The Art of Molecular Programming

Draft Table of Contents

The textbook is divided into three parts. 'Part 0' is preliminary material, followed by Part I: Structures, Part II: Circuits and Part III: Interfaces. Each part is divided into chapters (numbered), and chapters into sections (lettered). Bullets represent keywords, or otherwise general guidance on the content of the section.

- *Asymmetric PCR*
- *Cell culture*

ii. **Purification**

- *Agarose/PAGE purification*
- *Chromatography (HPLC, LC, maybe capillary?)*
- *Centrifugation (filter, gradient, ultracentrifugation)*

e. **Experimental Methods for Analysis:**

i. **Gel electrophoresis (focus on densitometry etc)**

- *Intercalating dyes vs fluorophores or radiolabeling*
- *Separation quality as a function of gel % and DNA length?*
- *Role of buffer conditions*
- *Pulse field gradient electrophoresis?*

ii. **Microscopy (AFM, TEM/cryoEM/SEM, fluorescence/superresolution)**

- *Contrast mechanisms and their drawbacks (stains being inconsistent, AFM always convolutes tip, etc)*

iii. **Spectroscopy (fluorescence/CD/etc)**

- *Relative sensitivities*
- *Background reporter signal?*

iv. **Sequencing**

f. **Example: Experiments for thermodynamics and kinetics of hybridization**

Part I Structures

1. **"From molecules to variables"** ***(Chemistry)***

a. **What are information bearing molecules?** ***(biopolymer building blocks)***

- *nucleic acids (basic features and geometry)*
- *proteins/peptides; Protein geometry (alpha helix, ...)*
- *Other (PNA, L-DNA, new nucleotides, polysaccharides, block-copolymers, other supramolecular programmable molecules)*

b. **How do molecules interact** ***(the glue)***

- *Chemical bonds and interactions (a primer in chemistry)*
- *Intramolecular forces via covalent bonds: peptide bonds, ester bond*
- *Intermolecular forces and non-covalent bonds: Van der Waals, hydrogen bond, stacking interactions, electrostatic interactions, Debye screening. (Table with relative bond strengths)*
- *Entropic forces: depletion, steric, fractionation, crowding, polymers*

c. **How can we represent molecules digitally?**

- *Structure abstraction layers: Primary, Secondary, Tertiary structure*
- *Sequence abstraction layers: domain level vs nucleotide level*

- *Notations: Dot parens plus notation, DU+ notation, ...*
- *Notions of "valid" conformations (nearest neighbor model)*
- *Intuition behind coarse-grained representations and macrostates*
- *How can we represent molecules graphically*

2. **Molecules as construction material**

   a. **DNA properties**
   - *data sheet to find useful parameters.*
   - *Crossover motifs (anti-parallel and parallel)*
   - *Special motifs (I motif, G-quads, kissing loops, aptamers)*
   - *Biophysical influence of buffer conditions on DNA structure*

   b. **RNA properties**
   - *data sheet to find useful parameters.*
   - *Special motifs (I motif, G-quads, kissing loops, aptamers)*
   - *Biophysical influence of buffer conditions on RNA structure*

   c. **Protein properties**
   - *data sheet to find useful parameters.*
   - *Enzymatic activity and binding pockets, ...*
   - *Special motifs (DNA binding)*
   - *Biophysical influence of buffer conditions on enzyme activity and protein structure*

3. **Introduction to self assembly. Basics of molecular self-assembly *(theory)***

   a. **Sequence based models of intramolecular self assembly**
   - *Levinthal's paradox. Why can we engineer DNA/RNA but proteins are still hard?*
   - *DNA/RNA Sequence design. Given a sequence, what is the structure? Given a structure, design a sequence?*
   - *Proteins engineering*

   b. **Principles of natural multi-component self-assembly processes**
   - *Examples (viruses, lipid micelles, nanowires, crystals, self-assembled monolayer, hydrogels, lipids, block co-polymers)*
   - *Biophysical models: The role of thermal energy driving self-assembly, nucleation, kinetic/thermodynamic control, stoichiometry, cyclisation*
   - *Stochasticity — having equilibrium close to thermal energy (kT) means there will be equilibrium population of defects*

   c. **Tile self-assembly**
   - *Tiling theory (Wang tiles, Penrose tiles, ...)*
   - *What shapes can be assembled in which tile model?*
   - *aTAM, kTAM*

- *Sequence design (Box OK)*

d. Deeper into computation and self-assembly (Fundamental limits of computation — reference chapter 8)

- *Finite state automata*
- *Tiles and algorithmic self assembly*

4. Programmed molecular self-assemblies *(experiments)*

a. Scaffold-less DNA assemblies

- *ss Tiles,*
- *dx Tiles*
- *HCR*
- *Simple polyhedra: Seeman cube, tetrahedron, Yamuna's icosahedron, Mao's Bucky ball and octahedron.*

b. DNA Origami

- *The concept*
- *Design principles, cooperativity*
  - *Sequence design constraints*
- *Design and simulation tools*
- *"Wireframe origami" examples 1D / 2D / 3D structures*
- *Production and purification*
  - *Custom scaffold design*
  - *Thermo/stoichiometry*

c. Periodic and Multi component assemblies [Higher complexity structures, DNA LEGO, Conformational changes of origami]

- *Shape complementary, base stacking*
- *Self-limiting assemblies (rings)*
- *Fractal assemblies*
- *Lattices, ribbons, nanotubes, and crystals*
- *Interlocked assemblies (e.g. origami rotaxanes)* $\rightarrow$ *connection to mechanics*

d. Dynamic rearrangements of structures

- *DNA tweezers, DNA Walkers, burnt-bridge motor*
- *Nanomechanical devices. Mechanical constructs / active components/ machines / walkers (comparison with molecular motors/enzymes)*
  - *Castro's work*[1]
- *Nanomotors*

e. Other phases of DNA structures (Physical properties)

- *DNA hydrogels*
- *DNA liquids*

[1] https://www.pnas.org/content/112/3/713/tab-figures-data

f. **Hybrid DNA objects**
   - *Colloids*
   - *Gels*
   - *Polymer/DNA brushes (surface/colloid coatings)*
   - *Coacervates*
   - *Grafted nanoparticles*

g. **Other biopolymer building materials — Beyond DNA**
   - *RNA*
   - *Protein engineering (David Baker, U Washington)*
   - *Polysaccharide, spider silk, ...*
   - *DNA-chimera assemblies (protein bioconjugation, eg the work of Nicholas Stephanopolous)*

Part II Circuits

1. **Introduction to Computation**
   a. **Intro to computation / information processing**
   b. **Conventional computation**
   c. **Background: Boolean logic, Turing machines, register machines**
   d. **Examples of non-conventional computing**
   e. **Examples of natural computing**
   f. **Guiding principles? Molecular programming "languages", modular/circuits intro**
   g. **Advanced: limits on computation / computability frameworks?**

2. **Programming molecular behaviors over time (CRNs)**
   a. **Introduction**
   b. **Computing with Stochastic CRNs**
      - *Theory example and analysis: min/max and Boolean logic programming with CRNs*
      - *Biology example: something simple*
      - *Computational Power of Stochastic CRN*
      - *Time complexity of stochastic CRN*
   c. **Compute with distributions**
      - *Computing with Deterministic CRNs*
      - *Theory examples: circuits, boolean circuits, oscillators, bistability, etc.*
      - *Well-mixed CRNs as example of analog computing*
      - *Computing functions (e.g. $y = kx$)*
      - *Approximate majority*

- *Dynamic system: Oscillators, bistability*
- *Biology example: (predator prey / ecology models)*
- *Computational power of deterministic CRNs*

d. Advanced topics in CRNs

3. Nucleic acids as a universal substrate for molecular programming

a. Introduction (reference Foundations for overlapping topics)

- *Reference to Structures section on background: nucleic acid hybridization and thermodynamics*
- *Background: nucleic acid branch migration*
- *Toehold-mediated strand displacement*
- *Experiment examples: toehold exchange reaction*

b. Strand displacement cascades

- *Theory & experiment: DNA universal substrate for CRNs*
- *Boolean circuits: seesaw circuit theory*
- *Box: discuss considerations and strategies in e.g. designing mutually orthogonal sequences*

c. Theory & Experiment examples: 3wDSD circuits

- *Oscillators*
- *Bistable memory*
- *Approximate majority*
- *Amplification, catalysis, and thresholding*
- *Seesaw circuits*
- *Neural networks*

d. Advanced designs in DNA strand displacement cascades

- *Toehold activation*
- *Mutation detection; use mismatch to control kinetics?*
- *Theory & Experiment: 4wDSD circuits*

e. Leakless circuits

- *Why do circuits leak and why leaks are problematic*
- *How to avoid leak. Examples of leakless circuits*

4. Incorporating enzymes into nucleic acid cascades

a. Introduction

- *Background: enzymatic behaviors*
- *Early example: traveling salesman?*

b. DNA polymerase-based circuits

- *PEN circuits, predator prey*
- *APR, PER circuits (whiplash PCR?)*

- *Shah et al. work on logic gates*

c. **Transcribed RNA-based circuits**
   - *Toehold switches*
   - *Conditional CRISPR*
   - *Oritatami*
   - *Genelets*

d. **Translated protein-based circuits**
   - *Repressilator*
   - *References to good syn bio resources?*

5. **Spatially-Organized Circuits**

   a. **Introduction: advantages of spatial structures**
   b. **Background: compartmentalization in biology**
   c. **Surface CRNs**
      - *Surface DNA circuits and DNA walkers*
   d. **Droplet-based computing / "synthetic compartmentalization"**
   e. **Reaction diffusion circuits**
   f. **Microfluidics Breakout Box**

6. **Advanced topics in tile assembly (algorithmic self-assembly)**

   a. **Computational complexity**
   b. **Wang tile Turing machine implementation?**
   c. **CRN-Tile hybrid systems?**
   d. **Relationships between DSD and TAM etc, e.g. TBNs (Thermodynamic Binding Networks)**

7. **Conclusion**

Part III Interfaces

1. **Introduction to Interfaces & Applications**

   a. **Intro**
   b. **General preview of interfaces (and the techniques used) with other fields through**
      - *Electronic interfacing*
      - *Chemical and physical interactions*
      - *Interacting with biology*

2. **The Interface between Traditional and Molecular Computers**

   a. **Background/Foundations**

- *Sequencing & Synthesis (reference to Foundations); NGS*
- *Nanopores*
- *Microarrays (mention how big of an industry this is...multi billion...e.g. 23 and Me uses it exclusively)*

b. DNA Computation

- *Why use molecules to do traditional computations?*
  - *Pros: lower energy usage, parallelism , physical stability, data density, etc.*
  - *Cons: expensive at high throughput, etc*

c. DNA Data Storage

- *Encoding data from digital data to DNA*
- *Synthesizing DNA*
- *Storing DNA*
- *Recovering stored DNA*
- *Sequencing the DNA*
- *Decoding the DNA back to digital data*

3. Chemical and Physical Interactions

a. Material science

- *Patterning non-DNA molecules*
- *Surface assisted methods*

b. Chemical control

- *pH [primer on DNA triplexes]*
- *Ions*
- *crowding agents*

c. Magnetic control

- *swimmers*
- *beads (for purification and/or computation)*

d. Optical control

- *DNA plasmonic and photonic circuits*
- *DNA scaffolding of non-DNA reaction networks*
- *Dynamic control of optical DNA devices*

e. Thermal control

- *Vinothan Manoharan's work*[2]
- *Jaeseung Hahn's work*[3]
- *More examples*[4][5]

[2] https://science.sciencemag.org/content/347/6222/639.abstract
[3] https://academic.oup.com/nar/article/47/20/10968/5580905
[4] https://onlinelibrary.wiley.com/doi/abs/10.1002/adfm.201706410
[5] https://pubs.acs.org/doi/10.1021/nn4030543

f. **Electronic driven redox power sources**
   - *Electrically driven process*[6]

g. **Future**

4. **Interacting with Biology, Medical Diagnostics and Therapeutics**

   a. **Biology toolbox primers — topics not included in Foundations primer**
      - *Cutting DNA (e.g., restriction enzymes, CRISPR, ribozymes?; talk about synbio gene editing using CRISPR and other restriction enzymes)*
      - *DNA Synthesis (e.g., ligation, overhang PCR)*

   b. **Diagnostics**
      - *How do current diagnostics work [primer on various techniques, biomarkers]*
        - *Which particular diseases will benefit from in-depth diagnostics? Why do we need to know about so many biomarkers?*
          * *Shotgun / Microarrays*
      - *Why are some molecules difficult to detect? What are the advantages of DNA nanotech methods over the standard techniques?*
        - *Methods*
        - *Aptamers — How can we program DNA to bind with targets other than DNA/RNA? [primer/external reference on selex]*
        - *toeholds*
        - *electrochemical/fluorescence*
        - *SPR (surface plasmon resonance) and QCM (quartz crystal microbalance) and FET (field effect transistor)*
        - *etc.*

   c. **Therapy**
      - *Programmable nucleic acids-based therapeutics*
        - *Protein production silencing or activation — mRNA mimics, siRNA etc. (Pfizer/Moderna vaccines are real-world examples of this) [primer/external reference on transcription/translation]*
        - *Aptamer therapeutics*
      - *Protection of therapeutic agents e.g. tetrahedron structures.*
        - *Different protection methods for nucleic acid nanostructures (e.g., polylysine, protein, lipid, etc.)*
      - *Examples of approved or under-trial DNA nanotech drugs.*
      - *Drug Delivery*
        - *Why is it beneficial to have targeted delivery of drugs in the body?*
        - *How are drugs loaded onto DNA structures? Various examples available.*
        - *Mechanisms for release of drugs in particular areas of the body.*

[6] http://www.francescoriccilab.com/wp-content/uploads/92.pdf

d. Burgeoning biological interfaces

- *Cellular interactions Communicating with cells and cell surfaces*
  - *Why do we want to create artificial cell signal controllers? [primer/external reference on cell signalling systems]*
  - *DNA nanotech channel receptor mimics (various versions using DNA origami; include lipid bilayer section here)*[7]
- *Replicating cellular components*
- *Proteinosomes (?)*

e. Challenges in working in a biological medium

- *Dealing with toxicity*
- *Preventing degradation — [primer or external links to well-known DNA modifications and their functionality, use of XNAs?]*
- *Examples of Chemical modifications, ligand attachments and encapsulating protection (nanoparticles)*
- *How to scale up "bio factories", using cells to manufacture molecules is difficult*
- *Working with RNA (mRNA, RNA origami, etc.)*
  - *Why RNA?*
  - *What's different from DNA?*

f. Regulation & Ethics — breakout box

- *High level overview of the general international medical regulatory processes, with some specific examples from, e.g., the FDA approval process*
- *Maybe briefly mention other regulatory processes for, e.g., synthetic crops*
- *DNA Data Storage — chemical waste*

[7] https://www.cell.com/trends/biotechnology/fulltext/S0167-7799(21)00034-2