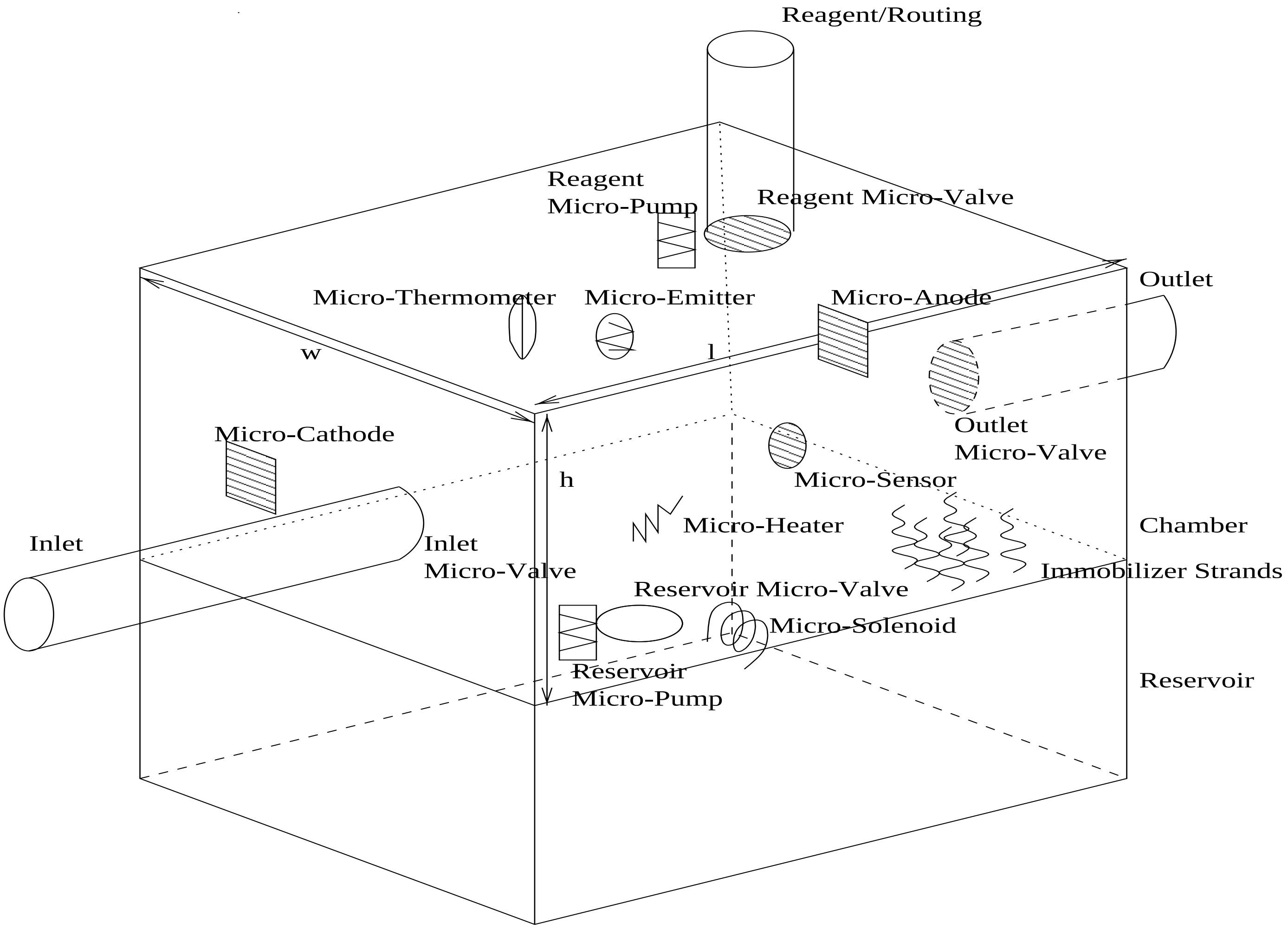
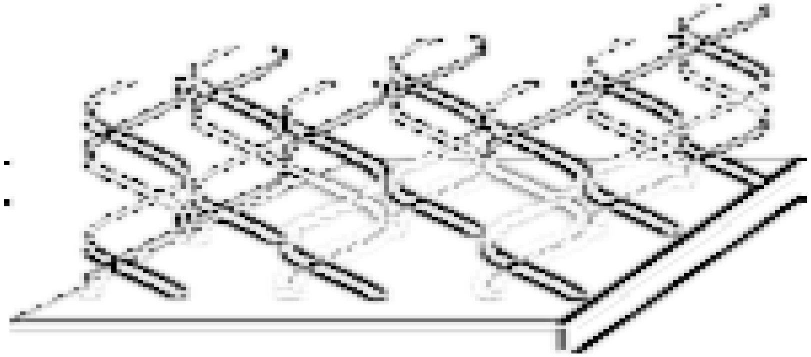


Figure 5: Prototypical chamber

Adapted Layers

RDNA enabling components not needed in specialized layers:

- Routing layer chambers only serve as intermediate storage - this layer's purpose is to facilitate fluid transport between chambers
- Material layer chambers adapted for input and output:
  - (i) Chambers store input/output sequences, reagents, etc.
  - (ii) Replace chambers with 2D DNA Chips:
    - ∗ Can be used for strand generation, or
    - ∗ Can be used for sequence determination



Adleman and Lipton Model Operations in a Chamber

Can execute operations in Adleman and Lipton models:

- Merge - use micro-valves and micro-pumps
- Copy - add salt buffer, DNTPs, primers, polymerase, ligase - use heater and thermometer to cycle through 95-55-75 deg C
- Detect - use optical emitter and sensor for density reading
- Separate - add strands complementary to anchors and target
- Ligate - add ligase

Reif RDNA Model Operations in a Chamber

Can execute operations in Reif's RDNA model:

- Denature - add buffer, activate micro-heater as needed
- Anneal - effected as chamber cools
- Cleave - add nicking enzyme
- Select - difficult to effect, potential methods:
  - Add low melting temperature agarose gel to chamber between source and destination chambers,
use electrodes to effect selective transport
  - Use difference in flow rates induced by micro-pumping
  - Apply membrane ultrafiltration

Routing

- Strand routing within a chamber
- Efficiency of strand routing
- Routing along a linear array of chambers
- Grid routing

Intrachamber Strand Routing

- "Routing" is algorithm specified interaction between strands
- Routing used by operations Anneal, PA-Match, etc.
- Given collection of N distinct strands
- Aim for one or more molecules of product
- Aim for production probability independent of N
- Using $O(N^2)$ volume, there are $O(N^2)$ potential reactant pairs
  Probability of mismatch for all pairs is $O((1-\frac{1}{N^2})^{N^2}) = O(1)$
  Goal achieved in $O(1)$ time
- Using $O(N)$ volume, there are $O(1)$ potential reactant pairs
  Probability of mismatch N times is $O((1-\frac{1}{N})^{N}) = O(1)$
  Goal achieved in $O(N)$ time

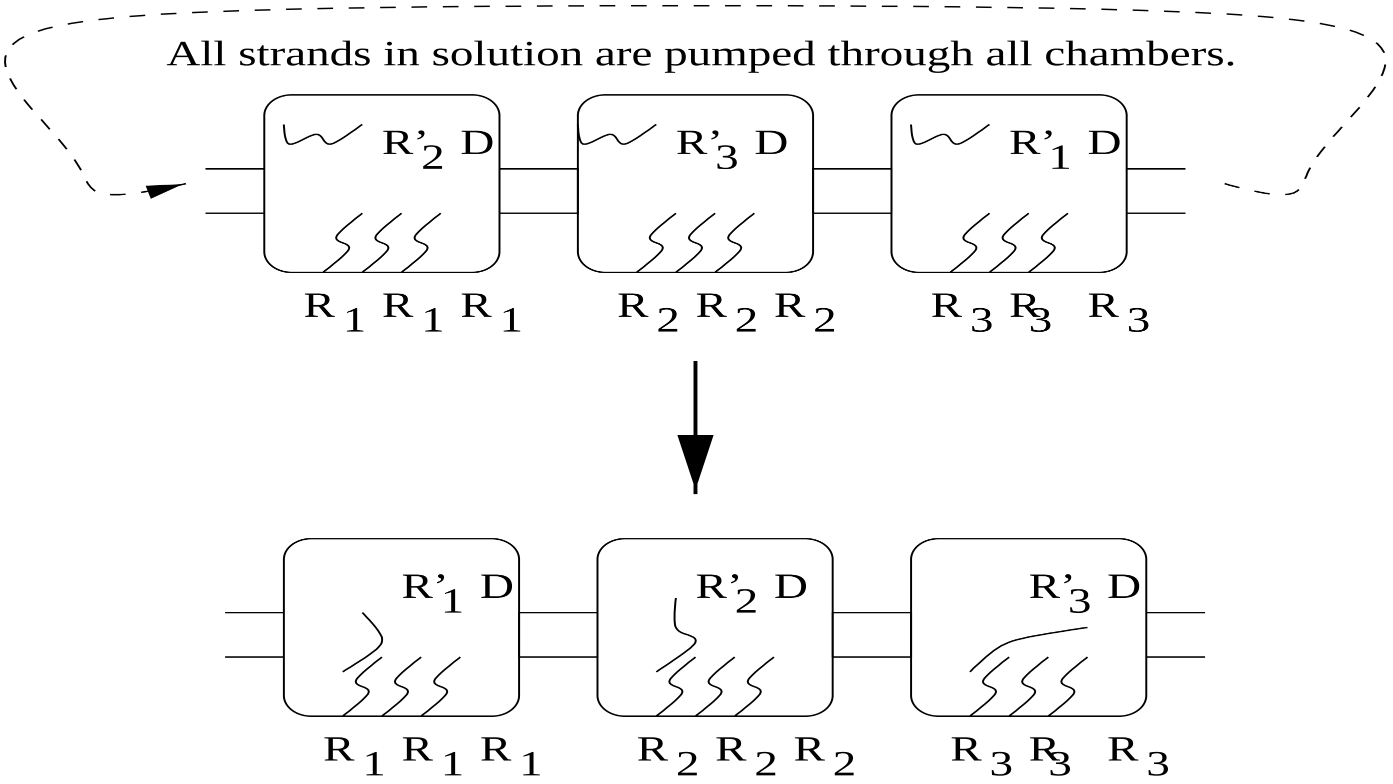
- Previous result in agreement with kinetics:
  - Routing rate is proportional to reactants' concentrations
  - Production rate $\propto \frac{1}{N} \cdot \frac{1}{N}$
  - Use volume inversely proportional to forward reaction rate
  - Volume $\propto \frac{1}{\frac{1}{N} \cdot \frac{1}{N}} = N^2$
  - Obtain small quantity of product in $O(1)$ time

Efficiency of Strand Routing

- Given N distinct strands, wish interaction in $O(1)$ time
- Volume needed is $O(N^2)$
- If possible, divide N strands among n chambers
- Now volume needed in individual chamber is $O((\frac{N}{n})^2)$
- Total volume is $n \cdot (\frac{N}{n})^2 = \frac{N^2}{n}$
- Cumulative volume for intrachamber strand routing drops by n
- Useful for large n, say $n = 10^6$
- n chosen to balance interchamber and intrachamber routing

Strand Routing on a Linear Array

- Each chamber has distinct anchor strands embedded
- Data strands' prefix complementary to destination's anchor
- All strands micro-pumped repeatedly through all chambers
- Strands attach at their destinations
- Isolate all chambers using micro-valves
- Heat chambers to free data strands from anchors
- With n chambers, interchamber routing takes $O(n)$ time



All strands in solution are pumped through all chambers.

Data strands with a prefix that matches the chamber address (encoded in the immobilized strands attached in the chamber) remain in their destination chamber since the prefix anneals to the immobilizer strands of the destination chamber.

Figure 6: Strand routing on a linear array

Routing N Strands on a n Node Grid

- Each chamber has distinct anchor strands embedded
- Anchor strand encodes row as prefix and column as suffix
- Data strands' prefix complementary to destination's anchor
- Routing to correct row:
  - Columns of chambers isolated using micro-valves
  - $\frac{N}{\sqrt{n}}$ strands micro-pumped into each column of chambers
  - Strands attach in chambers corresponding to correct rows
  - Isolate all chambers using micro-valves
  - Heat chambers to free data strands from anchors
  - Strands are now in chambers in destination row

- Routing to correct column:
  - Rows of chambers are isolated using micro-valves
  - Strands circulated among chambers in respective rows
  - Temperature raised, stable matches remain immobilized
  - Stable matches anneal at anchor's row and column encoding
  - Isolate all chambers using micro-valves
  - Heat chambers to free data strands from anchors
- Interchamber routing to correct row and column in $O(\sqrt{n})$ time

Applications

- BMC useful for handling large amount of information
  Natural application would be database implementation
  Merging records done using Join operation
  Join operation implemented using PA-Match
- Set of RDNA operations in chamber are Turing complete
  Apply current parallel algorithms using chambers as processors
- PRAM simulation using PA-Match previously shown
  Using $O(\sqrt{n})$ time interchamber and intrachamber routing,
  Volume required is $O(\frac{N^2}{n^{\frac{3}{2}}})$ compared to previous $O(N^2)$ volume

Lower Bounds

- With finite strand length, system has B bit memory capacity
- $B = \Omega(\log d)$, where:
  MF-BMC has d distinct configurations
- $BT = \Omega(|L|)$, where:
  L is largest input/output encoding,
  T is time to solve problem
- $BT^2 = \Omega(I^2)$, where:
  I is information content of problem as defined in VLSI models