Module 2 overview

lecture	*lab*
1. Introduction to the module	1. Start-up protein eng.
2. Rational protein design	2. Site-directed mutagenesis
3. Fluorescence and sensors	3. DNA amplification
4. Protein expression	4. Prepare expression system
SPRING BREAK	
5. Review & gene analysis	5. Gene analysis & induction
6. Purification and protein analysis	6. Characterize expression
7. Binding & affinity measurements	7. Assay protein behavior
8. High throughput engineering	8. Data analysis

Lecture 8: High throughput engineering

I. Paper discussion

II. General requirements for HT engineering
   A. High throughput *vs.* rational design
   B. Generating libraries

III. Selection techniques
   A. Phage display and related techniques
   B. Selection for properties other than affinity

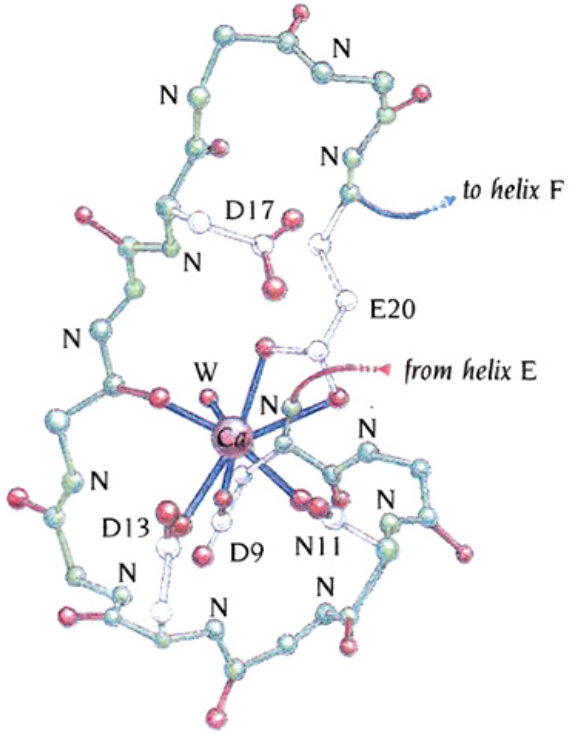
Module 2 assignment

Protein engineering research article

1. Abstract
2. Introduction
3. Materials and Methods
4. Results
5. Discussion
   - Begin with brief summary of rationale and results
   - Pick out a few high level topics to discuss (3-4 pp. DS)
   - Relate to the major themes/questions from your introduction
   - How reliable were your results, and what are caveats?
     (but do not dwell on minor technical problems)
   - Did results confirm hypothesis and why/why not?
     (be specific and relate to protein structure)
   - How do results relate to literature or other class findings?
   - What follow-up experiments could be performed?
   - What is the overall significance/impact of your results?
6. References
7. Figures

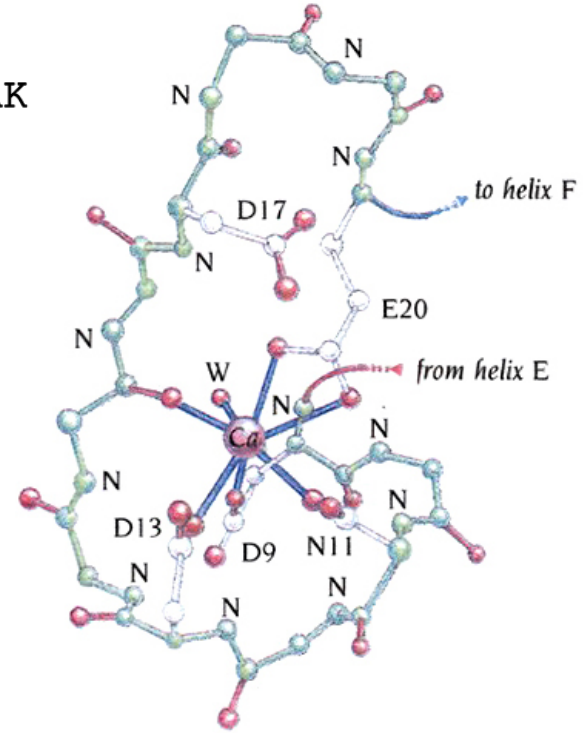
class mutants
2013

```
1         .         .         .         .         .
TDQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTEAELQD
                           E   Q
                           3x  2x

51        .         .         .         .         .
MINEVDADGNGTIYFPEFLTMMARKMKDTDSEEEIREAFRVFDKDGNGYI
T                    R       E   R
2x                   S       2x

101       .         .         .         .
SAAQLRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQMMTA
                              P                   R
```



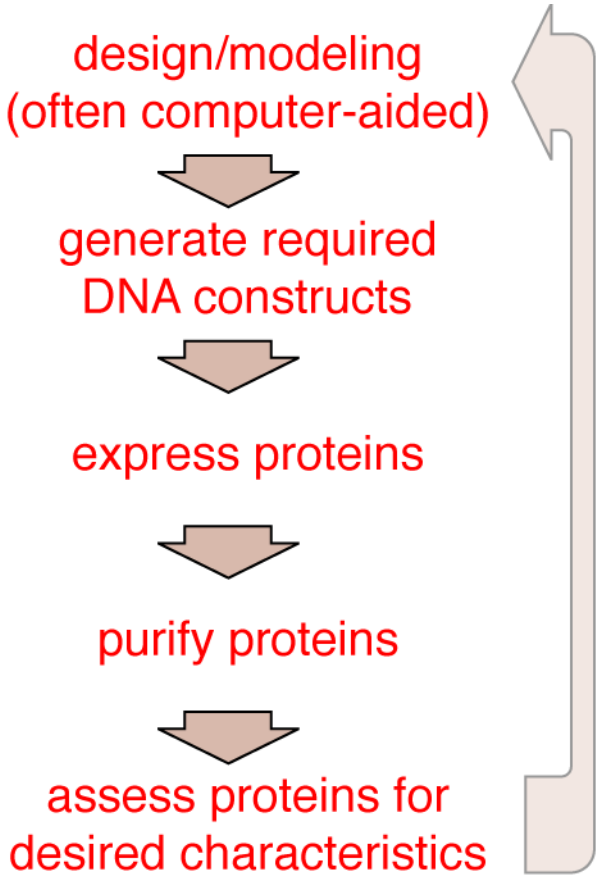
```
1           .           .           .           .           .
TDQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTEAELQD
                      P  WP  P   D   ID
                      R  HW      R   R
                      R

51          .           .           .           .           .
MINEVDADGNGTIYFPEFLTMMARKMKDTDSEEEIREAFRVFDKDGNGYI
      H  HPE    D   K              P    G           PHHD  D
      I    K    E                        K            A  K  P
                H                                          P
                K                                          W
                D
101         .           .           .           .
SAAQLRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAK
L   RS                      S   K  P  G  R    F       D
T                                      W
L
```

class mutants
2008-10



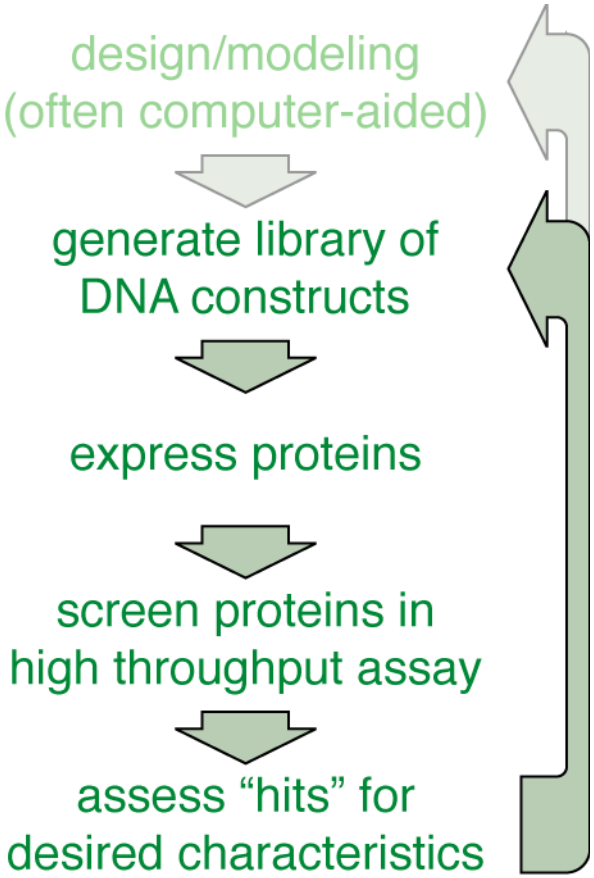
Rational protein design:

Knowldege-based, deterministic engineering of proteins with novel characteristics



"Irrational" high throughput protein engineering:

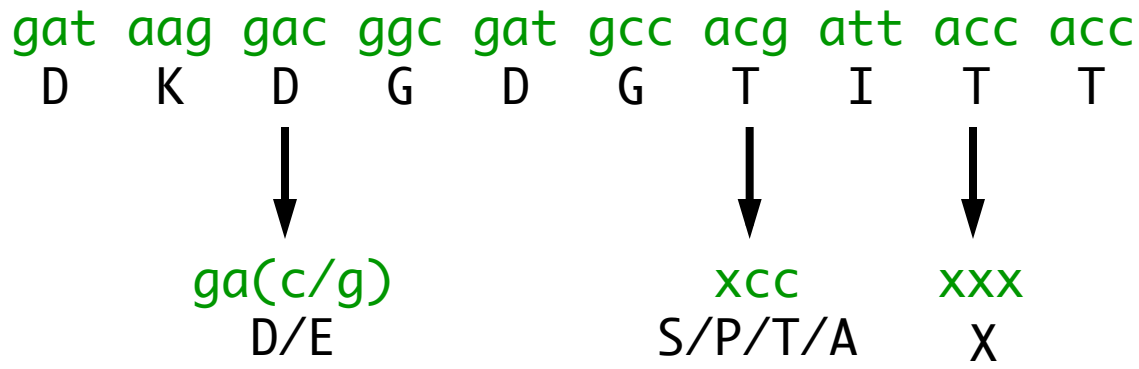
Selection for desired properties from libraries of random variants



Methods for generating mutant genetic libraries:

- site-directed mutagenesis with degenerate primers
- error-prone PCR
- gene shuffling

Degenerate primers



- not all combinations of AA's possible at each position
- number of combinations expands exponentially
- degenerate primers synthesized by split-pool method
- standard primer design criteria must be considered

PCR polymerase and conditions may be chosen to promote mutations

Polymerase	Template doublings (d)[a]	*lacI*$^-$ plaques[b] (% ± SD)	Mutation load[c] (per kilobase) (±SD)	Error rate[d] (per base) ($\times 10^{-6}$ ± SD)
Pfu-Pol (exo$^+$)	12.3	0.61 ± 0.09	0.017 ± 0.002	1.4 ± 0.2
Pfu-Pol (exo$^-$)	11.8	20 ± 1.7	0.58 ± 0.05	49 ± 4
Taq-Pol	11.6	3.9 ± 0.16	0.12 ± 0.006	10 ± 0.5

error rate = mutation load ÷ template doublings

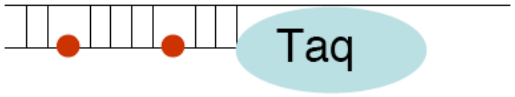
normal PCR

dCTP, dTTP

dGTP, dATP

Mg^{2+}

error-prone PCR



dCTP, dTTP ↑

dGTP, dATP ↓

Mg^{2+} ↑

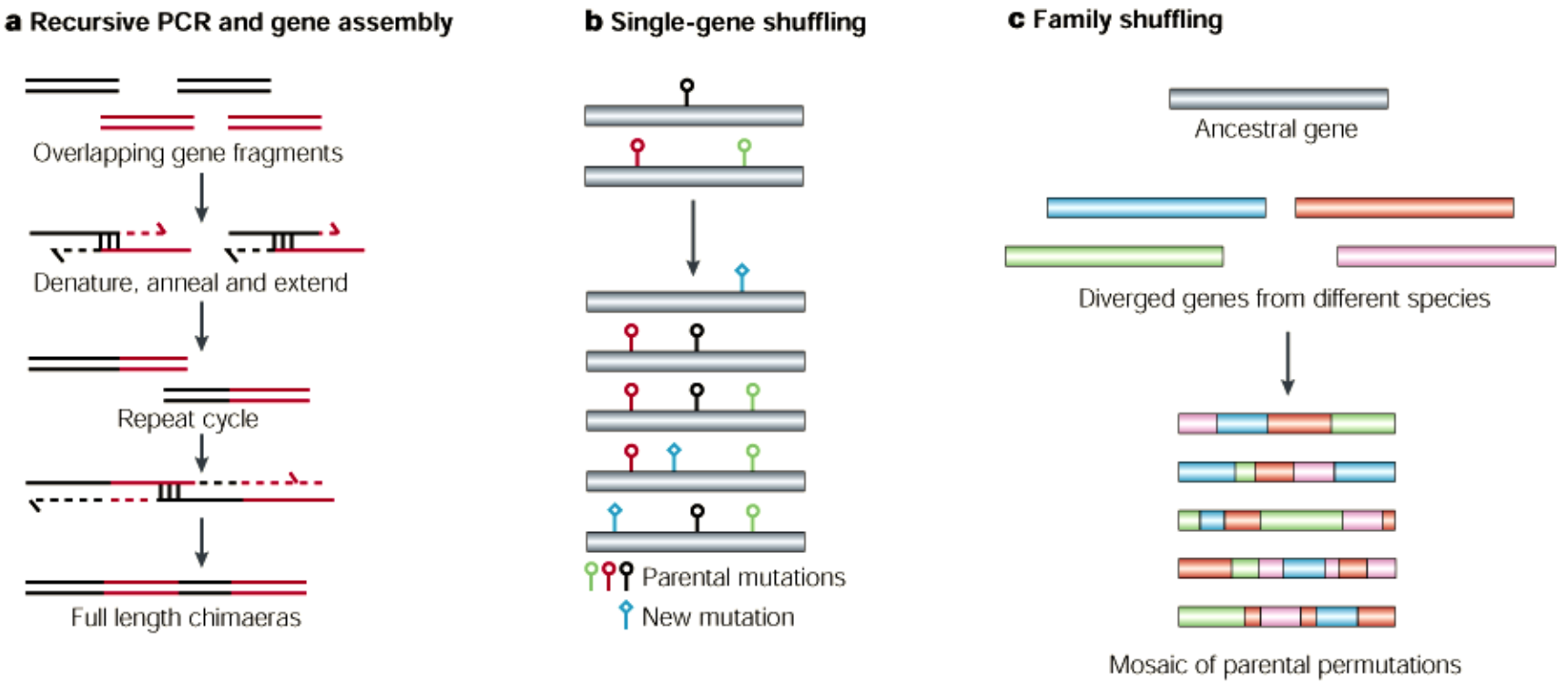
Mn^{2+}

some mutations are more likely than others

Mutation	Pfu-Pol(exo$^-$) D473G[a]	Taq-Pol (Mn^{2+}/ unbalanced dNTPs)[b]	Taq-Pol (Mn^{2+}/ unbalanced dNTPs)[c]	Taq-Pol (unnatural mutagenic bases)[d]
A→T/T→A	28	40.9	11.4	0.2
A→C/T→G	7.4	7.3	3.3	8.4
A→G/T→C	19.2	27.6	60.9	78.3
G→A/C→T	22	13.6	18.1	13.2
G→C/C→G	7.3	1.4	4.3	0.7
G→T/C→A	10.3	4.5	1.8	0.0
Insertion	2.9	0.3	Not given	~0
Deletion	2.9	4.2	Not given	~0

Gene shuffling techniques mimic diversity due to meiotic recombination:

- fragments of homologous genes combined using "sexual PCR"
- diversity may arise from error prone PCR or multiple genes



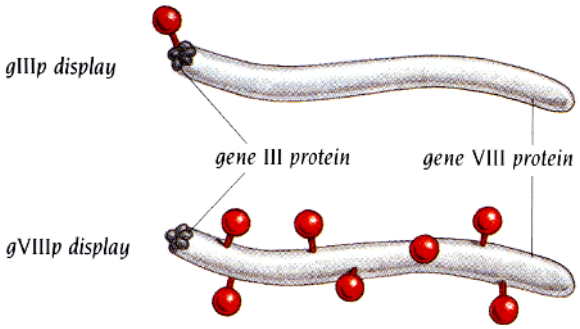
Brannigan & Wilkinson (2002) *Nat. Rev. Mol. Cell. Biol. 3:* 964-70

How are libraries of mutant proteins screened?

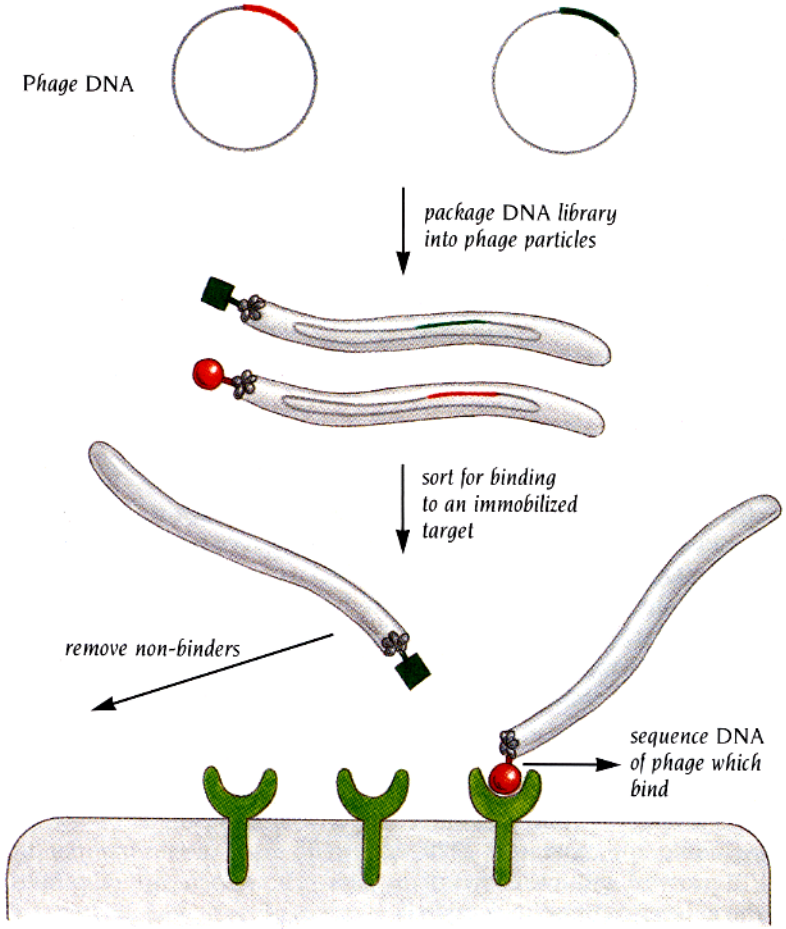
All major methods include a strategy to keep DNA sequence info associated with the proteins that are being screened.

Phage display is a versatile high throughput method to do this:

protein “displayed” on the coat of a bacteriophage, by fusing to a natural phage coat protein



M13 phage (contains DNA)

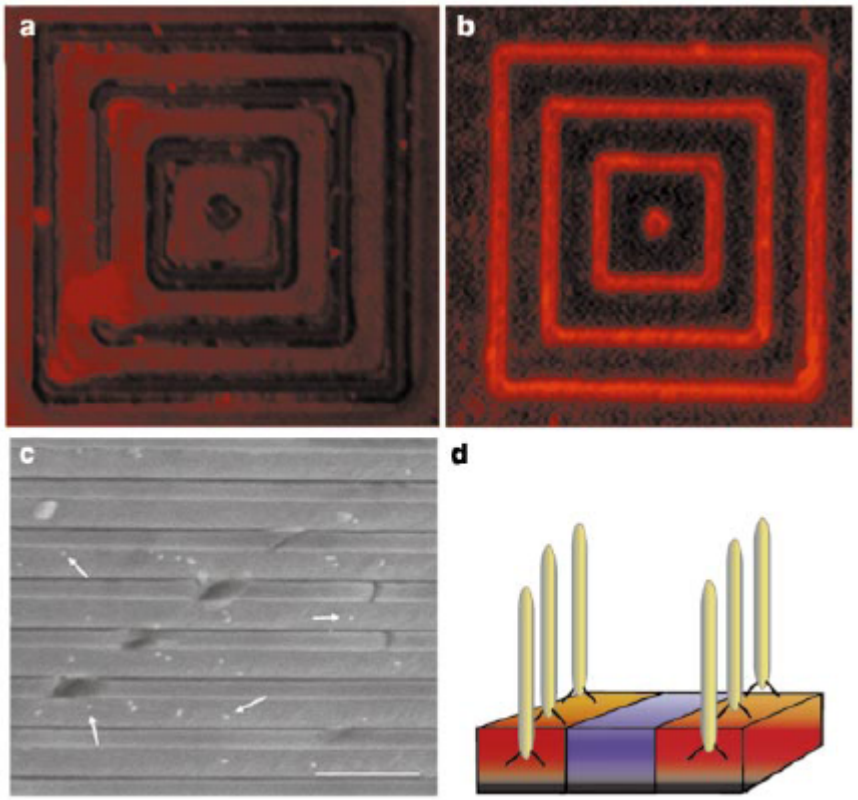


Application: phage-displayed peptides that bind to GaAs

selected sequences

G13-5				V	T	S	P	D	S	T	T	G	A	M	A
G12-5				A	A	S	P	T	Q	S	M	S	Q	A	P
G12-3				A	Q	N	P	S	D	N	N	T	H	T	H
G1-4	A	S	S	S	R	S	H	F	G	Q	T	D			
G12-4		W	A	H	A	P	Q	L	A	S	S	S	T		
G14-3		A	R	Y	D	L	S	I	P	S	S	E	S		
G7-4				T	P	P	R	P	I	Q	Y	N	H	T	S
G15-5				S	S	L	Q	L	P	E	N	S	F	P	H
G14-4					G	T	L	A	N	Q	Q	I	F	L	S	S
G11-3				H	G	N	P	L	P	M	T	P	F	P	G
G1-3				R	L	E	L	A	I	P	L	Q	G	S	G

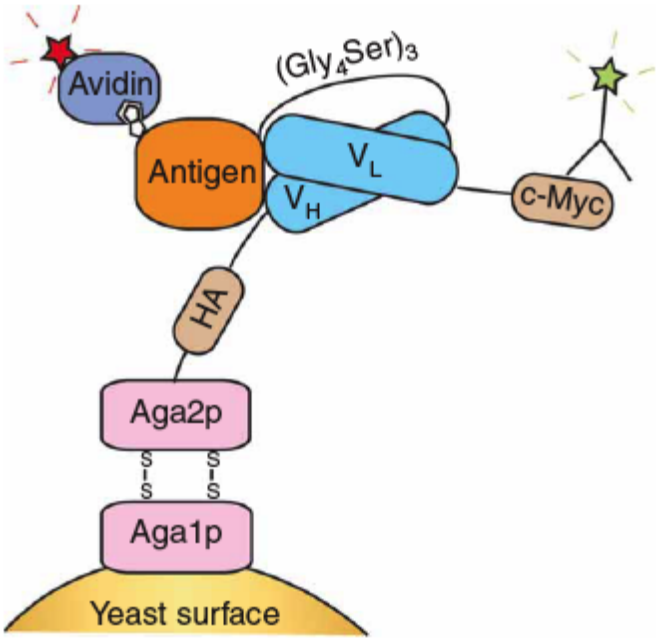
phages patterned on target substrate



Whaley *et al.* (2000) *Nature 405:* 665-8.

Yeast display: similar to phage display, but with proteins fused to a *Saccharomyces* cell wall protein (DNA in yeast)

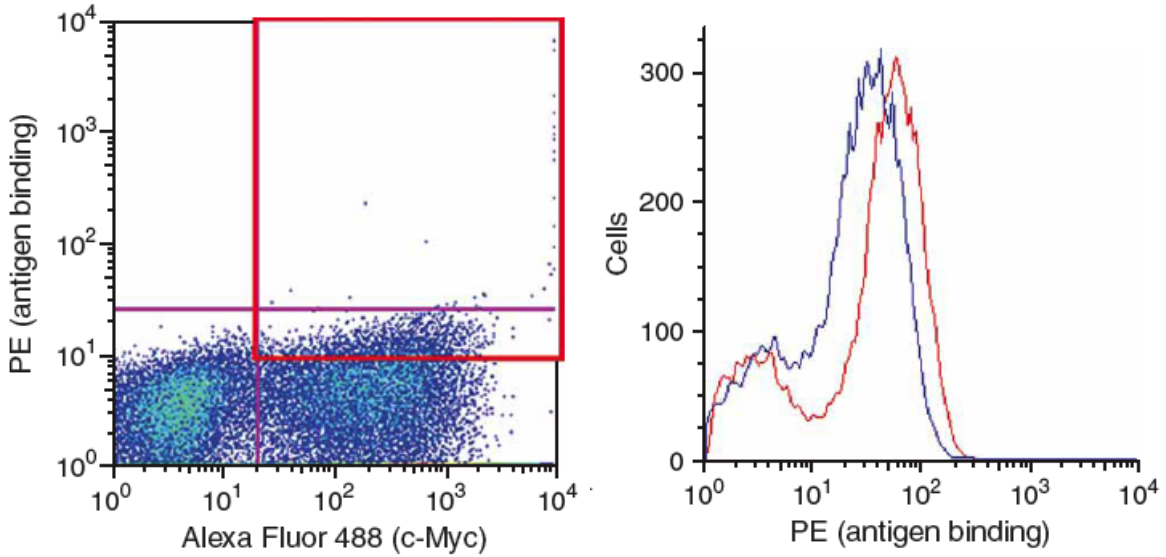
What would you expect advantages to be, compared with phage display?



In this example, a population of scFvs was screened for binding to an antigen

left: selection criterion for FACS assay

right: comparison of wt (blue) and selected (red) scFv binding

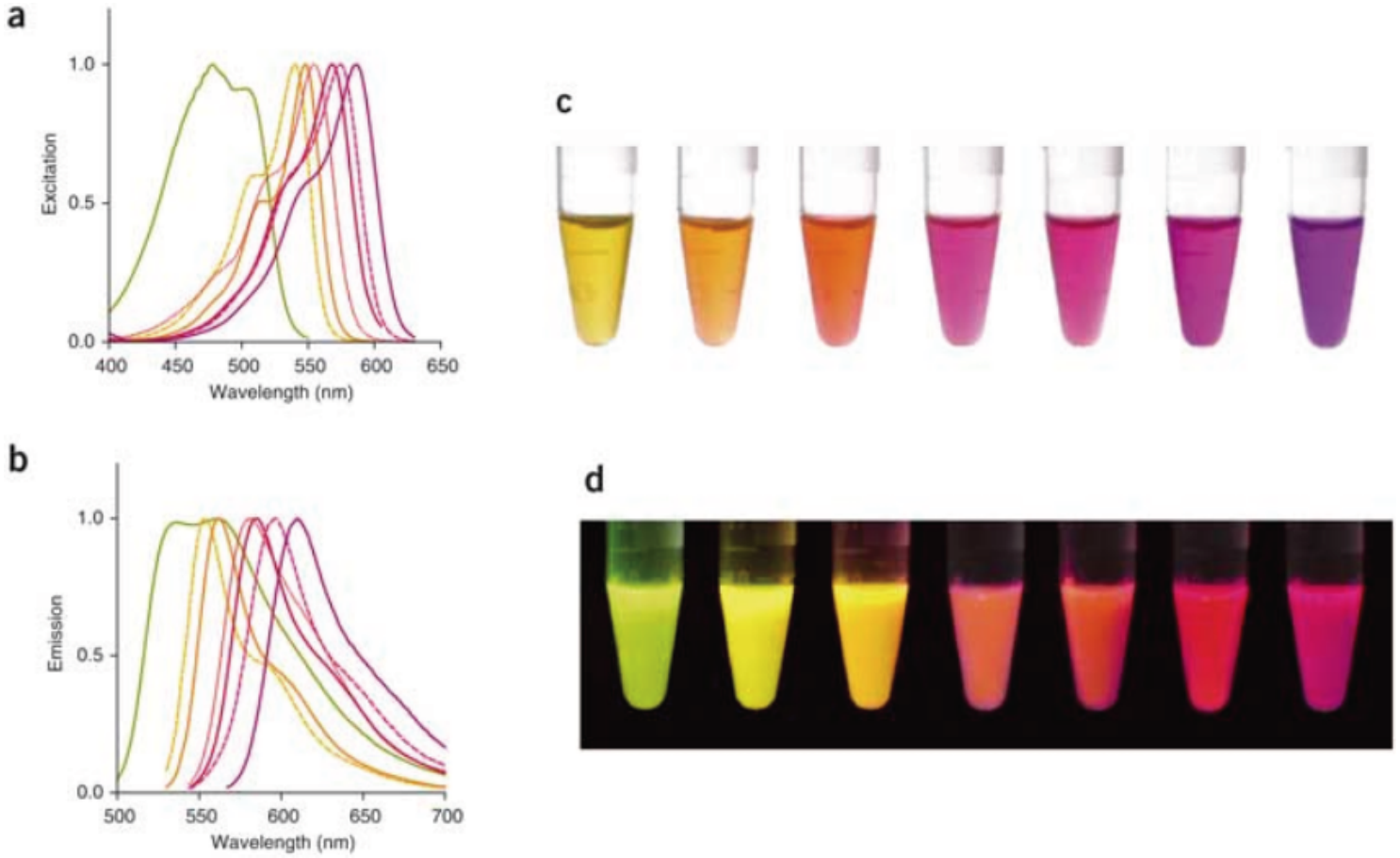


Why not more stringent selection?

Chao *et al.* (2006) *Nat. Protoc. 1:* 755-68

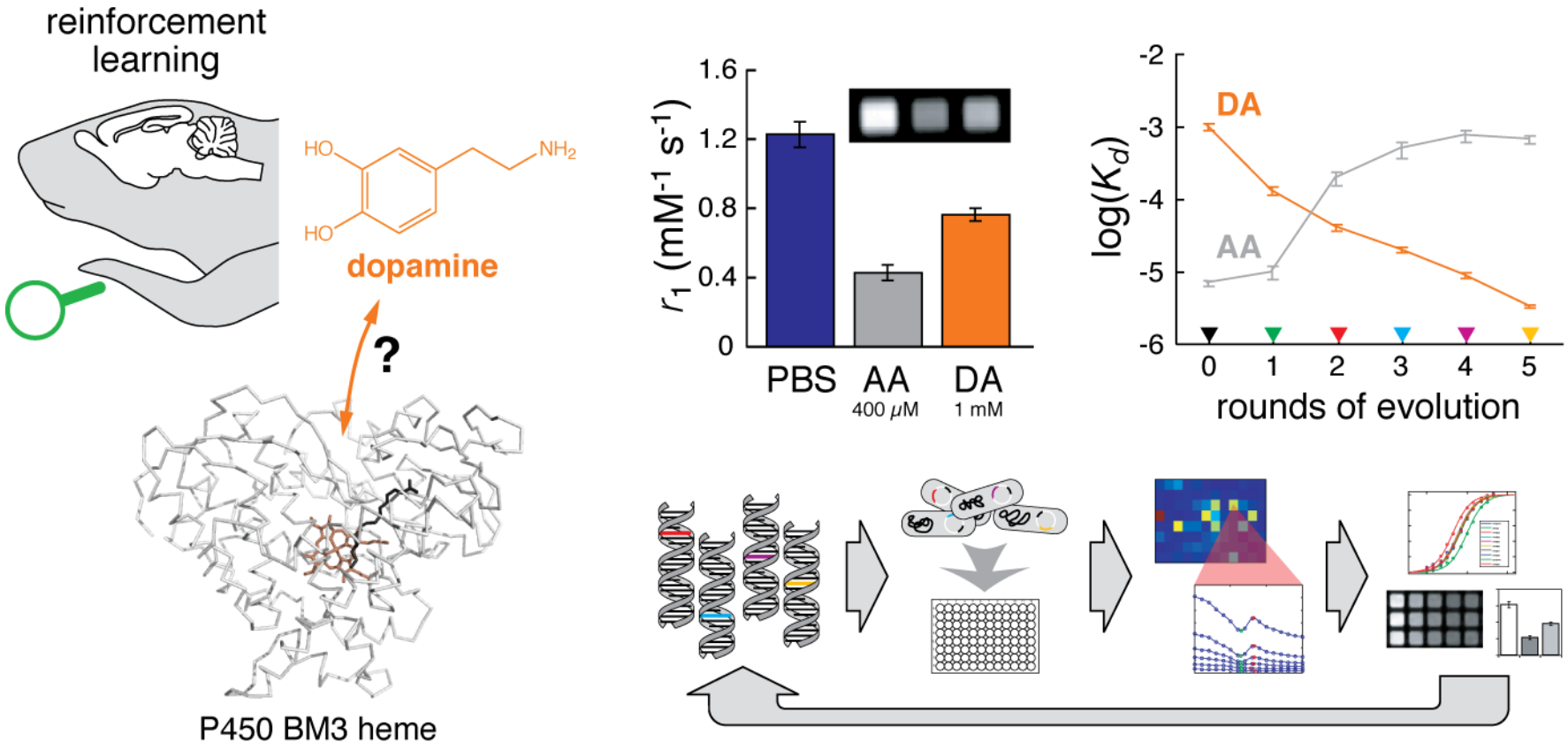
What about properties other than affinity?

A simple example: screen for DsRed variants with different excitation and emission wavelengths–how could this be done?



Shaner *et al.* (2004) *Nat. Biotechnol. 22:* 1567-72

another example: neurotransmitter sensor for MRI



This screen only involved ~500 variants/round; under what circumstances would you expect this level of throughput to be successful?

Shapiro, Westmeyer *et al.* (2010) *Nat. Biotechnol.*

Which type of screening method to use?

screen method	*throughput*	*other notes*
SELEX	10^{15}	selection of DNA/RNA
ribosome display	10^{15}	*in vitro* protein synthesis
phage display	10^{11}	best for small proteins/peptides
yeast display	10^{8}	compatible w/eukar. proteins
plate assays	$< 10^{5}$	versatile but more complex

number of variants in a protein library

x residues	= 20^{x} possible variants
12 residues	= 4 x 10^{15} variants

lesson: impossible to cover sequence space except with short sequences (or few positions) and only the most high throughput techniques

Good luck with your papers!

http://www.youtube.com/watch?v=qbtgmuyo1B8