



Programmable Synthetic Systems and Materials in Synthetic Biology

Randall Hughes

Andrew Ellington Laboratory

The University of Texas at Austin

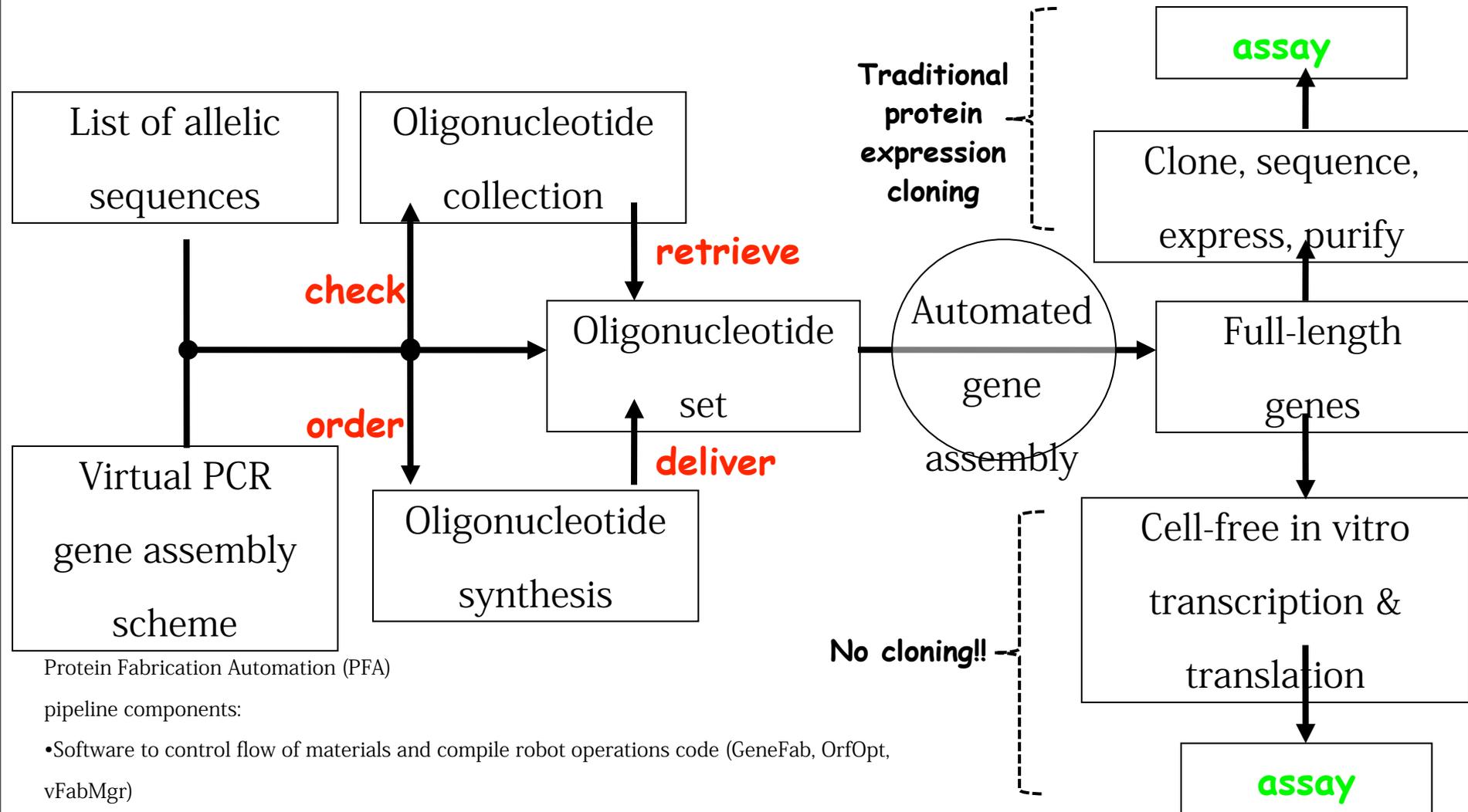
Institute for Cellular and Molecular Biology and

Applied Research Laboratories

Outline

- Gene Fabrication-Protein 'Write' Facility
- Examples of what we do with our Fab.
- Nucleic Acid diffusion controlled systems.
- Programmable Nucleic Acid-Protein Mimics.

Synthetic Gene Construction-Texas Style



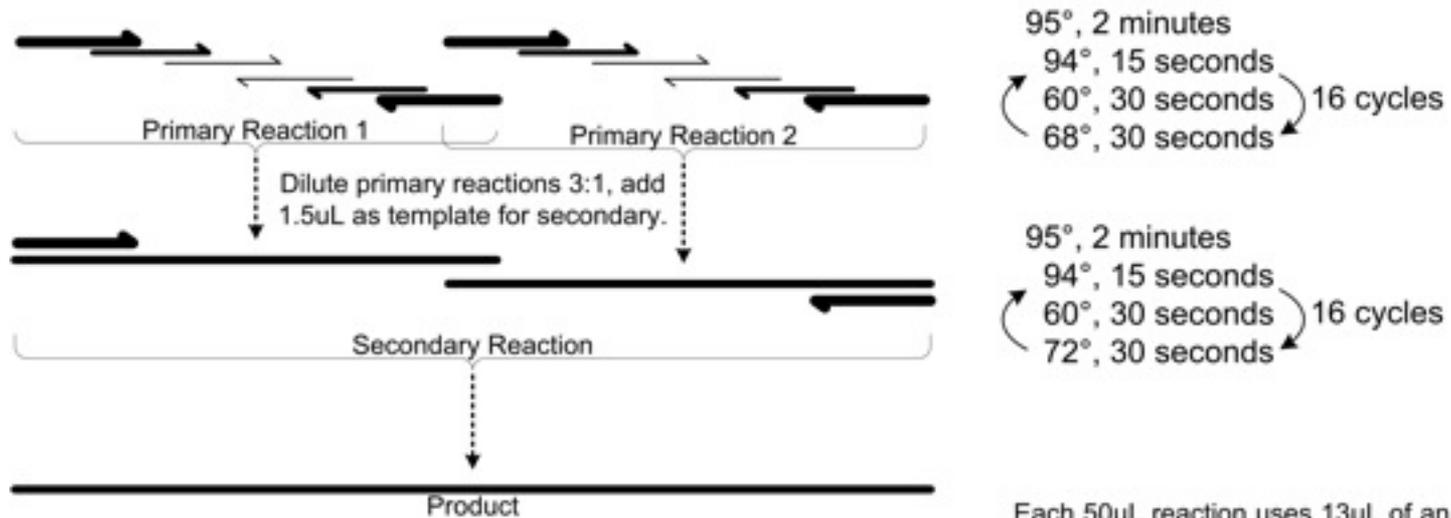
Protein Fabrication Automation (PFA)

pipeline components:

- Software to control flow of materials and compile robot operations code (GeneFab, OrfOpt, vFabMgr)
- Robust PCR gene assembly scheme
- Full-length ORF genetic selection
- Optimized in vitro expression
- Small-scale functional assays

Cox et al. 2007 *Prot. Sci.*, 16:379

Automated PCR Gene Assembly



$$[P_i] = \left(\frac{[P]_T}{\sum_{j=0}^{n-1} c^j} \right) c^{n-i}$$

$[P_i]$ = concentration of i^{th} primer pair
 c = constant, $(0.65 \leq c \leq 0.75)$
 $[P]_T$ = total [primer] (=600nM)

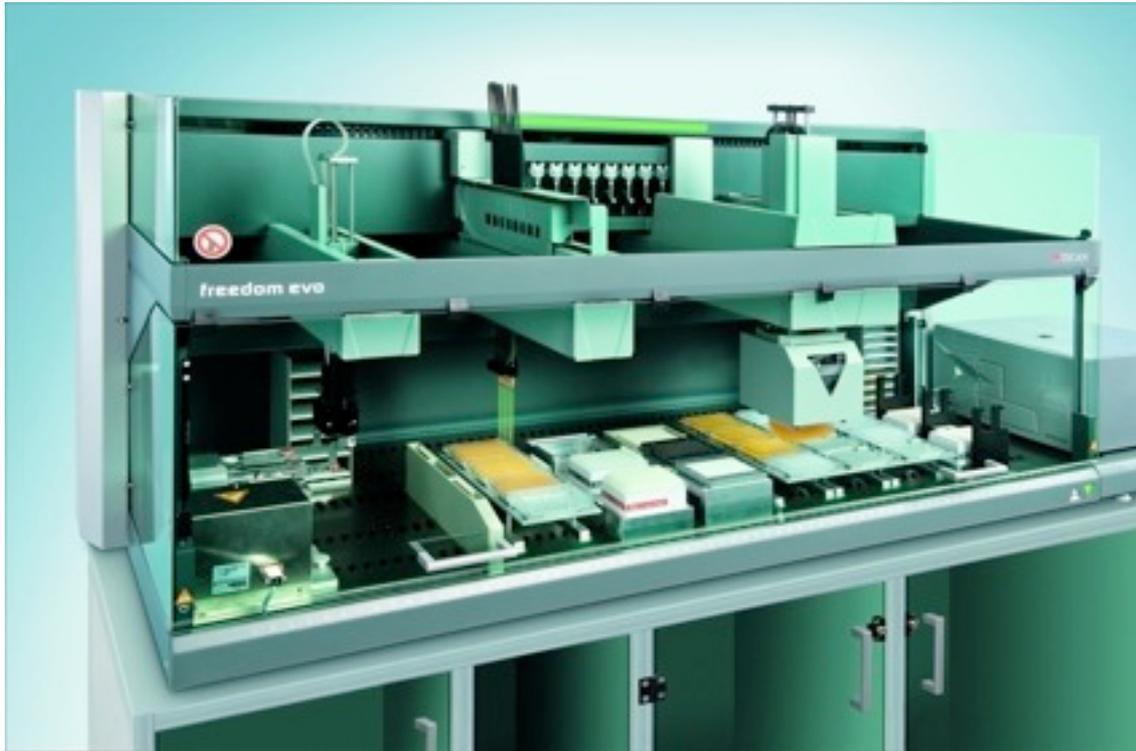
⇌ 67.1 nM
 ⇌⇌ 95.9 nM
 ⇌⇌⇌ 137 nM

Each 50 μ L reaction uses 13 μ L of an amplification master mix comprised of materials from the NovaGen KOD Hot-Start polymerase kit:

500mL 10X KOD Buffer
 500mL dNTP mix
 300mL 25 mM MgSO₄
 100mL Hot-Start KOD

Hughes, RA *et. al.* 2010 Meth. Enzymol. *In press.*

ARL-UT Gene Fabrication Facility

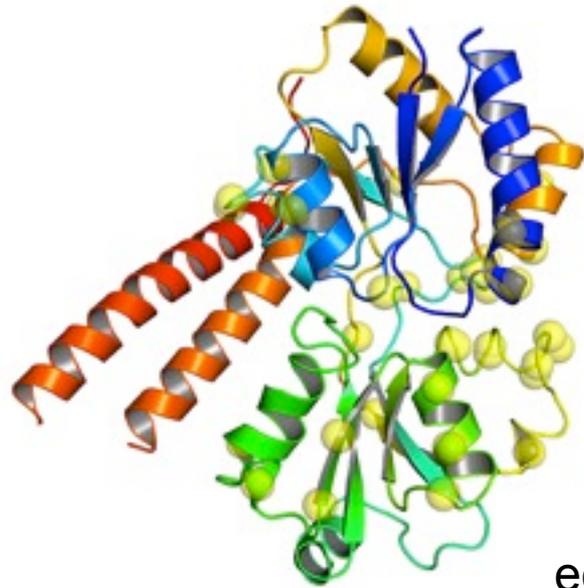


Design of synthetic schemes, oligonucleotide synthesis and databasing, and generation of robotic operations scripts are all automated in custom software.

Total throughput with existing equipment (limited by oligonucleotide synthesis):
Approximately 100-150 kilobases per week (100-150 completely unique 1,000 bp genes).

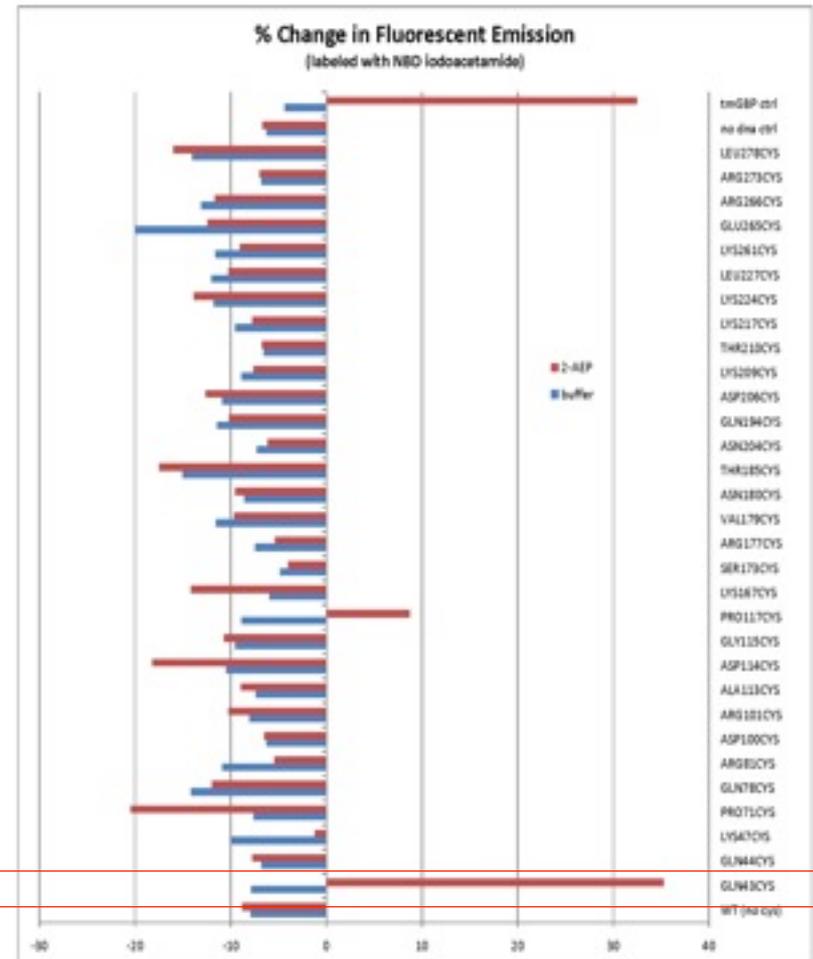
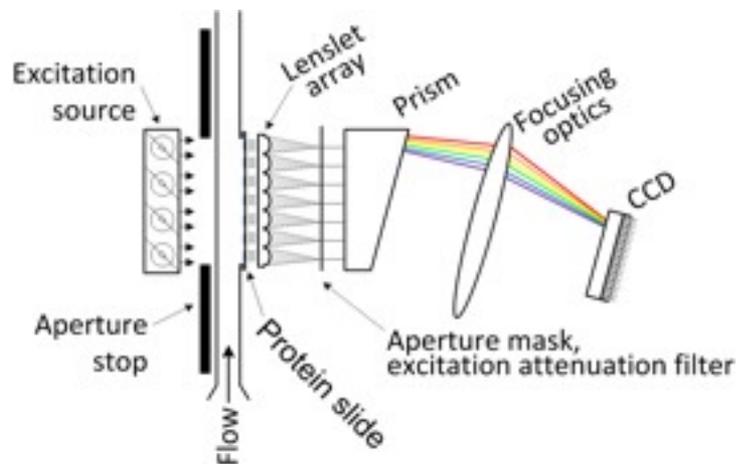
Substantial efficiencies are achieved when building gene variants (re-use of oligonucleotides)

Biosensor Discovery and Development

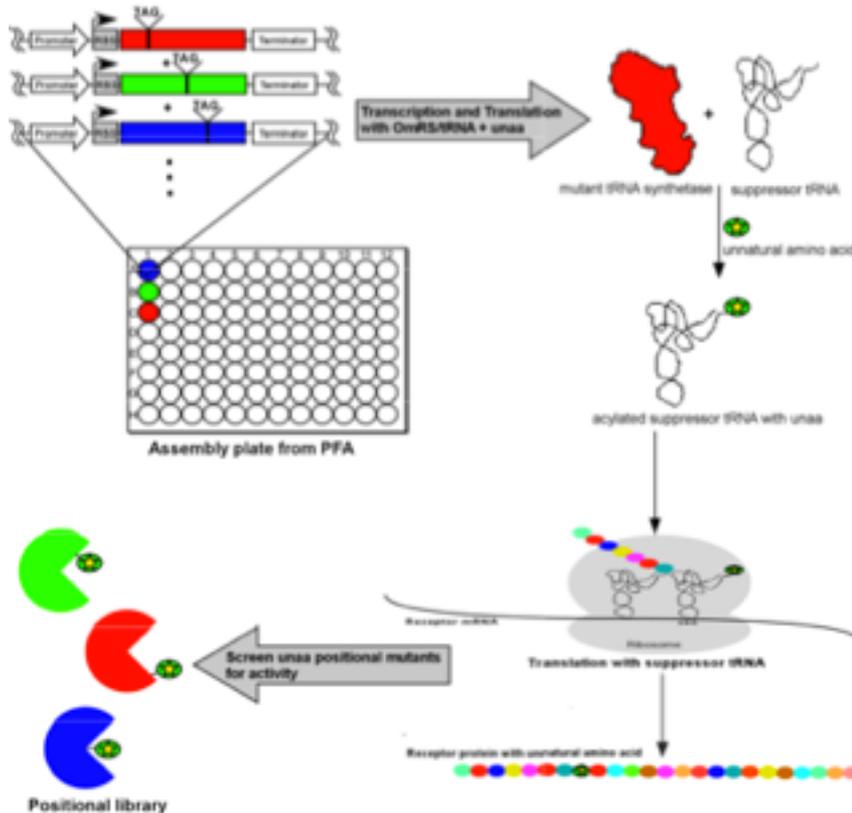


ec-phnD

Device Development

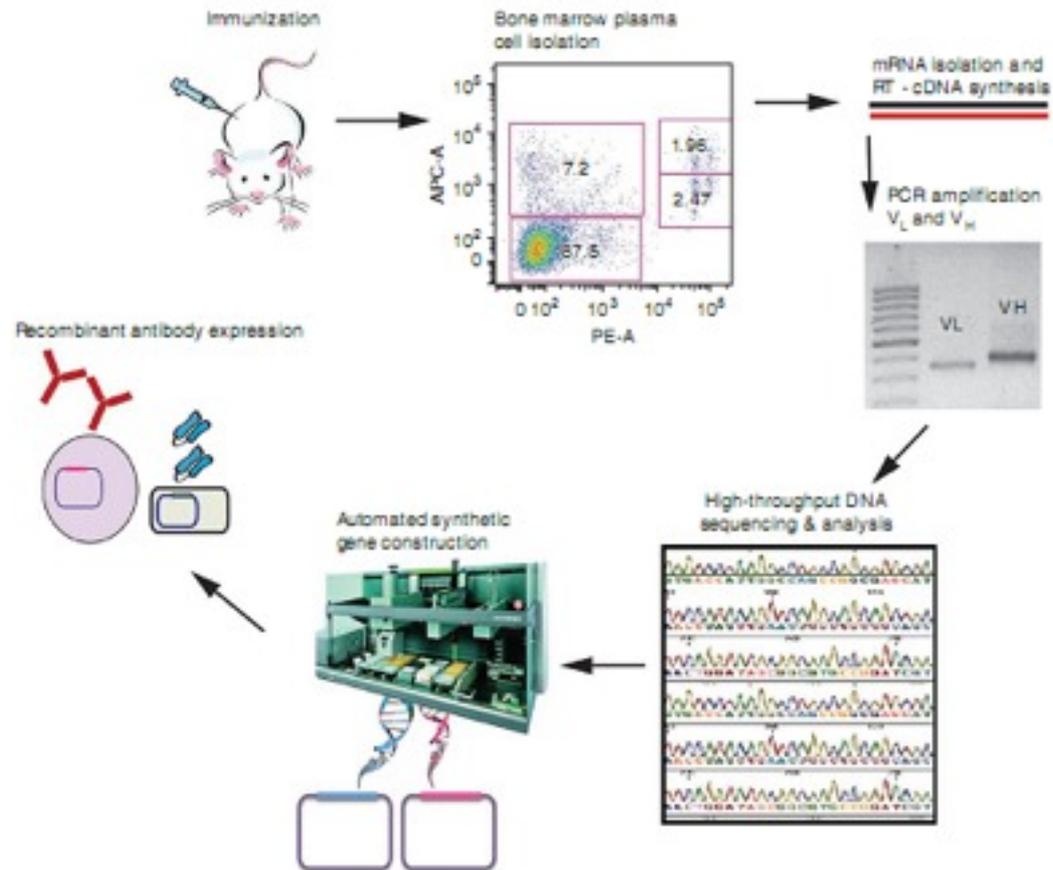


Application of PFA: unnatural amino acid scanning.



- Genes encoding TAG stop codons at various positions in the coding sequence are assembled via PFA.
- The genes are transcribed and translated in the presence of an engineered tRNA synthetase and suppressor tRNA orthogonal pair with the unnatural amino acids.
- The unnatural amino acid “labeled” proteins are then screened for desired activity.
- Can be used to stream-line biosensor production and discovery, by eliminating post-translational labeling of the protein with extraneous fluorophores.

Application of Bio-prospecting and synthetic gene construction-Antibody repertoire discovery



- ~200 recombinant antibody genes were built
- Synthesis to screen took ~3 weeks

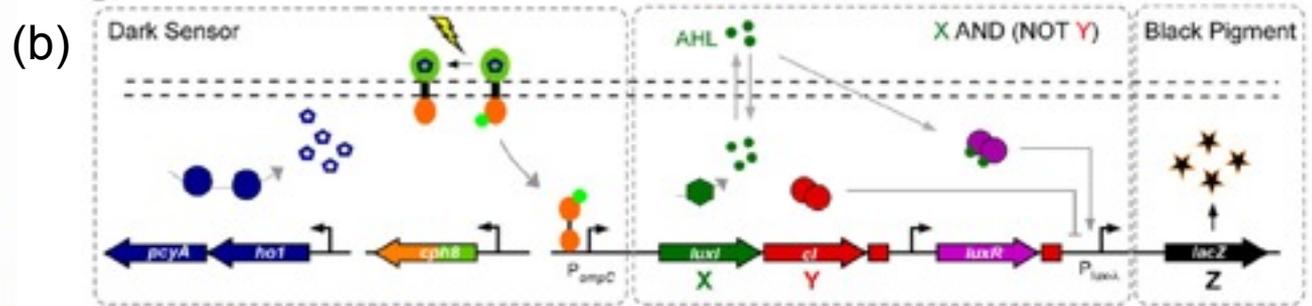
Reddy, S *et. al.* **2010** *Nature Biotech* 28:965-969

Nucleic Acid Operating Systems

What kind of operating system makes sense for biological systems? Amorphous Computing

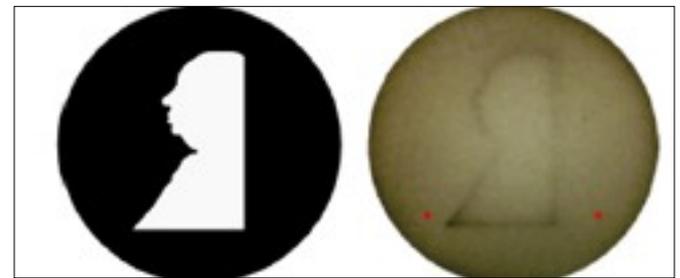
- All units (cells) hold and execute the same program in ROM (DNA)
- Units have limited memory
 - phenotypic state: transcription, post translational modifications, etc.
- Units have limited bandwidth
 - transcription and translation are not particularly fast
- No knowledge of position; local communication only
- Some units are faulty
- Abelson, Knight, Sussman @ MIT coined the term “Amorphous Computing” to describe this pattern.
- Hardware agnostic. AC is the study of algorithms, not hardware.

We have previously built amorphous computers ...
 but the address space is ridiculously small.



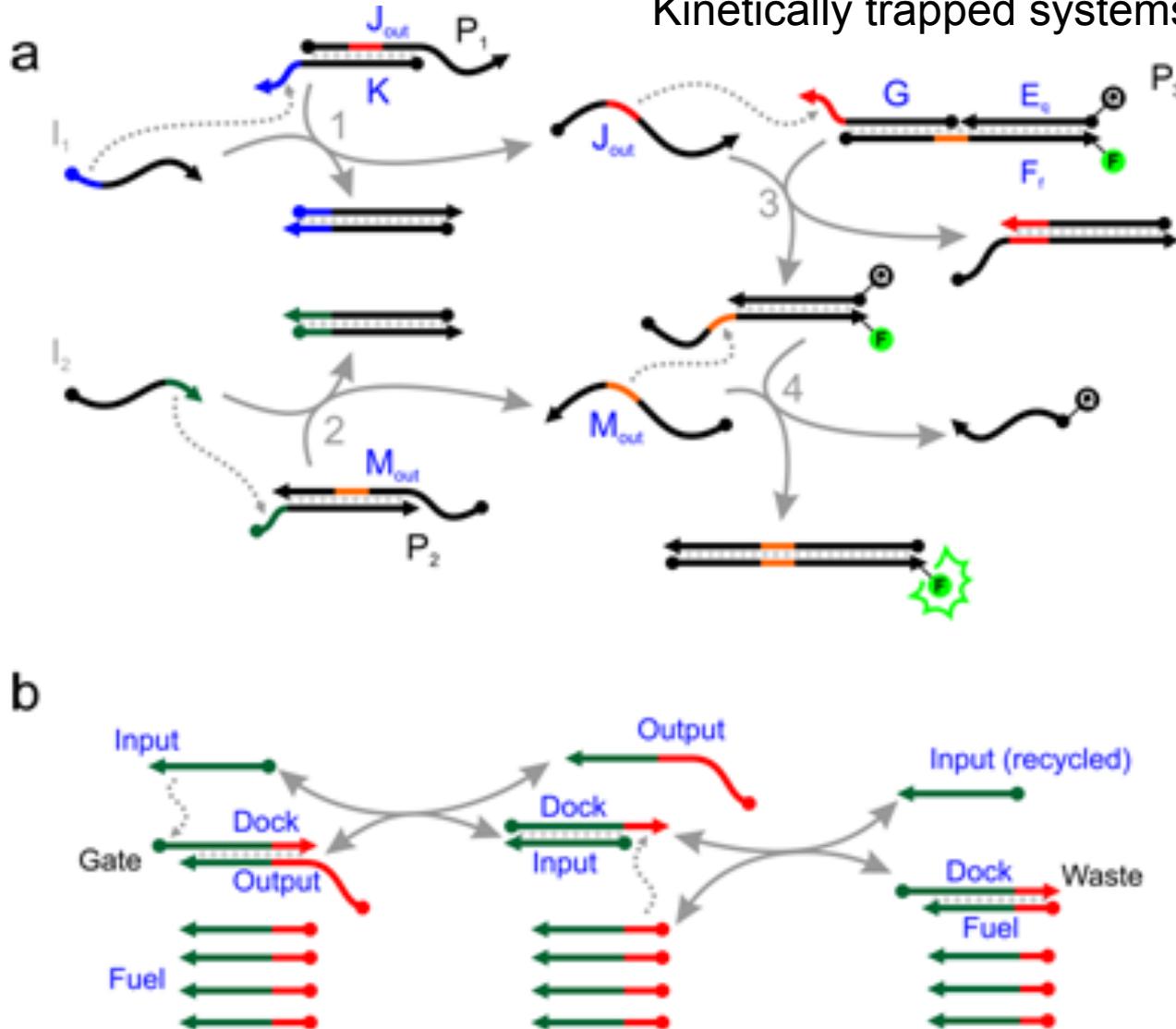
(c)

Light-directed genetic circuitry. (a) Output of bacterial photography circuit, showing resolution possible. (b) Edge detector circuit. (c) Computed output of edge detector circuit with complex mask input



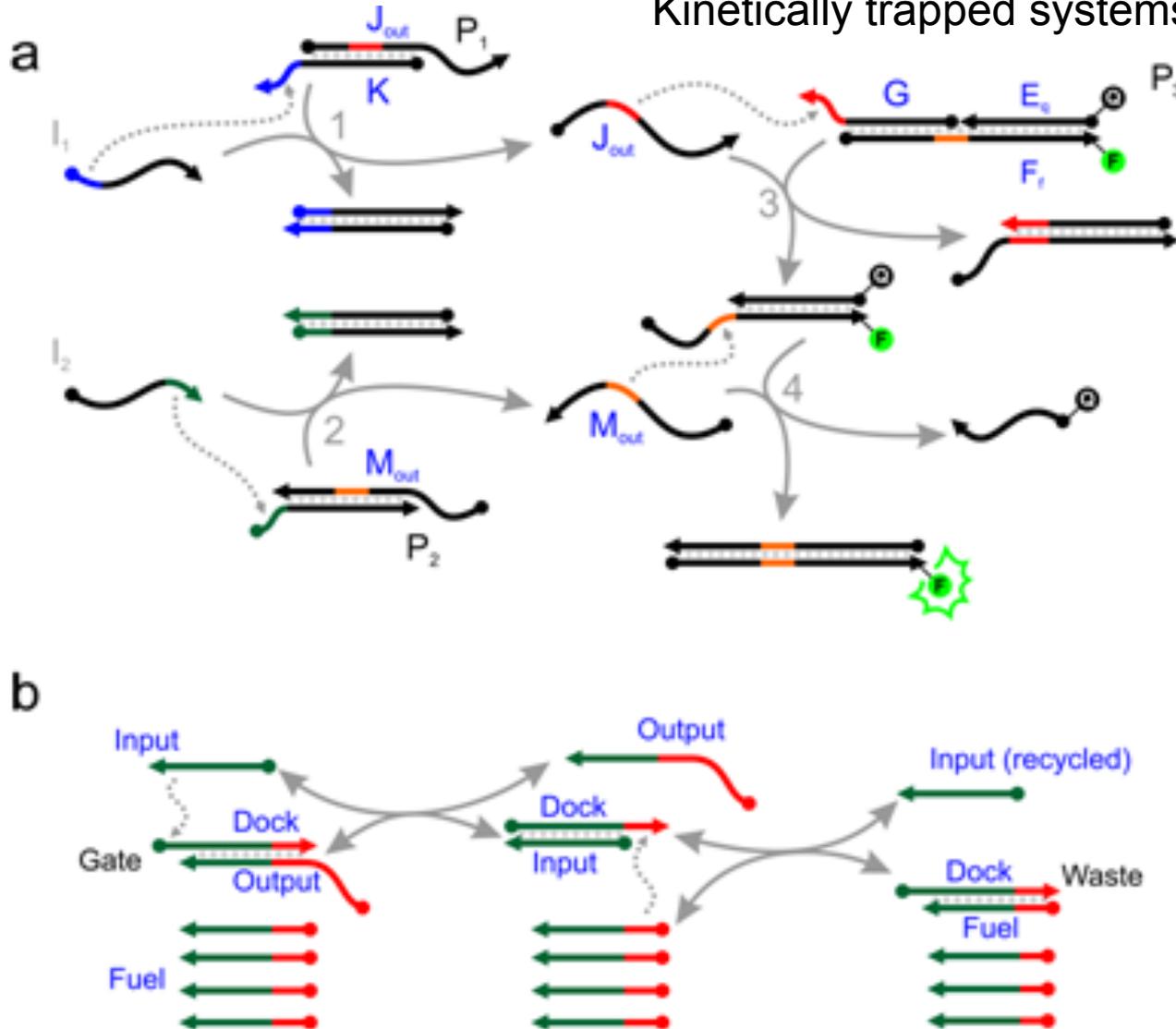
The large address space available to nucleic acids can be exploited to execute interesting algorithms

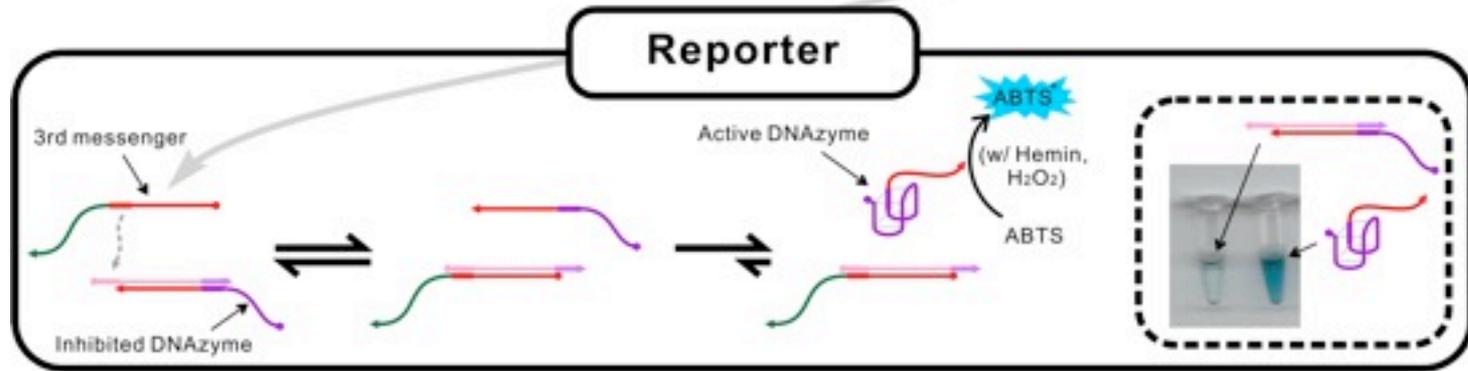
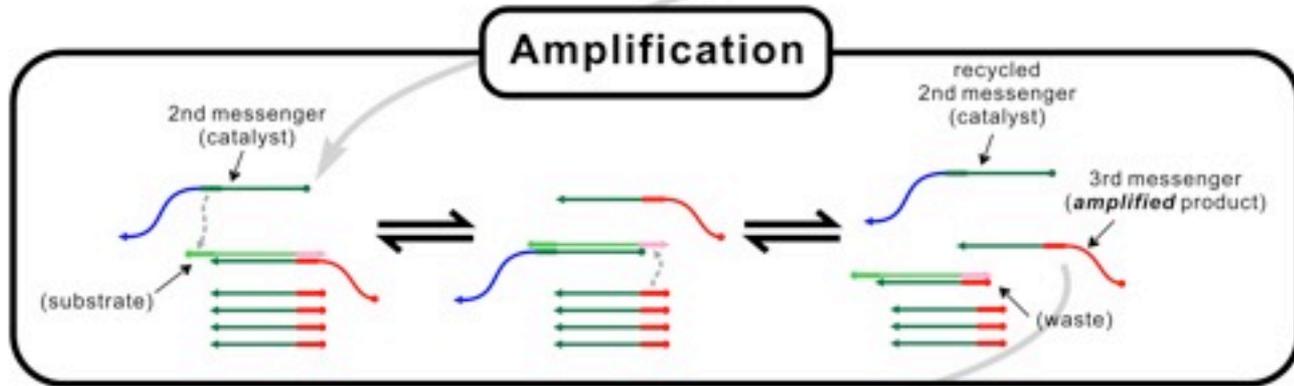
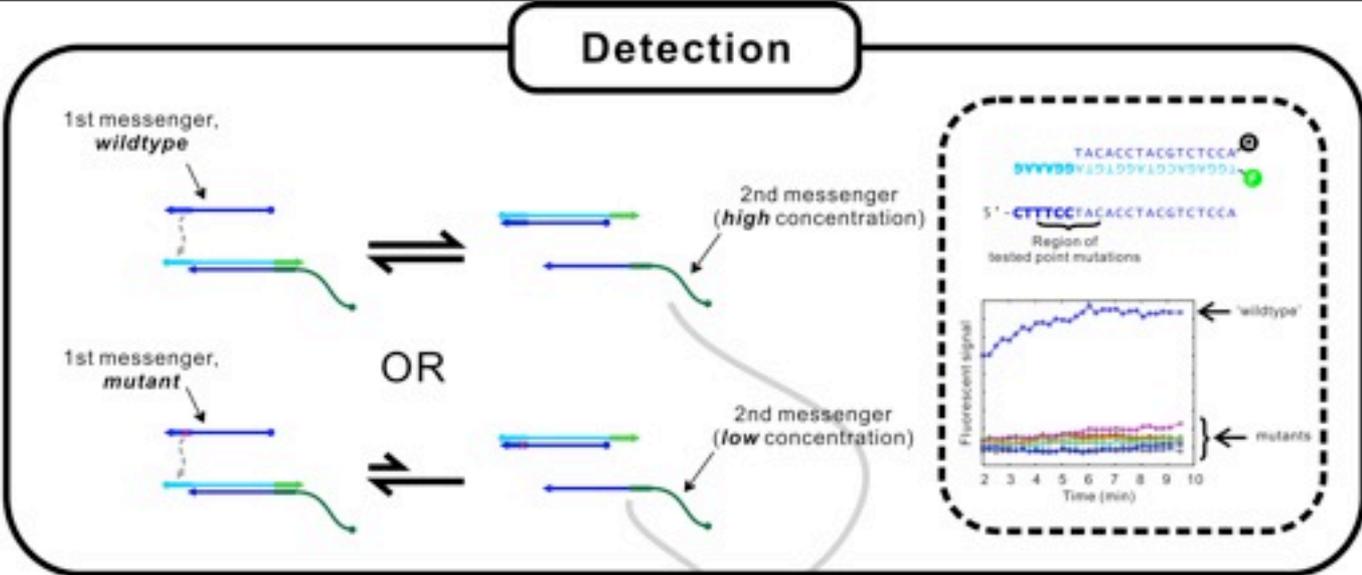
Kinetically trapped systems



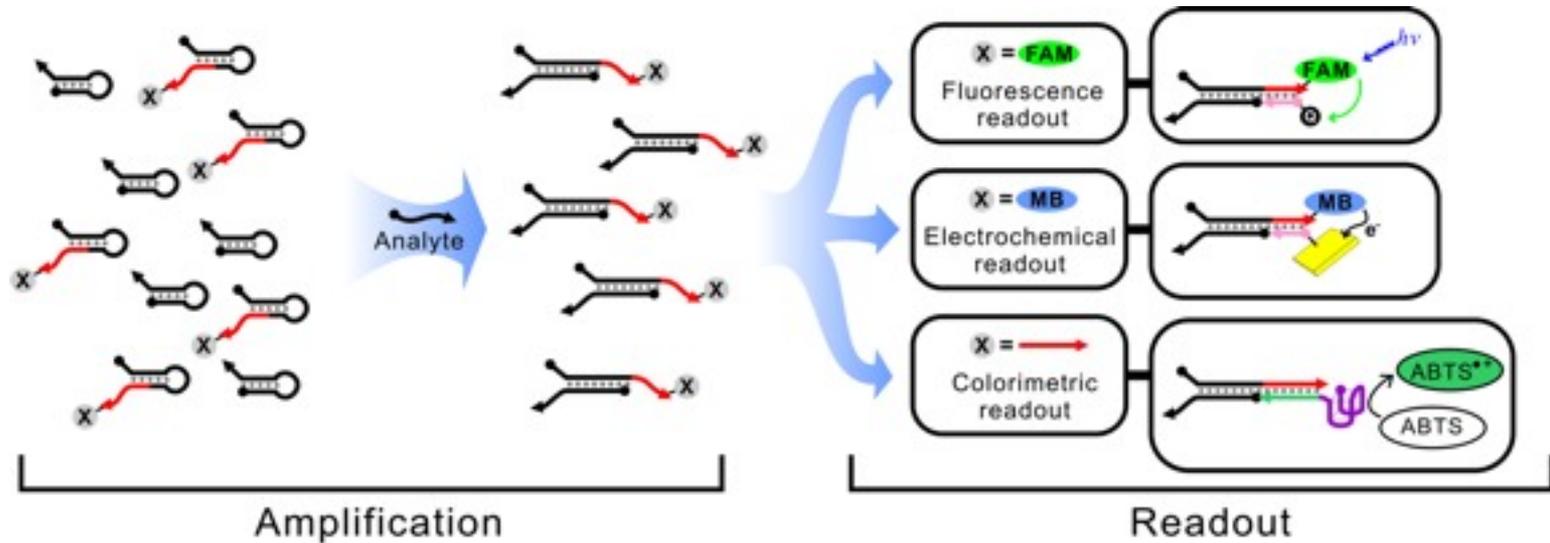
The large address space available to nucleic acids can be exploited to execute interesting algorithms

Kinetically trapped systems

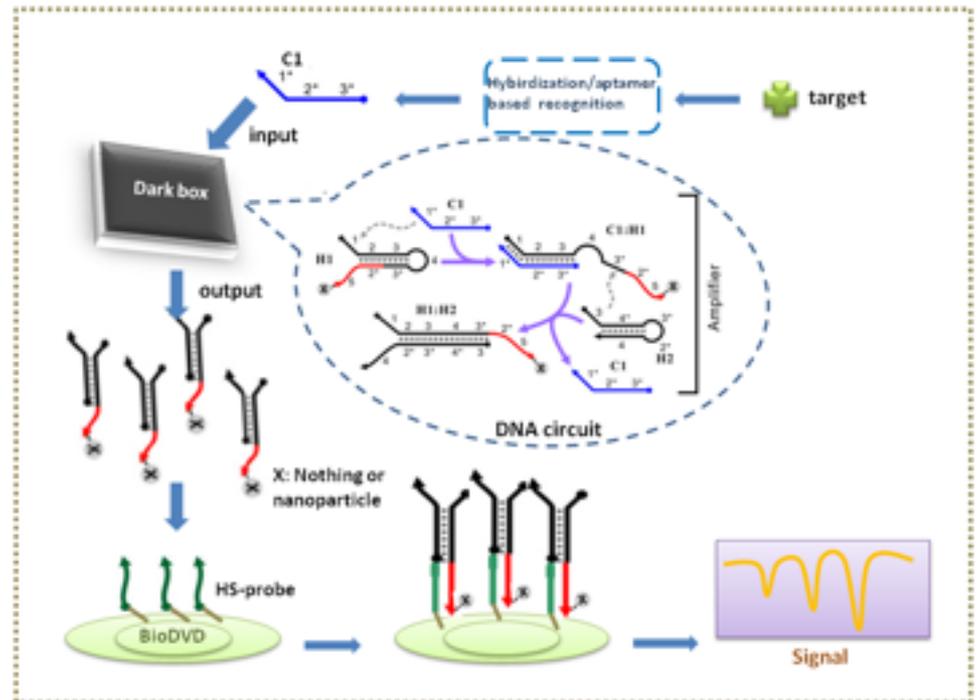




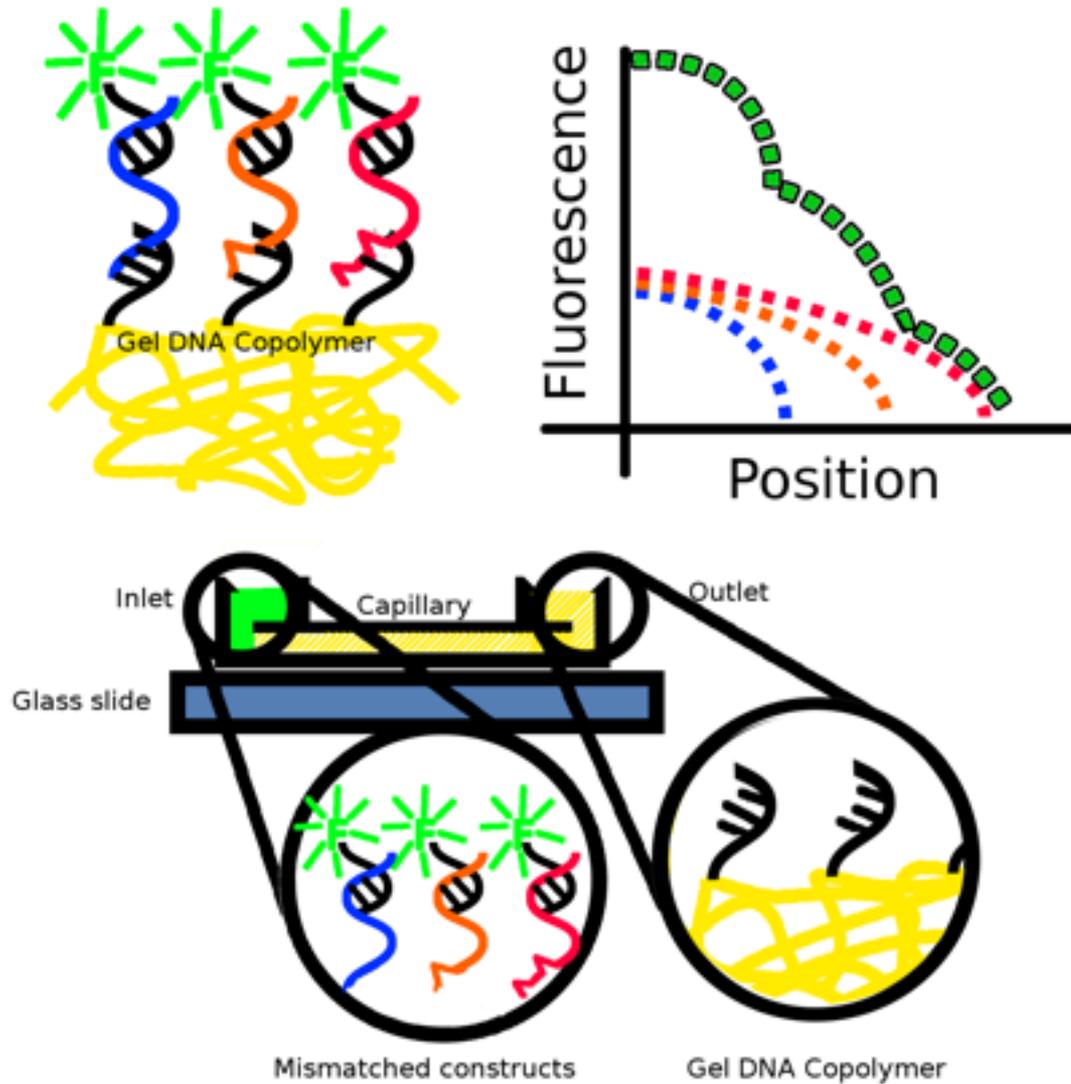
We are also adapting these circuits to standard analytical formats:



Work by Xi Chen
and Bingling Li

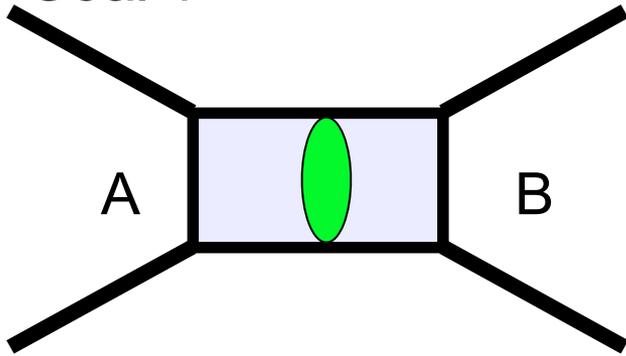


However, these circuits are *not* amorphous
Instead: diffusion control by hybridization

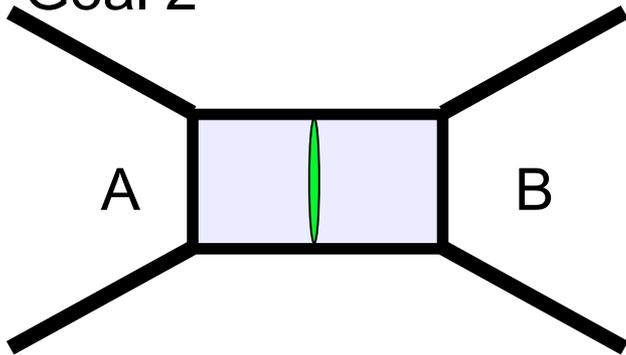


Diffusion Patterning

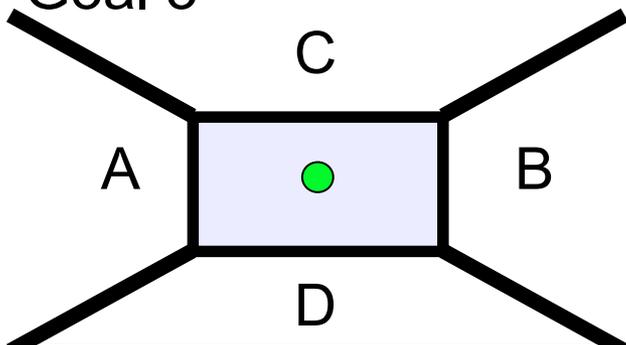
Goal 1



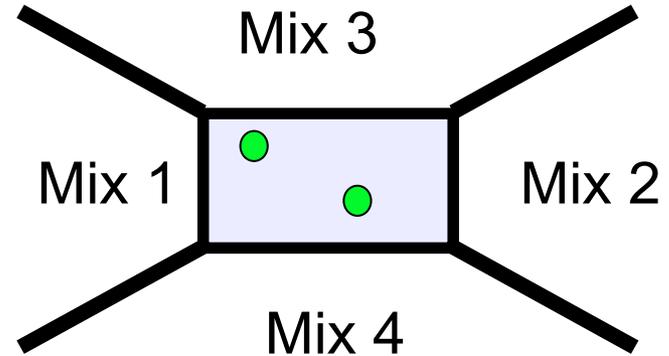
Goal 2



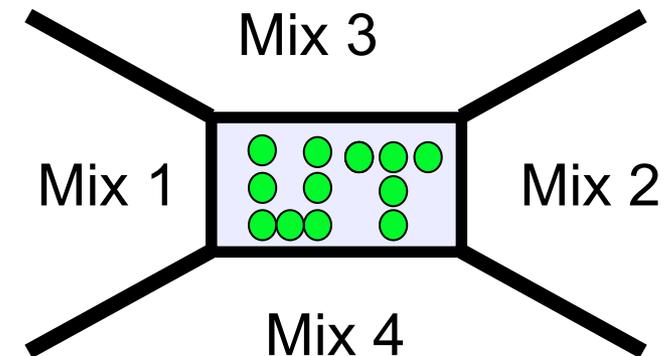
Goal 3



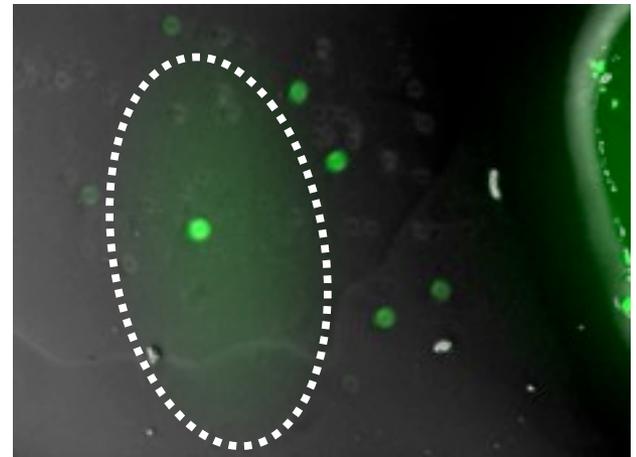
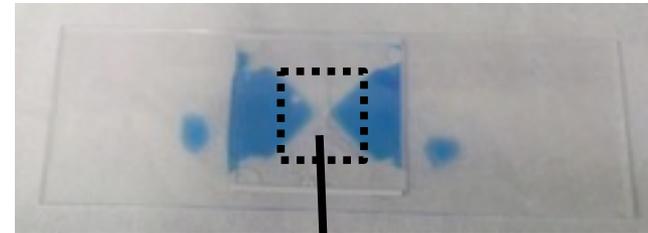
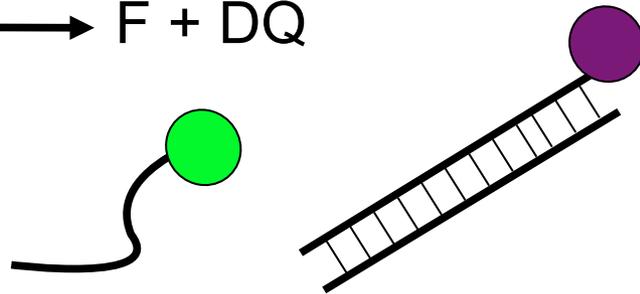
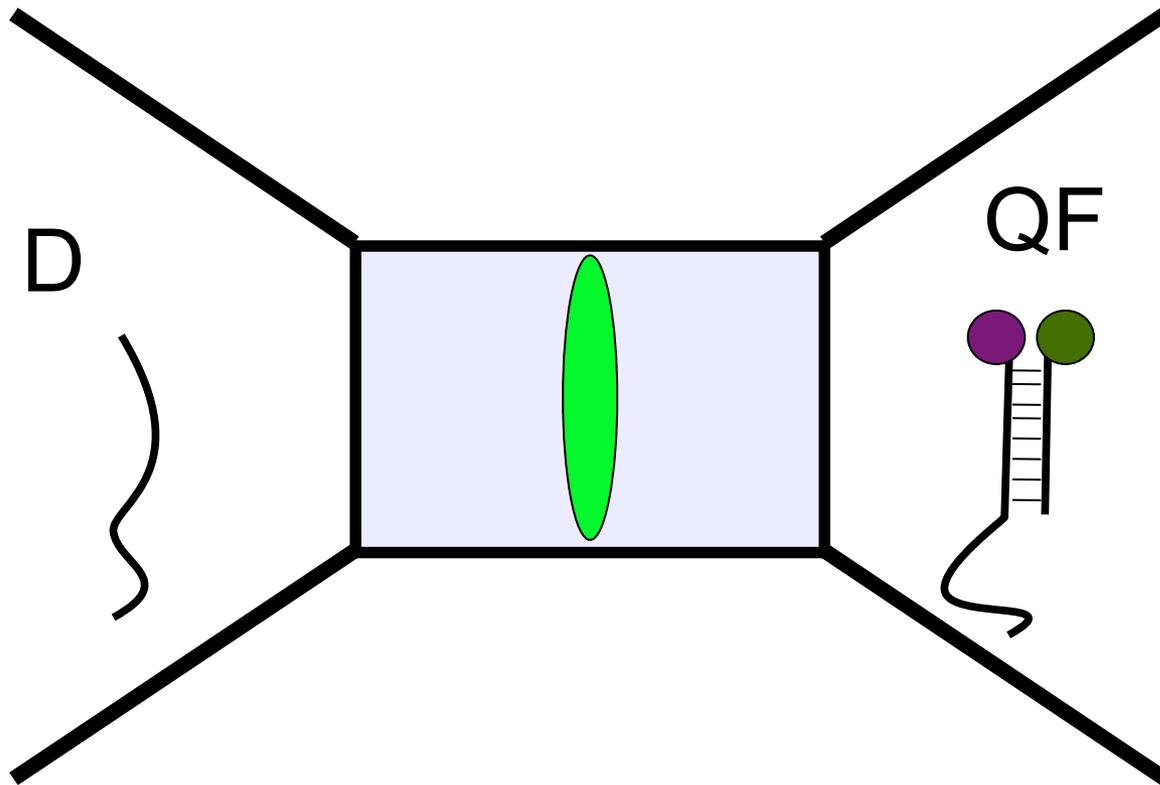
Goal 4



Goal 5



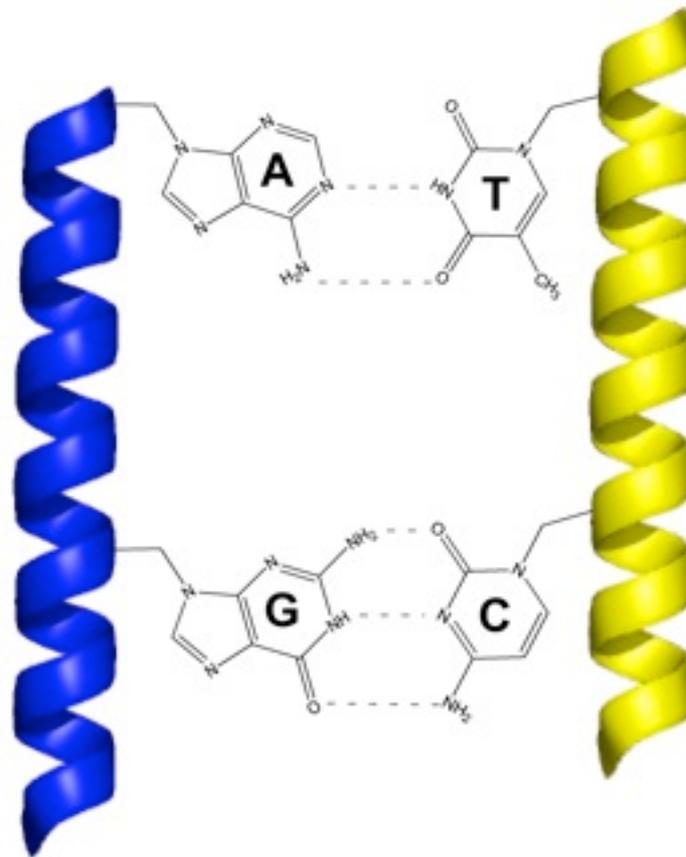
Goal 1: monitor diffusion reactions



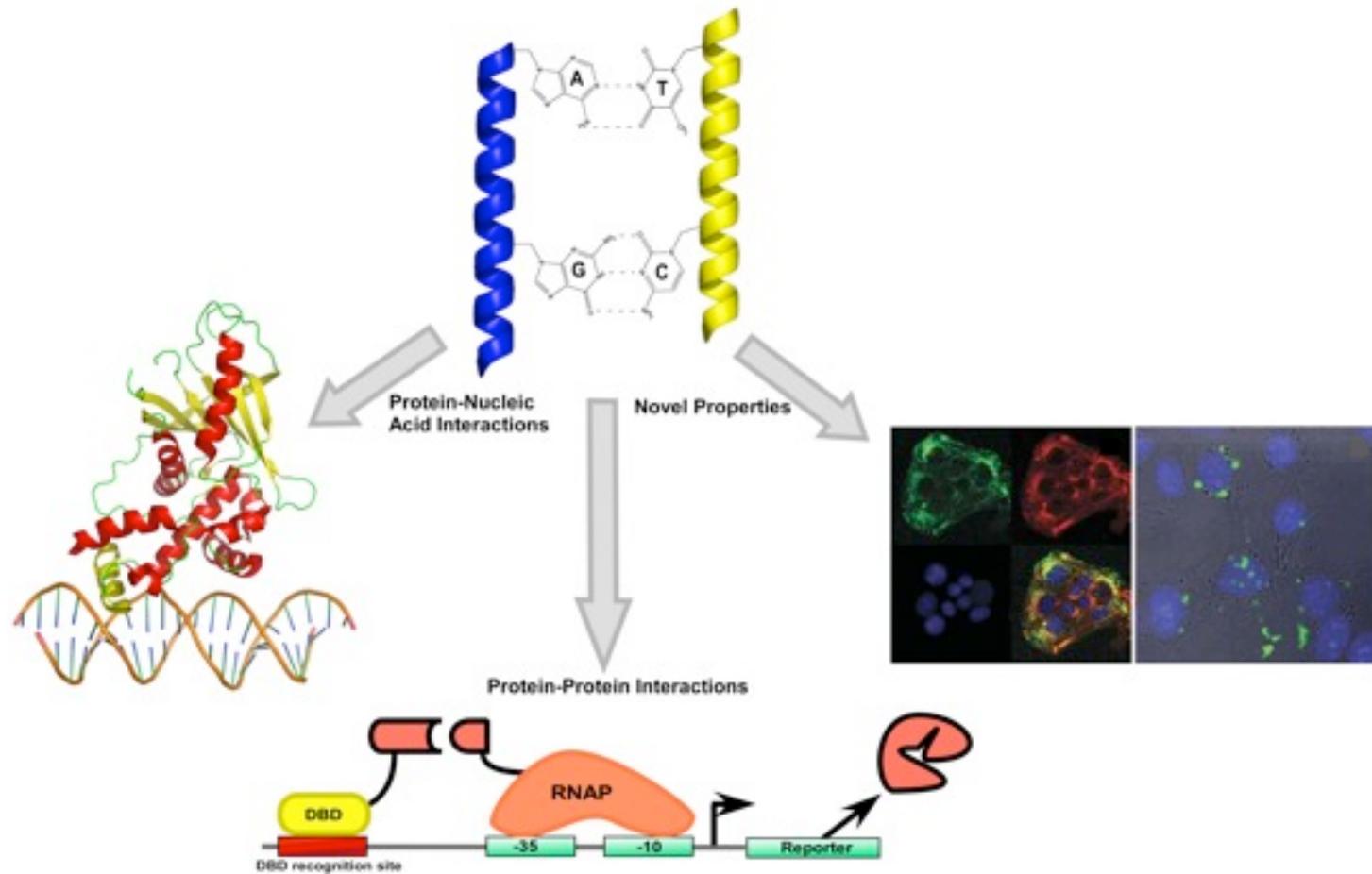
- From a **synthetic biology** perspective, the only parts that are truly modular, composable, and scalable are nucleic acids (but **not** in their capacity as protein-encoding molecules)
- From an **operating system** perspective, the only biologically relevant system that makes any sense for implementation is nucleic acids (again, not as protein-encoding molecules)
- Unfortunately, nucleic acids **are** typically only information-carrying molecules in biology, and we actually want to do things with cells (the current unit of biological replication)

Nucleic Acid Mimics

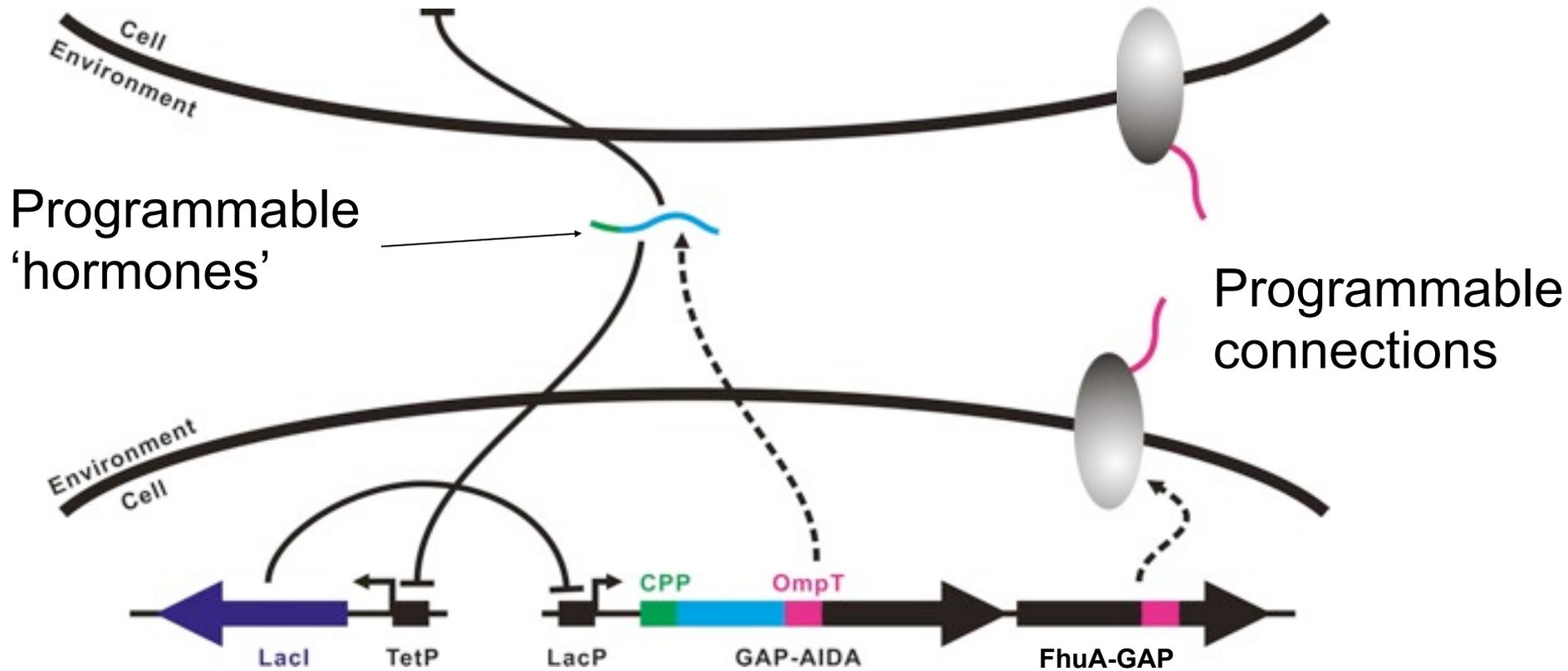
Protein-Nucleic Acid Mimics as Programmable Interaction Tools



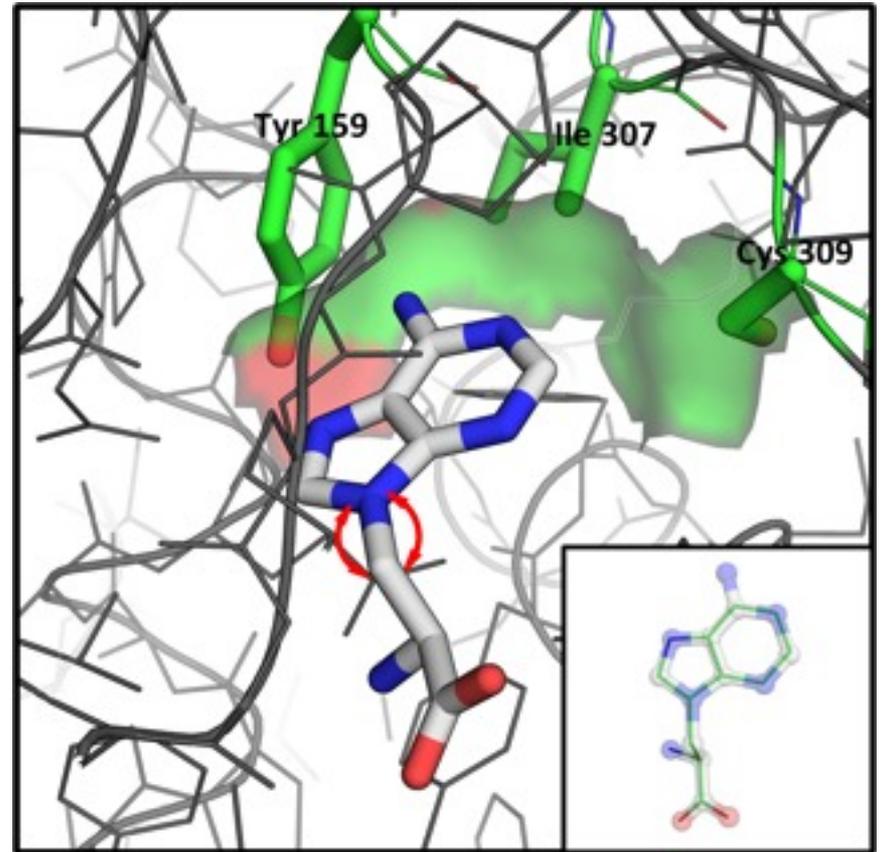
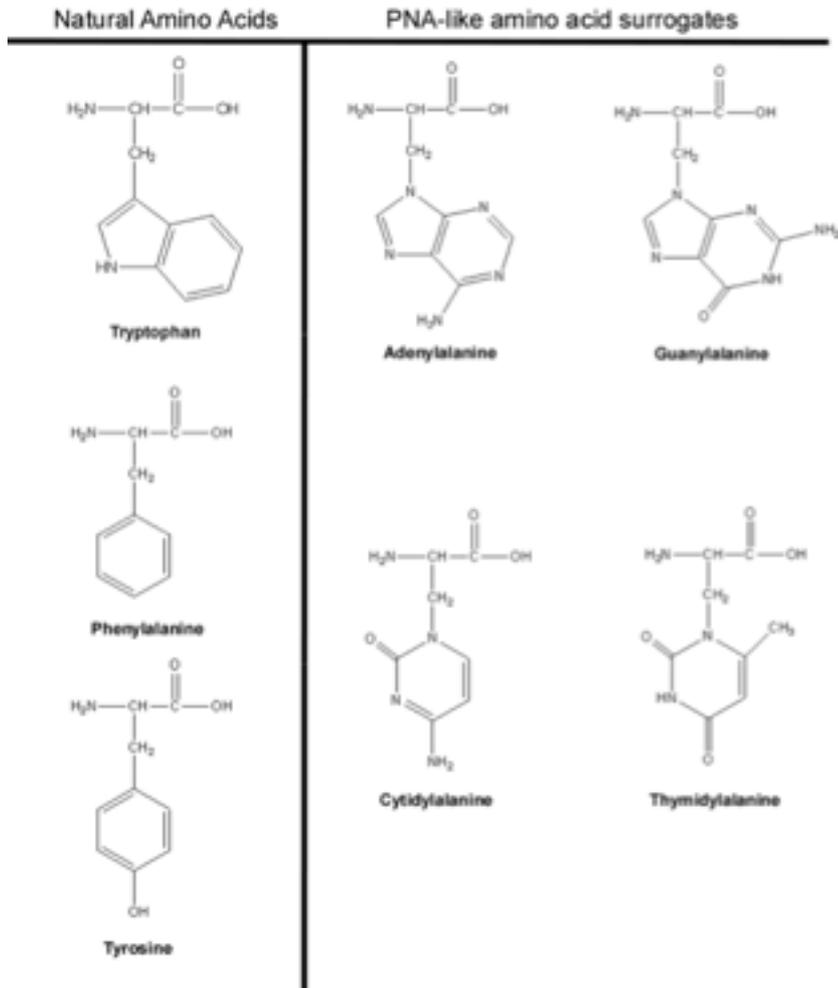
Nucleobase directed interactions



Nucleobase amino acids → Proteins with nucleic acid-like properties → Programming



Making proteins work like nucleic acids

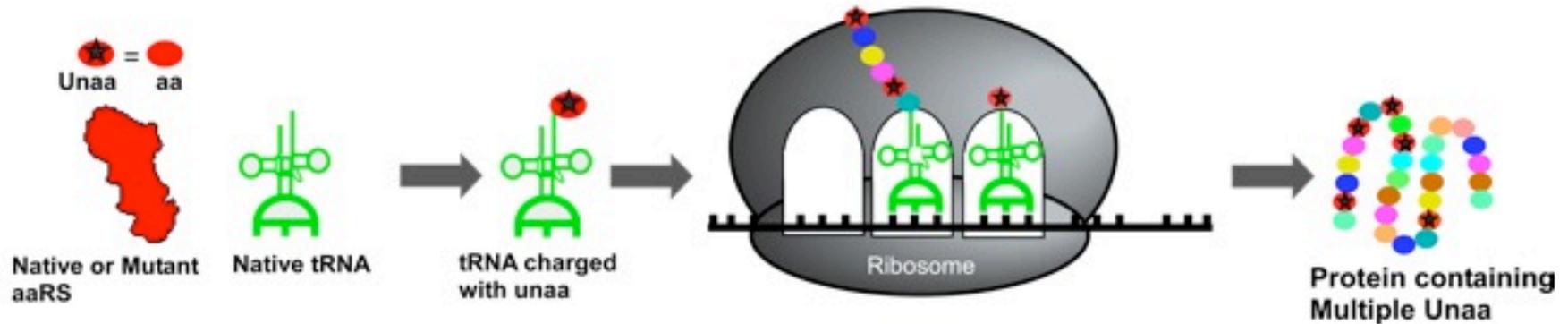


Adenyl alanine docked into the active site of tryptophanyl tRNA synthetase.

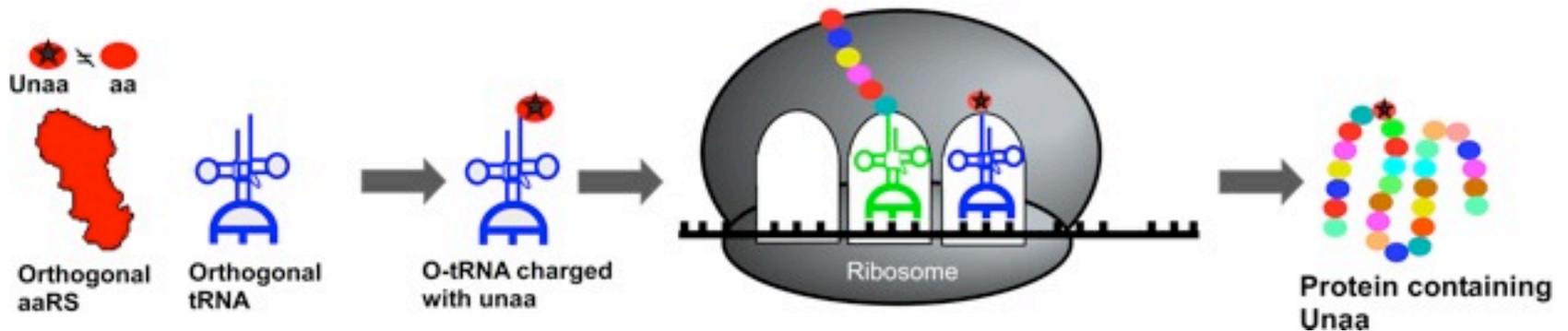
Whatever works: design, selection, and synthesis

Unnatural Amino acid Incorporation

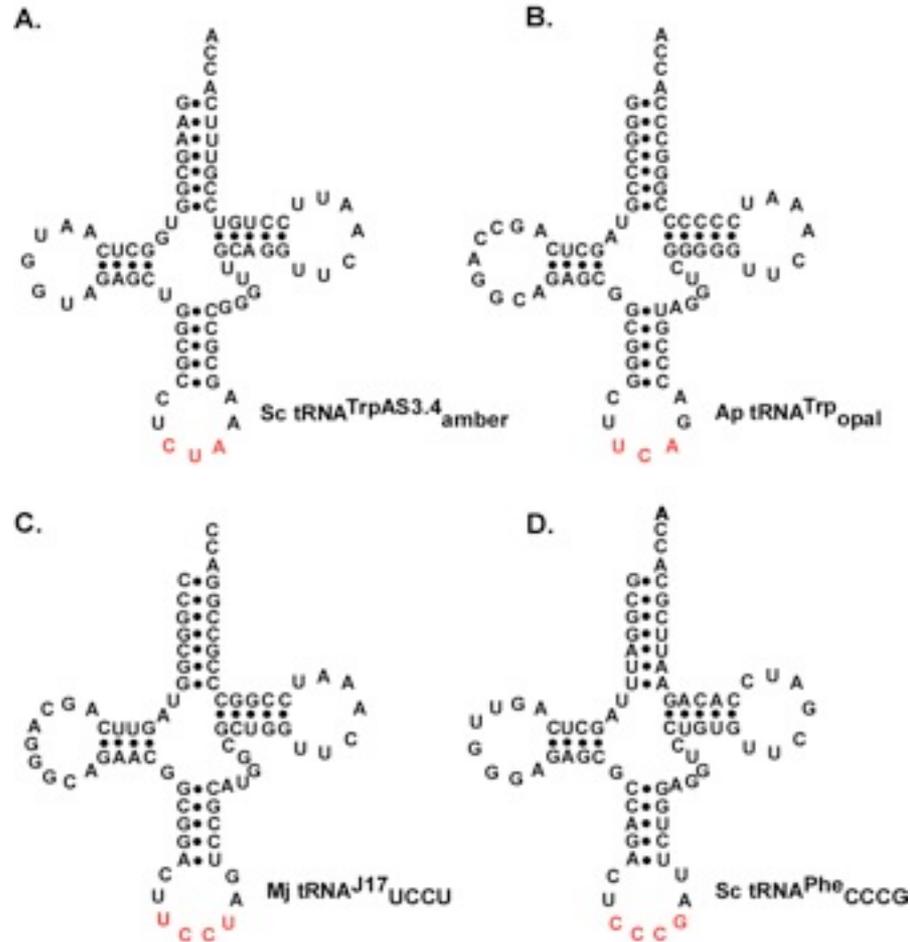
Global Replacement of a Natural Amino Acid with an Unnatural Amino Acid (UnAA)



Site-Specific Incorporation of a Natural Amino Acid with an Unnatural Amino Acid (UnAA)

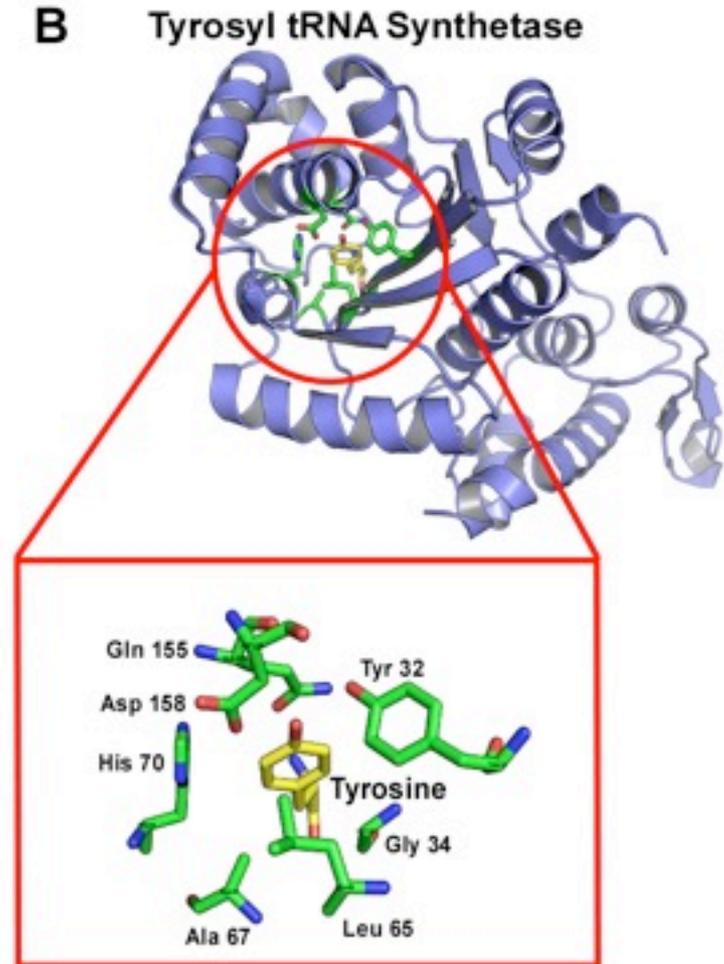
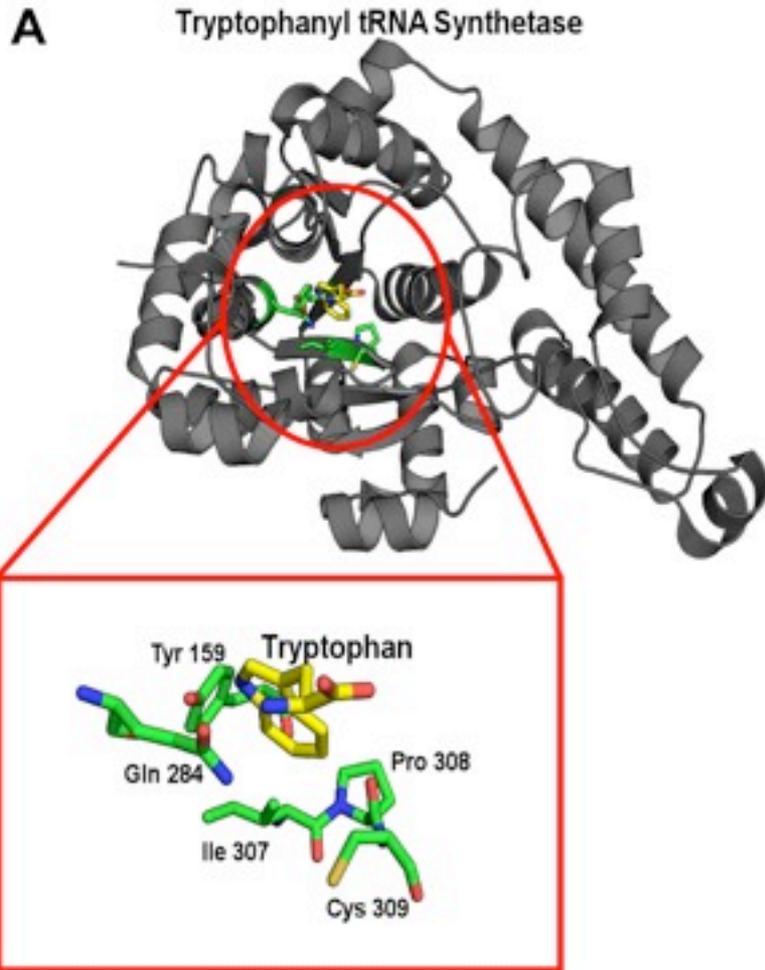


Proposed Suppressor tRNAs for NBA incorporation



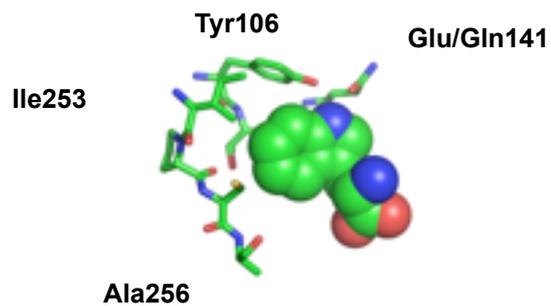
A: Hughes, RA and Ellington AD 2010 NAR 38:6813-30

Directed Evolution of aaRSs-Library Design

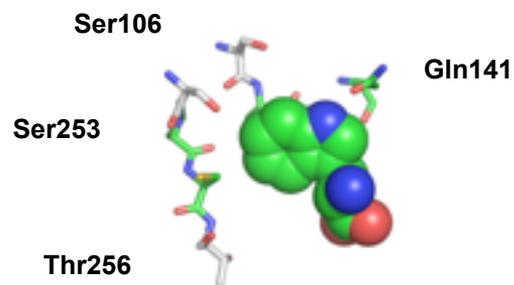


Ade-Ala RS Rational Designs

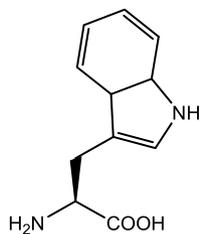
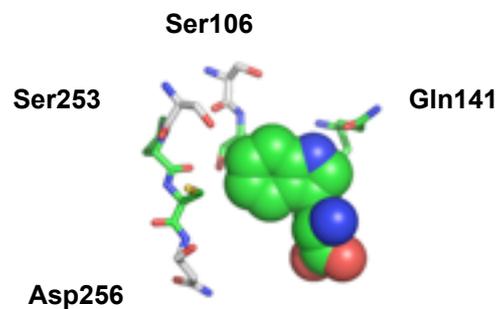
Hs-WRS



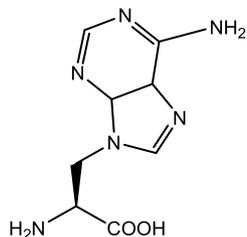
Raa2



Raa1



Tryptophan



Adenyl-alanine

Where are we going?

- Learn to build biological software based on amorphous algorithms and nucleic acids
- Create organisms with 4 additional nucleobase amino acids
- Program the organisms to do what we want
- Attempt to study modified proteins' ability to directly communicate with nucleic acids and other proteins via a programmable interface.
- Join Sarah Connor in Mexico for the inevitable Apocalypse between Hyper-intelligent Robots and **NeoLife 2.0**.

