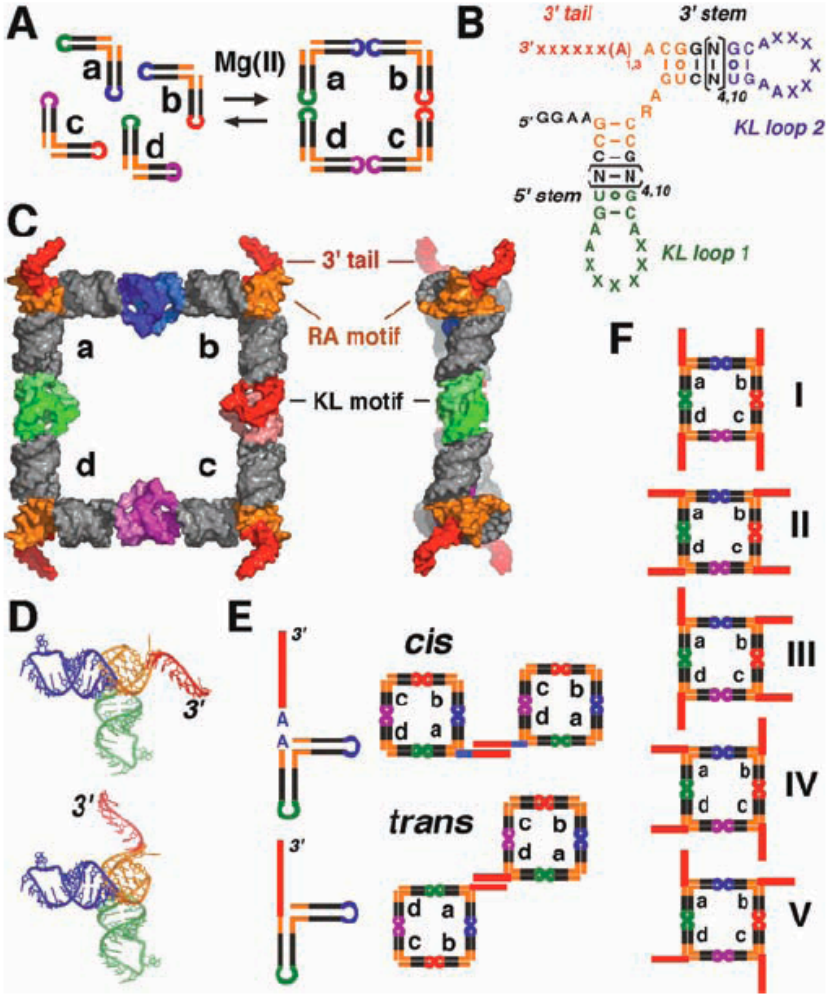
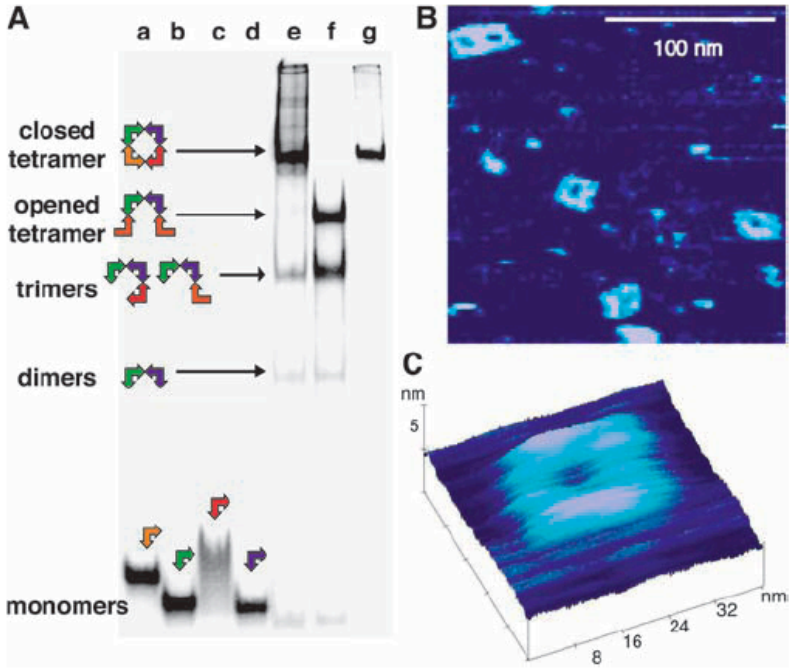
RNA Self-assembly

Building Programmable Jigsaw Puzzles with RNA

Arkadiusz Chworos,[1] Isil Severcan,[1] Alexey Y. Koyfman,[1,2] Patrick Weinkam,[1,4] Emin Oroudjev,[3] Helen G. Hansma,[3] Luc Jaeger[1,2*]

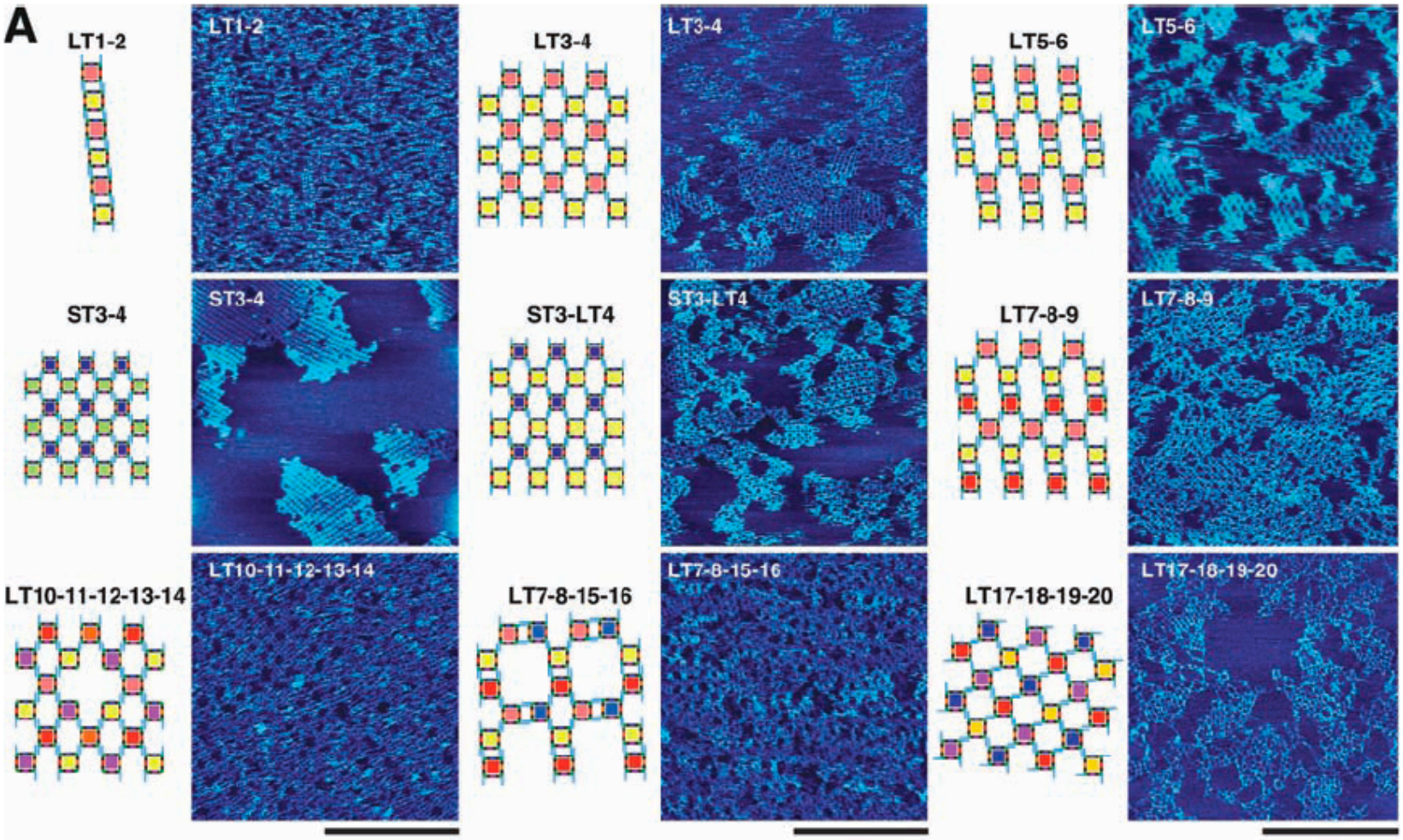
(2004) *Science* **306**, 2068-2072





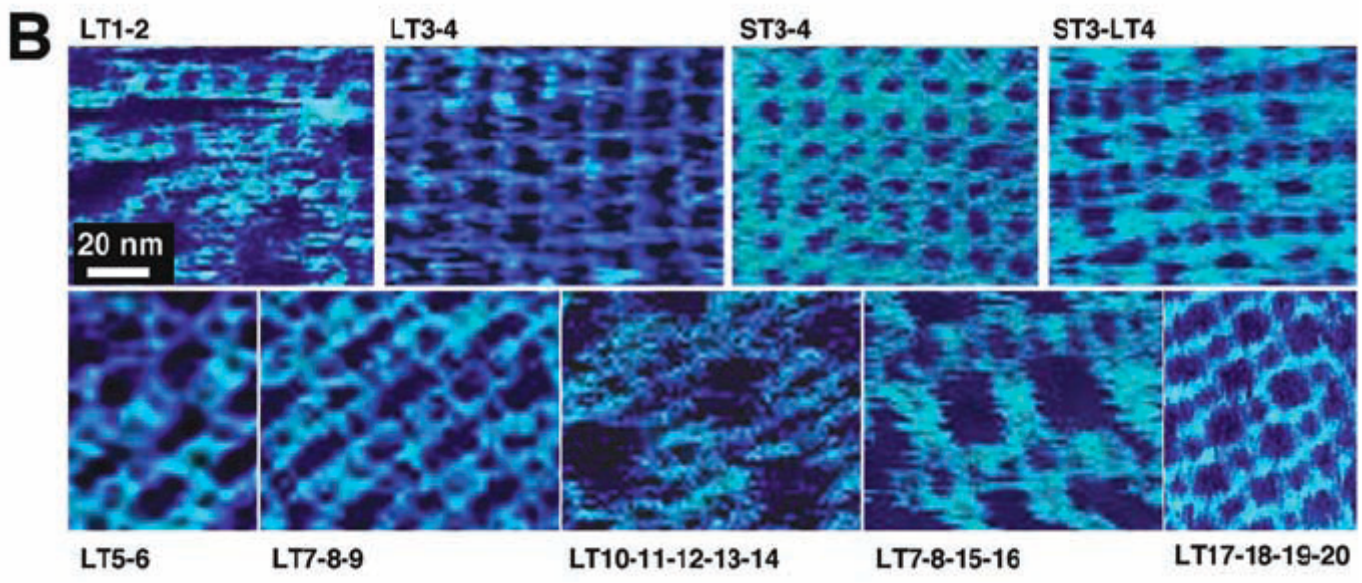
RNA Self-assembly

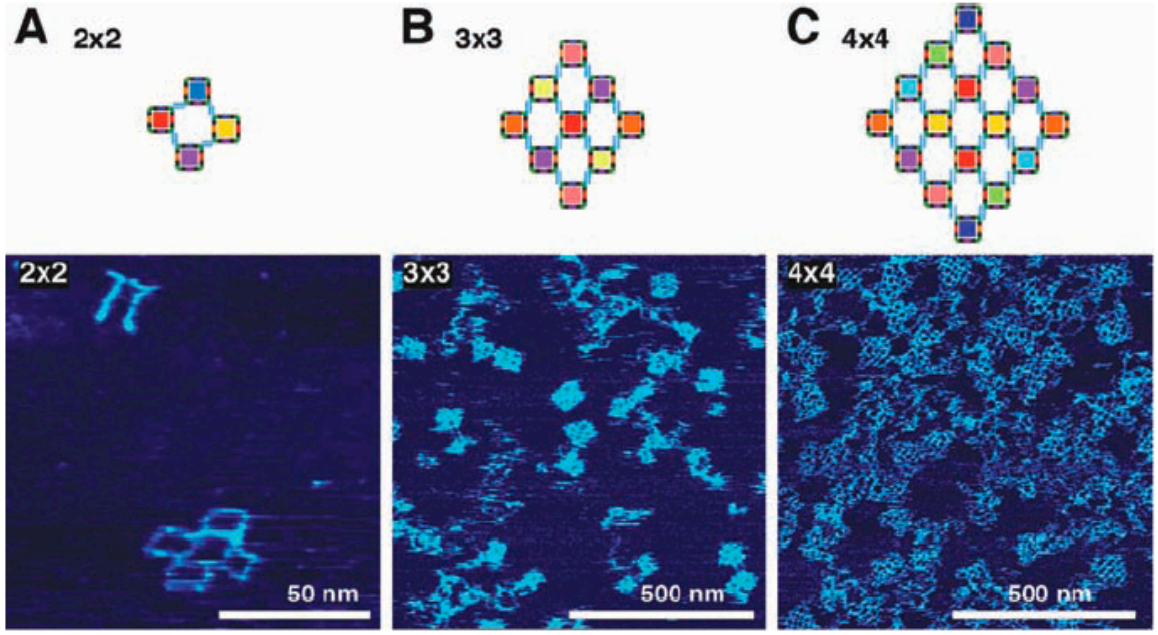
(2004) *Science* **306**, 2068-2072

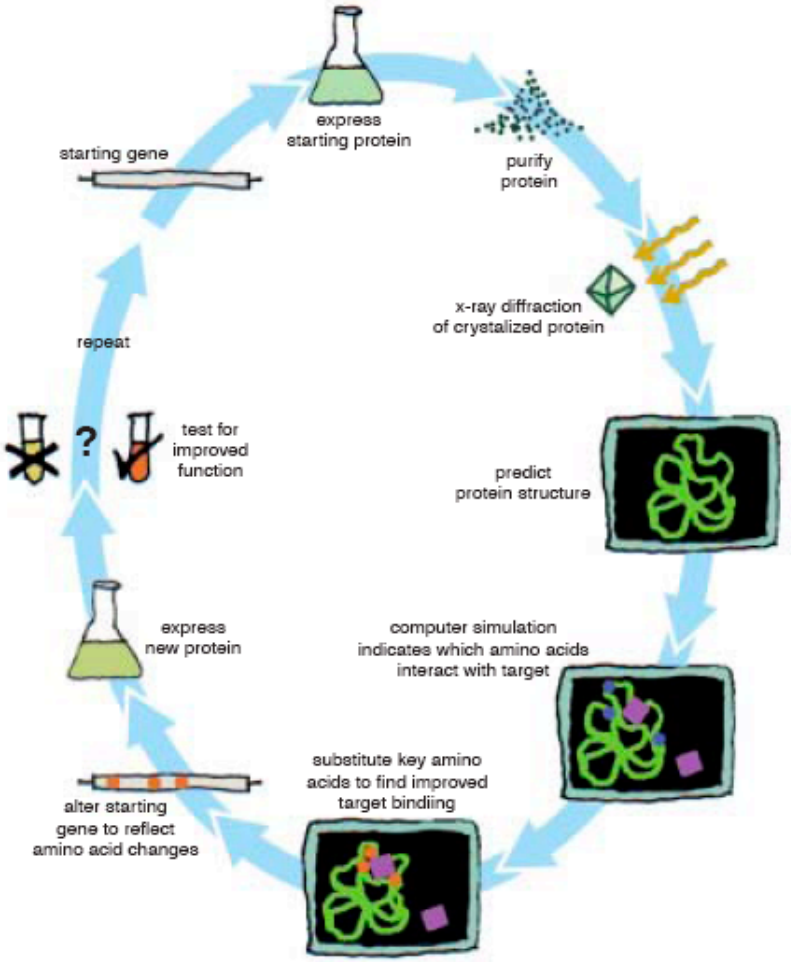


RNA Self-assembly

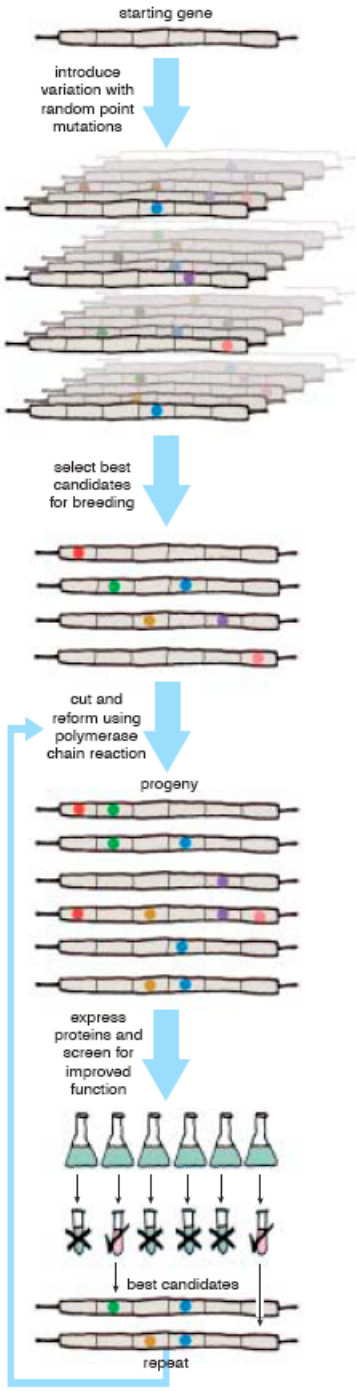
(2004) *Science* **306**, 2068-2072







Rational design



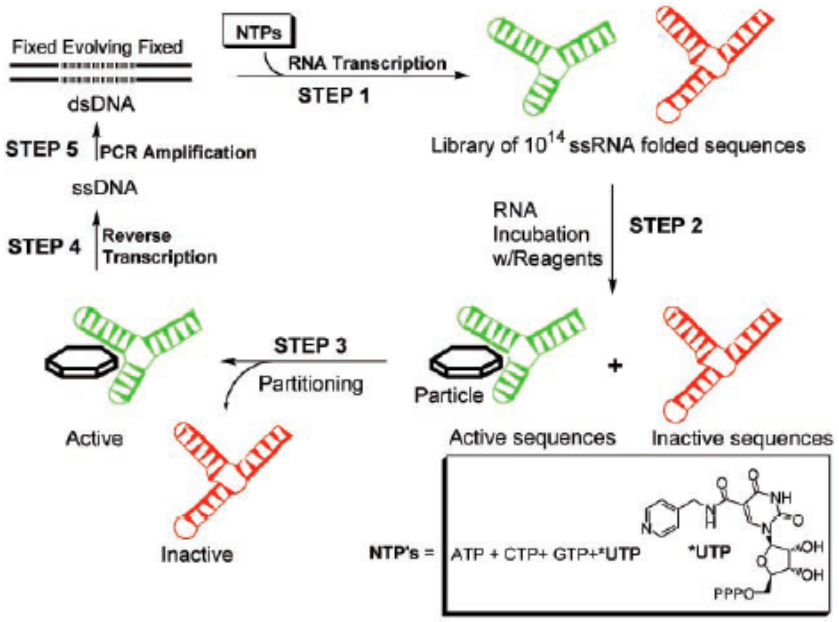
Directed evolution

RNA Physical Catalyst

RNA-Mediated Metal-Metal Bond Formation in the Synthesis of Hexagonal Palladium Nanoparticles

Lina A. Gugliotti, Daniel L. Feldheim,* Bruce E. Eaton*

(2004) *Science* **304**, 850-852



Scheme 1. Steps of the RNA in vitro selection cycle.

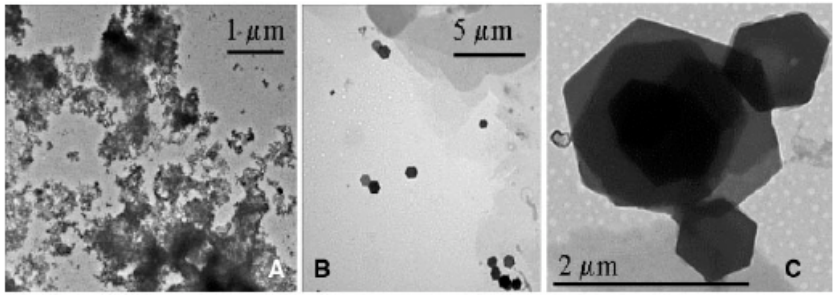
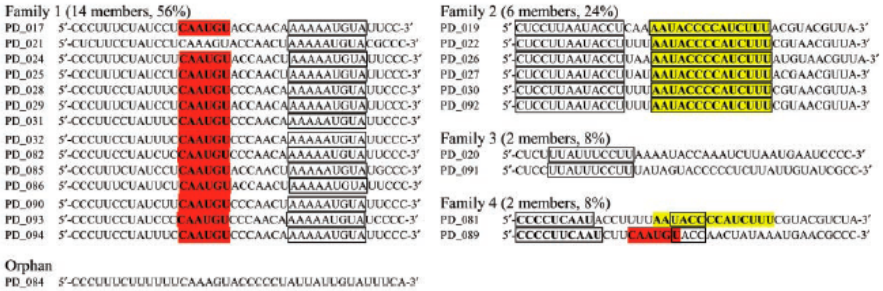


Fig. 1. Transmission electron micrograph images of palladium particles formed in the presence of cycle 0 pool modified RNA (A) and the cycle 8 RNA pool (B and C).



*UTP is 5-(4-pyridylmethyl)-uridine 5' triphosphate
Pd2(DBA)3 is dibenzylide-acetone palladium (0)

RNA Physical Catalyst

RNA-Mediated Control of Metal Nanoparticle Shape

Lina A. Gugliotti,† Daniel L. Feldheim,*,† and Bruce E. Eaton*,†,‡

J. AM. CHEM. SOC. 2005, *127*, 17814−17818

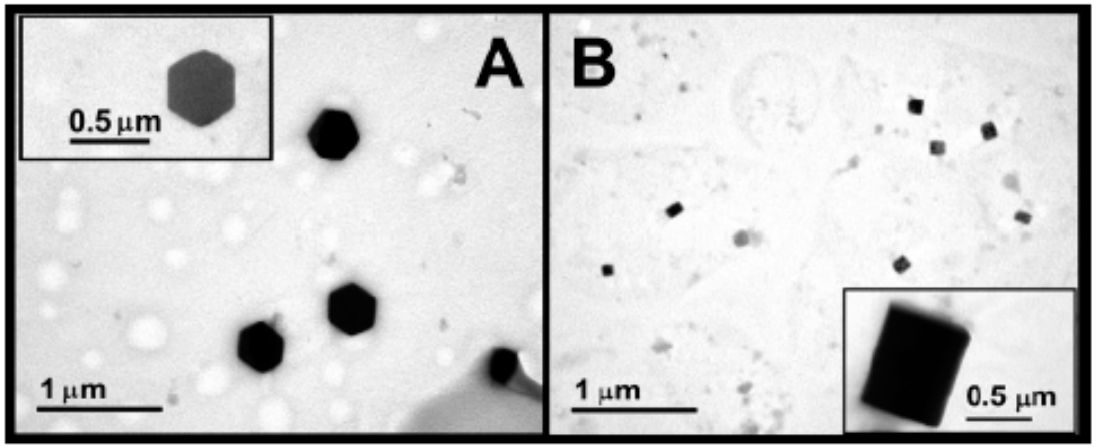


Figure 2. TEM micrographs of (A) hexagonal palladium particles formed in the presence of Pdase 17 and (B) cubic palladium particles formed in the presence of Pdase 34.

Table 1. Metal Precursor Effects on Particle Shape and Size

RNA isolate	metal precursor	particle shape	size (μm)	% population
17	$Pd_2(DBA)_3$	hexagonal	1.24 ± 0.57	100
	$Pt_2(DBA)_3$	spherical	0.10 ± 0.06	14
		cubic	0.22 ± 0.13	16
		hexagonal	0.46 ± 0.22	69
	$Pd(PPh_3)_4$	spherical	0.27 ± 0.16	100
	$Pt(PPh_3)_4$	spherical	0.007 ± 0.003	100
	$Ni(PPh_3)_4$	spherical	0.27 ± 0.10	100
34	$Pd_2(DBA)_3$	cubic	0.10 ± 0.05	100
	$Pt_2(DBA)_3$	spherical	0.16 ± 0.05	2
		cubic	0.14 ± 0.07	29
		hexagonal	0.37 ± 0.16	68
	$Pd(PPh_3)_4$	spherical	0.31 ± 0.29	100
	$Pt(PPh_3)_4$	spherical	0.009 ± 0.007	100
	$Ni(PPh_3)_4$	spherical	0.24 ± 0.10	100

*UTP is 5-(4-pyridylmethyl)-uridine 5' triphosphate

Pd2(DBA)3 is dibenzylide-acetone palladium (0)

Ellington & Szostak

Nature. 1990 Aug 30;**346**(6287):818-22.

In vitro selection of RNA molecules that bind specific ligands.

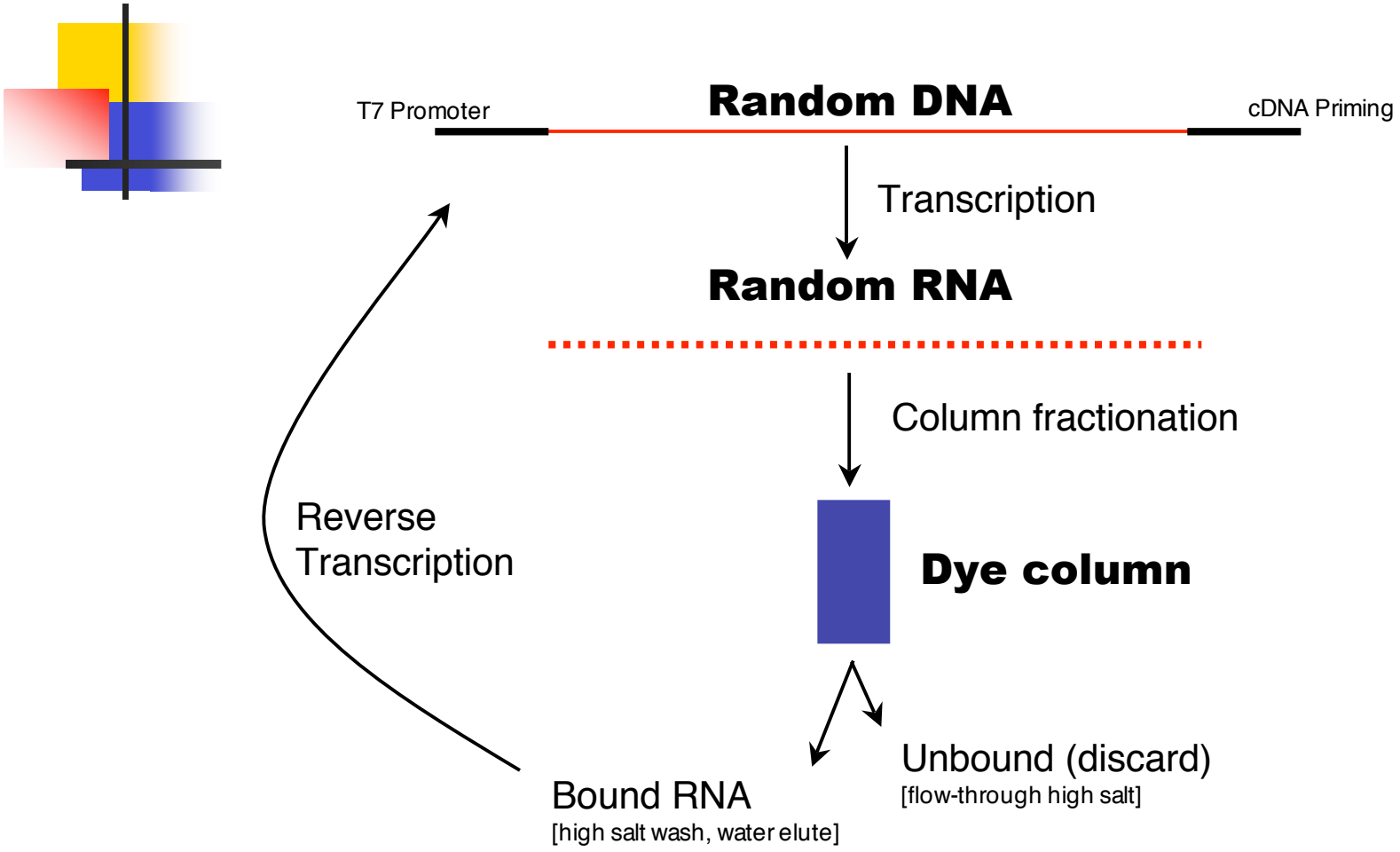
Ellington AD, Szostak JW.

Department of Molecular Biology, Massachusetts General Hospital, Boston 02114.

Subpopulations of RNA molecules that bind specifically to a variety of organic dyes have been isolated from a population of random sequence RNA molecules. Roughly one in 10^{10} random sequence RNA molecules folds in such a way as to create a specific binding site for small ligands.

Ellington & Szostak

Nature. 1990 Aug 30;**346**(6287):818-22.



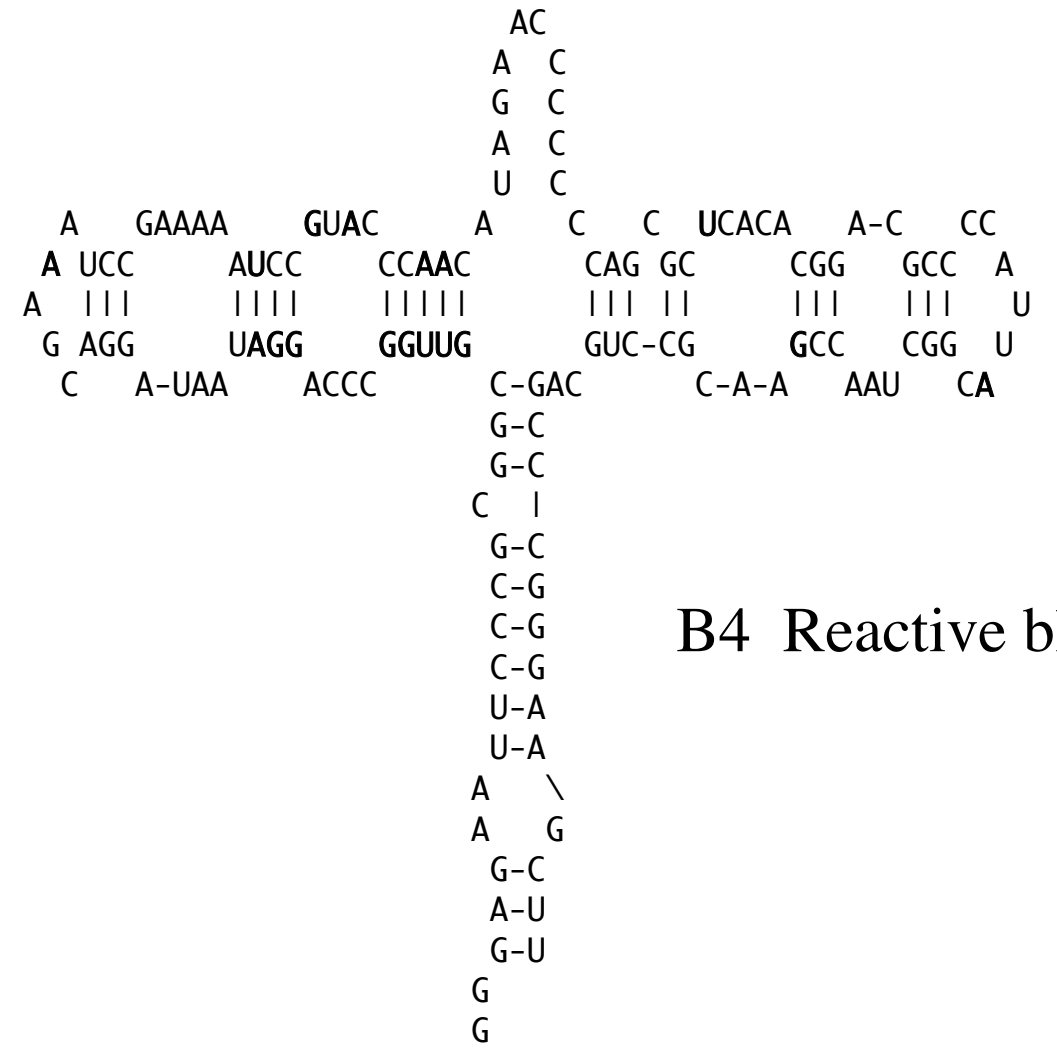
Ellington & Szostak

Nature. 1990 Aug 30;**346**(6287):818-22.

- 100 random nucleotides (+ constant regions)
- Complexity of initial pool ~10^{15} (10^{60} possible)
- Poor transcription; PCR amp; ~10^{13} diversity
- ~$4x10^{14}$ RNA molecules applied to column
- <0.1% of initial RNA library bound then after 4 to 5 cycles of bind, RT, PCR, Transcribe, rebind ~50% of RNA molecules bound.
- Binding is specific for dye selected.
- For 1 column 20 clones were sequenced ~16-20 nucleotides were invariant over the set of clones.
- APTAMER from Latin ‘aptus’, to fit.

Ellington & Szostak

Nature. 1990 Aug 30;**346**(6287):818-22.



B4 Reactive blue binder

Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase

CRAIG TUERK AND LARRY GOLD

Science, New Series, Vol. 249, No. 4968 (Aug. 3, 1990), 505-510.

- SELEX- systematic evolution of ligands by exponential enrichment.
- Selected RNA for binding to T4 DNA polymerase.
- 65,536 possible sequences.
- Wild-type T4 bacteriophage mRNA sequence found and a variant with 50% identity to wild-type.

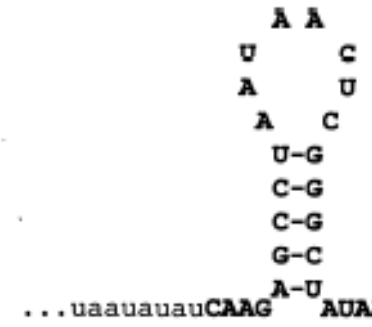
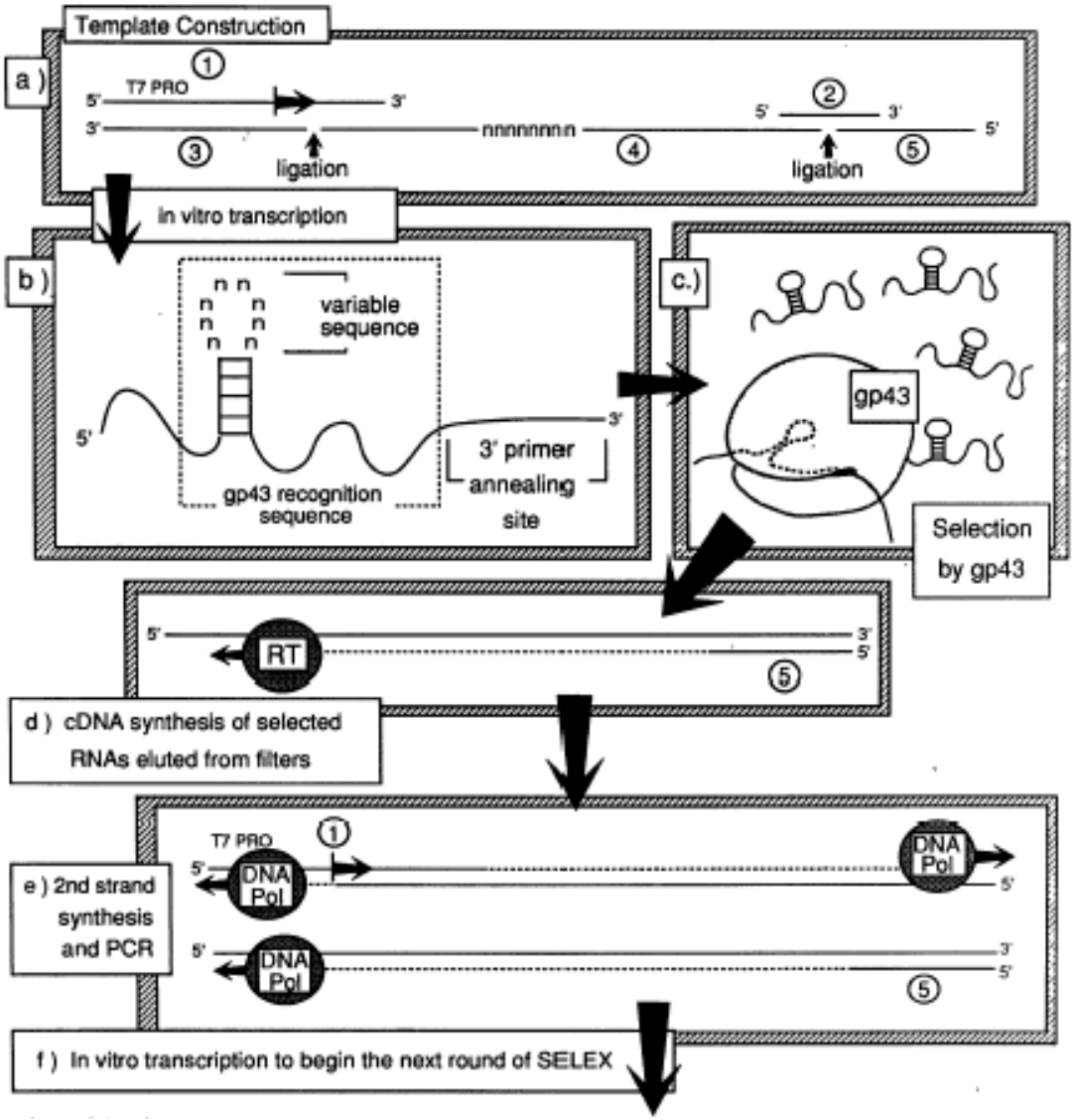


Fig. 1. The gene *43* translational operator in bacteriophage T4. Shown is the sequence at the ribosome binding site of gene *43*. The bold-faced capitalized letters indicate the extent of the information required for binding of gp43 (*6*).

Tuerk & Gold

Science, New Series, Vol. 249, No. 4968 (Aug. 3, 1990), 505-510.



Tuerk & Gold

Science, New Series, Vol. 249, No. 4968 (Aug. 3, 1990), 505-510.

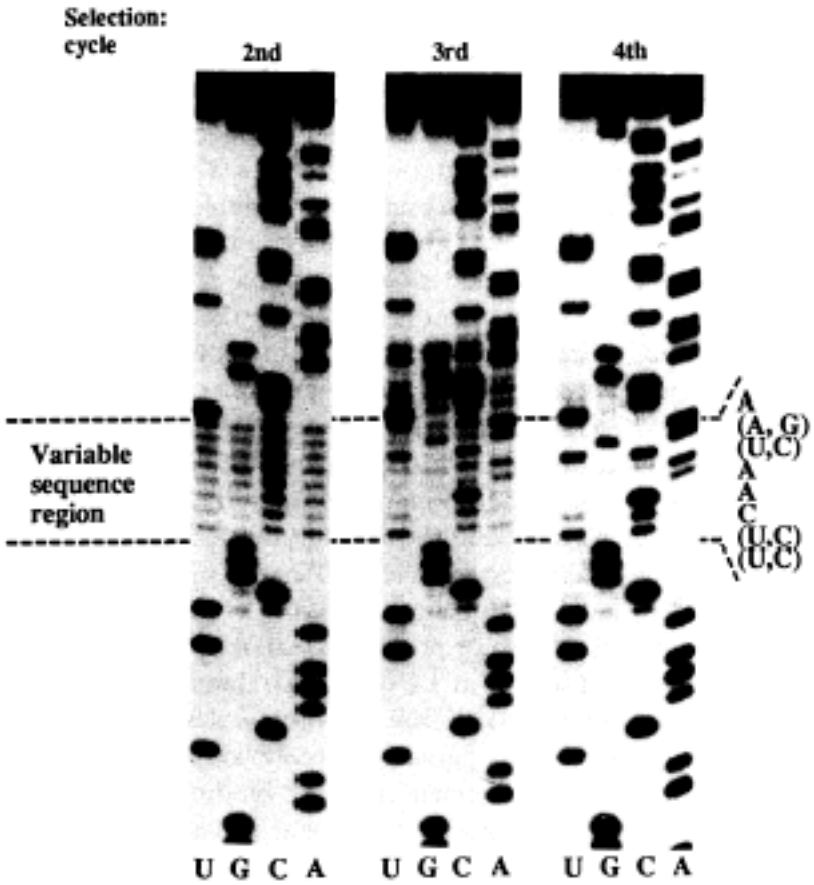


Fig. 4. Batch RNA sequences through rounds of selection. Sequencing and labeling were done as in Fig. 3. The transcripts were derived from amplified products of selections 2, 3, and 4 for experiment B.

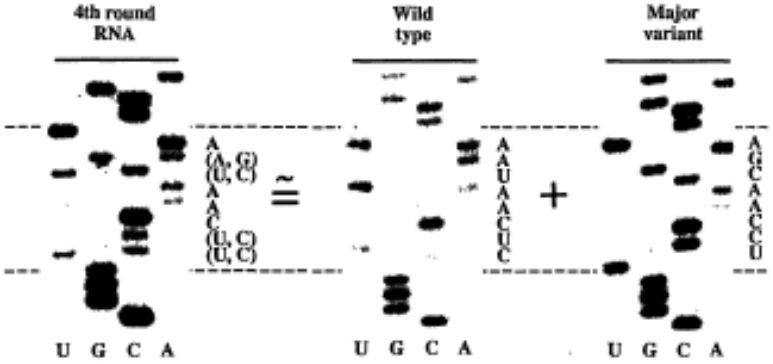


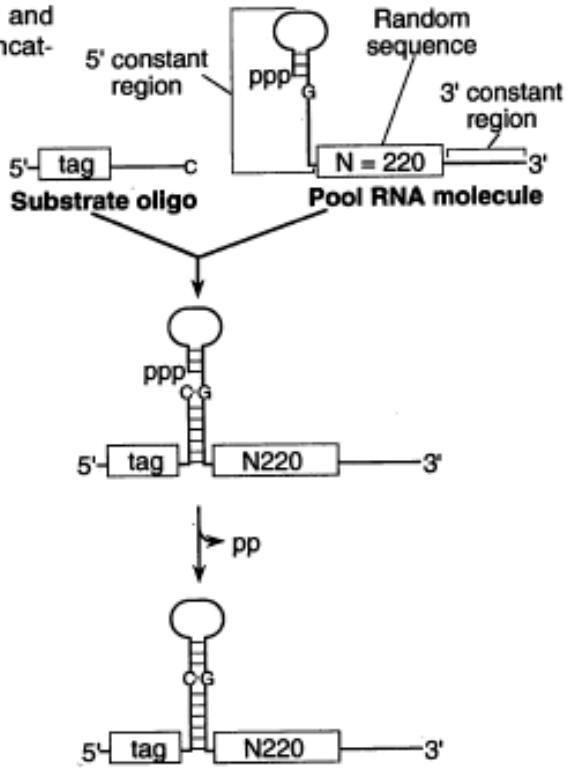
Fig. 6. The clonal composition of the "consensus" sequence. Of the three sequencing gels, the first is the batch sequencing of selected RNA's as shown in Fig. 4 for experiment B; the second and third gels show sequencing products derived from plasmid templates (*26*) of representative clonal isolates of the batch population. These two individual sequences alone yield the batch "consensus" that is observed.

Science, New Series, Volume 261, Issue 5127 (Sep. 10, 1993), 1411-1418.

Isolation of New Ribozymes from a Large Pool of Random Sequences

David P. Bartel and Jack W. Szostak

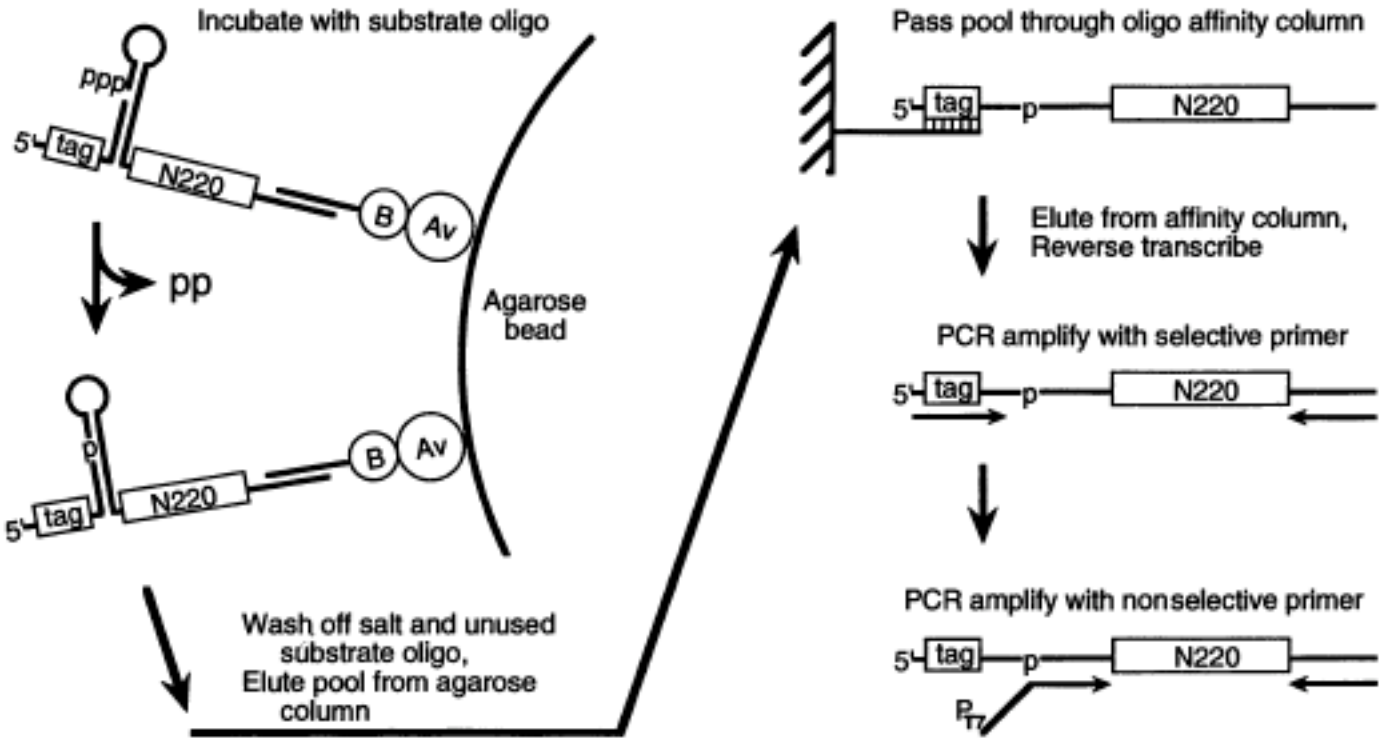
An iterative in vitro selection procedure was used to isolate a new class of catalytic RNAs (ribozymes) from a large pool of random-sequence RNA molecules. These ribozymes ligate two RNA molecules that are aligned on a template by catalyzing the attack of a 3′-hydroxyl on an adjacent 5′-triphosphate—a reaction similar to that employed by the familiar protein enzymes that synthesize RNA. The corresponding uncatalyzed reaction also yields a 3′,5′-phosphodiester bond. In vitro evolution of the population of new ribozymes led to improvement of the average ligation activity and the emergence of ribozymes with reaction rates 7 million times faster than the uncatalyzed reaction rate.



- 220 random positions
- 1.6×10^{15} diversity

Science, New Series, Volume 261, Issue 5127 (Sep. 10, 1993), 1411-1418.

Bartel & Szostak



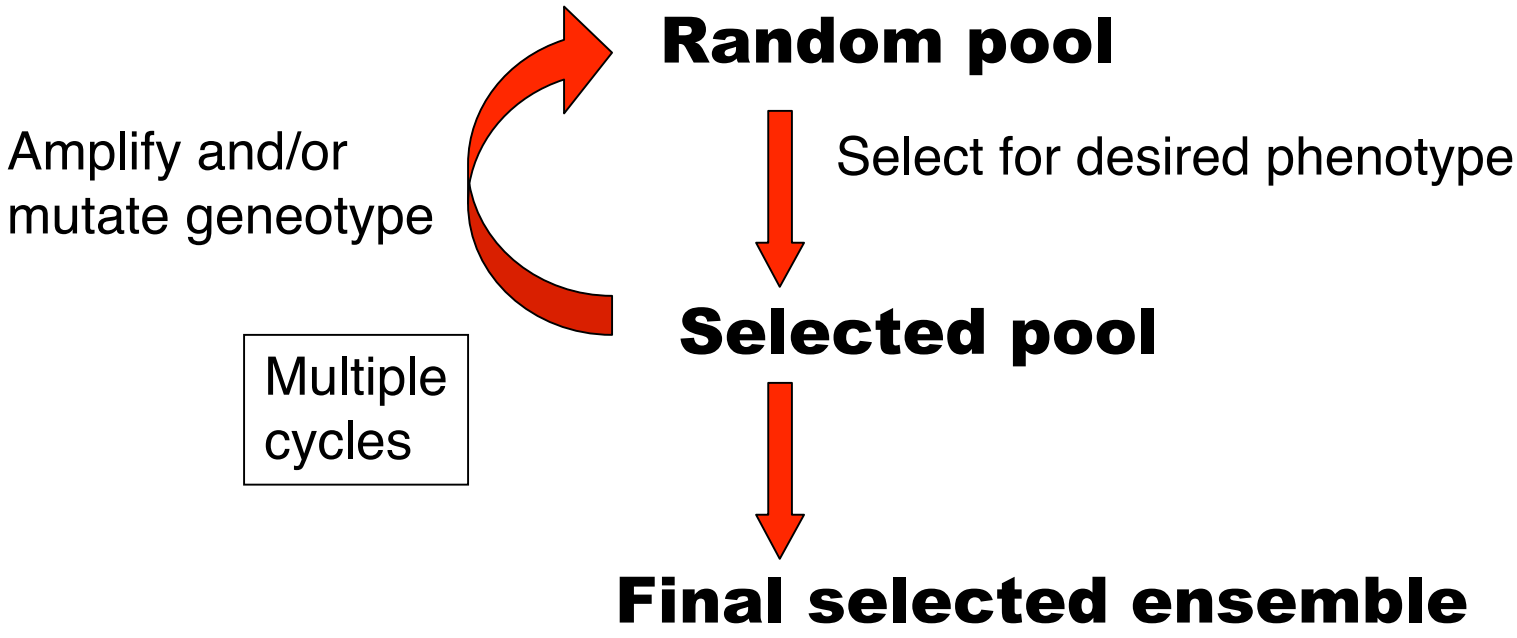
50 60 70 80 90 100 110 120
AAAGCGTACGAACTAAGTGATAAGGAATTCAGATTAAAGCGAGCGGGGTTCCGCACGTGGTTTGCAGGGGTTGCCTTGG

130 140 150 160 170 180 190 200
CGACACTGTTCCTCTCACTCTGTTCAAAAGCCGATTGCTGAACGTGTTAAAACGCCATCACAATATGTGCCCGTACGGCACC

210 220 230 240 250 260 270
GAGAAAGGCATAGGGTGACGTCGGTGGAGATGTATAGTCTTAGGGTGAGGCTGGTATGACCGGTAACGGCTC

Fig. 9. Sequence of one of the isolated ribozymes. Members of pool 10 DNA were cloned (TA Cloning, Invitrogen) and sequenced. The sequence of the central 233 nucleotides of clone 4 is shown. Sequences of the 5′ constant region (bases 1 to 42) and the 3′ constant region (bases 276 to 295) are described in (*29*). Sty I and Ban I restriction sites used in pool construction are underlined. The 10 bases that differ from the consensus sequence of the dominant pool 10 sequence family are also underlined. Consensus bases at these positions are 56A, 63A, 90A, 94T, 96T, 131C, 188A, 195A, 217A, and 258A.

In vitro evolution of nucleic acids.

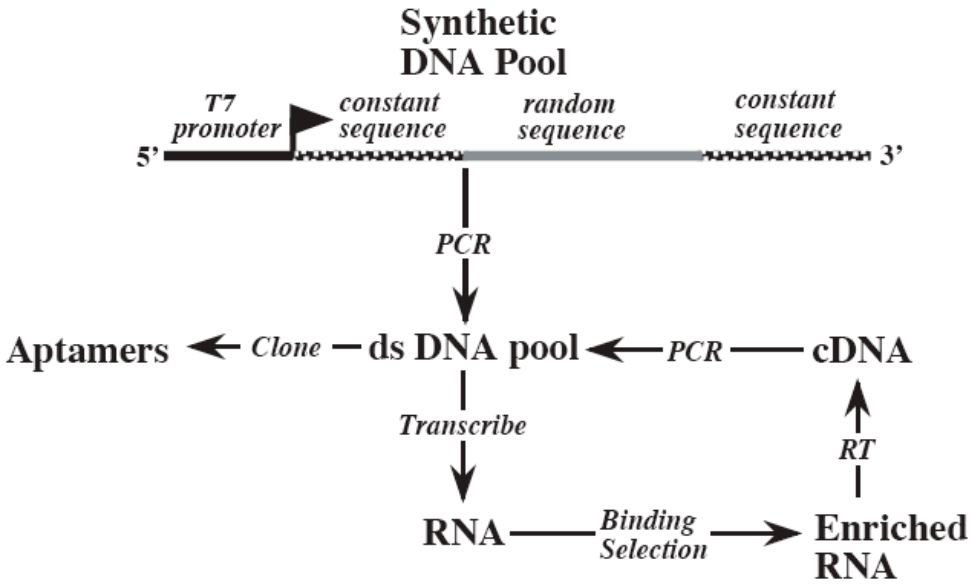


In Vitro Selection of Functional Nucleic Acids

David S. Wilson and Jack W. Szostak
Howard Hughes Medical Institute and Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114-2696;

Annu. Rev. Biochem. 1999. 68:611–647

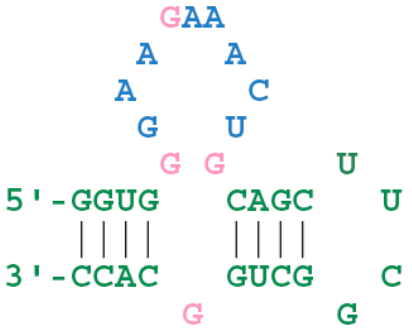
- In vitro selection requires linking of genotype and phenotype.
- Binding targets: simple ions, small molecules, peptides, proteins, organelles, and cells.



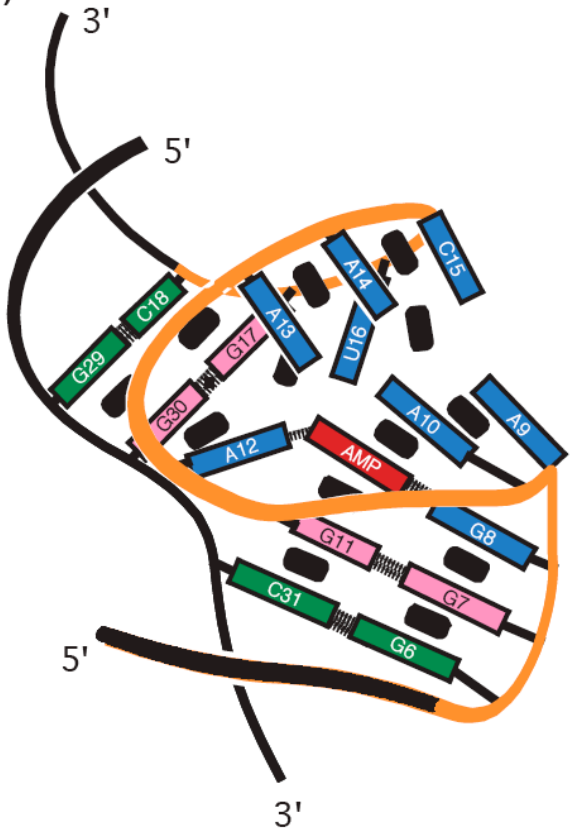
Annu. Rev. Biochem. 1999. 68:611–647

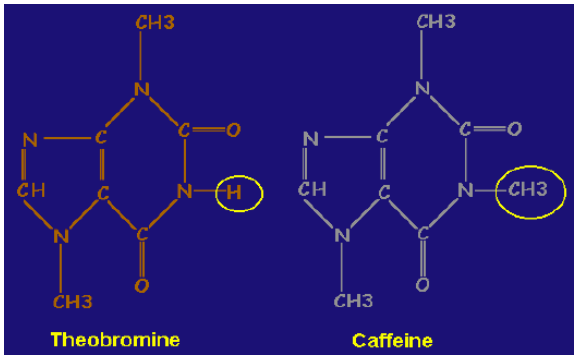
AMP binding aptamer

a)



b)





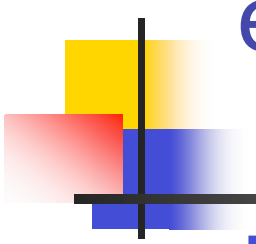
- Selectivity: theophylline aptamer shows 10,000-fold decrease in binding to caffeine which differs by one additional methyl group.

Jenison RD, Gill SC, Pardi A, Polisky B. 1994. *Science* 263:1425–29

TABLE 1 Aptamers for small molecules

Target	Estimated K_D (μm)
Nucleotides and nucleobases	
ATP/adenosine	1
ATP/adenosine (DNA)	6
Guanosine	32
Guanine/xanthine	1.8
7-Methyl-GTP	~0.5
Theophylline	0.11
Amino acids	
Arginine	0.33
Citrulline	62
Valine	12,000
Tryptophan	18
Cofactors	
Cyanocobalamin	0.09
N-methylmesoporphyrin IX	~14
N-methylmesoporphyrin IV (DNA)	~0.5
Flavin	0.5
NAD	2.5
RMP-botin	2
Antibiotics	
Tobramycin (aminoglycoside)	0.0008
Neomycin (aminoglycoside)	0.1
Lividomycin (aminoglycoside)	<0.2
Kanamycin (aminoglycoside)	<0.2
Streptomycin (aminocyclitol)	~1
Viomycin (basic peptide)	12
Chloramphenicol (small, neutral)	2.1
Transition state analogs	
Diels-Alder reaction	3,500
Bridged biphenyl isomerization	542
Other	
Dopamine	2.8
Peptide (substance P)	0.19
Divalent metals	~1

In vitro selection and directed evolution of peptides and proteins

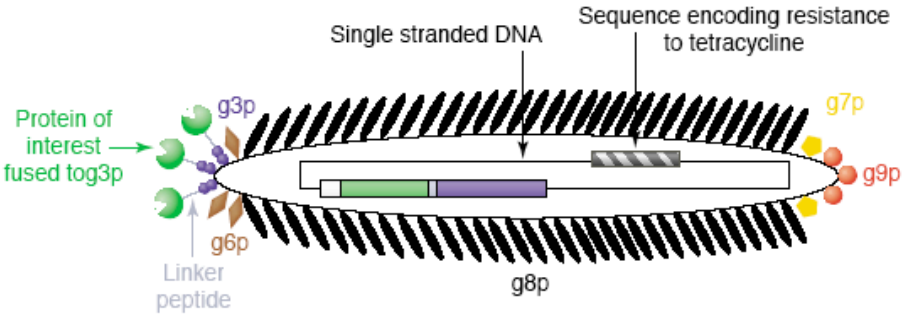


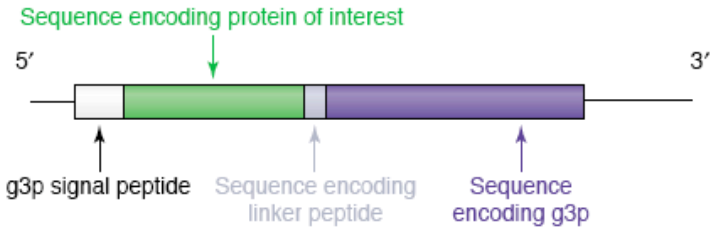
- Goal: find protein/peptide with desired function (binding, catalyst, stability, etc.)
- *In vivo* vs *In vitro* selection.
- Rational design vs. Library methods.
- Combinatorial peptide libraries.
  - Geysen. Discontinuous epitope mapping. Solid phase peptide synthesis. Birth of combinatorial chemistry.
- Phage display.
- mRNA fusion. Ribosome stalling.
- Antibody engineering and evolution.
- Primerless PCR. DNA shuffling. Sexual PCR.

Searching for Peptide Ligands with an Epitope Library

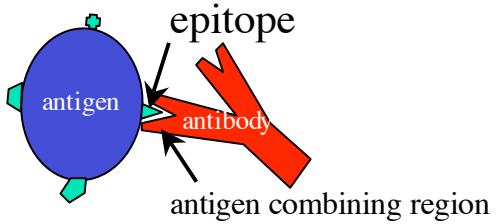
JAMIE K. SCOTT AND GEORGE P. SMITH*

Science, New Series, Vol. 249, No. 4967 (Jul. 27, 1990), 386-390.
Science, New Series, Vol. 228, No. 4705 (Jun. 14, 1985), 1315-1317.

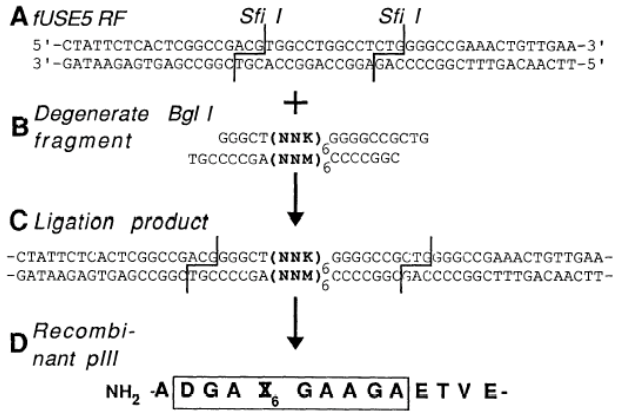




Filamentous bacteriophage (M13, fd)



Cloning into p3 (pIII) coat protein



NNK codons: N=ACTG, K=GT. Encodes 32 triples, all 20 amino acids and only one stop codon.

Scott & Smith

Science, New Series, Vol. 249, No. 4967 (Jul. 27, 1990), 386-390.

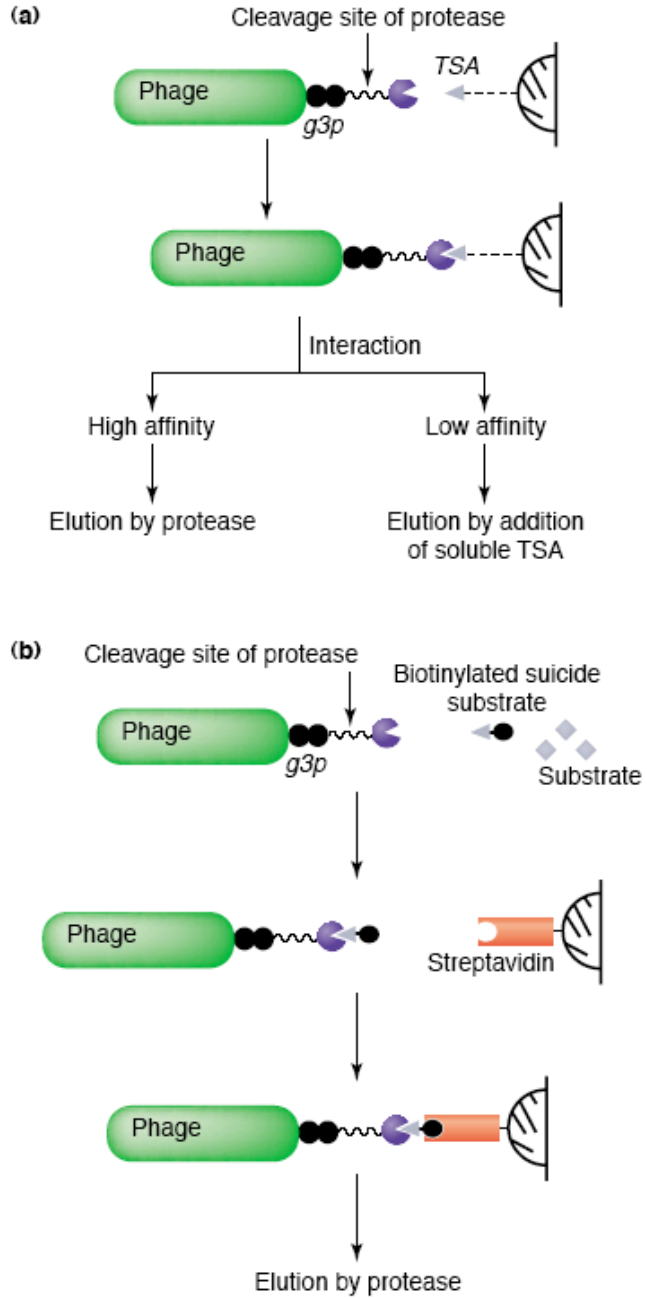
- 64 million possible hexapeptides
- Library contained 1.3 x 10^{14} phage from 2 x 10^{8} clones

Displayed peptides		Number of clones isolated with		Binding (OD × 10^3)			
				MAb M33		MAb A2	
No.	Sequence	MAb M33	MAb A2	Fab	IgG	Fab	IgG
MHr	DFLEKI			72	352	136	235
	Third round with 0.1 nM MAb						
1	----WL*	16		203	551	229	233
2	----ML	1	2	92	558	269	236
3	---AWL		14	14	242	139	273
	Second round with 0.1 nM MAb						
4	----R-	1		94	318	118	290
5	CRFVWC	2		65	331	−4	−5
6	CE---C	1		18	294	−1	123
7	CR---C	1		21	205	4	127
8	--M-WL*	1		35	288	1	125
9	---VQL†	1		5	127	23	147
10	---AIV†		4	15	43	97	142
11	----Y-		2	67	157	9	253
12	----IL		1	100	431	159	308
13	----IV		1	37	168	195	293
	Second round with 10 nM MAb						
14	----QL	1					
15	---HF-	1		6	108	−3	−2
16	AWERRG	1		4	8	−3	9
17	--F-I-		1	29	244	77	224
18	---MLV		1	3	18	17	50
19	Q-VFCW		1	3	16	−5	28

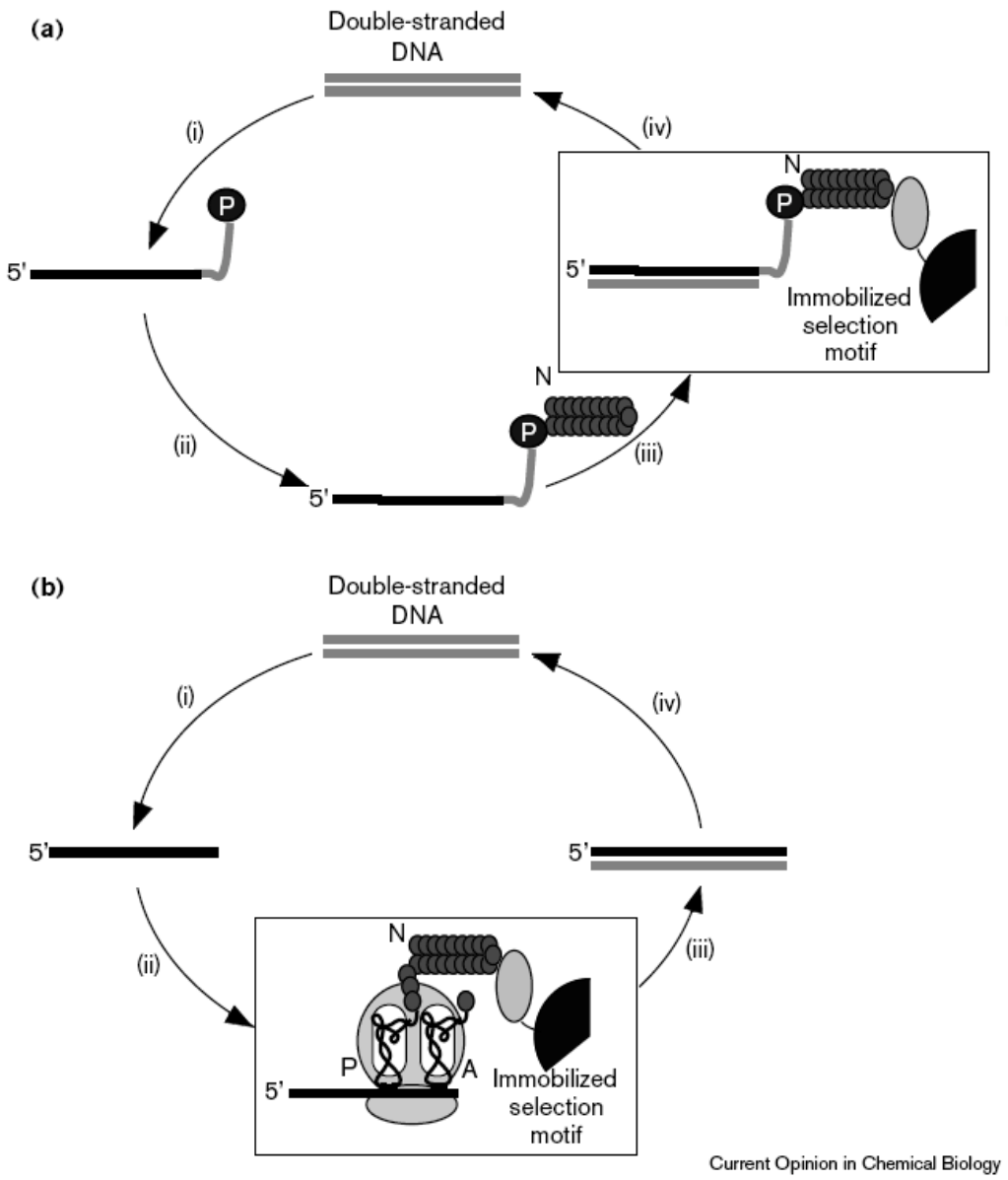
*Four clones with the peptide 1 sequence and one with the peptide 8 sequence were also isolated from the second round of AP with 10 nM MAb M33. †One clone each with the peptide 9 and 10 sequences were also isolated from the second round of AP with 10 nM MAb A2.

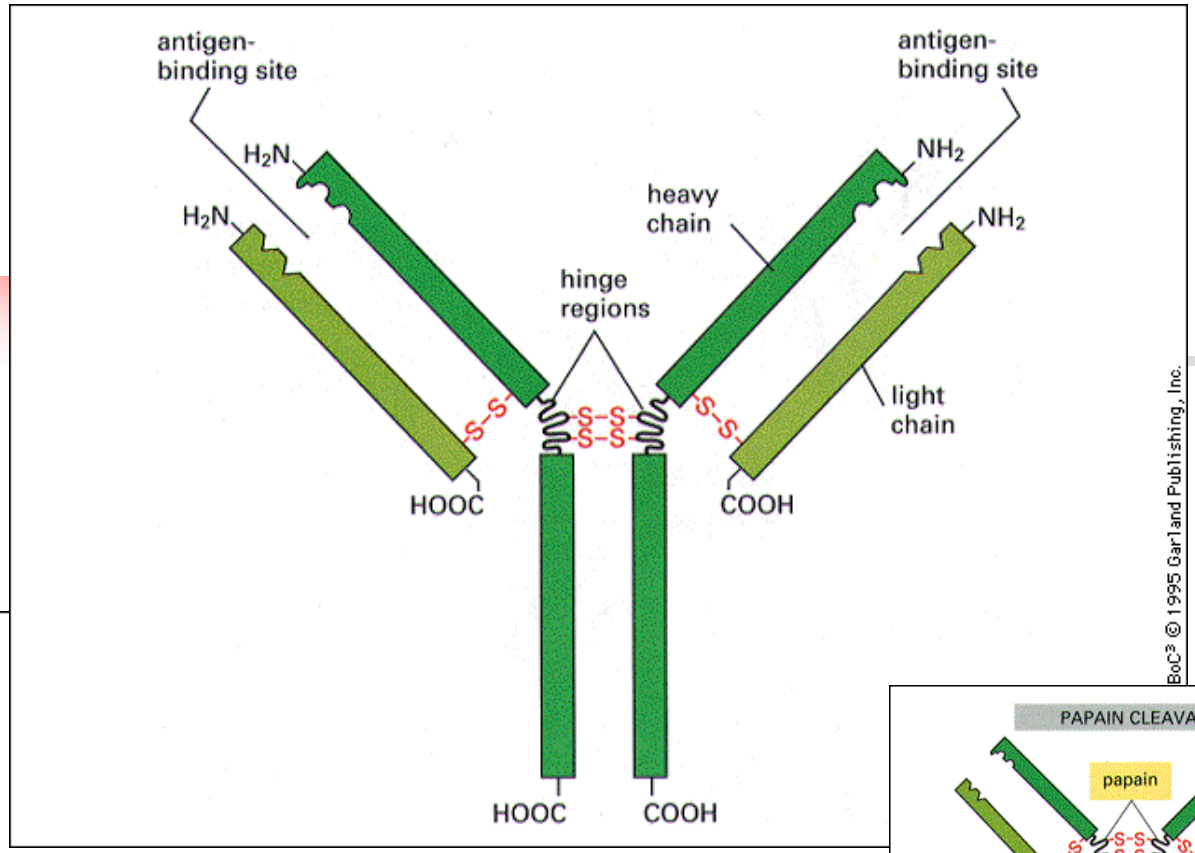
Phage display

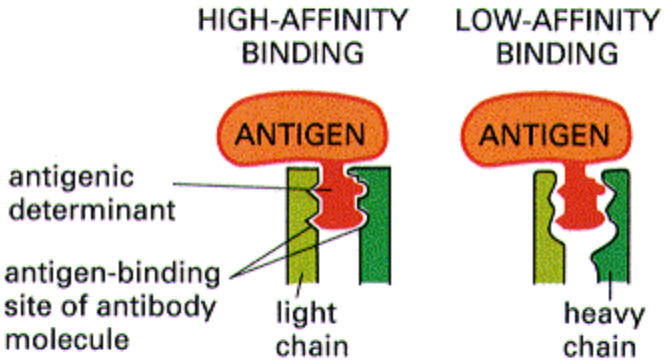
- Enzyme selection.
- Note: hidden *in vivo* selection too...

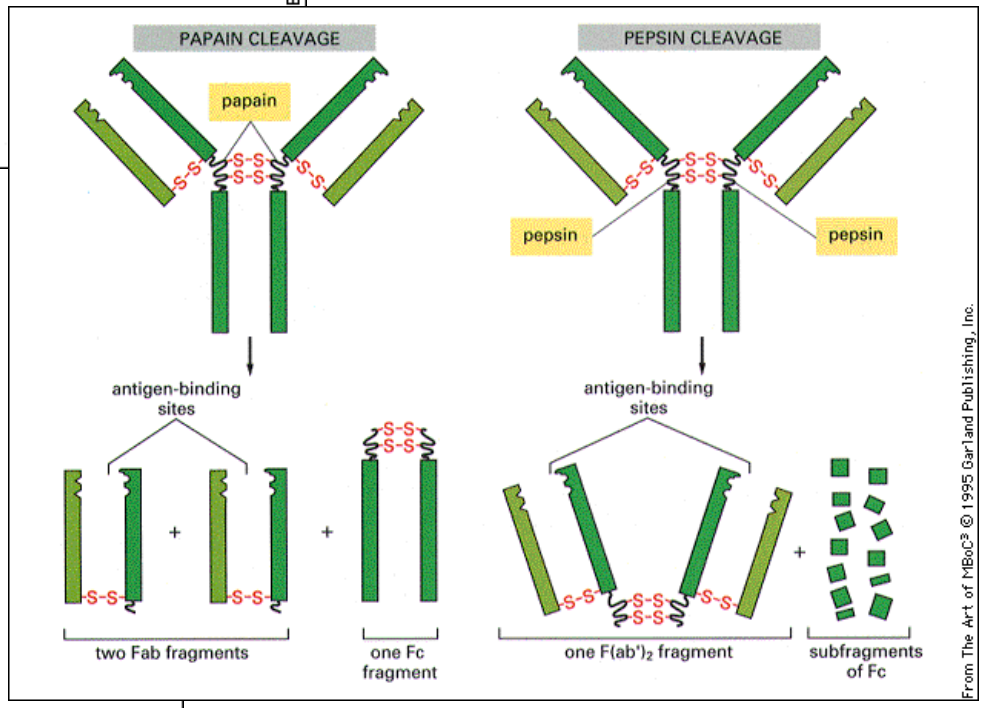


- mRNA-protein fusion
- Ribosome display
- Other. Cro?

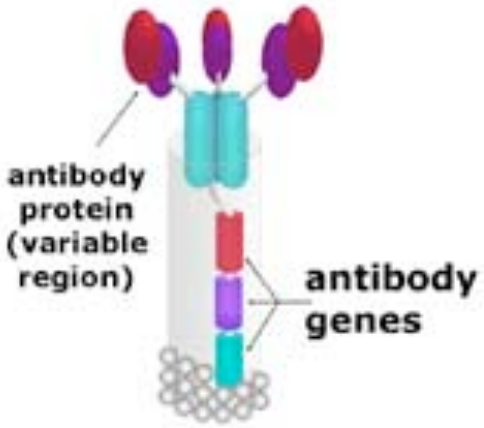


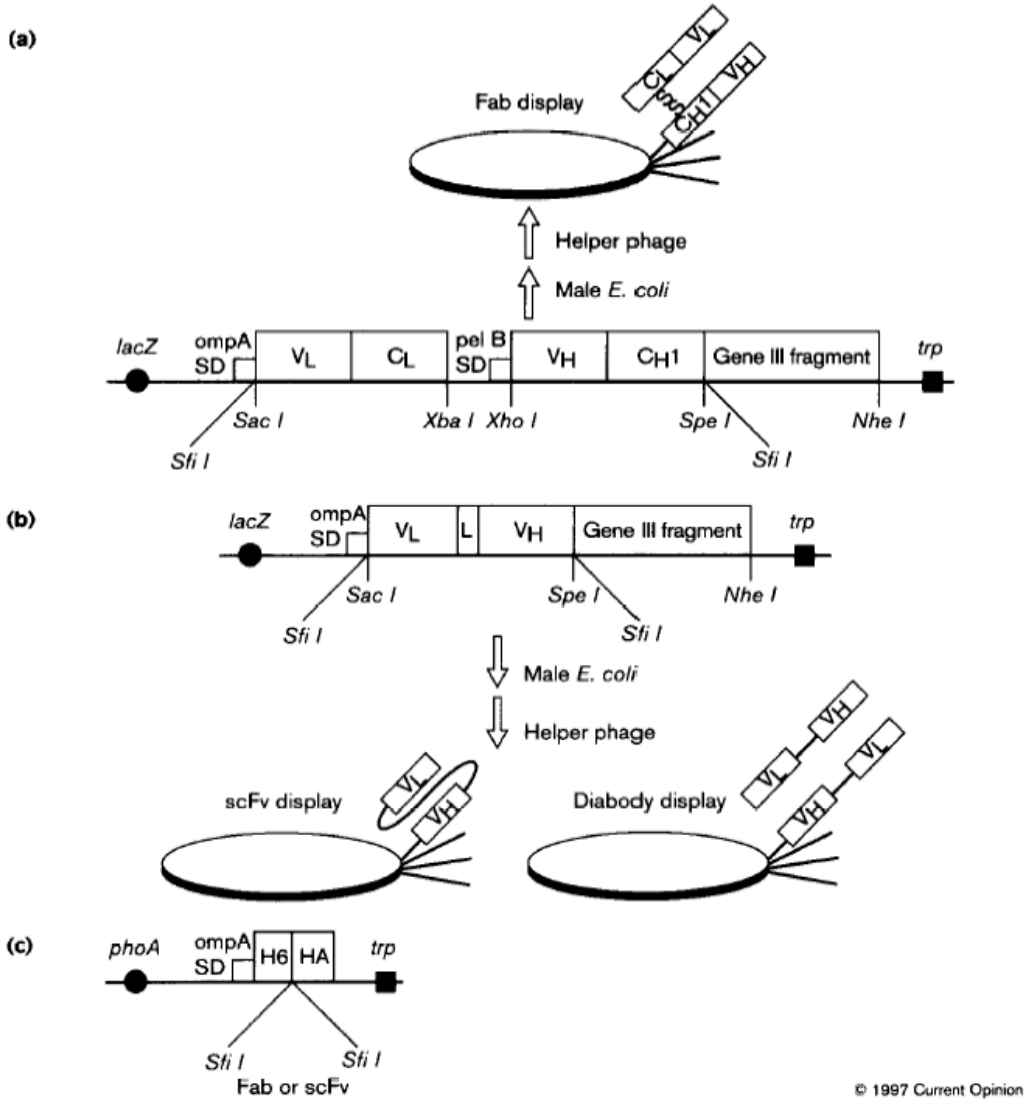






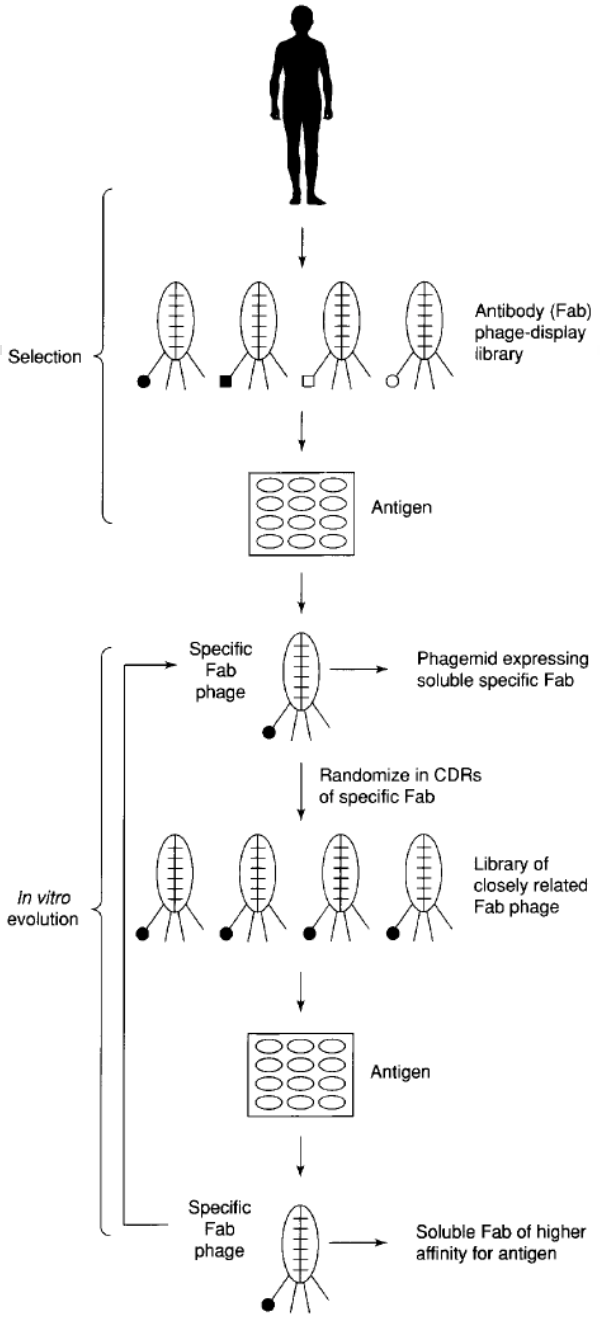
Phage display of Ab libraries





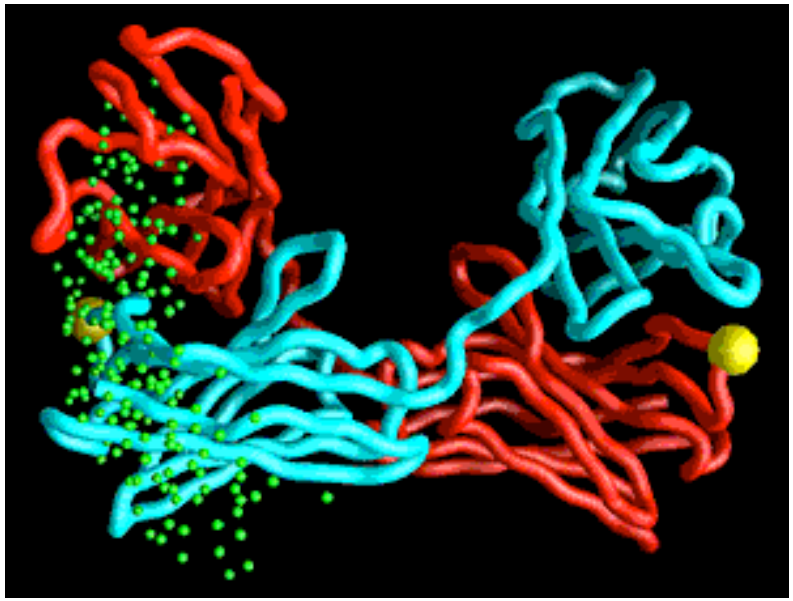
Artificial immune maturation

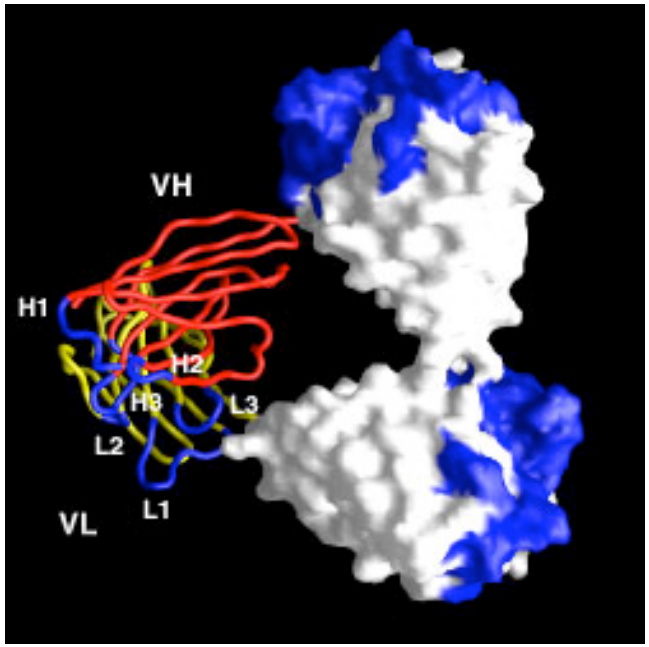
- Gene library from human, animal, or synthetic source.
- Maturation:
  - Saturation mutagenesis of CDRs.
  - Chain shuffling.
  - Growth in mutagenic *E. coli* strains.



Diabodies and triabodies

- In vitro selection/evolution of bifunctional antibody fragments.
- Universal protein-based adapter/address system.





DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution

WILLEM P. C. STEMMER

Affymax Research Institute, 4001 Miranda Avenue, Palo Alto, CA 94304

Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 10747–10751, October 1994

http://www.maxygen.com/science-pub.php

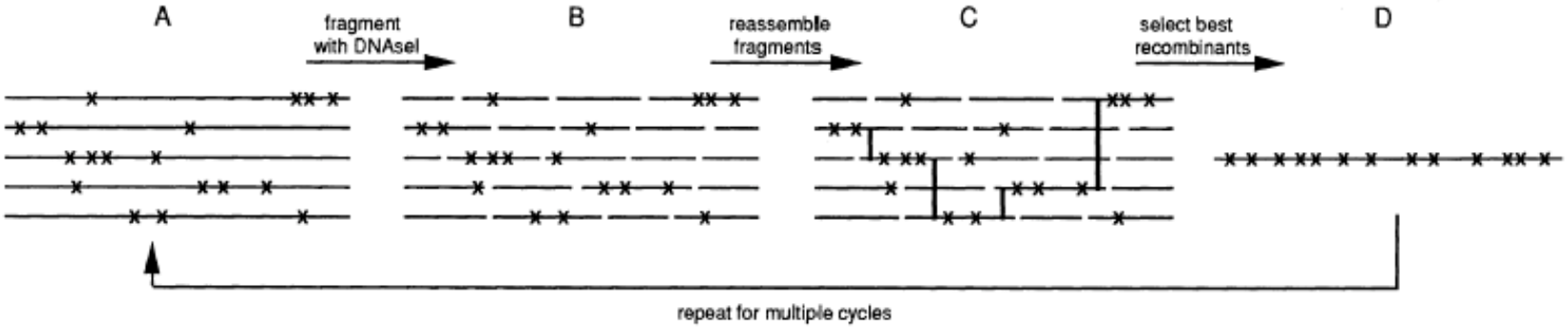


FIG. 1. (*A*) A pool of homologous genes with different point mutations is fragmented with DNase I. (*B*) For simplicity, all mutations shown are considered beneficial and additive. (*C*) Reassembly of the random fragments into full-length genes results in frequent template switching and recombination. A recombinant gene containing the four crossovers (thick lines) can be selected from the library of recombinants based on its improved function. (*D*) Selected pool of improved recombinants provides the starting point for another round of mutation and recombination. The recombination process alone causes a low level of point mutations but, if desired, additional mutations could be introduced by error-prone PCR or UV mutagenesis of the pool of genes.

- Gene reassembly
- Whole plasmid reassembly

Table 1. Mutations introduced by mutagenic shuffling

Transition	Frequency	Transversion	Frequency
G → A	6	A → T	1
A → G	4	A → C	2
C → T	7	C → A	1
T → C	3	C → G	0
		G → C	3
		G → T	2
		T → A	1
		T → G	2

A total of 4437 bases of shuffled *lacZ* DNA were sequenced. The mutation rate was 0.7%; 11/12 types of base substitutions were found, but there were no frameshifts.

DNA shuffling of a family of genes from diverse species accelerates directed evolution

NATURE | VOL 391 | 15 JANUARY 1998

Andreas Crameri, Sun-Ai Raillard, Ericka Bermudez & Willem P. C. Stemmer

Maxygen Inc., 3410 Central Expressway, Santa Clara, California 95051, USA

- Can natural diversity accelerate the evolution process?
- Goal: obtain moxalactamase activity from four cephalosporinase genes. DNA shuffling within genes or between genes...

Figure 1 Phylogenetic tree of the four cephalosporinase genes. The numbers on the vertical bars indicate the percentage of DNA sequence similarity.

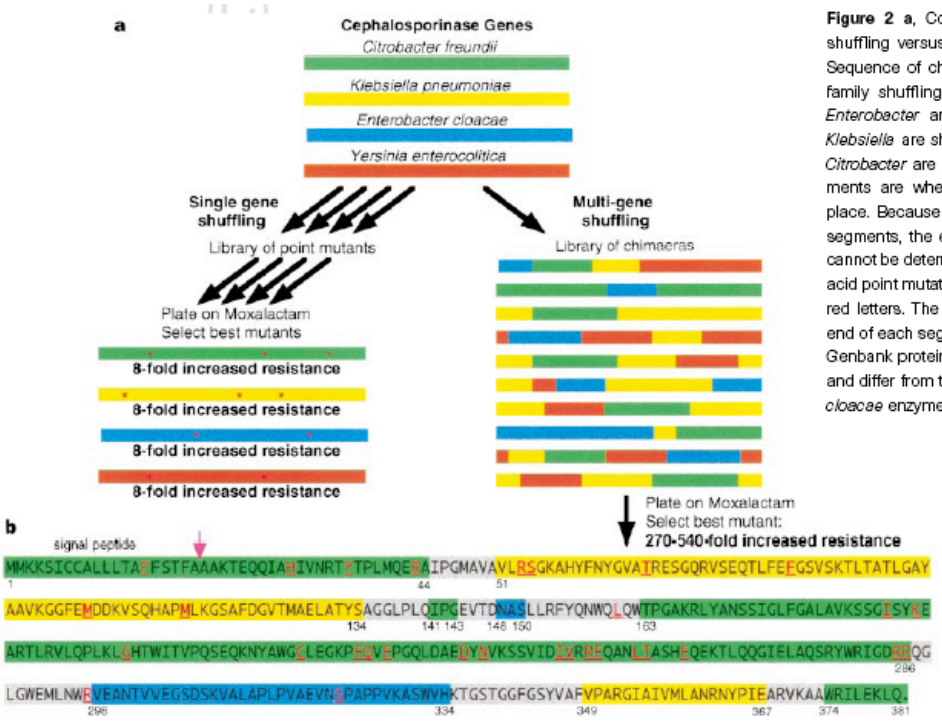
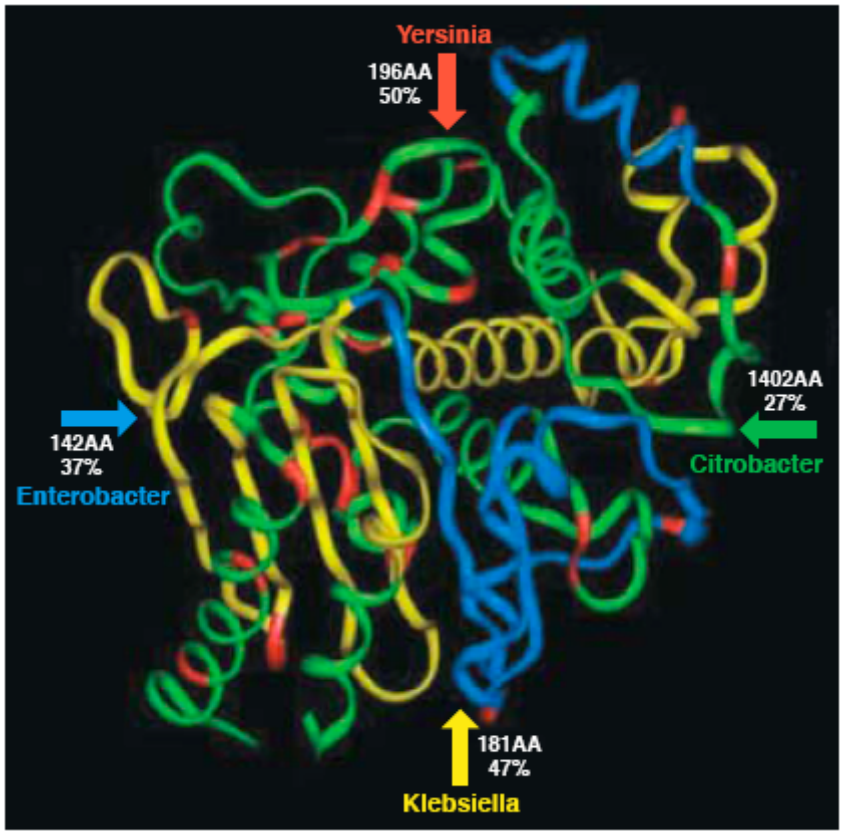


Figure 2 **a**, Comparison of single sequence shuffling versus sequence family shuffling. **b**, Sequence of chimaeric mutant A obtained by family shuffling. The segments derived from *Enterobacter* are shown in blue, those from *Klebsiella* are shown in yellow, and those from *Citrobacter* are shown in green. The grey segments are where the crossovers have taken place. Because of DNA homology in the grey segments, the exact location of the crossover cannot be determined more exactly. The amino-acid point mutations are shown with underlined red letters. The numbers at the beginning and end of each segment are the numbers from the Genbank protein files of the wild-type enzymes and differ from those used for the *Enterobacter cloacae* enzyme[14].



Stemmer

NATURE | VOL 391 | 15 JANUARY 1998

Stemmer

NATURE | VOL 391 | 15 JANUARY 1998

- Can natural diversity accelerate the evolution process?
- Goal: obtain moxalactamase activity from four cephalosporinase genes. DNA shuffling within genes or between genes...
- One cycle: ~8-fold improvements from separately evolved genes.
- One cycle: 270- to 540-fold improvement from the four genes shuffled together.
- Best clone contained 8 segments from 3 of the 4 genes & 33 amino-acid point mutations.

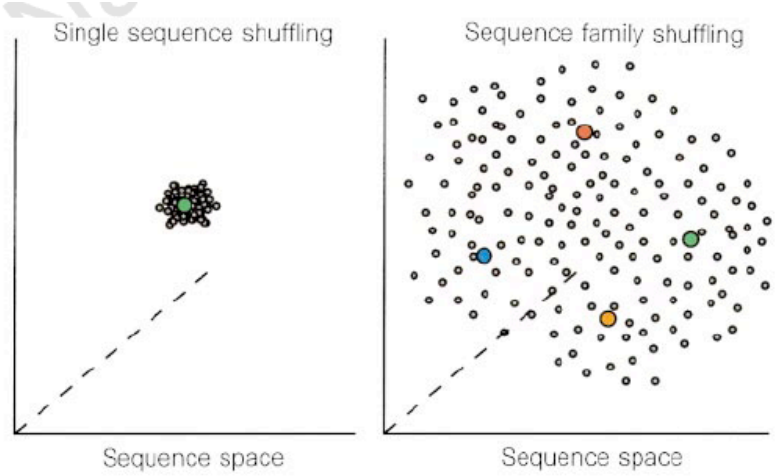


Figure 4 Searching sequence space by family shuffling versus by single sequence shuffling. Single sequence shuffling yields clones with a few point mutations and the library members are typically 97–99% identical. Family shuffling causes sequence block exchange which yields chimaeras that have greater sequence divergence. At equal library size, the increased sequence diversity of the chimaeric library results in sparse sampling of a much greater area of sequence space, allowing more promising areas to be found and subsequently explored at increased sampling density.

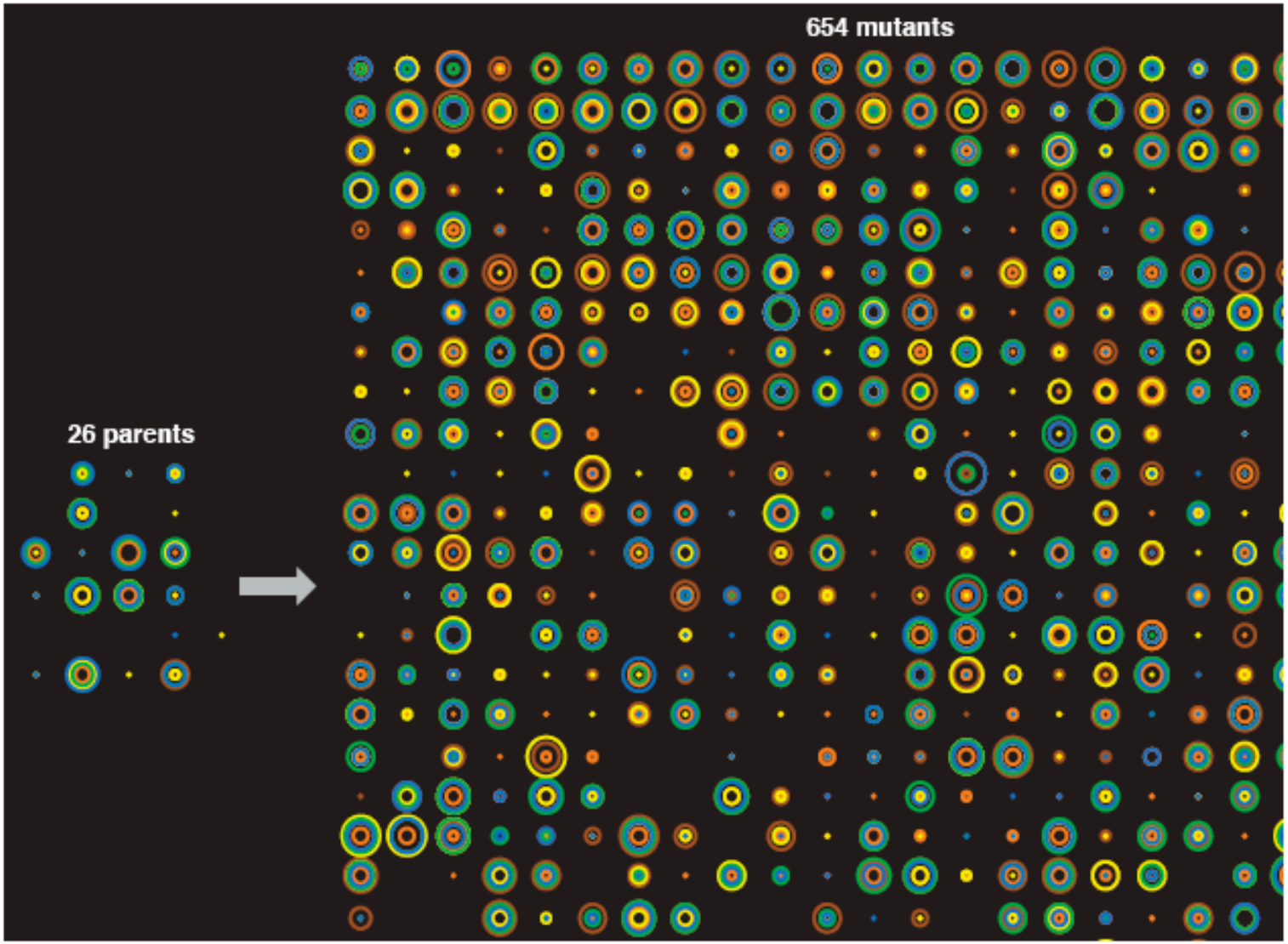


Figure 6. DNA family breeding of the subtilisin gene illustrates the potential for rapid improvement in protein function. Subtilisin is a protein-degrading enzyme used in laundry detergent and the most highly engineered enzyme in existence. Twenty-six subtilisin genes from various *Bacillus* microbes were shuffled to yield 654 progeny that were screened for enzyme activity at five commercially relevant conditions. Each clone is shown as a set of concentric circles, with the size of the circle indicating the activity of the enzyme and each assay condition represented by a different color. Several progeny showed simultaneous improvement in multiple properties over the best parental sequences.

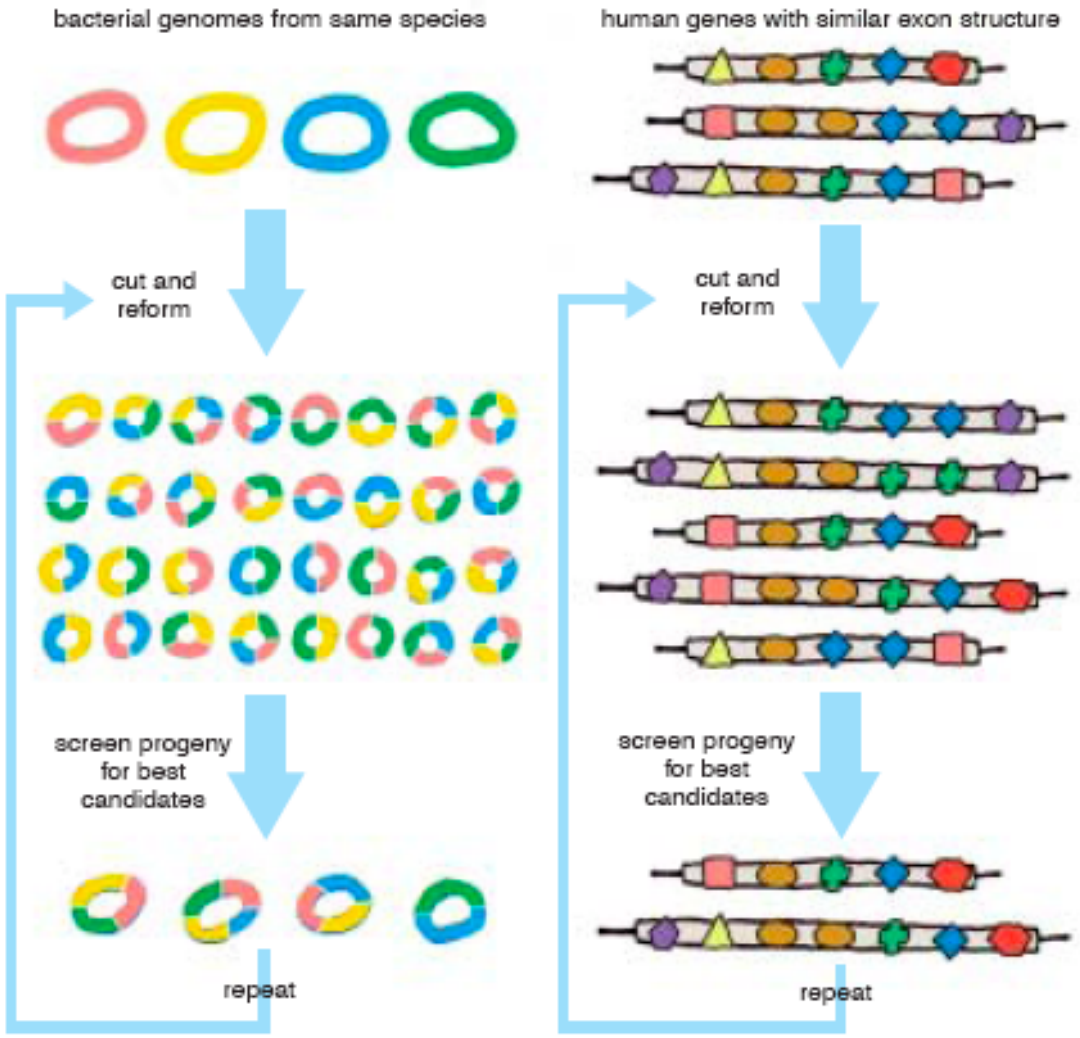


Figure 7. Genome shuffling *(left)* and exon shuffling *(right)* are two other applications of the DNA breeding technique. Genome shuffling is modeled after the natural evolution of prokaryotic organisms, whereas exon shuffling mimics eukaryotic protein evolution. In both cases the progeny are more likely to have useful properties because the breeding step starts with "proven diversity"—natural selection having already removed most deleterious point mutants.

Belcher

(2000) *Nature* **405**, 665-668

Selection of peptides with semiconductor binding specificity for directed nanocrystal assembly

Sandra R. Whaley*, D. S. English*, Evelyn L. Hu†, Paul F. Barbara*‡ & Angela M. Belcher*‡

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

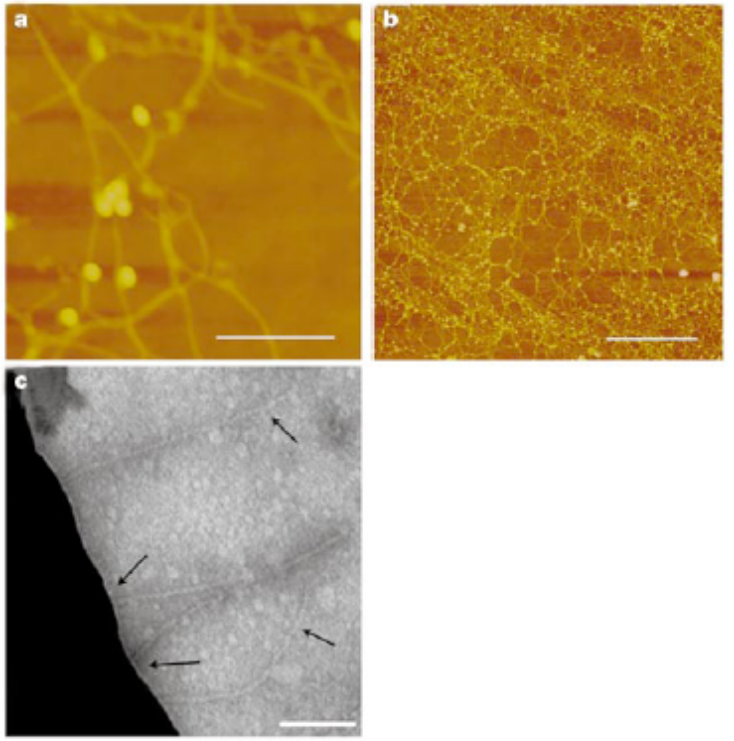


Figure 3 AFM and TEM analysis of peptide–semiconductor recognition. **a**,**b**, AFM images of G1-3 phage bound to an InP(100) substrate. **a**, Individual phage and their attached Au nanoparticles. Scale bar, 250 nm. **b**, Image showing the uniformity of phage coverage on the InP surface. Scale bar, 2.5 μm. **c**, TEM image of G1-3 phage recognition of GaAs. Individual phage particles are indicated with arrows. Scale bar, 500 nm.

Belcher

(2000) *Nature* **405**, 665-668

- GaAs (100), GaAs(111), InP, GaN, ZnS, CdS, Fe_3O_4, $CaCO_3$, etc.