Module Overview

Day	Lecture	Lab
1	Introduction	DNA library synthesis (PCR)
2	SELEX I: Building a Library	DNA library purification (agarose gel electrophoresis)
3	SELEX II: Selecting RNA with target functionality	RNA library synthesis (*In vitro* transcription = IVT)
4	SELEX III: Technical advances & problem-solving	RNA purification and heme affinity selection
5	Characterizing aptamers	RNA to DNA by RT-PCR
6	Introduction to porphyrins: chemistry & biology	Post-selection IVT Journal Club 1
7	Aptamer applications in biology & technology	Aptamer binding assay
8	Aptamers as therapeutics	Journal Club 2

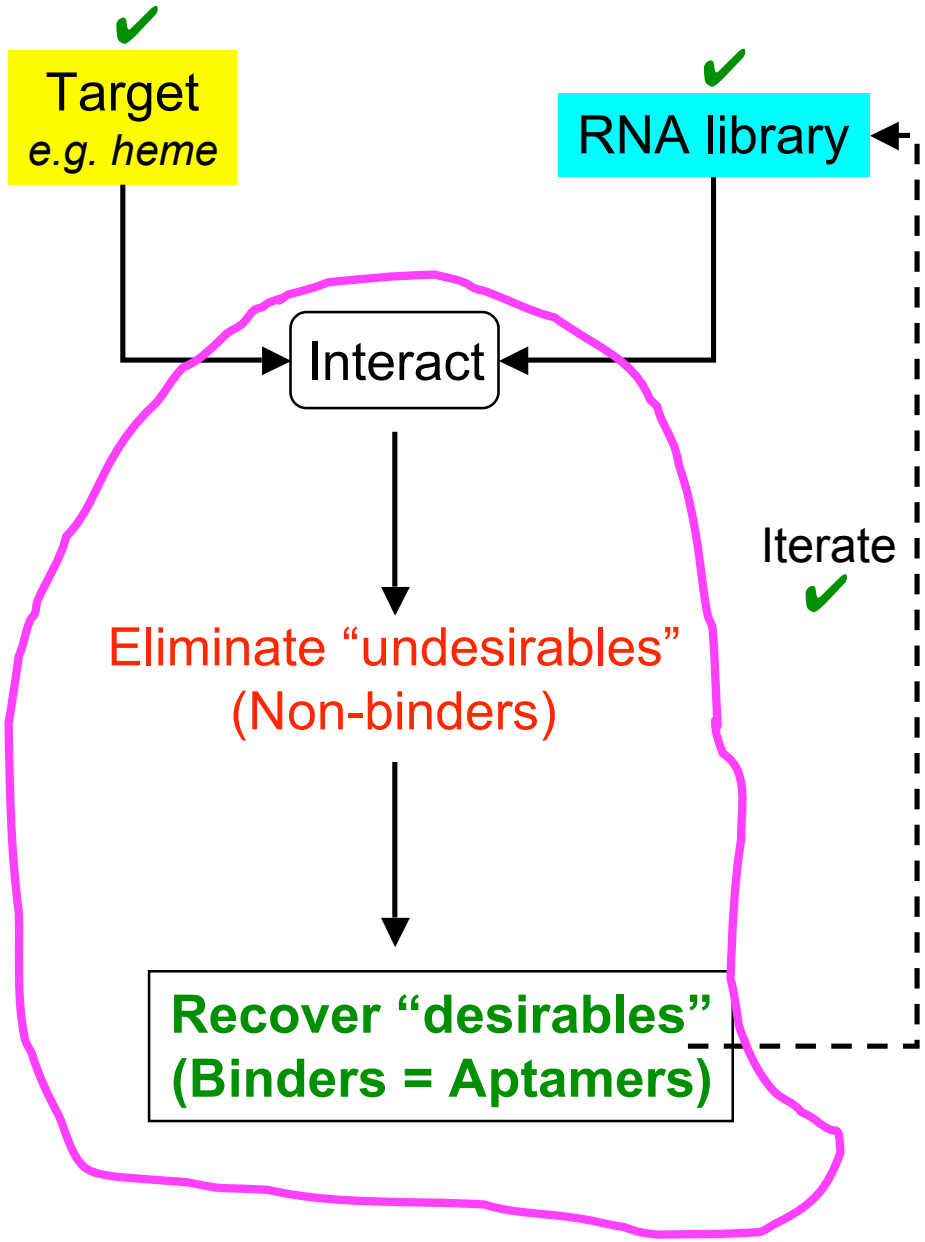
SELEX II

Selecting RNA with target functionality

20.109 Lecture 3

16 February, 2012

SELEX: The process (simply)



Now, that we understand:

- Target selection
- Library construction & manipulation

– How do we enrich for binders?

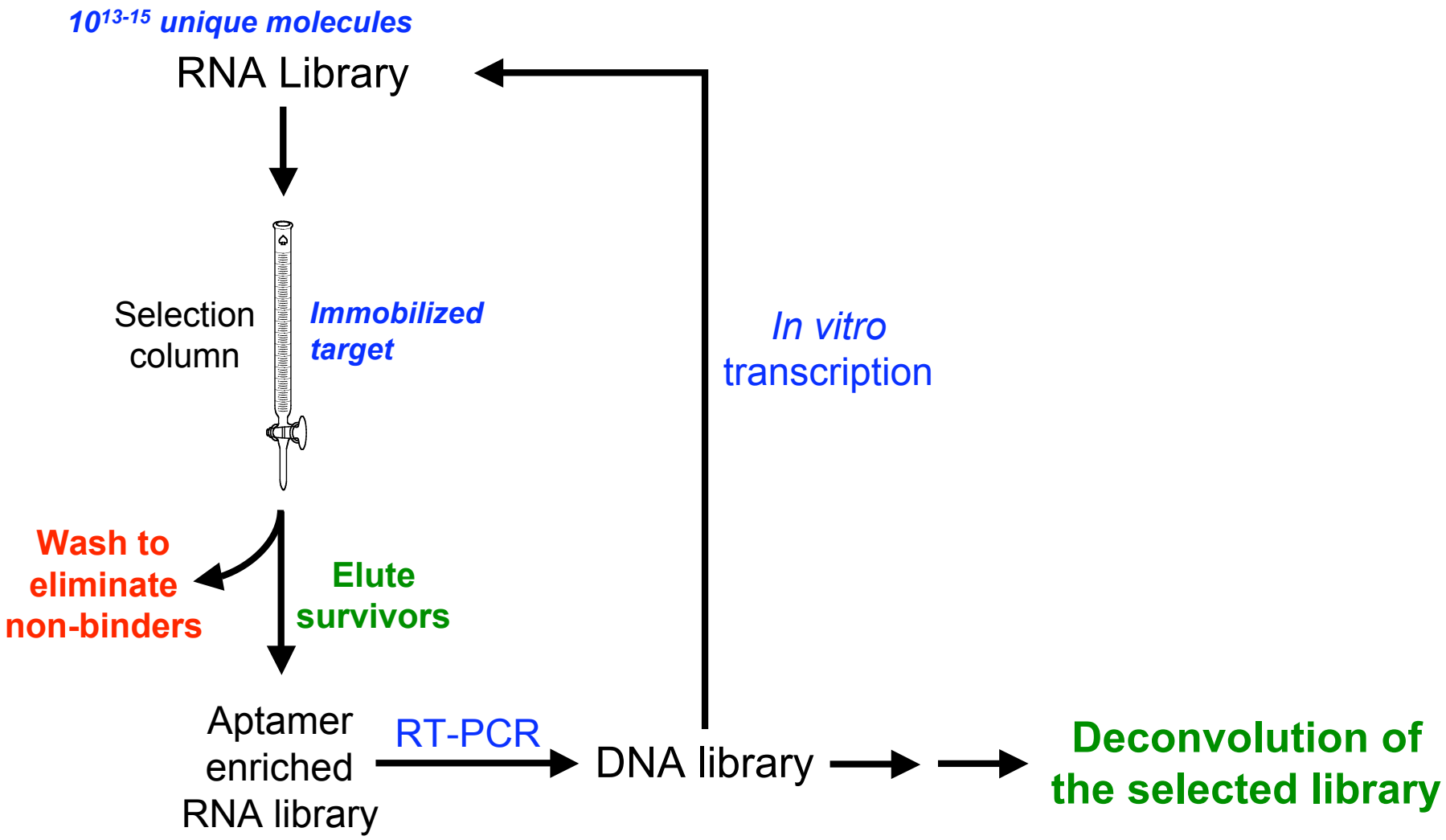
– *How do we put this all together into a workflow?*

Enriching your library for binders

(how would you do this?)

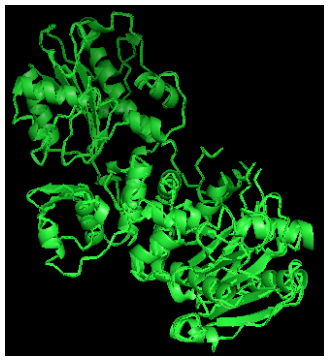
- Need a partitioning strategy:
  - Separate target bound RNA from unbound fraction
  - Selectively release target bound RNA
- Most commonly involves immobilizing target on:
  - A membrane (e.g. nitrocellulose)
  - Solid support (usually some kind of bead)
    - Column format
    - Magnetic separation
- Other approaches, for e.g.:
  - Electrophoretic methods to separate {RNA:target complex} from free RNA
  - Microfluidics devices

Putting it all together: A typical SELEX workflow



SELEX *à la* Tuerk & Gold

T4 DNA polymerase

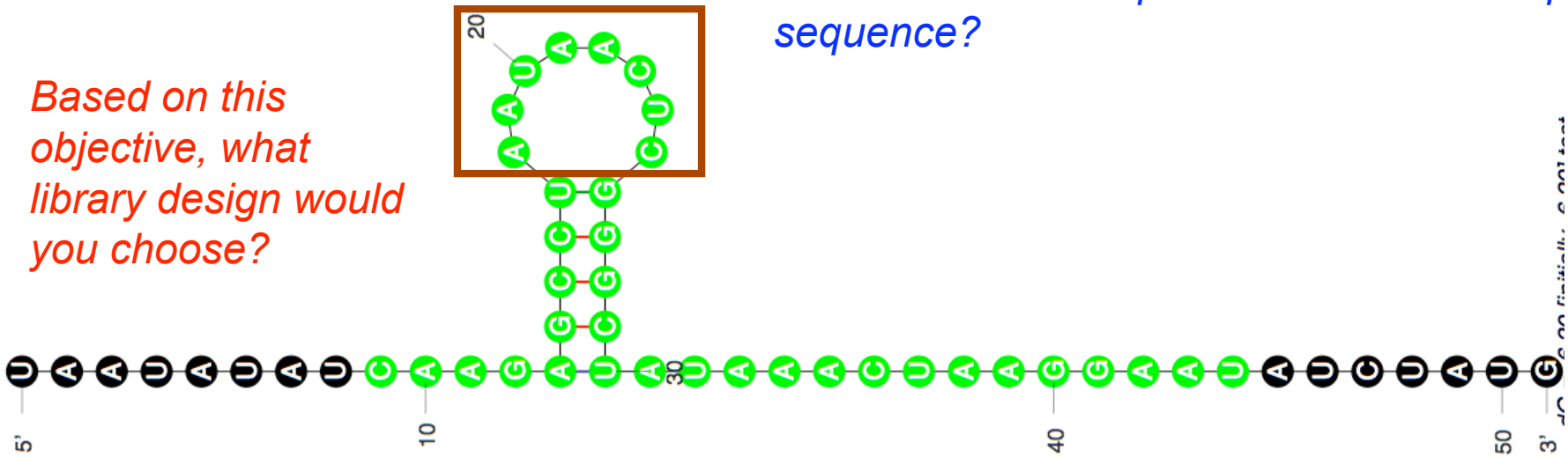


Structure for residues 1-388 from the PDB (www.rcsb.org)

Target

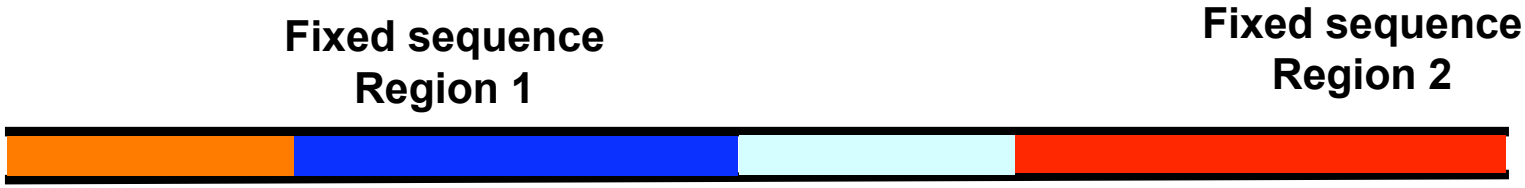
- Target known to interact with RNA from prior work
  - Sequence below found in the mRNA encoding the T4 DNA polymerase
  - Regulatory mechanism:
    - T4 DNA polymerase binds its own mRNA decreases its own synthesis
- 8 nucleotides [AAUAACUC] are critical for the interaction
  - *What underlies the preference for this loop sequence?*

Based on this objective, what library design would you choose?



C. Tuerk and L. Gold; ***Science***; 249 (4968), 505-510, 1990

SELEX *à la* Tuerk & Gold



RNA library

T7 promoter

Variable Region:
= 8 nucleotides

- **Total space = 6 x 10^{14} molecules**
- 8 nucleotide variable region:
  - Maximum Diversity = $(4)^8$ ~ 65, 556 unique sequences
  - Each sequence present @ (6 x 10^{14}/~6.6 x 10^4) ~ 1 x 10^{10} copies/library
  - The known RNA target present @ 2 in 10^5 molecules!

C. Tuerk and L. Gold; ***Science***; 249 (4968), 505-510, 1990

How do you co-optimize across these parameters

Scenario II

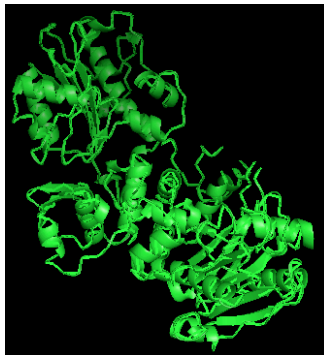
- Set space limit (i.e. reasonable cost, practicality)
  - 6×10^{14} molecules
- Maximize diversity within this limit*
  - Enough information available to enforce a boundary condition on diversity
  - Cognate recognition sequence = 8 nucleotides long
- Preserve representation at some acceptable (read: arbitrary) limit?
  - 2 potential solutions per 10^5 molecules

SELEX *à la* Tuerk & Gold

Immobilize on nitrocellulose

- Works well for many protein targets

T4 DNA polymerase



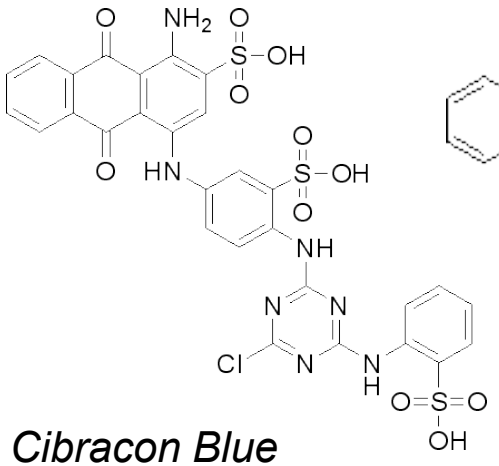
Structure for residues 1-388 from the PDB (www.rcsb.org)

- *Advantages*
  - Very easy and inexpensive!
  - Well-developed and straightforward protocols available
- *Disadvantages*
  - Protein can denature during immobilization step
    - Selected aptamers cannot recognize native protein
  - Not all proteins stick strongly enough to survive washing steps to remove unbound library

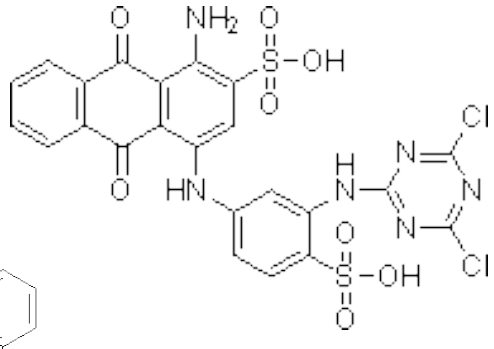
C. Tuerk and L. Gold; ***Science***; 249 (4968), 505-510, 1990

SELEX *à la* Ellington & Szostak

Target



Cibracon Blue



Reactive Blue 4

www.sigmaaldrich.com

- Discover RNA binding to small molecule organic dyes
  - No prior knowledge of their RNA binding capacity
- *Can RNA specifically interacting with these molecules be discovered?*

Based on this objective, what library design would you choose?

A.D. Ellington and J.W. Szostak; ***Nature***; 346 (6287), 818-822, 1990

SELEX *à la* Tuerk & Gold

RNA library

T7 promoter

Variable Region:
= 100 nucleotides!

- **Total space ~ 6 x 10^{14} molecules**
- 100 nucleotide variable region:
  - Maximum Diversity = $(4)^{100}$ ~ **2 x 10^{60}** unique sequences possible!
  - Each sequence present @ (6 x 10^{14}/~2 x 10^{60}): Absent or 1copy/library
  - The known RNA target present @ *??? frequency*

A.D. Ellington and J.W. Szostak; ***Nature***; 346 (6287), 818-822, 1990

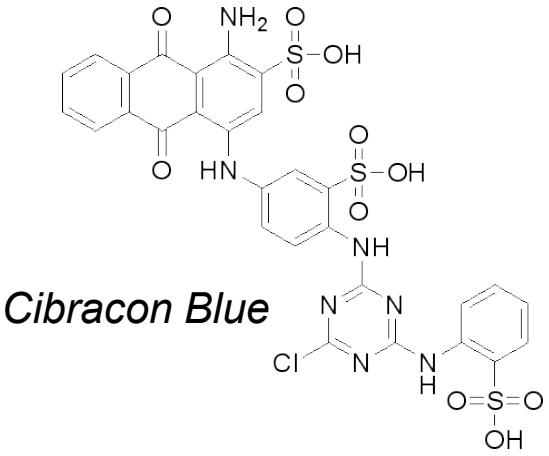
How do you co-optimize across these parameters

Scenario III

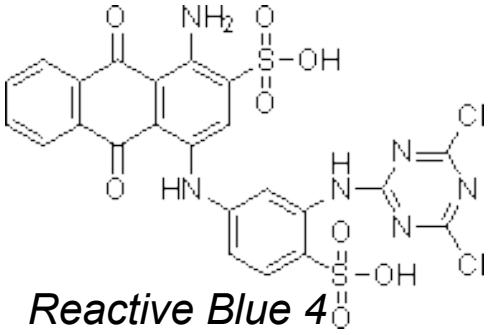
- Set space limit (i.e. reasonable cost, practicality)
  - 6×10^{14} molecules
- Maximize diversity
  - No information available to enforce a boundary condition on diversity required
  - Cognate recognition sequence = unknown!
- Sacrifice representation
  - A given 100 nt sequence available only once in library!
  - How would you avoid sampling without replacement?
  - How might you still have good functional representation with such a library?

SELEX *à la* Ellington & Szostak

Immobilize on agarose beads



Cibracon Blue



Reactive Blue 4

www.sigmaaldrich.com

- Very common strategy
  - Low molecular weight compounds
  - Macromolecules (e.g. proteins)

- *Advantages*
  - Extremely convenient and adaptable to many formats (e.g. column)
  - Better define how your target is displayed for binding (though not completely)

- *Disadvantages*
  - Not all immobilized molecules will be able to interact (even with its cognate RNA)
  - Immobilized form recognized is distinct from the free form of the target

A.D. Ellington and J.W. Szostak; ***Nature***; 346 (6287), 818-822, 1990

Summary

- Developed a conceptual framework for SELEX

- Library diversity
  - *Calculations*
  - *Maximizing diversity within technical constraints*
  - *Choosing the appropriate library for your needs!*

- Examined some key steps involved in the process:
  - *Target selection*
  - *RNA library construction*
  - *Partitioning strategies*

- SELEX can be successfully executed on:
  - *Very distinct targets*
  - *Using distinct library design (diversity, representation, etc)*
  - *Using distinct partitioning strategies*
  - *Fairly robust and generally applicable strategy*

Next time…

- Determining the sequence identity of individual aptamers in the selected library

- Determining that your library truly contains RNA with affinity for your target!

- Modifying your SELEX strategy to more efficiently achieve your desired outcome