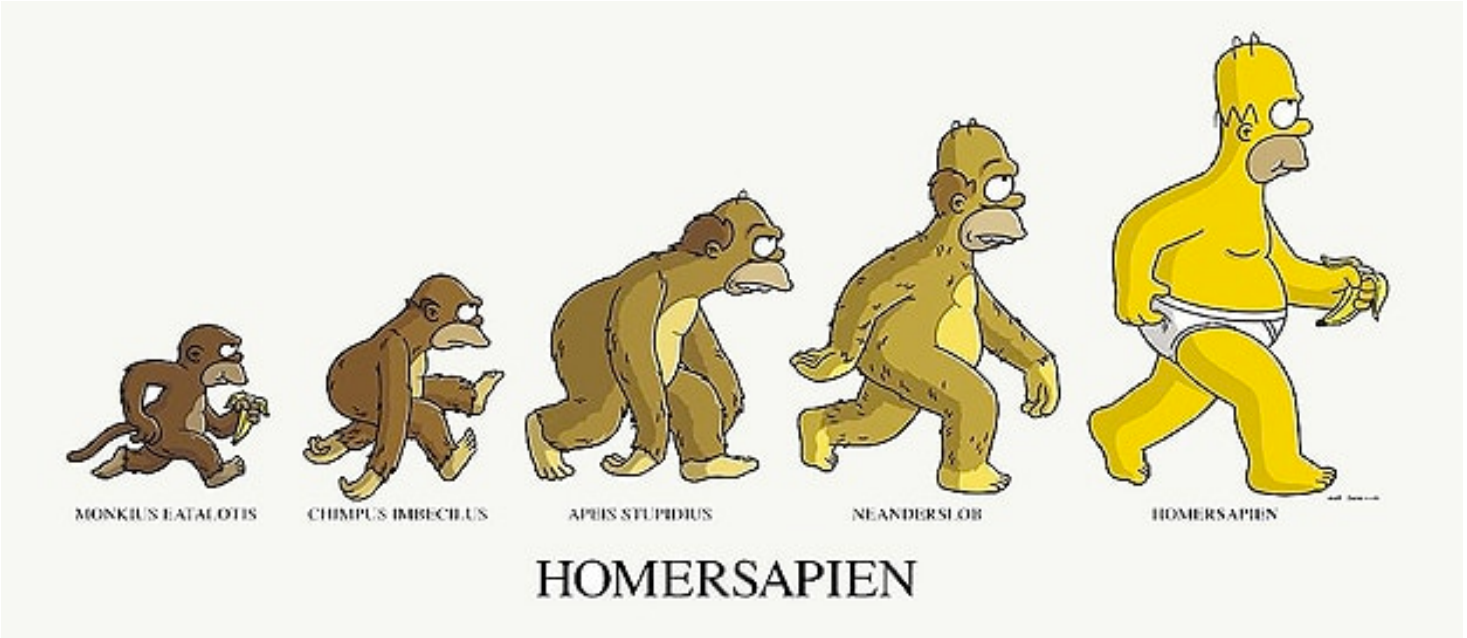
Applications of Directed Evolution in Organic Synthesis



Steven Loskot

Lit Talk

Dec. 7th 2018

Outline

- Background
  - Timeline

Outline

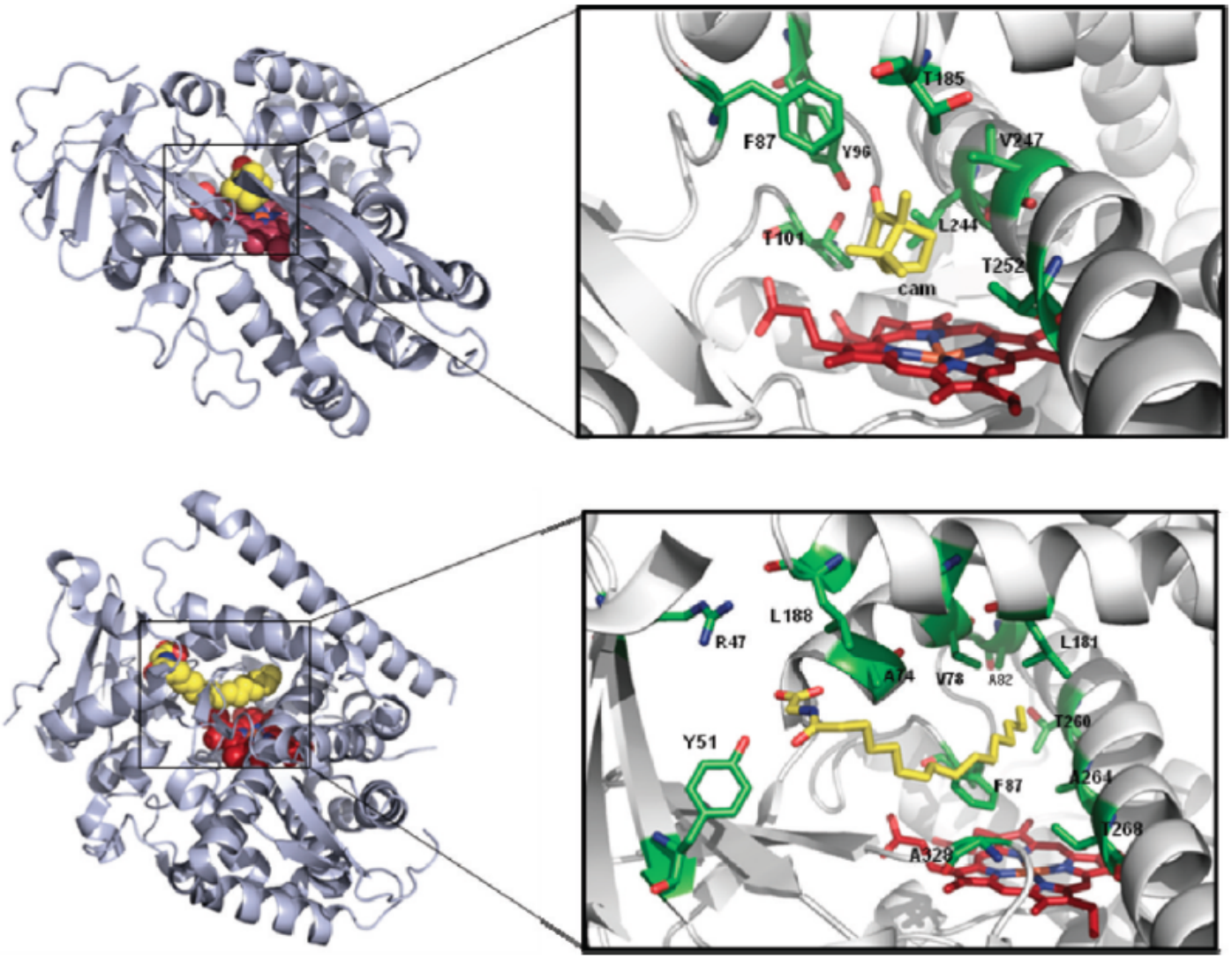
- Background
  - Timeline
- Directed Evolution Breakdown
  - Mutagenesis
  - Expression
  - Screening

Outline

- Background
  - Timeline
- Directed Evolution Breakdown
  - Mutagenesis
  - Expression
  - Screening
- Examples
  - Pharmaceutical/ Industrial examples
  - New to Nature Reactions

Background

Enzymes are proteins that posess complex structures

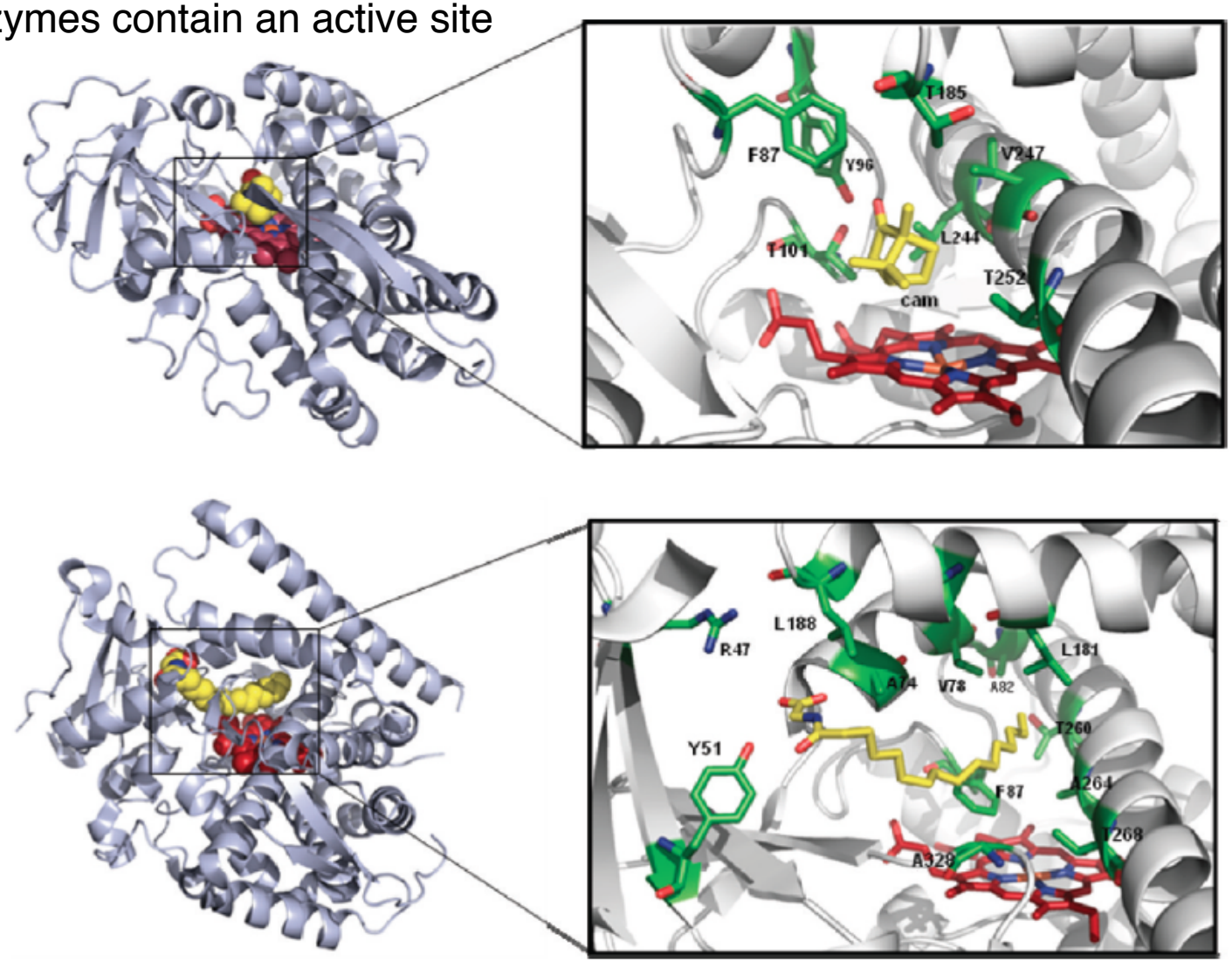


Fasan, *ACS Catal.* **2012**, *2*, 647–666.

Background

Enzymes are proteins that posess complex structures

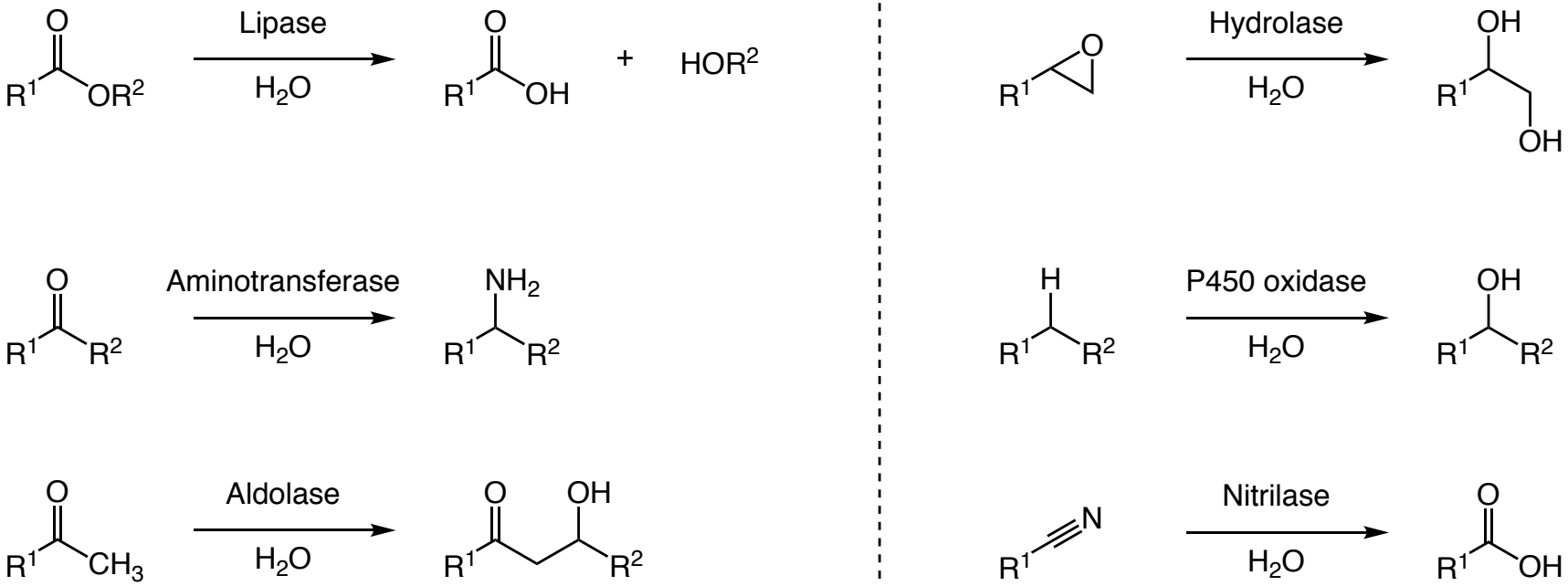
All Enzymes contain an active site



Fasan, *ACS Catal.* **2012**, *2*, 647–666.

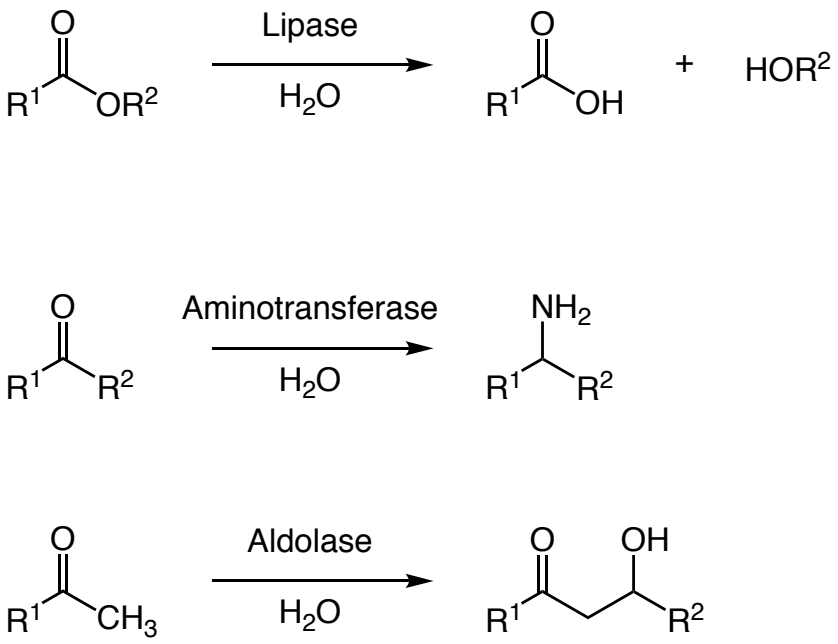
Background

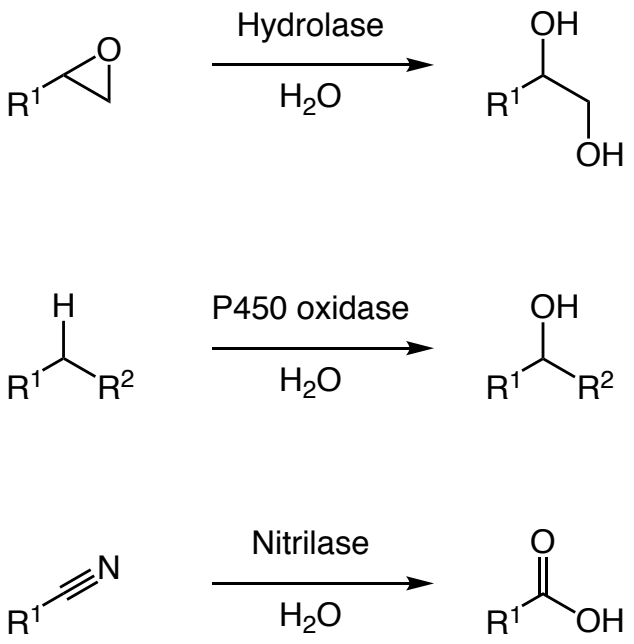
Common natural enzyme catalyzed reactions



Background

Common natural enzyme catalyzed reactions

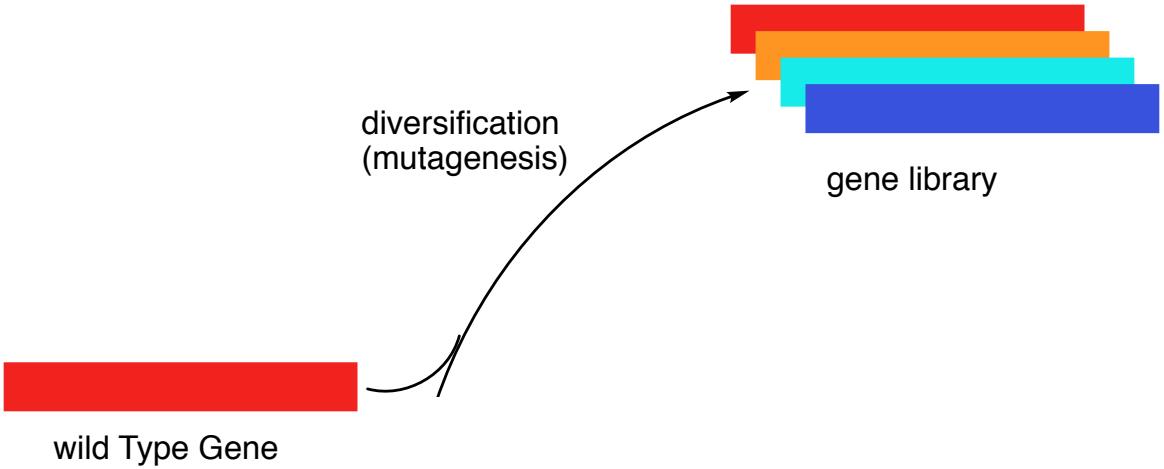




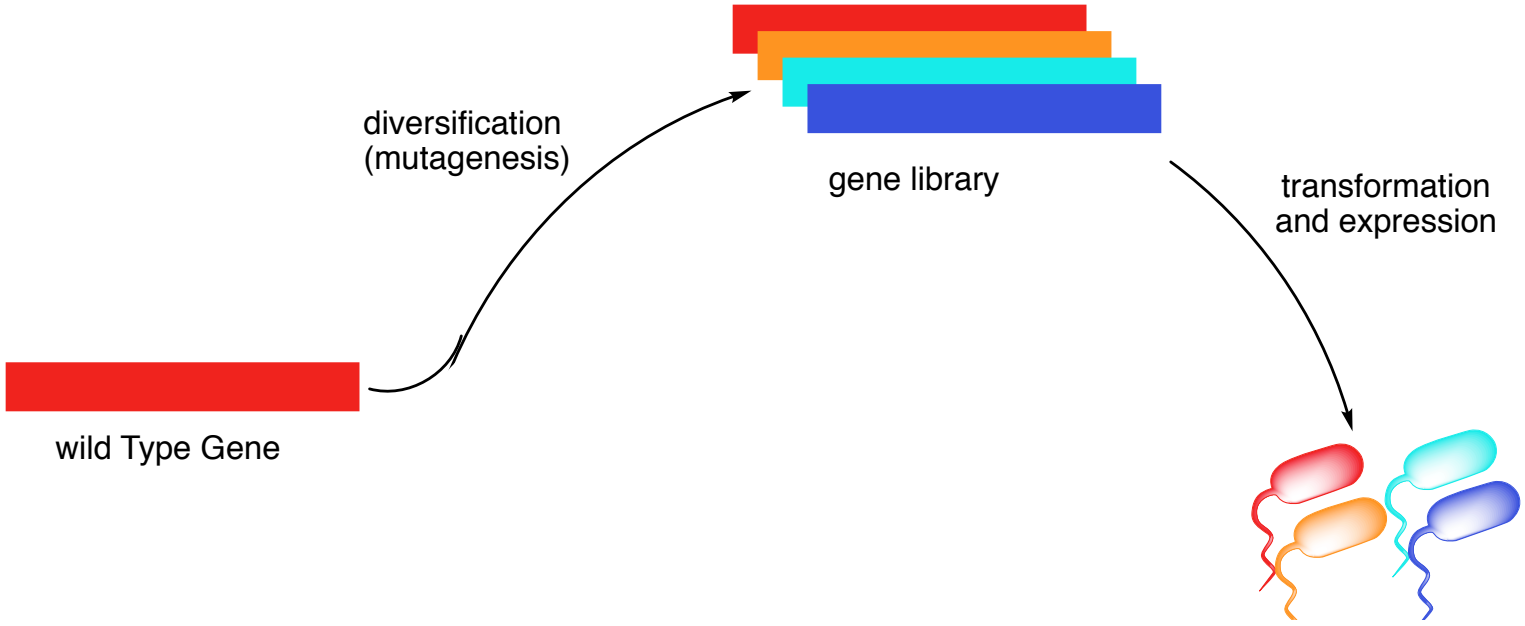
Limitations:

- Few natural enzymes show wide levels of promiscuity
- Elevated temperatures results in denaturation
- Water is the only viable solvent
- Substrate tolerance is limited by steric and electronic fators
- high EE's are only achievable from native substrate
- Millions of years of evolution required to reach this state

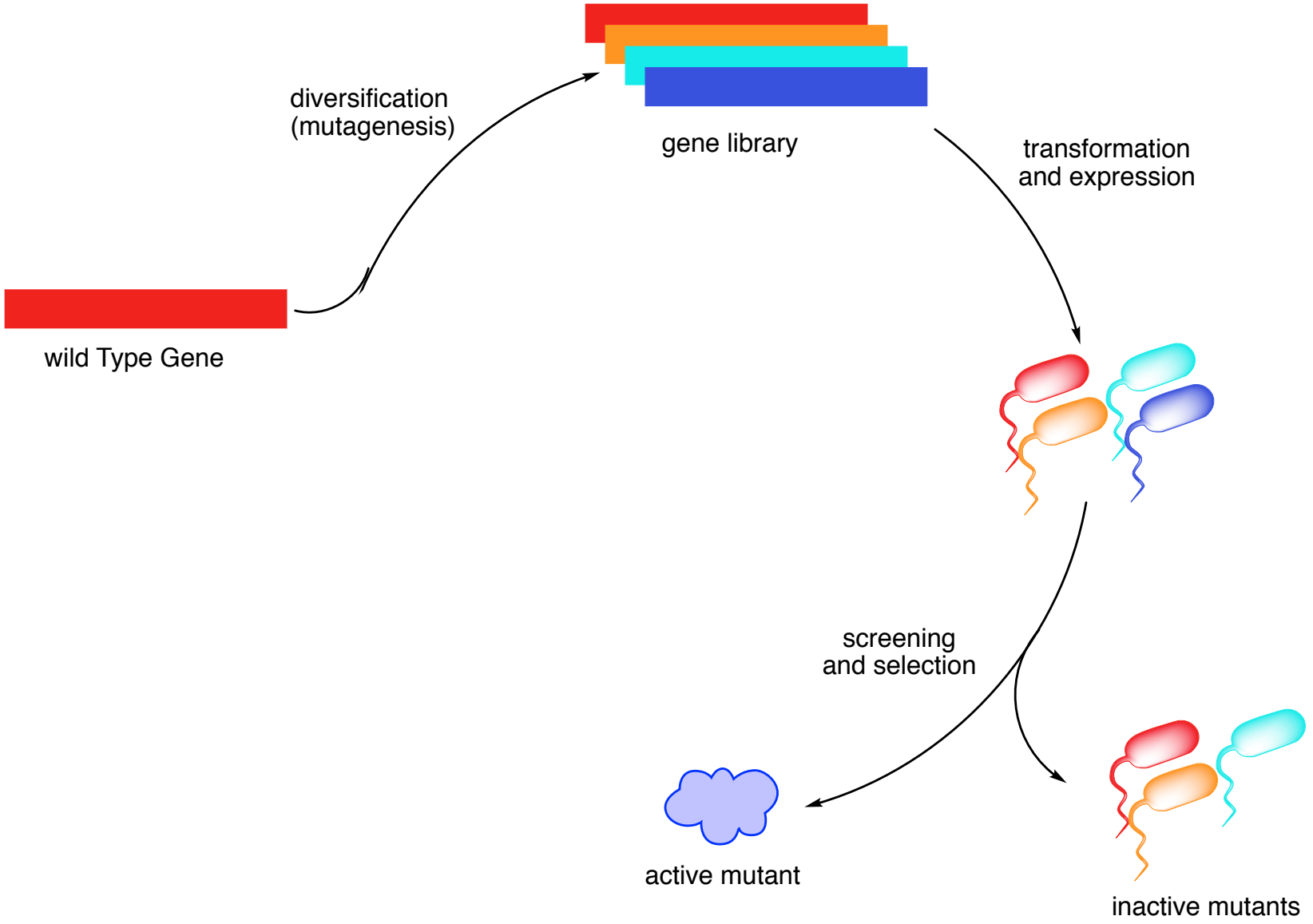
Directed Evolution



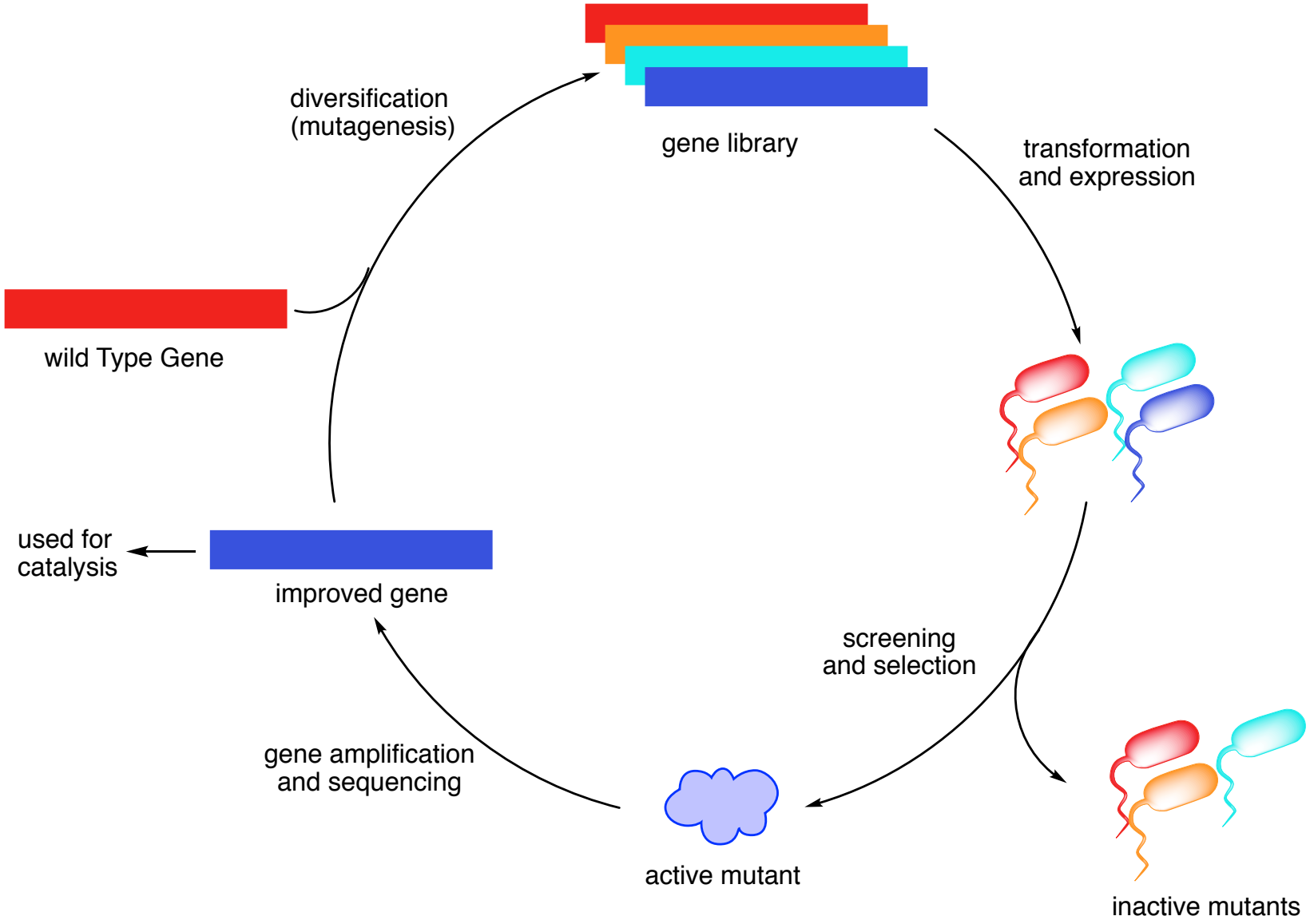
Directed Evolution



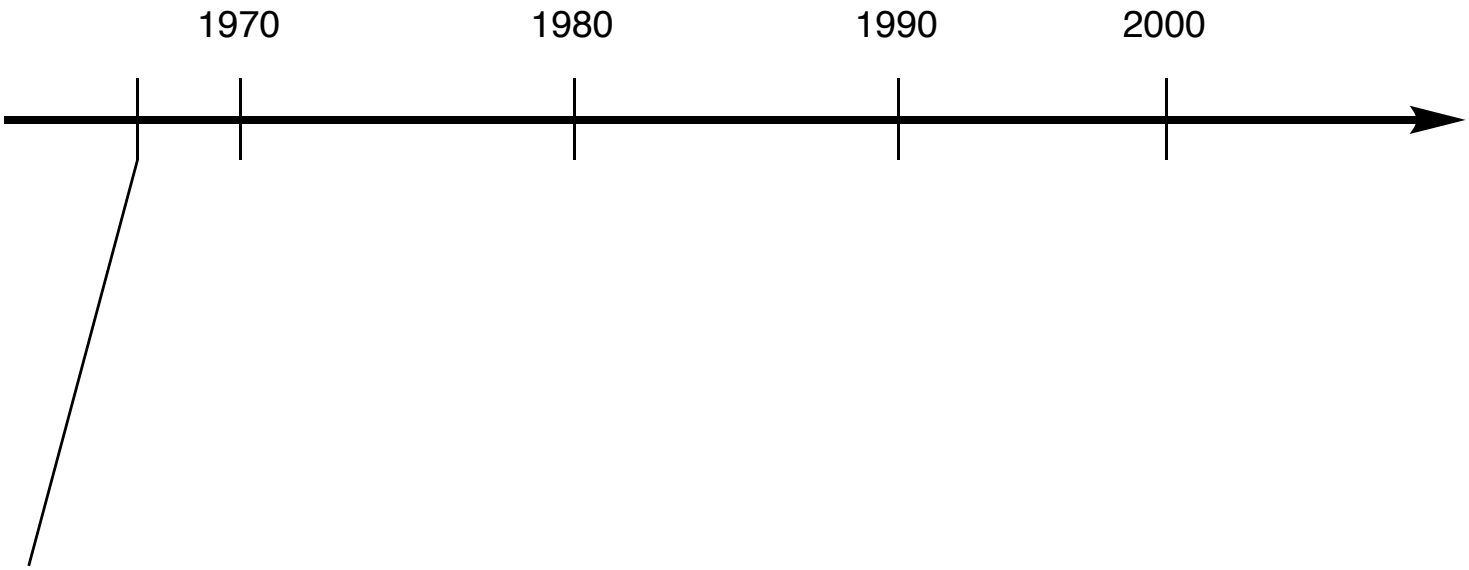
Directed Evolution



Directed Evolution



Timeline

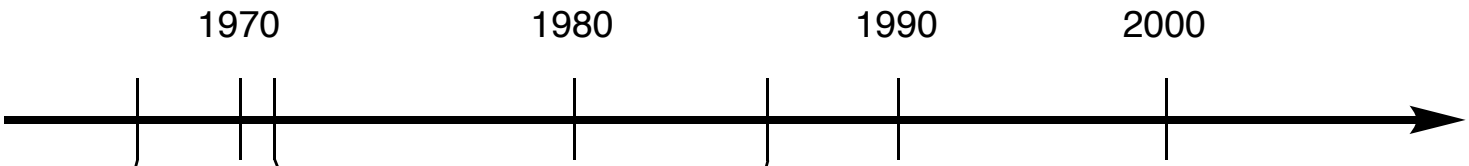

1970
1980
1990
2000
1967
S. Spiegelman reports an in vitro darwinian experiment using self-replication RNA
Spiegelman, *Proc. Natl. Acad. Sci. U.S.A.* **1967,** *58*, 217–224.

Timeline

1971
M. Eigen reports a thoery of evolution at the molecular level
Eigen, *Naturwissenschaften,* **1971,** 58, 465–523.

1967
S. Spiegelman reports an in vitro darwinian experiment using self-replication RNA
Spiegelman, *Proc. Natl. Acad. Sci. U.S.A.* **1967,** *58*, 217–224.

Timeline

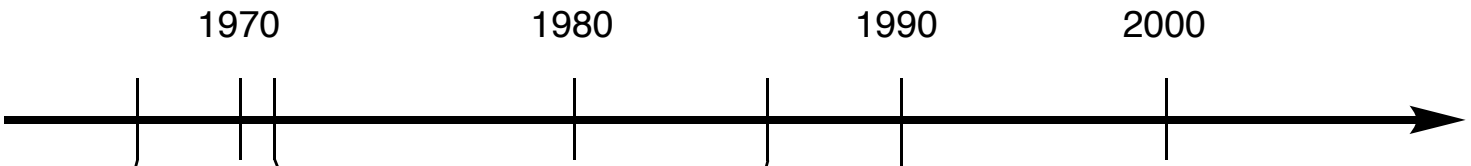


1971
M. Eigen reports a thoery of evolution at the molecular level
Eigen, *Naturwissenschaften,* **1971,** 58, 465–523.

1967
S. Spiegelman reports an in vitro darwinian experiment using self-replication RNA
Spiegelman, *Proc. Natl. Acad. Sci. U.S.A.* **1967,** *58*, 217–224.

1986
Hageman succeed in the first example of directed evolution by in vivo rational mutagenesis
Hageman, Proc. Natl. Acad. Sci. U.S.A. **1986,** *83*, 576–580.

Timeline



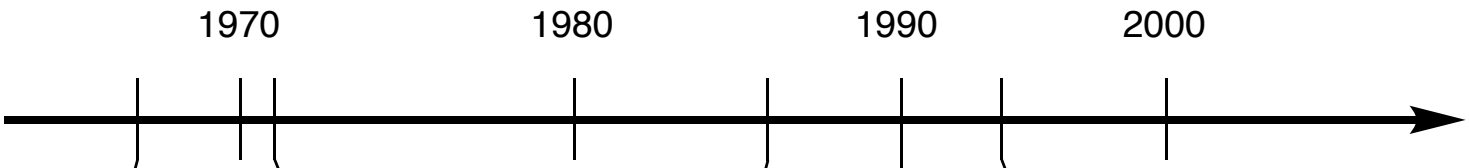
1967
S. Spiegelman reports an in vitro darwinian experiment using self-replication RNA
Spiegelman, *Proc. Natl. Acad. Sci. U.S.A.* **1967,** *58*, 217–224.

1971
M. Eigen reports a thoery of evolution at the molecular level
Eigen, *Naturwissenschaften,* **1971,** 58, 465–523.

1986
Hageman succeed in the first example of directed evolution by in vivo rational mutagenesis
Hageman, Proc. Natl. Acad. Sci. U.S.A. **1986,** *83*, 576–580.

J. R. Knowles reports the first true random mutagenesis by using the full sequence space
Knowles, *Proc. Natl. Acad. Sci. U.S.A.* **1990***, 87*, 696–700

Timeline



1967
S. Spiegelman reports an in vitro darwinian experiment using self-replication RNA
Spiegelman, *Proc. Natl. Acad. Sci. U.S.A.* **1967,** *58*, 217–224.

1971
M. Eigen reports a thoery of evolution at the molecular level
Eigen, *Naturwissenschaften,* **1971,** 58, 465–523.

1986
Hageman succeed in the first example of directed evolution by in vivo rational mutagenesis
Hageman, Proc. Natl. Acad. Sci. U.S.A. **1986,** *83*, 576–580.

J. R. Knowles reports the first true random mutagenesis by using the full sequence space
Knowles, *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 696–700

1993
F. H. Arnold reports the first use of iterative radnom mutagenesis
Arnold, *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5618–5622.

Timeline

1970 1980 1990 2000

1967
S. Spiegelman reports an in vitro darwinian experiment using self-replication RNA
Spiegelman, *Proc. Natl. Acad. Sci. U.S.A.* **1967,** *58*, 217–224.

1971
M. Eigen reports a thoery of evolution at the molecular level
Eigen, *Naturwissenschaften,* **1971,** 58, 465–523.

1986
Hageman succeed in the first example of directed evolution by in vivo rational mutagenesis
Hageman, Proc. Natl. Acad. Sci. U.S.A. **1986,** *83*, 576–580.

J. R. Knowles reports the first true random mutagenesis by using the full sequence space
Knowles, *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 696–700

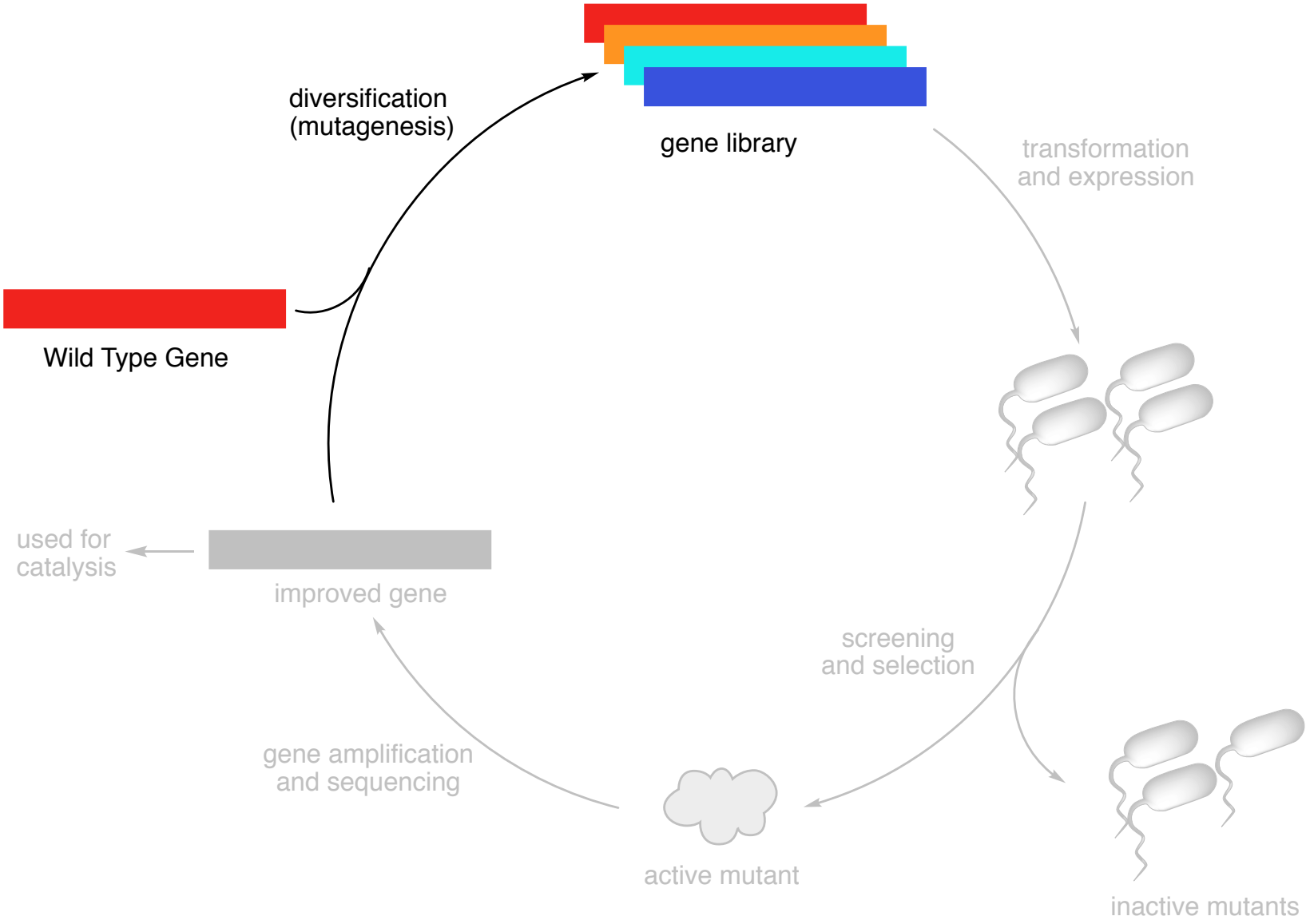
1993
F. H. Arnold reports the first use of iterative radnom mutagenesis
Arnold, *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5618–5622.

1997
M. T. Reetz and K.-E. Jeager use directed evolution to develop an enantioselective biocatalysis
Reetz, *Angew. Chem. Int. Ed.* **1997**, *36*, 2830–2832.

Outline

- Background
  - Timeline
- Directed Evolution Breakdown
  - Mutagenesis
  - Expression
  - Screening
- Examples
  - Pharmaceutical/ Industrial examples
  - New to Nature Reactions

Directed Evolution



Mutagenesis

In Vivo:

- Chemical mutagenesis
- Mutator Strains

In Vitro:

- Random mutagenesis
- Focused mutagenesis
- Recombinative mutagenesis

Random In Vivo Mutagenesis

Chemical mutagenesis:

– alkylating agents, UV light, nitrous acid, surrogate base pairs (2-aminopyridine)

Low mutation rate, high stress environment, uneven mutational spectrum

Jerpseth, *Mol. Biotechnol.* **1997**, *7*, 189–195.

Random In Vivo Mutagenesis

Chemical mutagenesis:

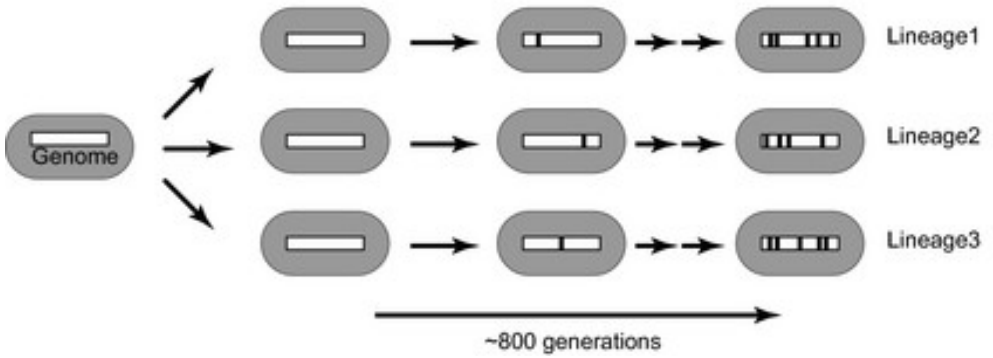
– alkylating agents, UV light, nitrous acid, surrogate base pairs (2-aminopyridine)

Low mutation rate, high stress environment, uneven mutational spectrum

Mutator Strain:

– XL1-red *E. coli* contains defective DNA proof-reading and repair mechanisms
Normal *E. coli* mutation rate 10^{-10}; XL1- red mutation rate 10^{-5}
8 h of growth (24 generations) accumulate ~ 1 mutation/gene

mutates whole genome decreasing desired mutations
limits the number of mutations/generations



Jerpseth, *Mol. Biotechnol.* **1997**, *7*, 189–195.

Mutagenesis

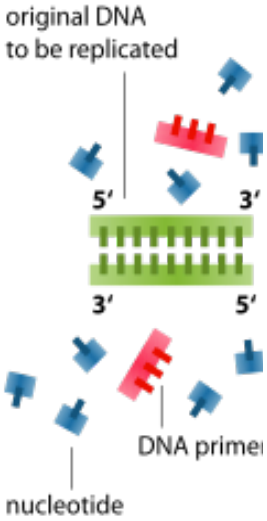
In Vivo:

- Chemical mutagenesis
- Mutator Strains

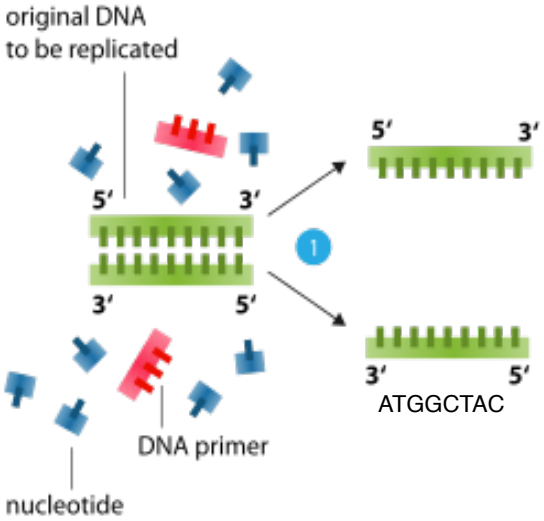
In Vitro:

- Random mutagenesis
- Focused mutagenesis
- Recombinative mutagenesis

Polymerase Chain Reaction (PCR)

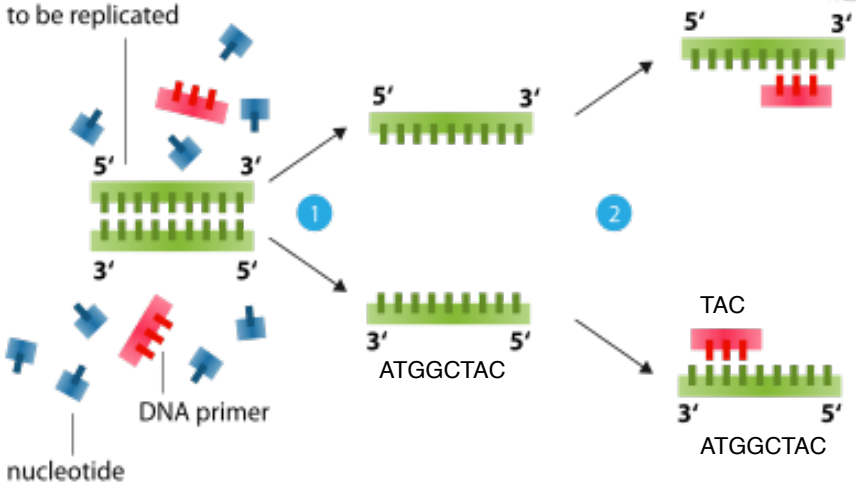


Polymerase Chain Reaction (PCR)



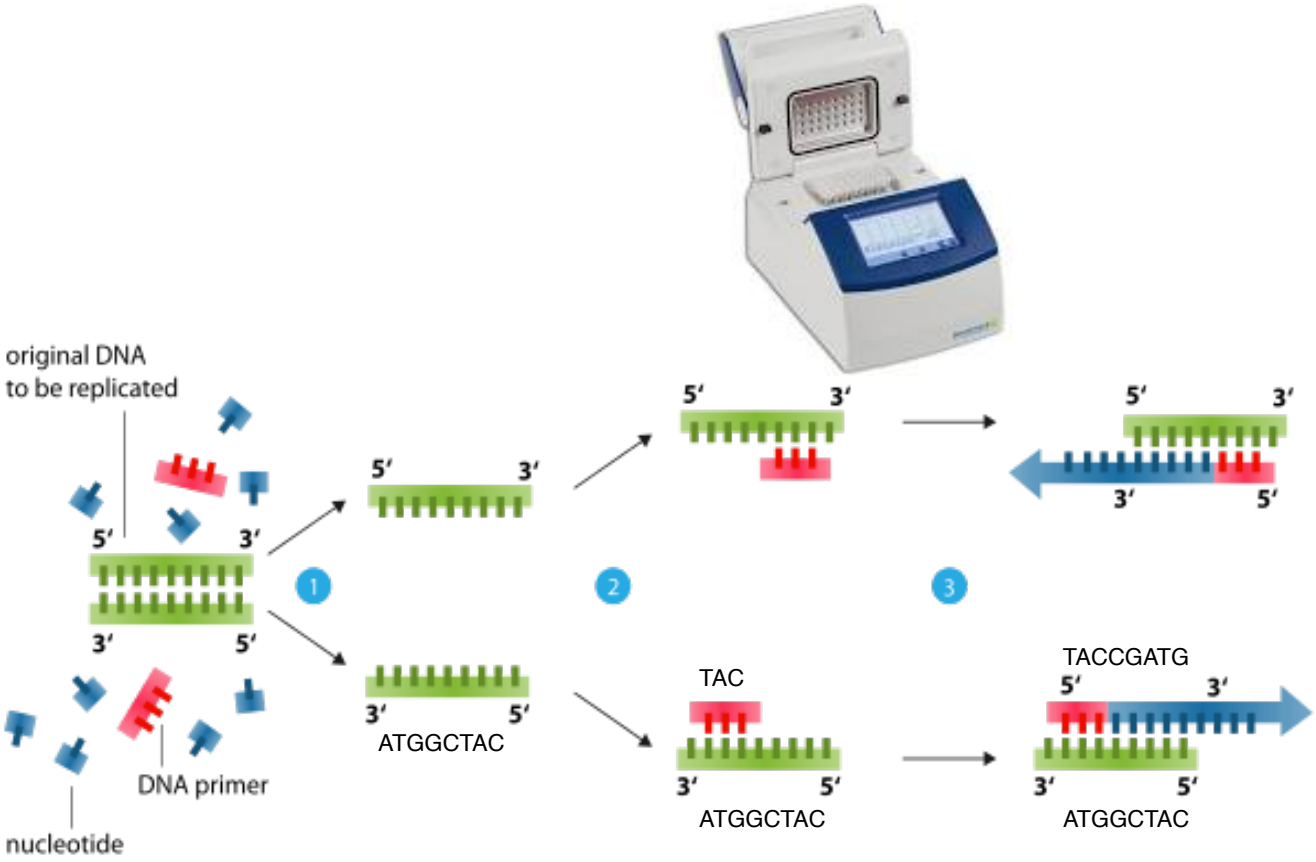
1. Denaturation 94–98 °C (20–30 sec)

Polymerase Chain Reaction (PCR)



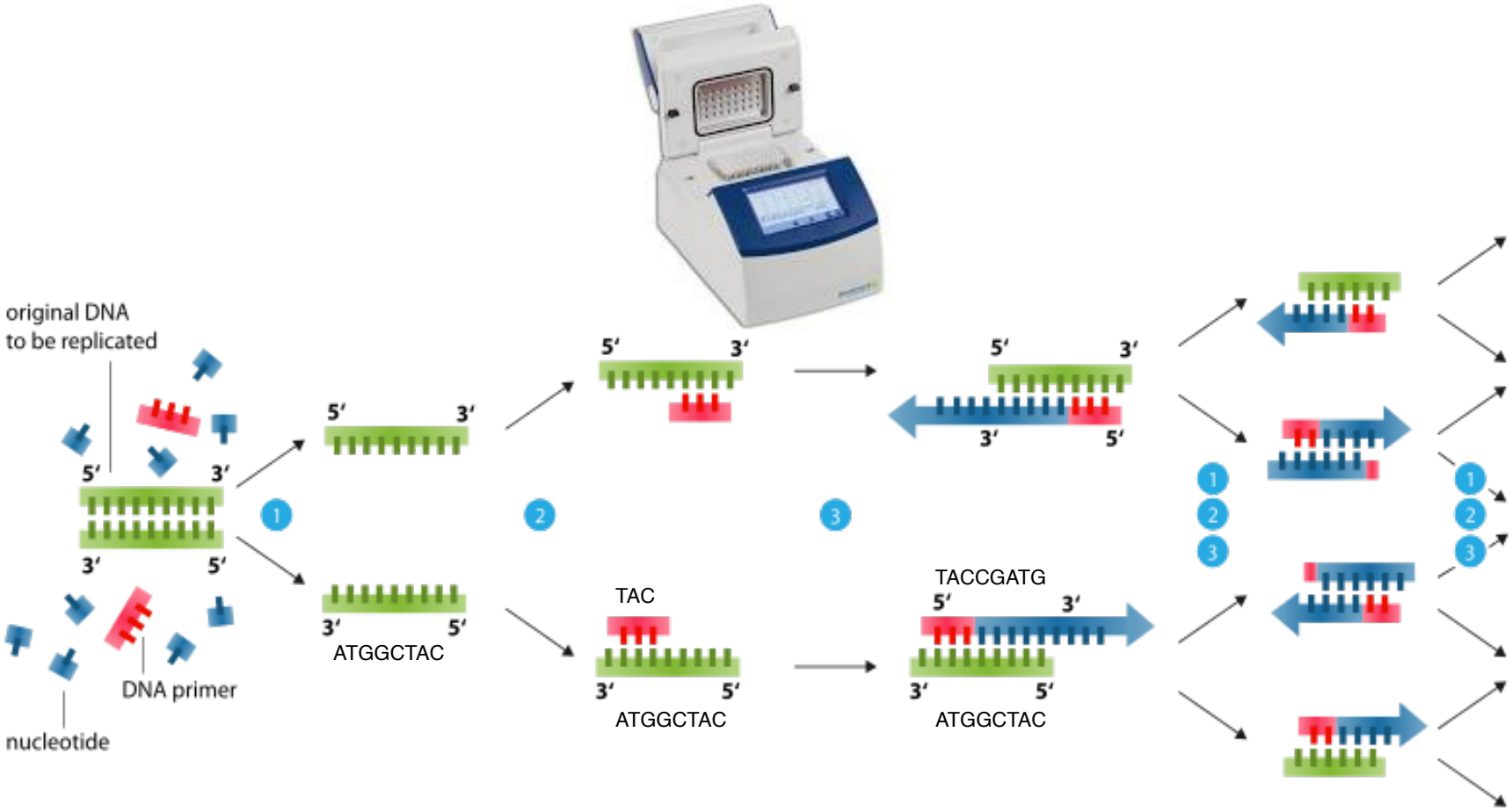
1. Denaturation 94–98 °C (20–30 sec)
2. Anneal 50–65 °C (20–40 sec)

Polymerase Chain Reaction (PCR)



1. Denaturation 94–98 °C (20–30 sec)
2. Anneal 50–65 °C (20–40 sec)
3. extension 75–80 °C (~ 1 kb/min)

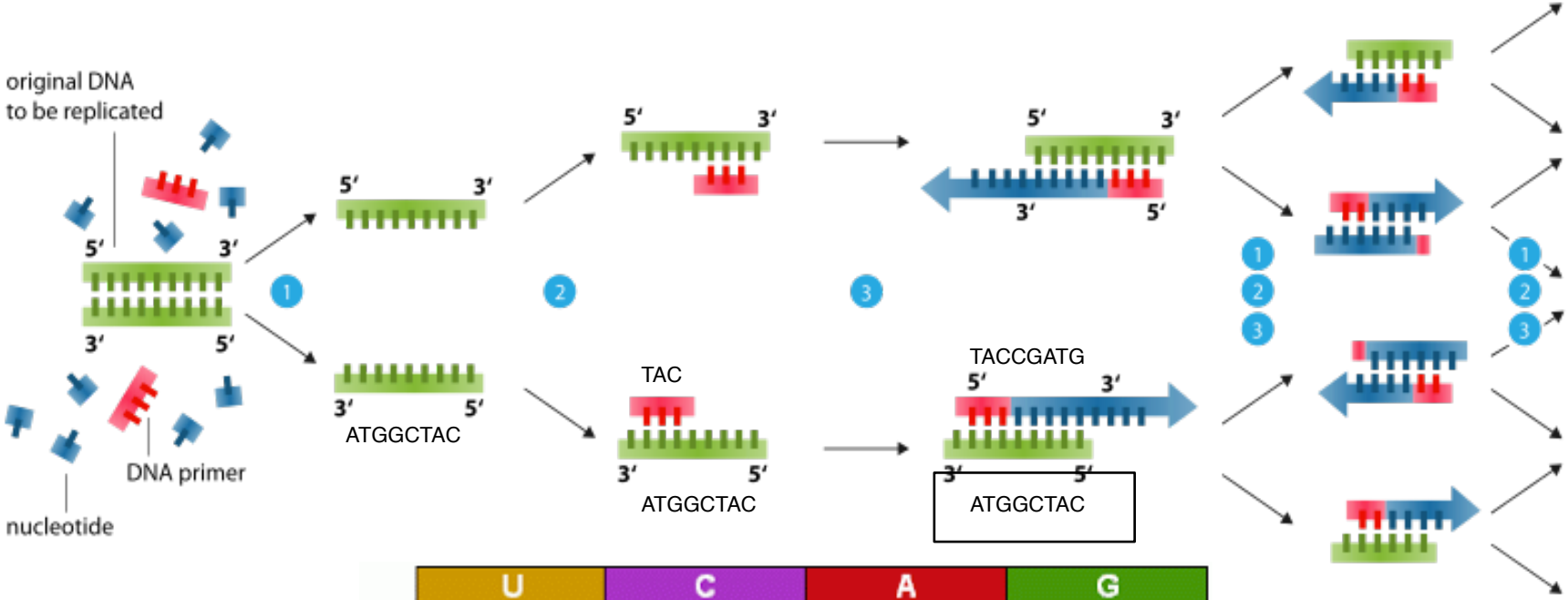
Polymerase Chain Reaction (PCR)



1. Denaturation 94–98 °C (20–30 sec)
2. Anneal 50–65 °C (20–40 sec)
3. extension 75–80 °C (~ 1 kb/min)

average gene length 10–15 kb
30 cycles generates 10^9 copies of DNA
5.5–8 hours

Polymerase Chain Reaction (PCR)



	U	C	A	G
U	UUU = phe UUC = phe UUA = leu UUG = leu	UCU = ser UCC = ser UCA = ser UCG = ser	UAU = tyr UAC = tyr UAA = stop UAG = stop	UGU = cys UGC = cys UGA = stop UGG = trp
C	CUU = leu CUC = leu CUA = leu CUG = leu	CCU = pro CCC = pro CCA = pro CCG = pro	CAU = his CAC = his CAA = gln CAG = gln	CGU = arg CGC = arg CGA = arg CGG = arg
A	AUU = ile AUC = ile AUA = ile AUG = met	ACU = thr ACC = thr ACA = thr ACG = thr	AAU = asn AAC = asn AAA = lys AAG = lys	AGU = ser AGC = ser AGA = arg AGG = arg
G	GUU = val GUC = val GUA = val GUG = val	GCU = ala GCC = ala GCA = ala GCG = ala	GAU = asp GAC = asp GAA = glu GAG = glu	GGU = gly GGC = gly GGA = gly GGG = gly

Random In Vitro Mutagenesis

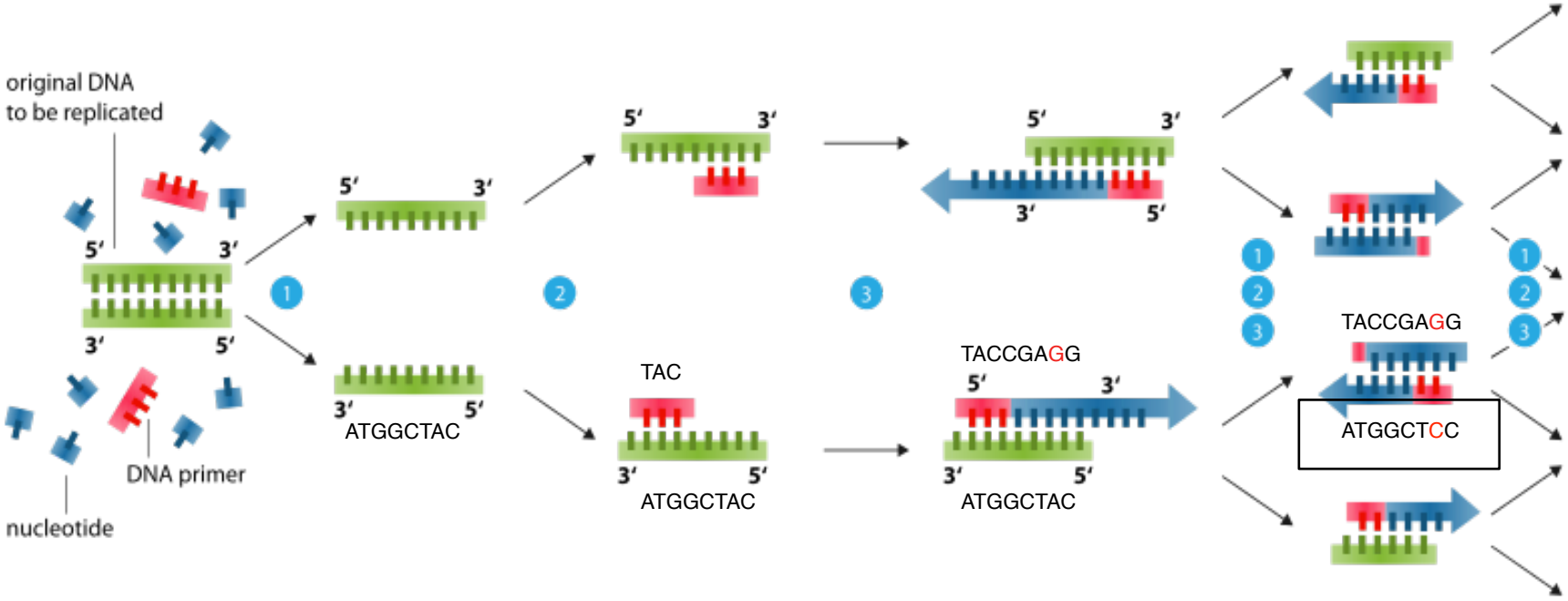
epPCR (error-prone PCR): Use of a lower fidelity DNA polymerase Taq with an increase Mg^{+2} and Mn^{+2} concentration

easy to use with a high mutation rate (1/1000 bp)

Random In Vitro Mutagenesis

epPCR (error-prone PCR): Use of a lower fidelity DNA polymerase Taq with an increase Mg^{+2} and Mn^{+2} concentration

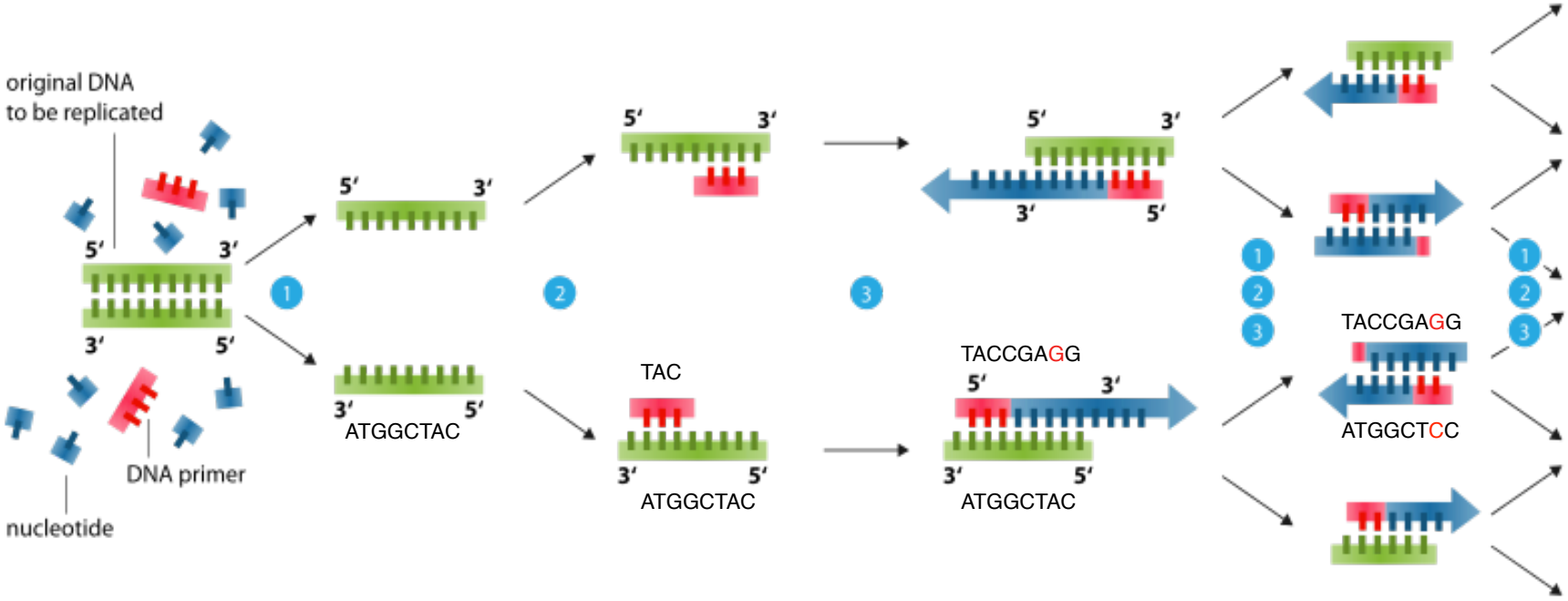
easy to use with a high mutation rate (1/1000 bp)



Random In Vitro Mutagenesis

epPCR (error-prone PCR): Use of a lower fidelity DNA polymerase Taq with an increase Mg^{+2} and Mn^{+2} concentration

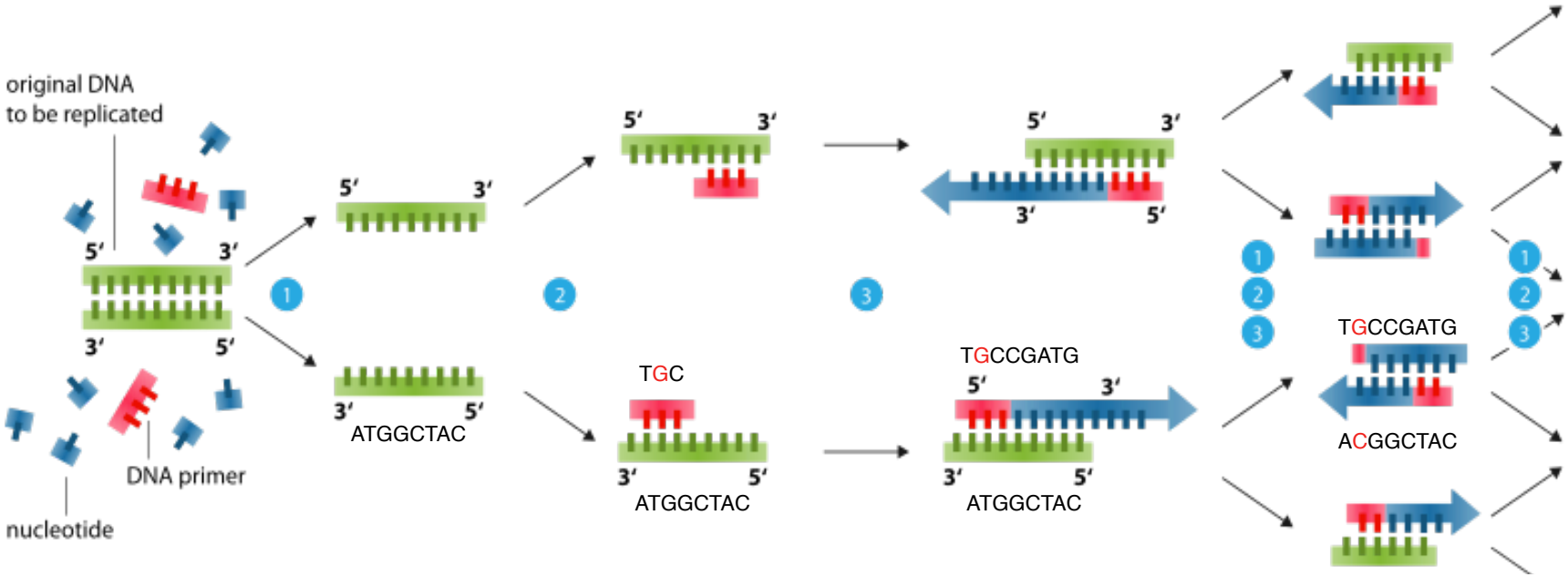
easy to use with a high mutation rate (1/1000 bp)



certain nucleotide substitutions are preffered over others
does not evenly sample the amino acid codon space

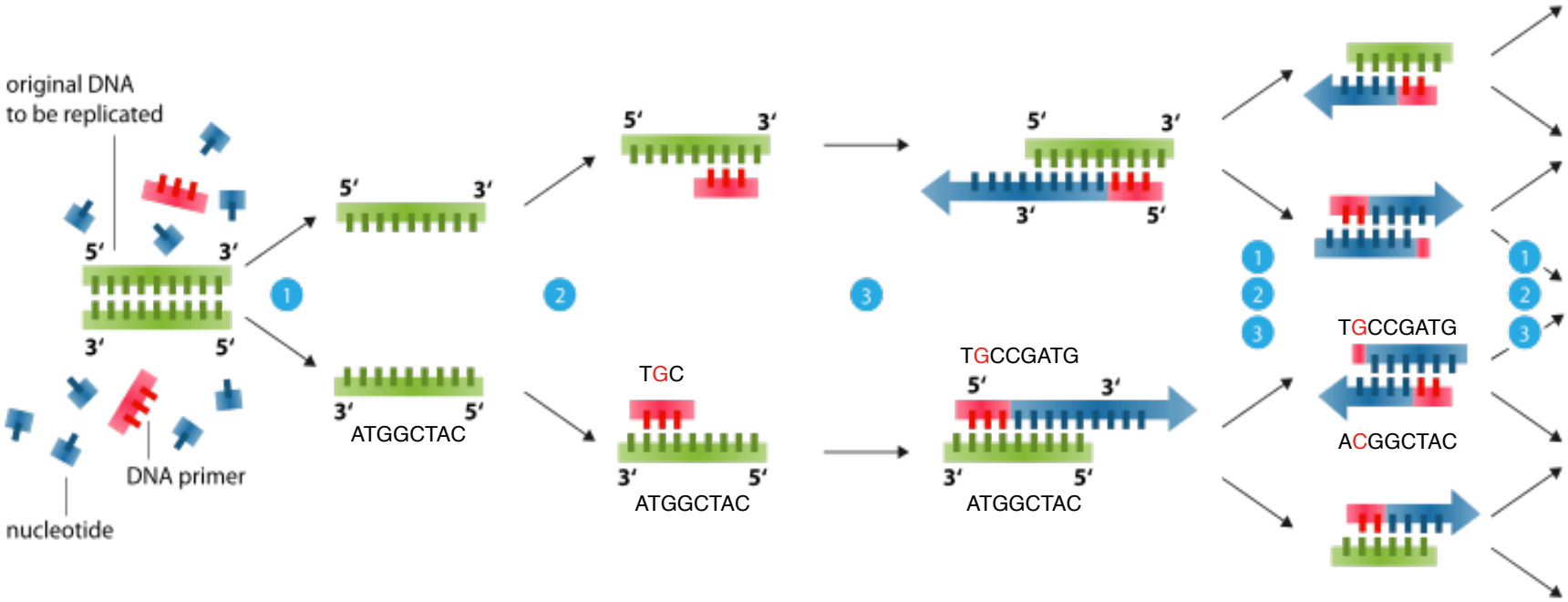
Focused In Vitro Mutagenesis

Site-Directed Mutagenesis: modified primers used to introduce a specific point mutation. capable of surveying all 20 amino acids at one position



Focused In Vitro Mutagenesis

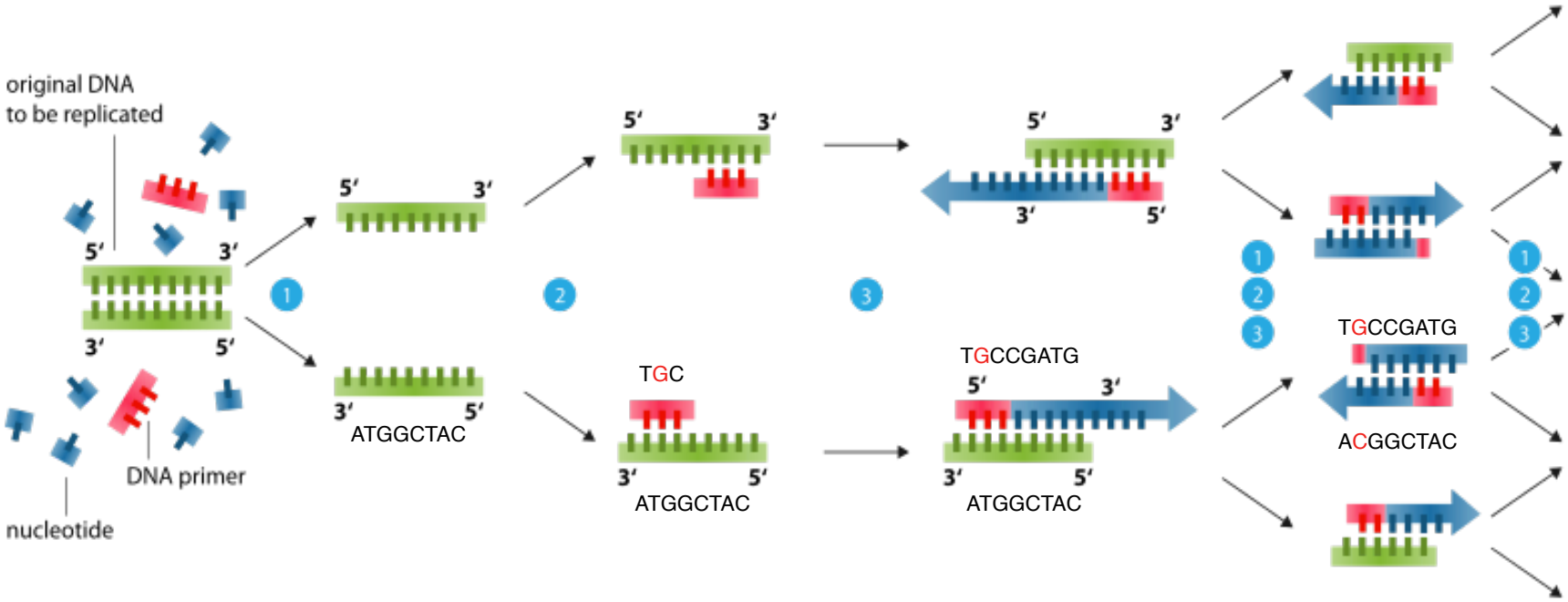
Site-Directed Mutagenesis: modified primers used to introduce a specific point mutation. capable of surveying all 20 amino acids at one position



requires structural or biochemical knowledge (know the “hot spots”)
excess of inactive clones

Focused In Vitro Mutagenesis

Site-Directed Mutagenesis: modified primers used to introduce a specific point mutation. capable of surveying all 20 amino acids at one position

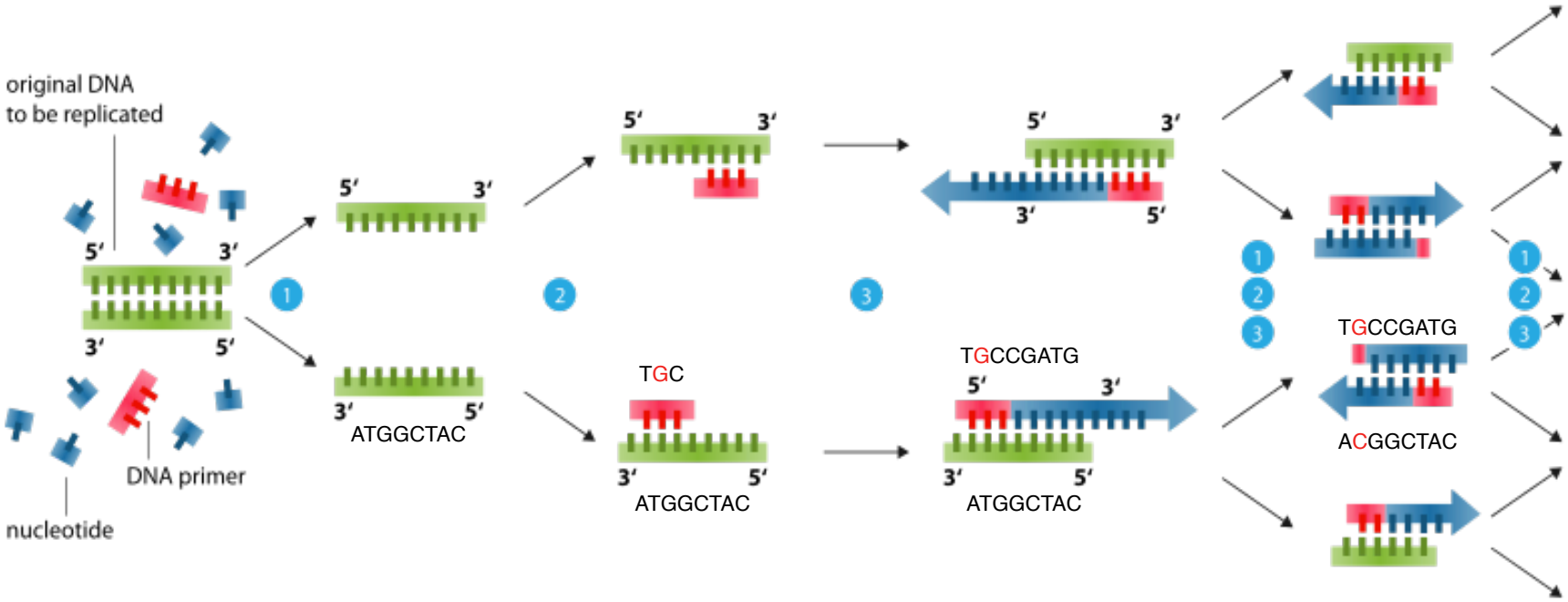


requires structural or biochemical knowledge (know the “hot spots”)
excess of inactive clones

Site-Directed Saturation Mutagenesis: modified primers replace a codon instead of a single point mutation

Focused In Vitro Mutagenesis

Site-Directed Mutagenesis: modified primers used to introduce a specific point mutation. capable of surveying all 20 amino acids at one position



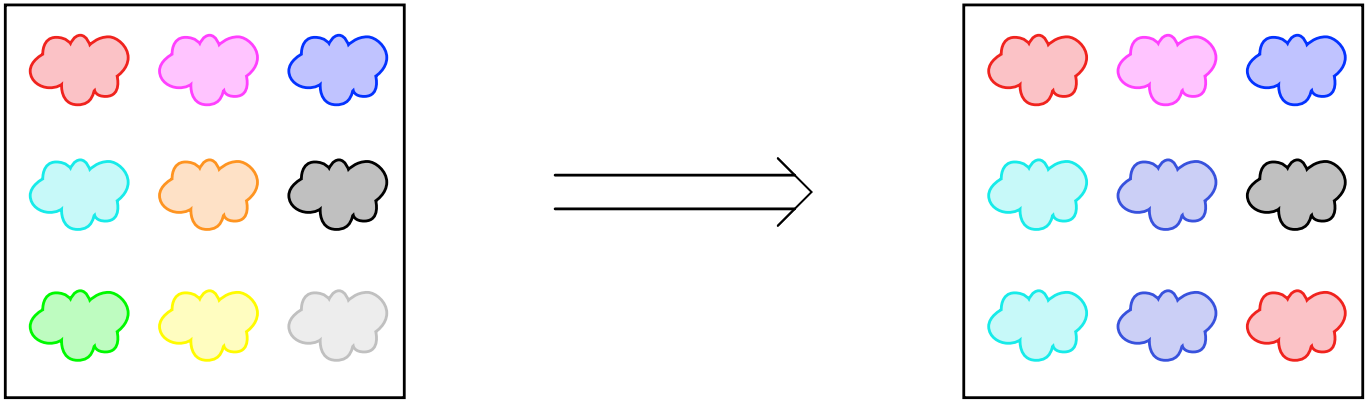
requires structural or biochemical knowledge (know the “hot spots”)
excess of inactive clones

Site-Directed Saturation Mutagenesis: modified primers replace a codon instead of a single point mutation

Cassette Mutagenesis: modified primers that replace a series of codons allowing for greater variation

The Number Problem with Saturation Mutagenesis

Saturation mutagenesis requires significant oversampling in activity screens



Ideal plate

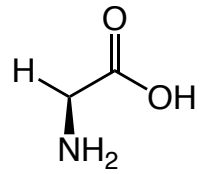
Real plate

20 AA

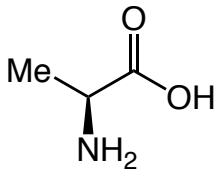
No. of AA to be Randomized	No. of mutants to be screened to cover 95% of possible combinations
1	94
2	3,066
3	98,163
4	3,141,251
5	100,520,093

Reetz, *ChemBioChem* ***2008***, 9, *1797–1804.*

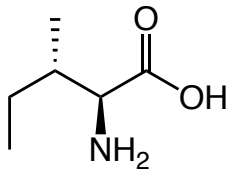
The 20 Amino Acids



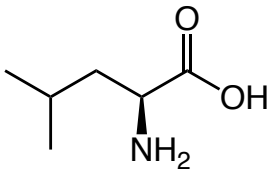
Glycine Gly (G)



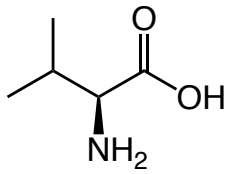
Alanine Ala (A)



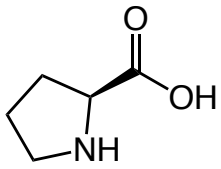
Isoleucine Ile (I)



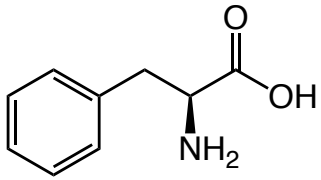
Leucine Leu (L)



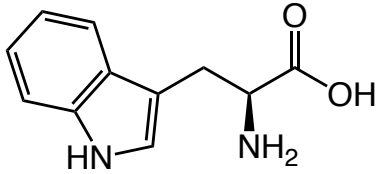
Valine Val (V)



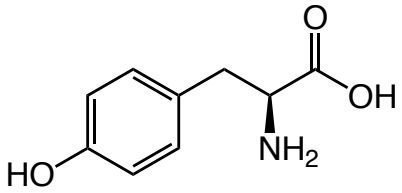
Proline Pro (P)



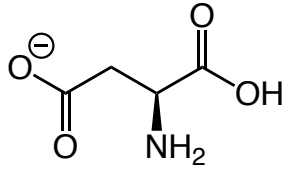
Phenylalanine Phe (F)



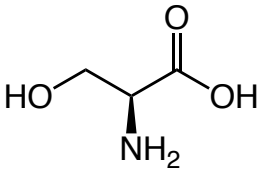
Tryptophan Trp (W)



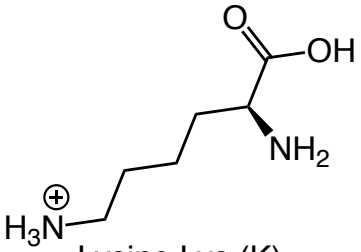
Tyrosine Tyr (Y)



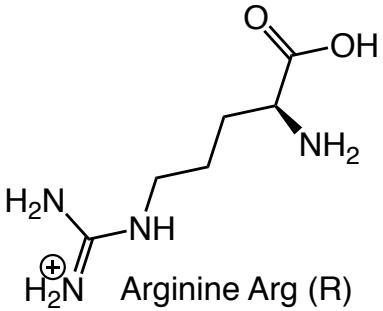
Aspartic Acid Asp (D)



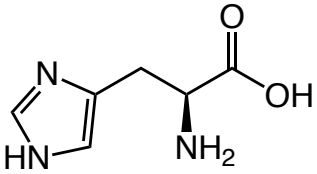
Serine Ser (S)



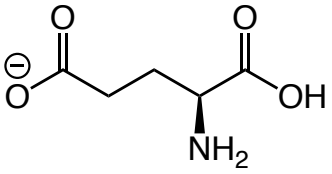
Lysine Lys (K)



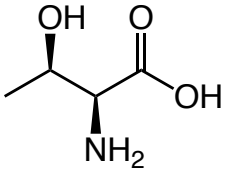
Arginine Arg (R)



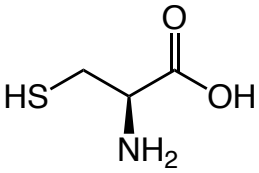
Histidine His (H)



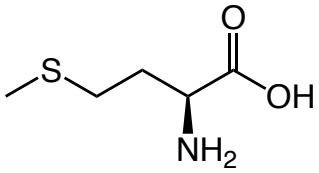
Glutamic Acid Glu (E)



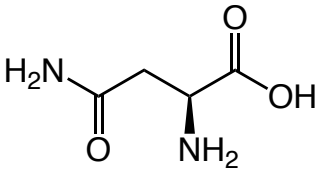
Threonine Thr (T)



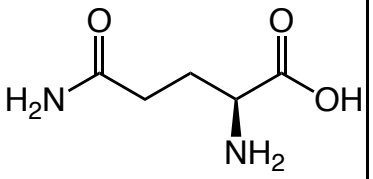
Cysteine Cys (C)



Methionine Met (M)



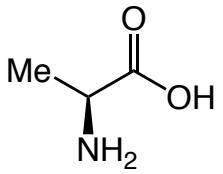
Asparagine Asn (N)



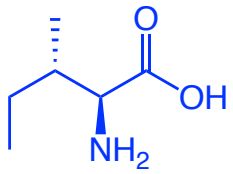
Glutamine Gln (Q)

The 20 Amino Acids

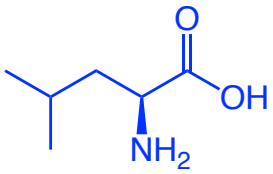
Glycine Gly (G)



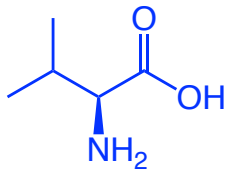
Alanine Ala (A)



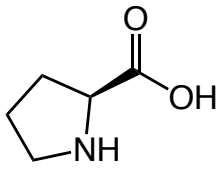
Isoleucine Ile (I)



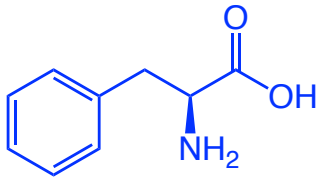
Leucine Leu (L)



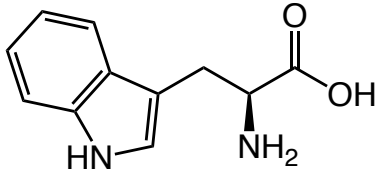
Valine Val (V)



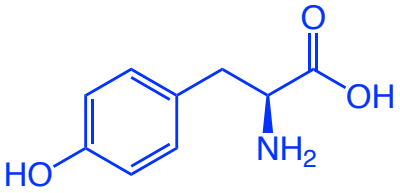
Proline Pro (P)



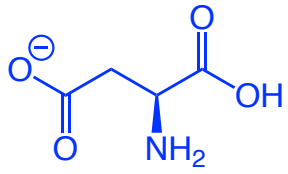
Phenylalanine Phe (F)



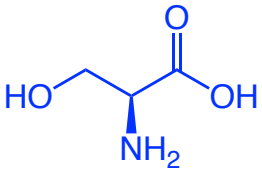
Tryptophan Trp (W)



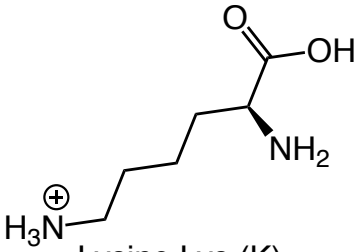
Tyrosine Tyr (Y)



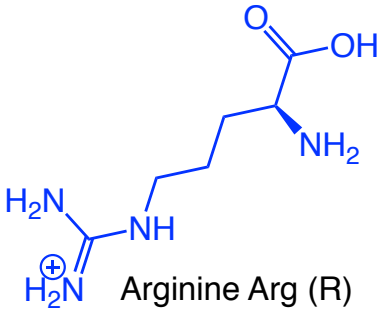
Aspartic Acid Asp (D)



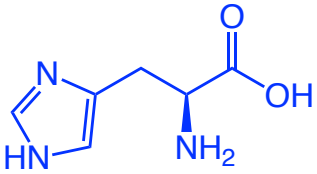
Serine Ser (S)



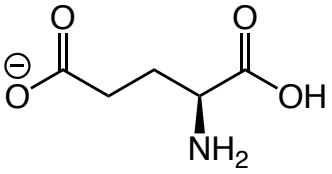
Lysine Lys (K)



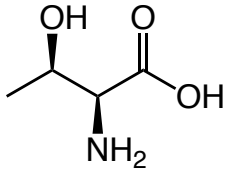
Arginine Arg (R)



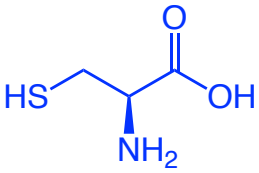
Histidine His (H)



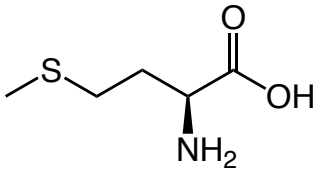
Glutamic Acid Glu (E)



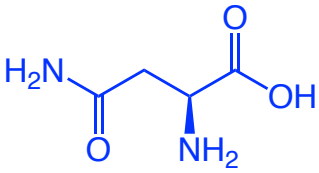
Threonine Thr (T)



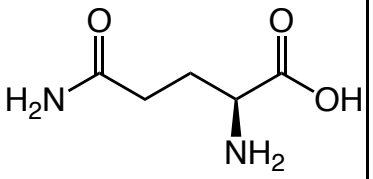
Cysteine Cys (C)



Methionine Met (M)



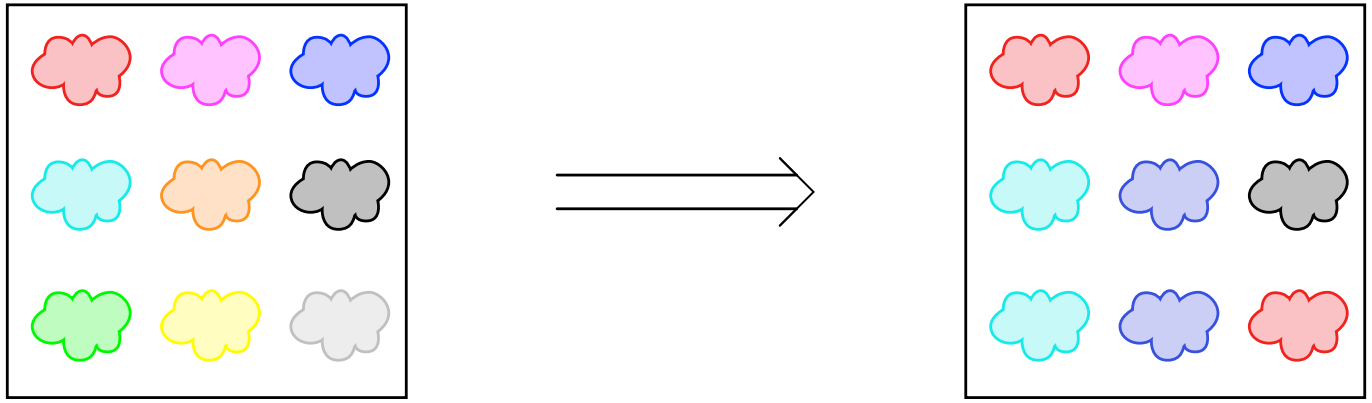
Asparagine Asn (N)



Glutamine Gln (Q)

The Number Problem with Saturation Mutagenesis

Saturation mutagenesis requires significant oversampling in activity screens



Ideal plate

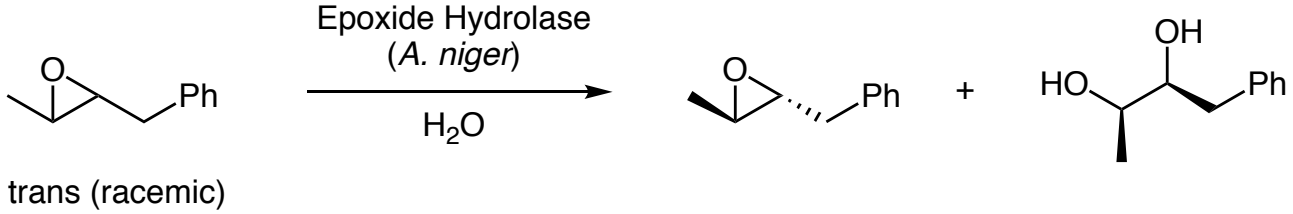
Real plate

20 AA	
No. of AA to be Randomized	No. of mutants to be screened to cover 95% of possible combinations
1	94
2	3,066
3	98,163
4	3,141,251
5	100,520,093

12 AA	
No. of AA to be Randomized	No. of mutants to be screened to cover 95% of possible combinations
1	34
2	430
3	5,175
4	62,118
5	745,433

Reetz, *ChemBioChem* ***2008***, 9, *1797–1804.*

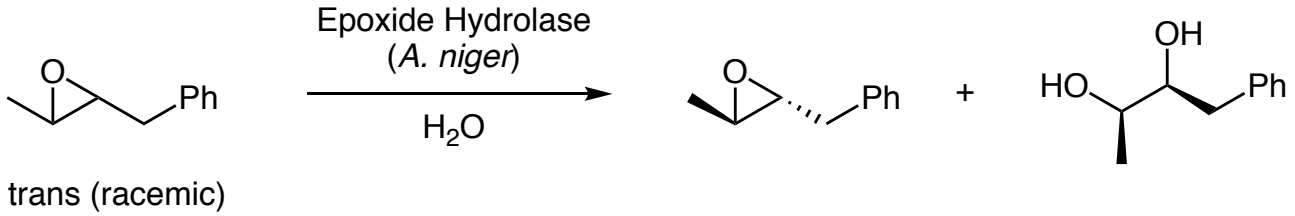
Example of Limiting the Amino Acid Space



trans (racemic)

Reetz, *ChemBioChem* ***2008***, 9*, 1797–1804.*

Example of Limiting the Amino Acid Space



trans (racemic)

Saturation mutagenesis of 3 positions based on 20 amino acids (15% coverage)

wild type
(no reaction)

Sat. Mut.

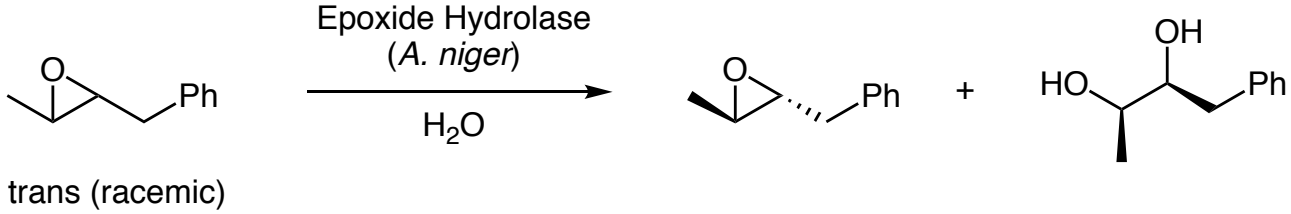
5000 mutants

38 active

best mutant
TOF 17
s = 101

Reetz, *ChemBioChem* ***2008***, 9, *1797–1804.*

Example of Limiting the Amino Acid Space



trans (racemic)

Saturation mutagenesis of 3 positions based on 20 amino acids (15% coverage)

wild type (no reaction) → Sat. Mut. → 5000 mutants → 38 active → best mutant TOF 17 s = 101

Saturation mutagenesis of 3 positions based on 12 amino acids (95% coverage)

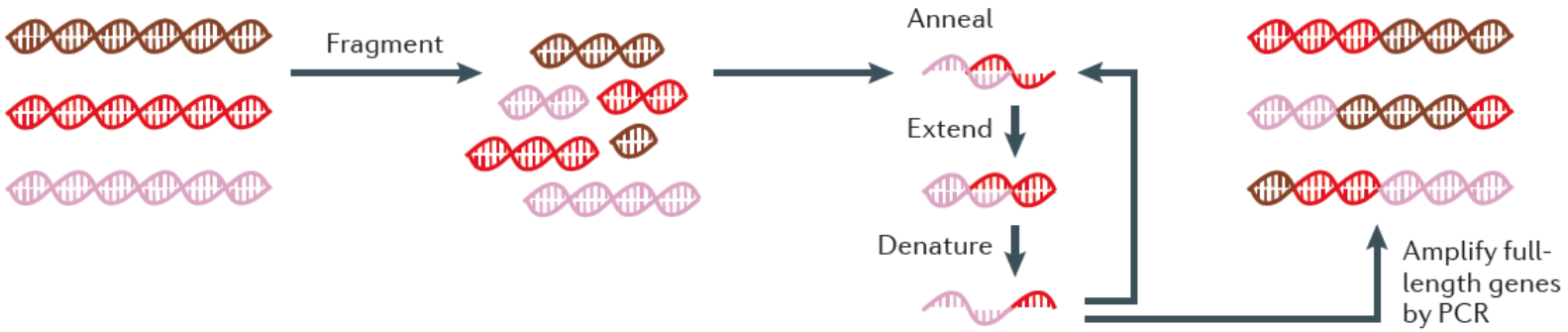
wild type (no reaction) → Sat. Mut. → 5000 mutants → 511 active → best mutant TOF 55 s > 200

Reetz, *ChemBioChem* ***2008***, 9, *1797–1804.*

Diversification by Recombination

Reassortment of mutations from random and focused mutegenesis procedures

DNA shuffling (Stemmer, 1994)

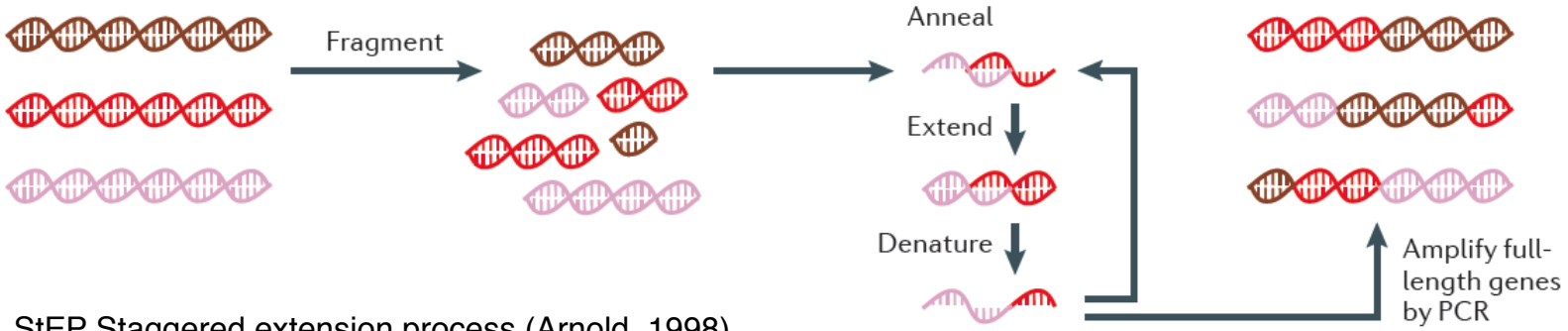


Lui, *Nat Rev. Genet.* **2015**, *16*, 379–394.

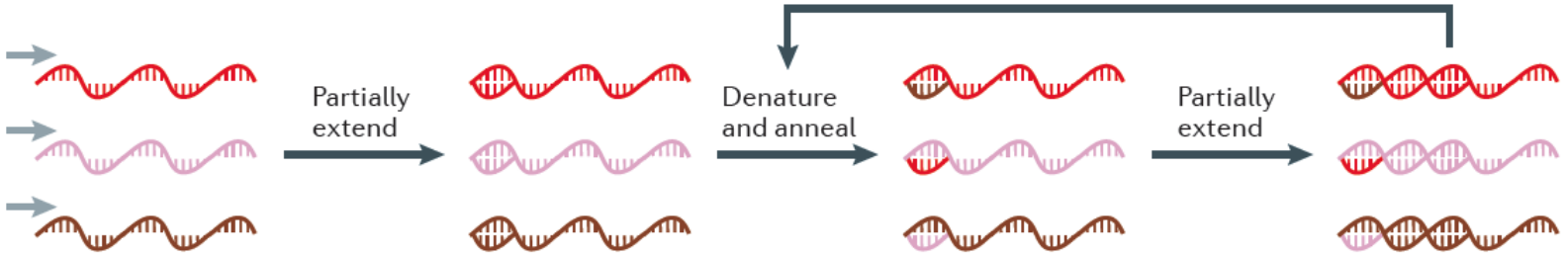
Diversification by Recombination

Reassortment of mutations from random and focused mutegenesis procedures

DNA shuffling (Stemmer, 1994)



StEP Staggered extension process (Arnold, 1998)

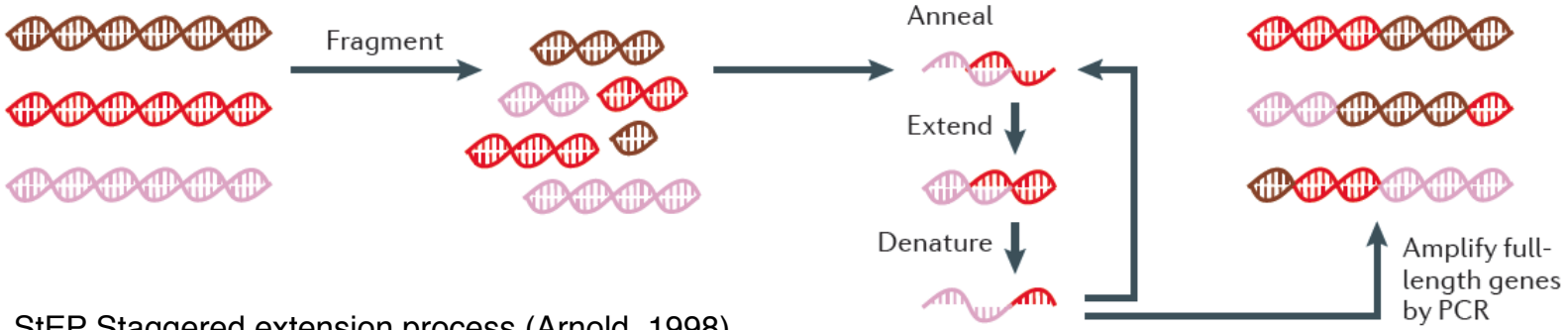


Lui, *Nat Rev. Genet.* **2015**, *16*, 379–394.

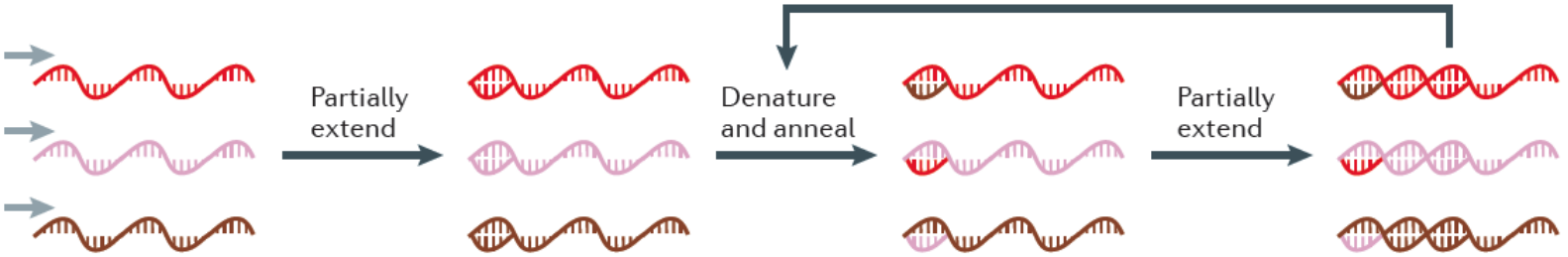
Diversification by Recombination

Reassortment of mutations from random and focused mutegenesis procedures

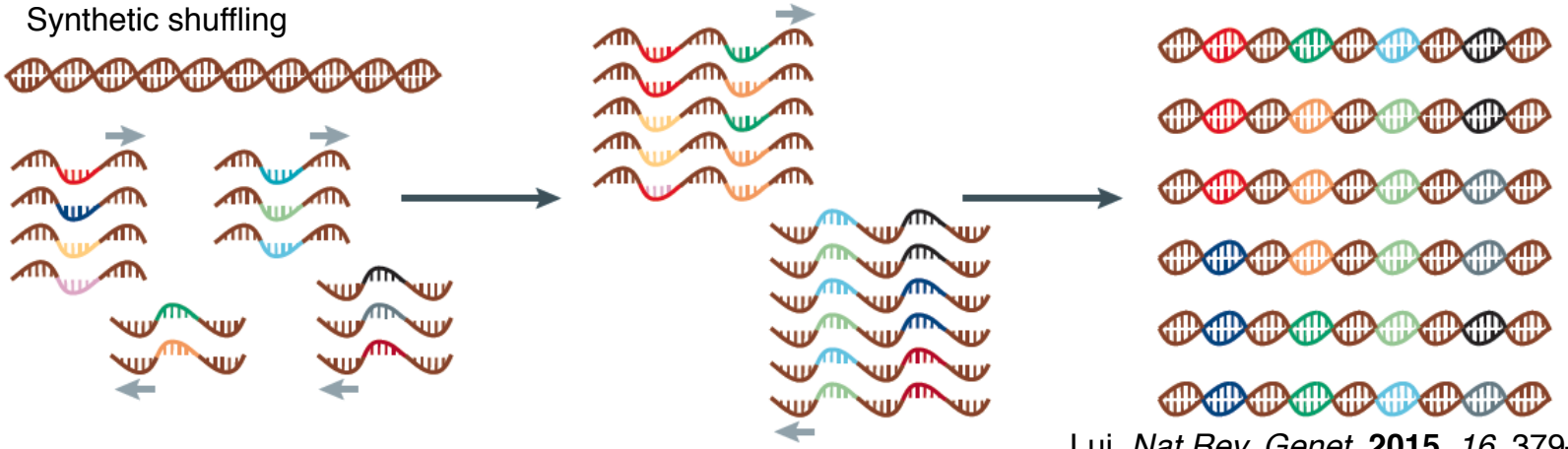
DNA shuffling (Stemmer, 1994)



StEP Staggered extension process (Arnold, 1998)



Synthetic shuffling



Lui, *Nat Rev. Genet.* **2015**, *16*, 379–394.

Mutagenesis Summary

Chemical mutagenesis:

- Not commonly used
- Inconsistent and biased mutations

Mutator Strains:

- Easy to use
- Low Mutation rates
- There is a mutational limit
- Somewhat used

epPCR:

- Easy to use
- High mutation rates
- Difficult to sample all AA combinations
- Very common

Site-directed saturation mutagenesis:

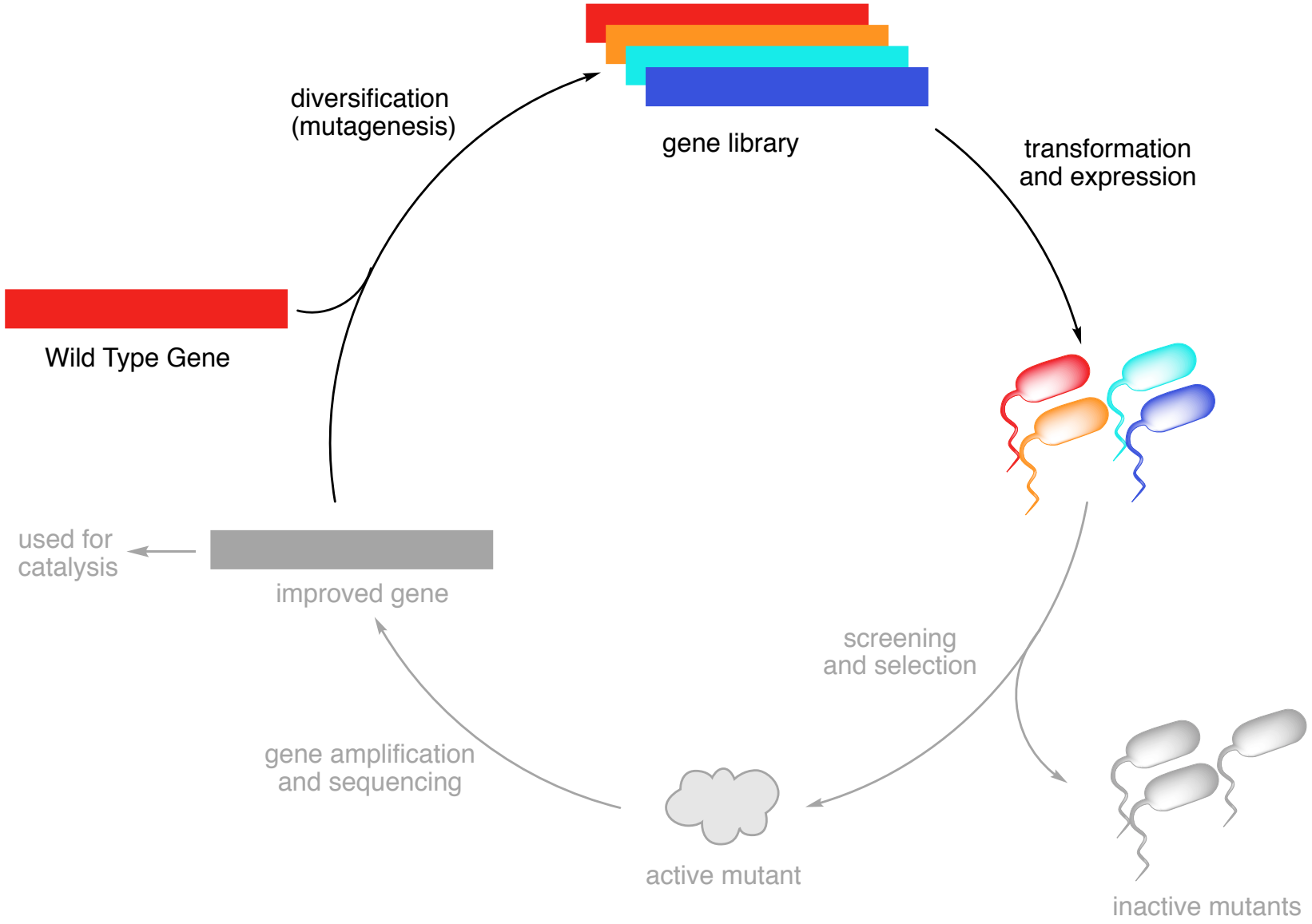
- Fully samples amino acid repertoire
- Capable of introducing specfic mutations
- Higher quality libraries
- Generally require larger screens
- Very common

Recombinative mutagenesis:

- Used in tadem with epPCR or Sat. Mut.
- Very common

Lui, *Nat Rev. Genet.* **2015**, *16*, 379–394.

Directed Evolution

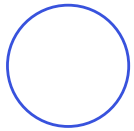


Transformation and Expression

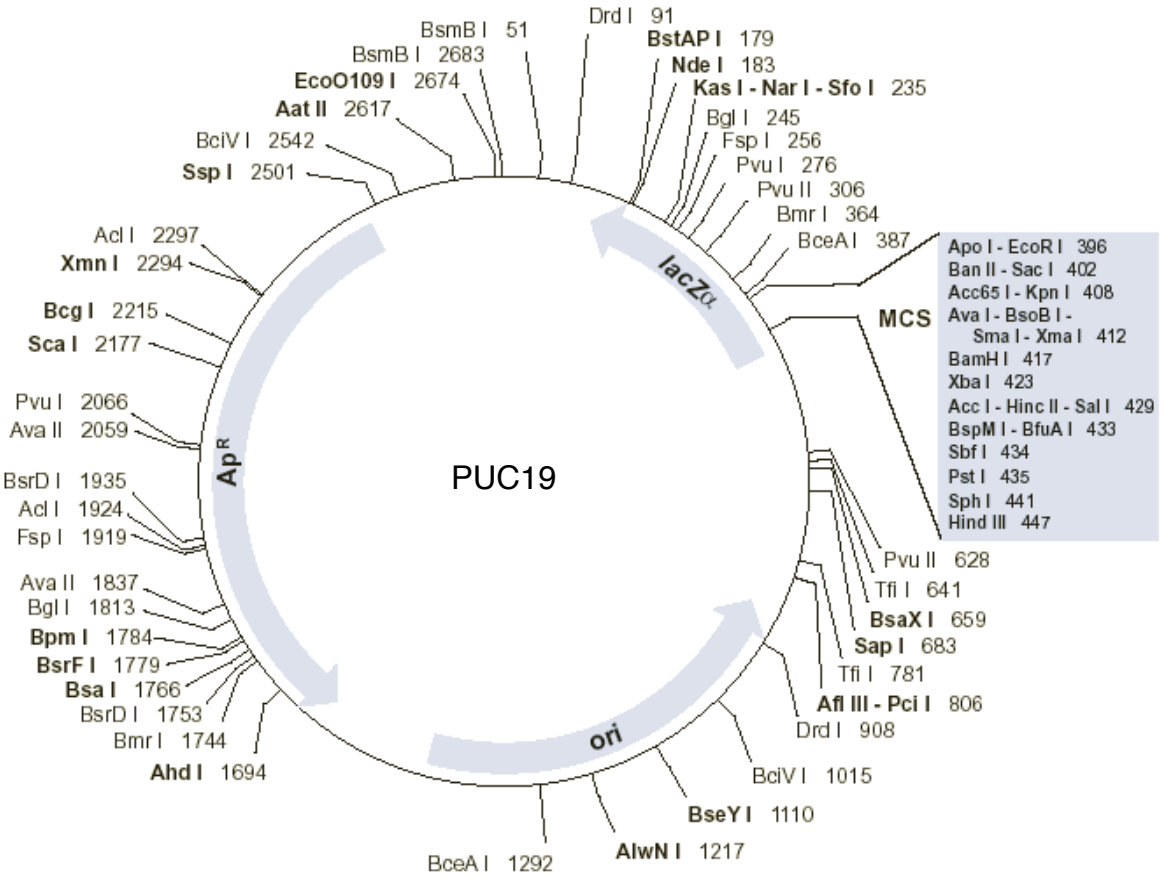
mutated genes

insert into plasmid

or perform mutagenesis on plasmid



mutated plasmids

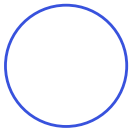


Transformation and Expression

mutated genes

insert into plasmid

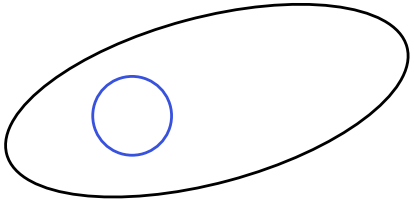
or perform mutagenesis on plasmid



mutated plasmids

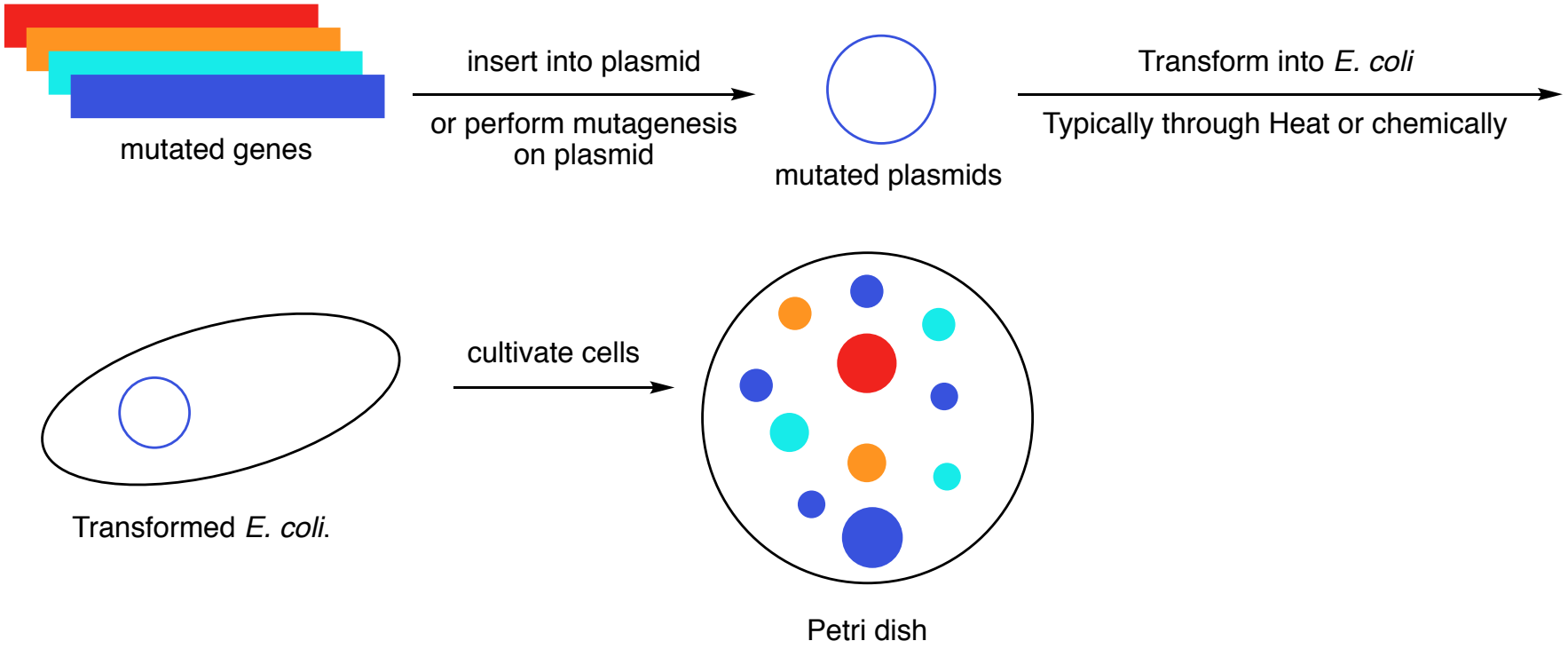
Transform into *E. coli*

Typically through Heat or chemically



Transformed *E. coli*.

Transformation and Expression

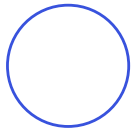


Transformation and Expression

mutated genes

insert into plasmid

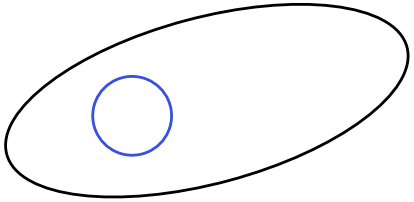
or perform mutagenesis on plasmid



mutated plasmids

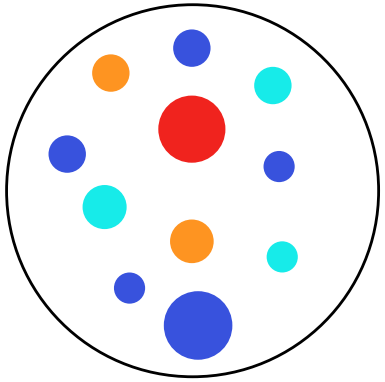
Transform into *E. coli*

Typically through Heat or chemically



Transformed *E. coli*.

cultivate cells



Petri dish

individually pick colonies

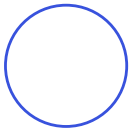
cultivate cells and induce expression

Transformation and Expression

mutated genes

insert into plasmid

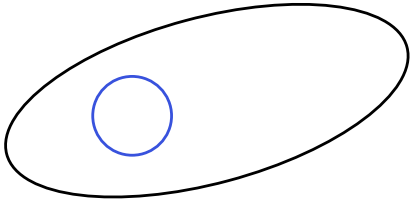
or perform mutagenesis on plasmid



mutated plasmids

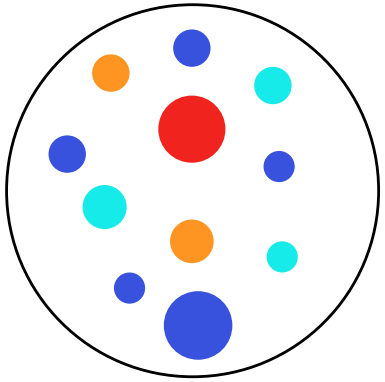
Transform into *E. coli*

Typically through Heat or chemically



Transformed *E. coli*.

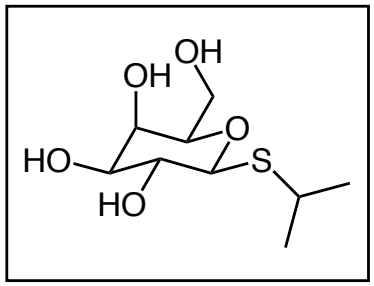
cultivate cells



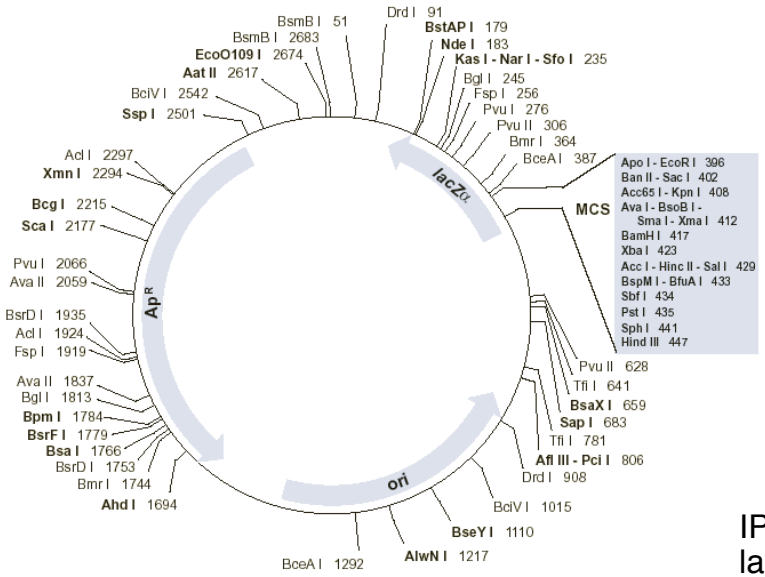
Petri dish

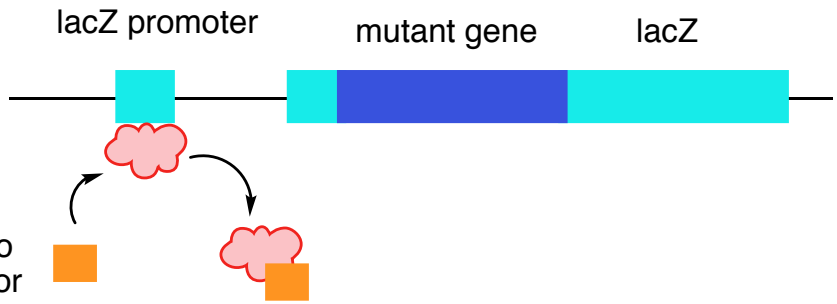
individually pick colonies

cultivate cells and induce expression



plasmid expression is induced with ITPG





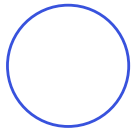
IPTG binds to lacZ repressor

Isolation

mutated genes

insert into plasmid

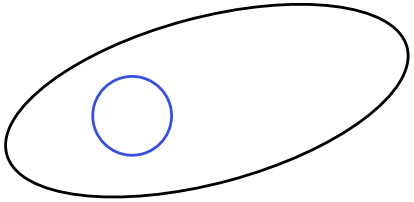
or perform mutagenesis on plasmid



mutated plasmids

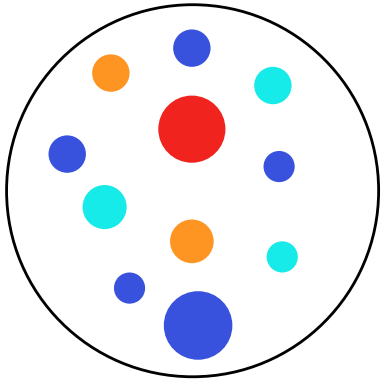
Transform into *E. coli*

Typically through Heat or chemically



Transformed *E. coli*.

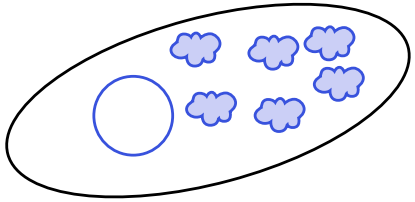
cultivate cells



Petri dish

individually pick colonies

cultivate cells and induce expression



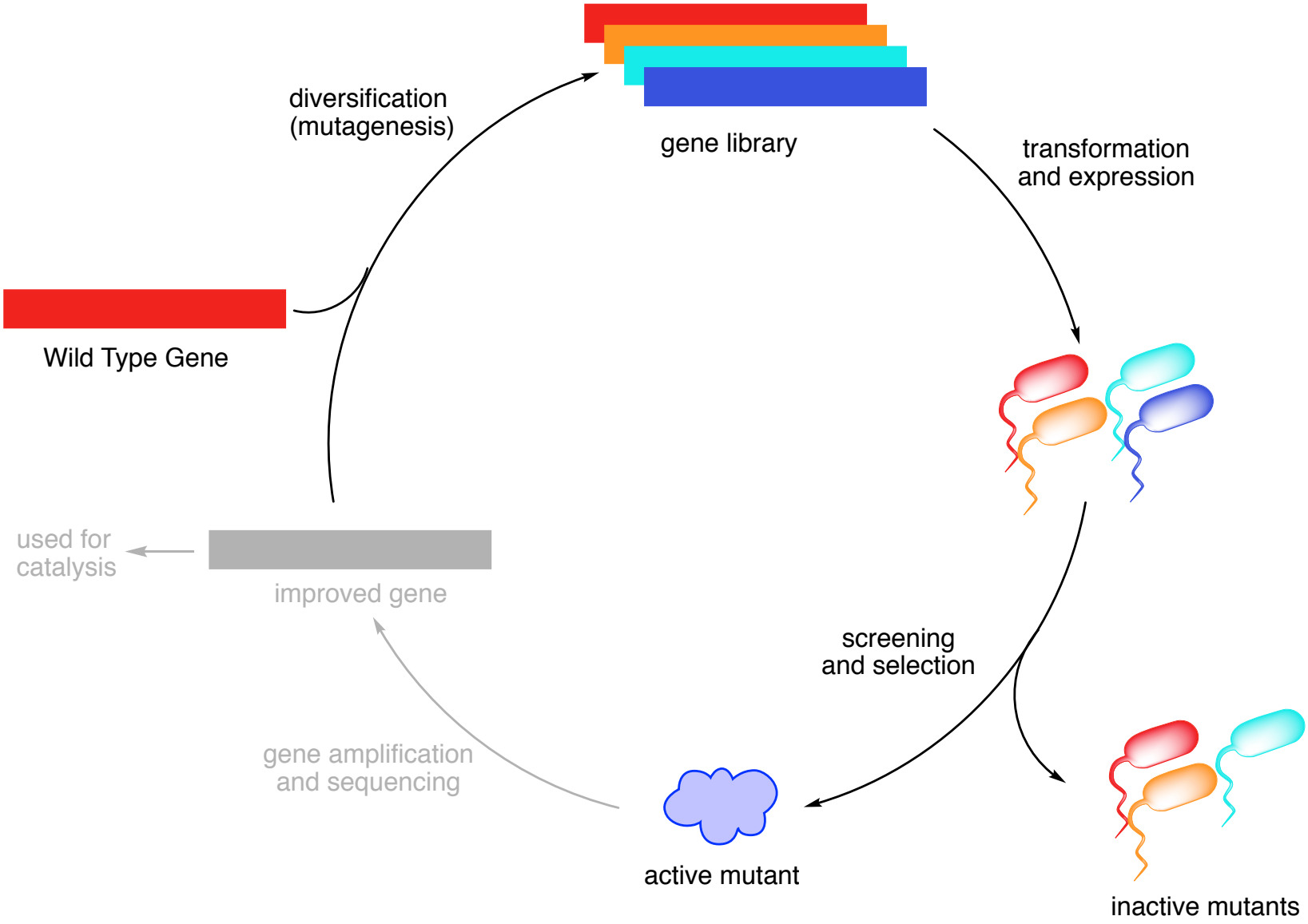
isolation

isolated mutant enzymes

– can use whole cell or lysate for catalysis

– isolated enzyme can be sequenced but not necessary

Directed Evolution



Screening and Selection

Screening: In Vitro analysis of all mutant enzymes

– High-throughput assays are desired (96-well plates)

– UV-Vis and fluorescnce quenching assays are the most common

Backvall, *Proc. Natl. Acad. Sci. U.S.A.* **2012,** *109*, 78–83
Maillard, *Chem. Commun.* **2009**, 34–46.

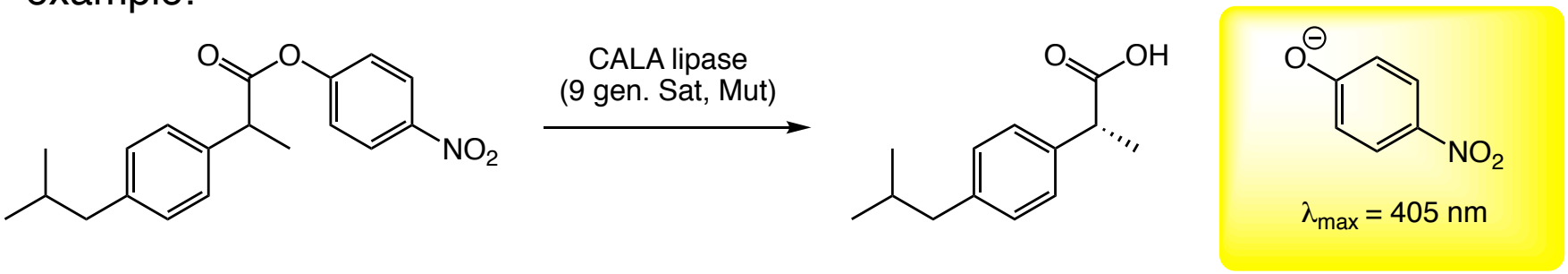
Screening and Selection

Screening: In Vitro analysis of all mutant enzymes

– High-throughput assays are desired (96-well plates)

– UV-Vis and fluorescnce quenching assays are the most common

example:



Backvall, *Proc. Natl. Acad. Sci. U.S.A.* **2012,** *109*, 78–83
Maillard, *Chem. Commun.* **2009**, 34–46.

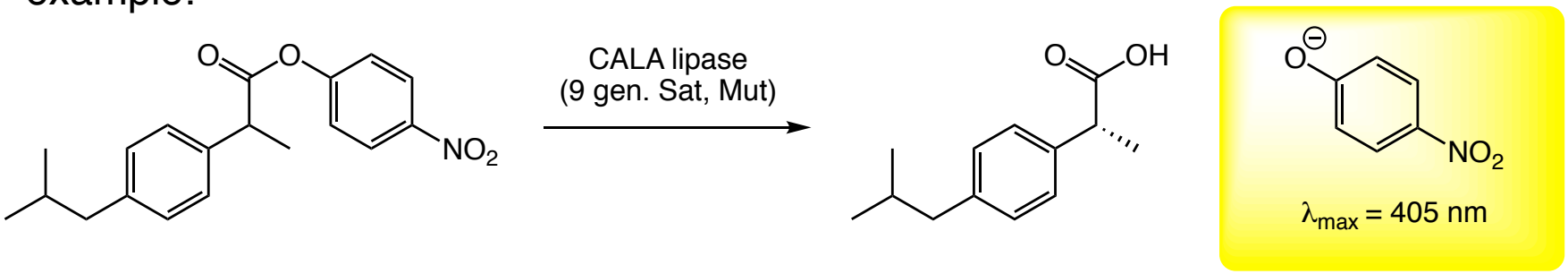
Screening and Selection

Screening: In Vitro analysis of all mutant enzymes

– High-throughput assays are desired (96-well plates)

– UV-Vis and fluorescnce quenching assays are the most common

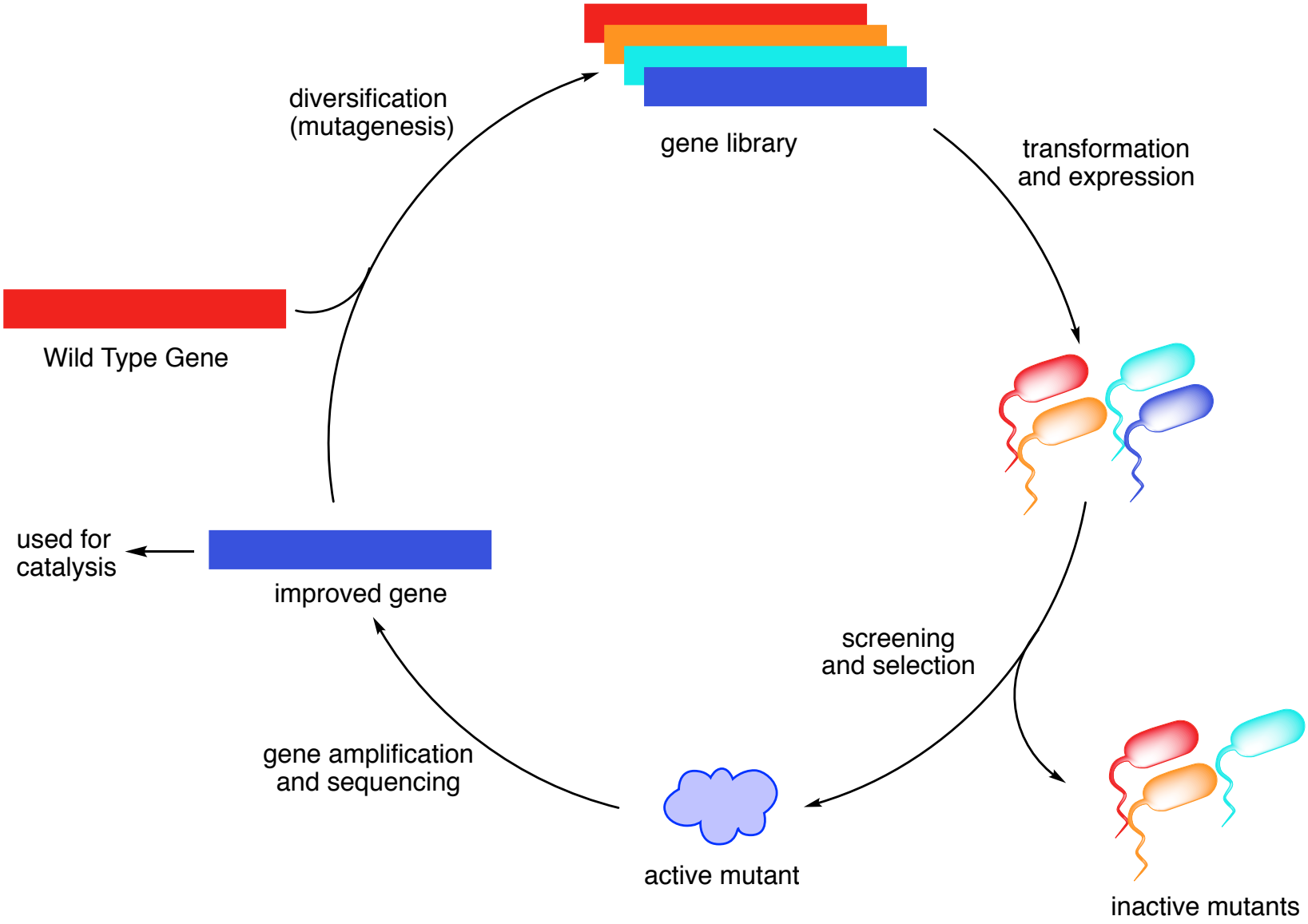
example:



– other methods include: pH, MS, NMR, IR, GC-MS, LC-MS

Backvall, *Proc. Natl. Acad. Sci. U.S.A.* **2012,** *109*, 78–83
Maillard, *Chem. Commun.* **2009**, 34–46.

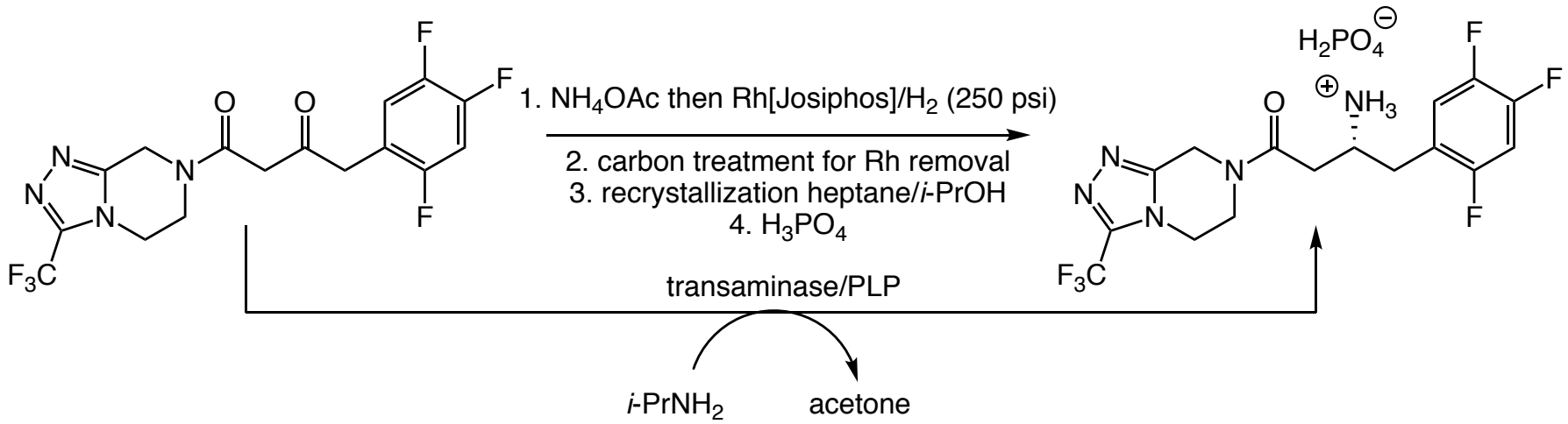
Directed Evolution



Outline

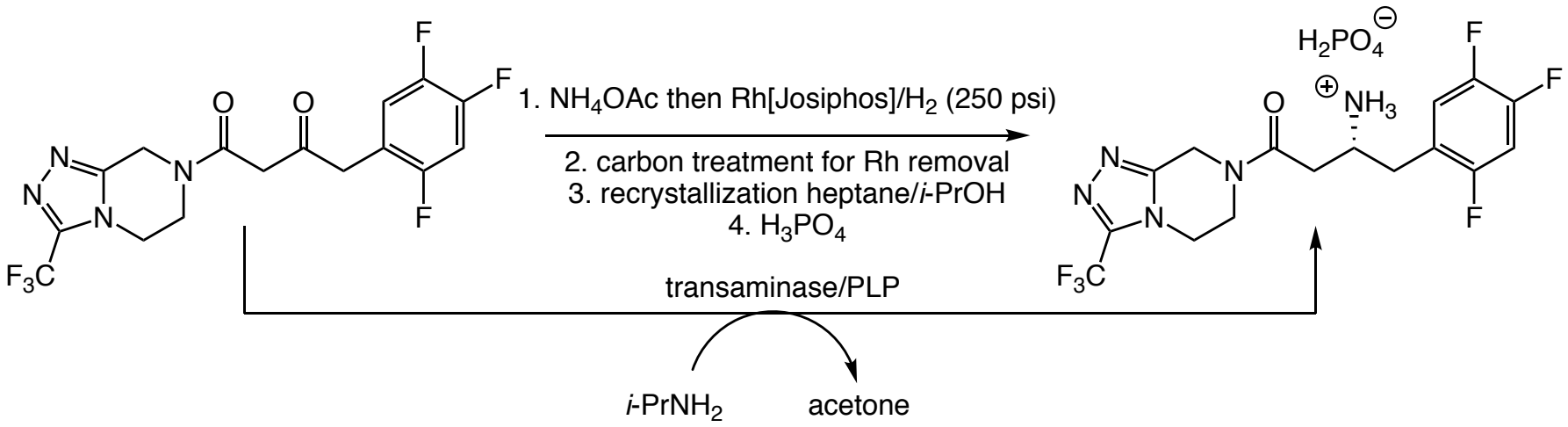
- Background
  - Timeline
- Directed Evolution Breakdown
  - Mutagenesis
  - Expression
  - Screening
- Examples
  - Pharmaceutical/ Industrial examples
  - New to Nature Reactions

Merck's Sitagliptin Synthesis



Savile, *Science*, **2010**, *329*, 305–309.

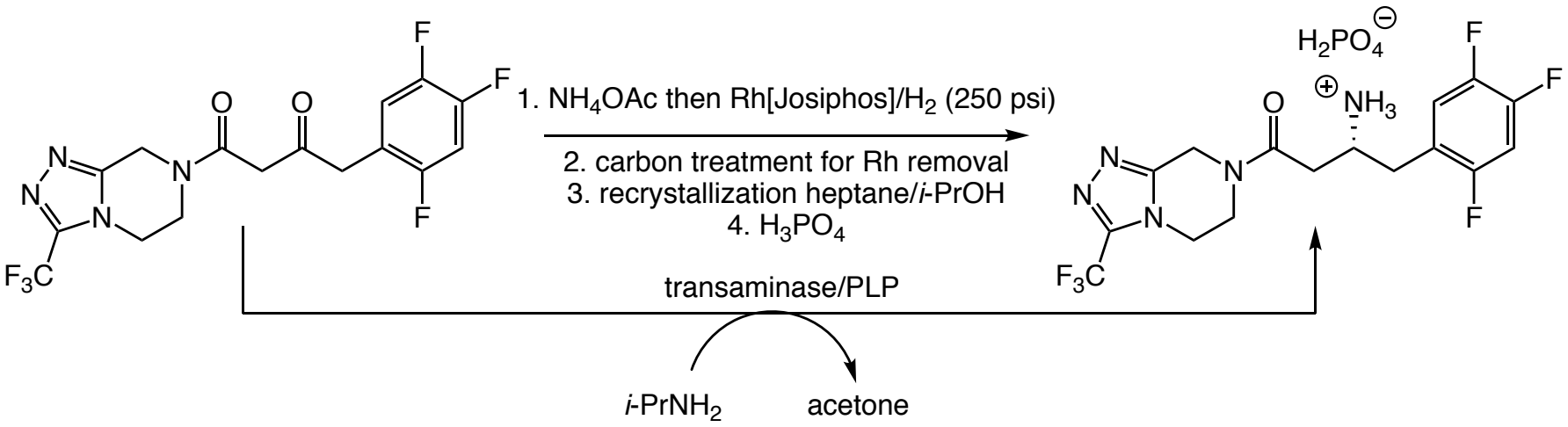
Merck's Sitagliptin Synthesis



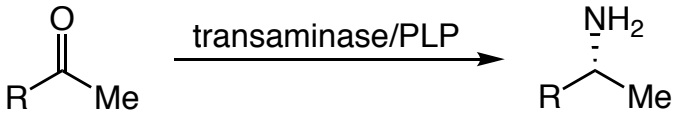
ezymatic screens of natural transaminases all showed no reaction

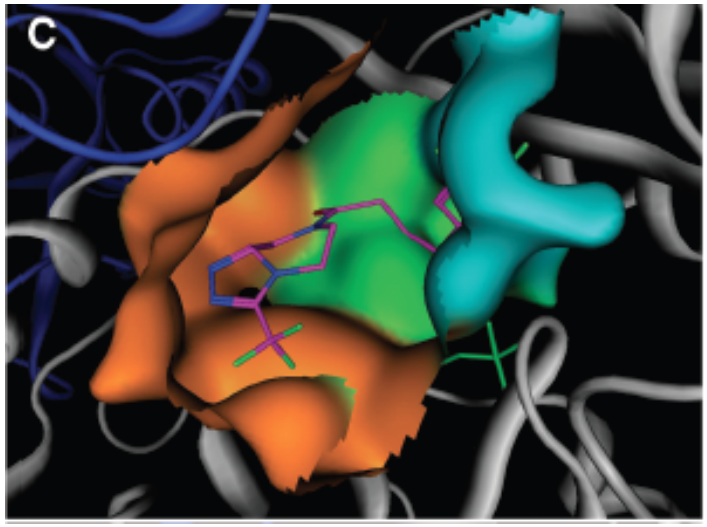
Savile, *Science*, **2010**, *329*, 305–309.

Merck's Sitagliptin Synthesis



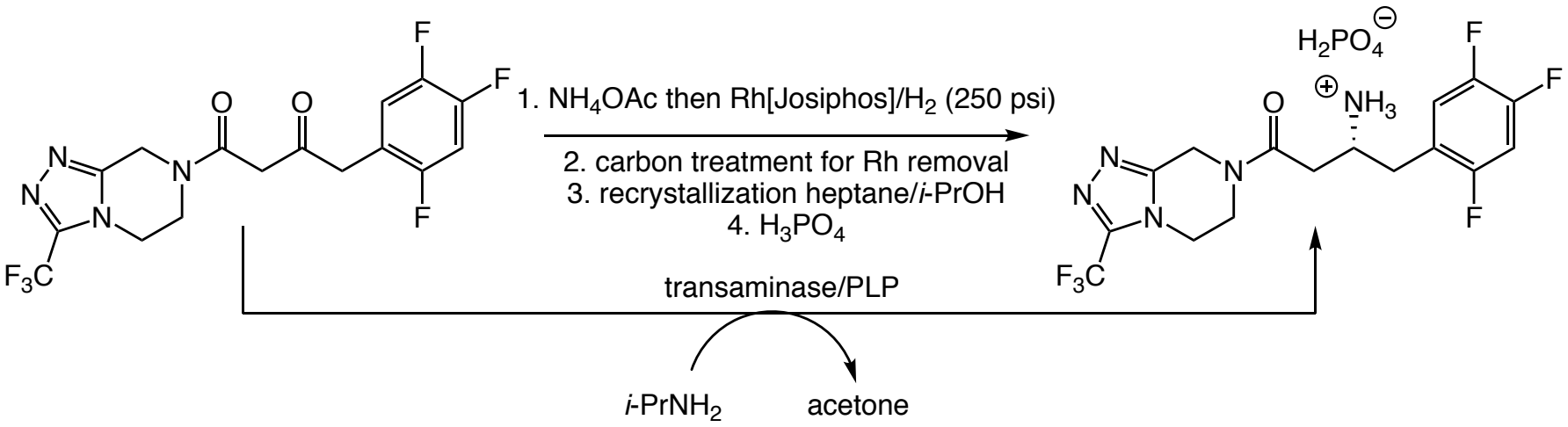
ezymatic screens of natural transaminases all showed no reaction



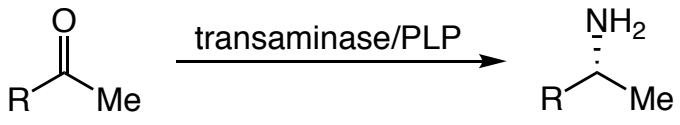


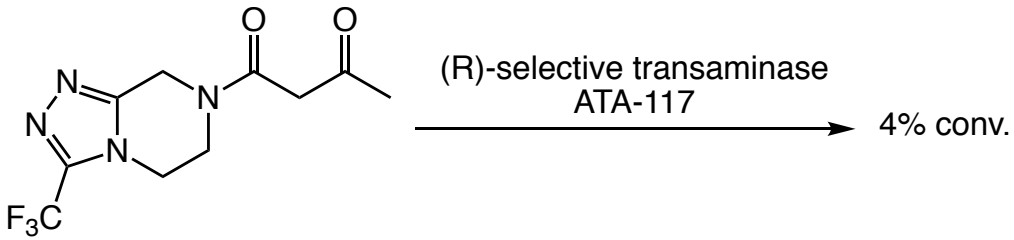
Savile, *Science*, **2010**, *329*, 305–309.

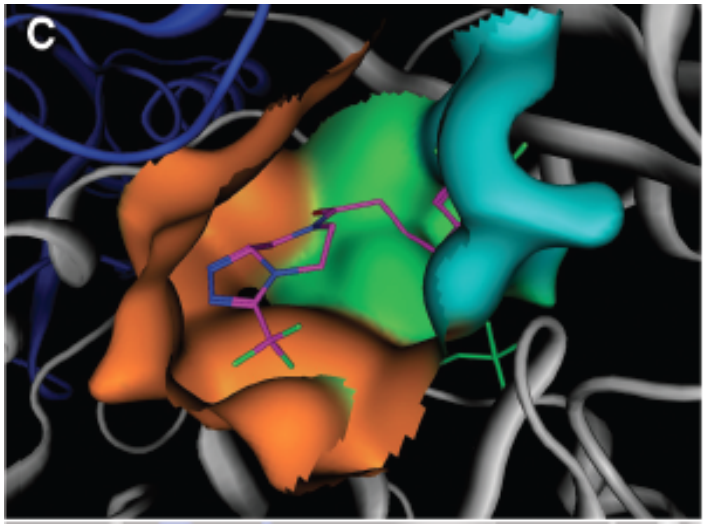
Merck's Sitagliptin Synthesis



ezymatic screens of natural transaminases all showed no reaction

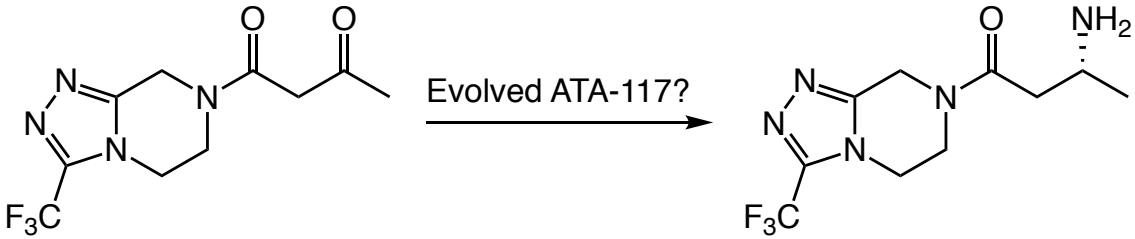






Savile, *Science*, **2010**, *329*, 305–309.

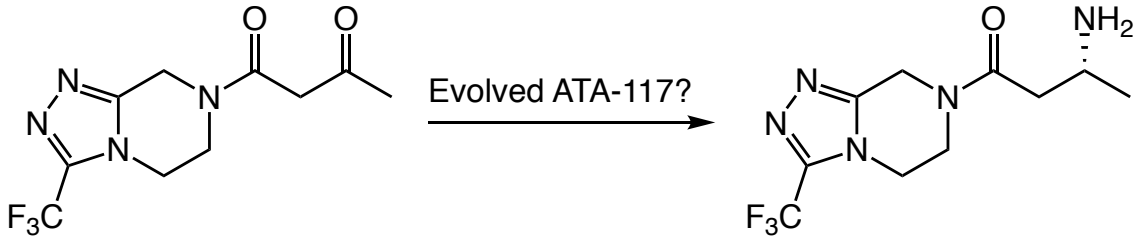
Merck's Sitagliptin Synthesis



ATA-117 —(homology analysis; in silico modeling for large pocket Ser223Pro)→

Savile, *Science*, **2010**, *329*, 305–309.

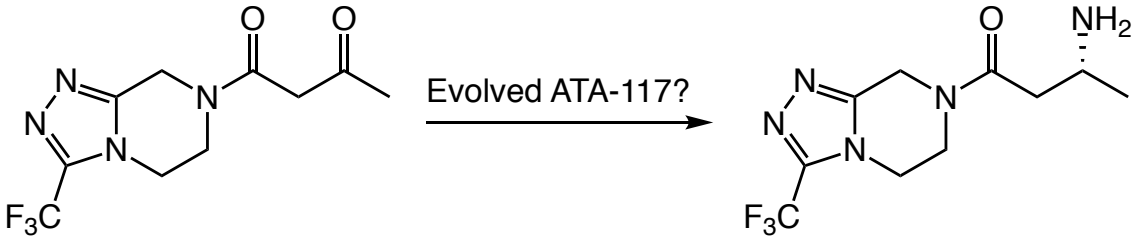
Merck's Sitagliptin Synthesis



ATA-117 —(homology analysis; in silico modeling for large pocket Ser223Pro)→ 11 fold increase toward model

Savile, *Science*, **2010**, *329*, 305–309.

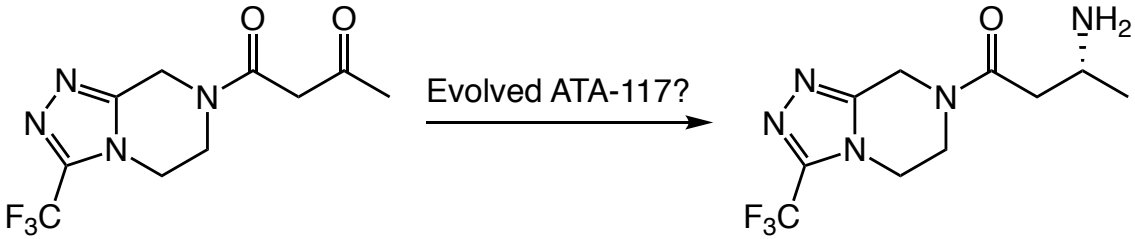
Merck's Sitagliptin Synthesis



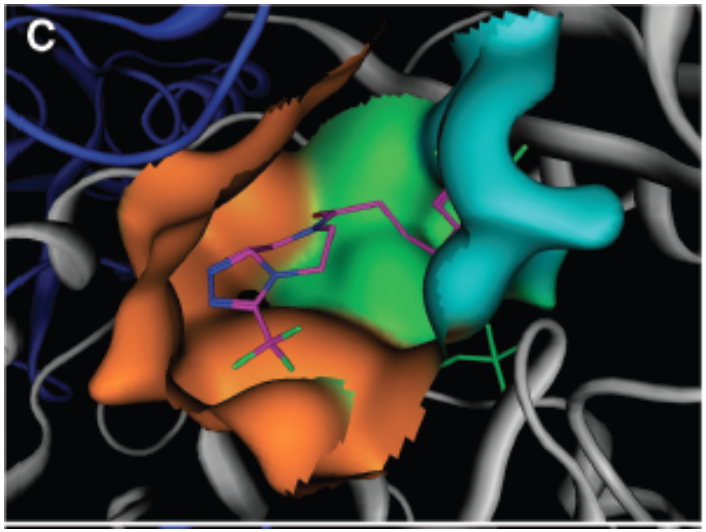
ATA-117 → (homology analysis; in silico modeling for large pocket Ser223Pro) → 11 fold increase toward model → (repeat for small pocket; Sat. Mut. 4 sites 5 different AAs) → ATA-117 Ser223Pro; Val69Gly; Phe122Ile; Ala284Gly; first detectable conversion of desired substrate 0.7%

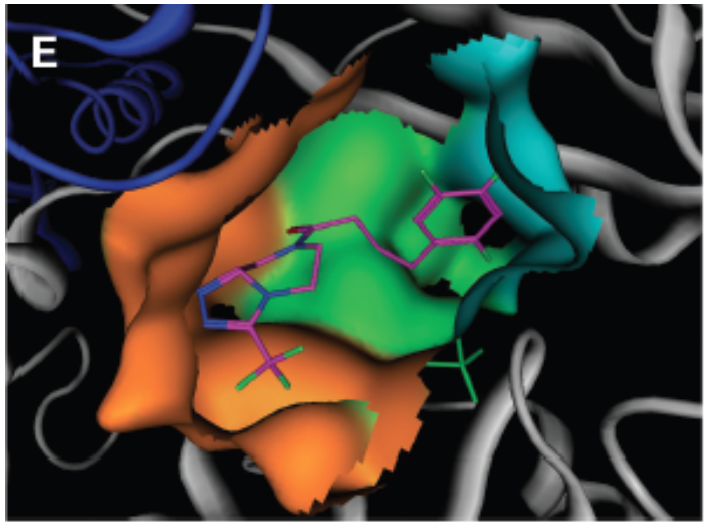
Savile, *Science*, **2010**, *329*, 305–309.

Merck's Sitagliptin Synthesis



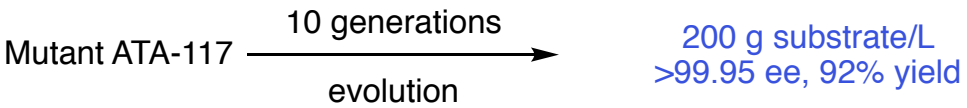
ATA-117 → homology analysis; in silico modeling for large pocket Ser223Pro → 11 fold increase toward model → repeat for small pocket; Sat. Mut. 4 sites 5 different AAs → ATA-117 Ser223Pro; Val69Gly; Phe122Ile; Ala284Gly; first detectable conversion of desired substrate 0.7%





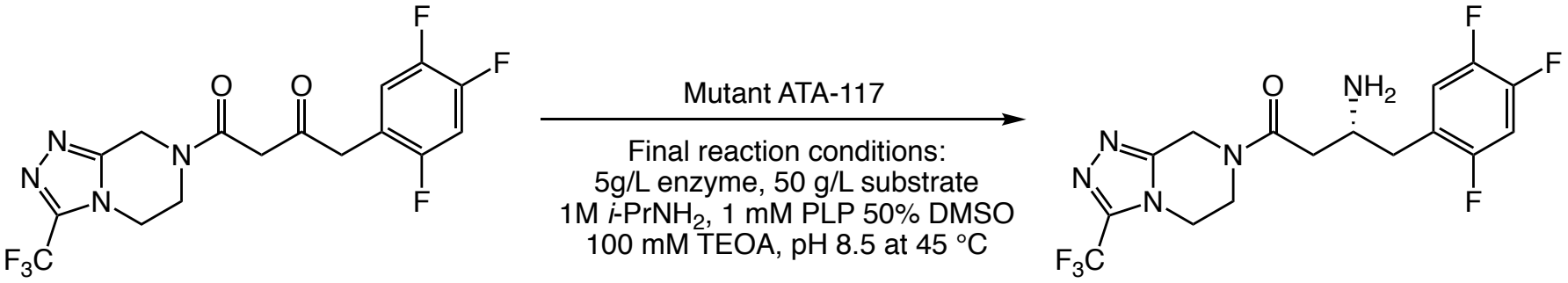
Savile, *Science*, **2010**, *329*, 305–309.

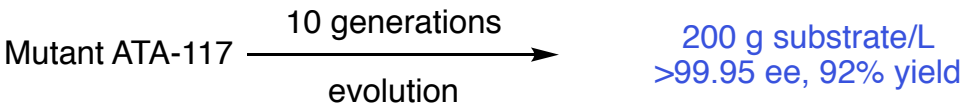
Merck's Sitagliptin Synthesis



Savile, *Science*, **2010**, *329*, 305–309.

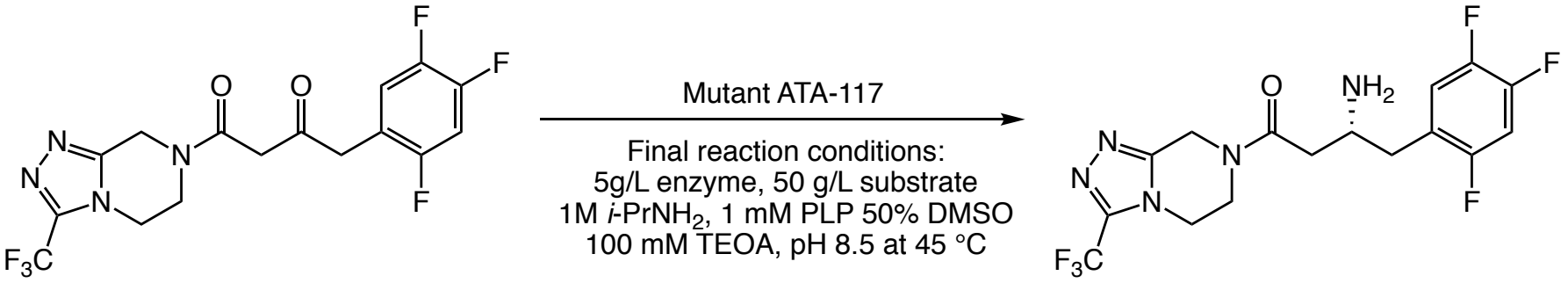
Merck's Sitagliptin Synthesis





Savile, *Science*, **2010**, *329*, 305–309.

Merck's Sitagliptin Synthesis



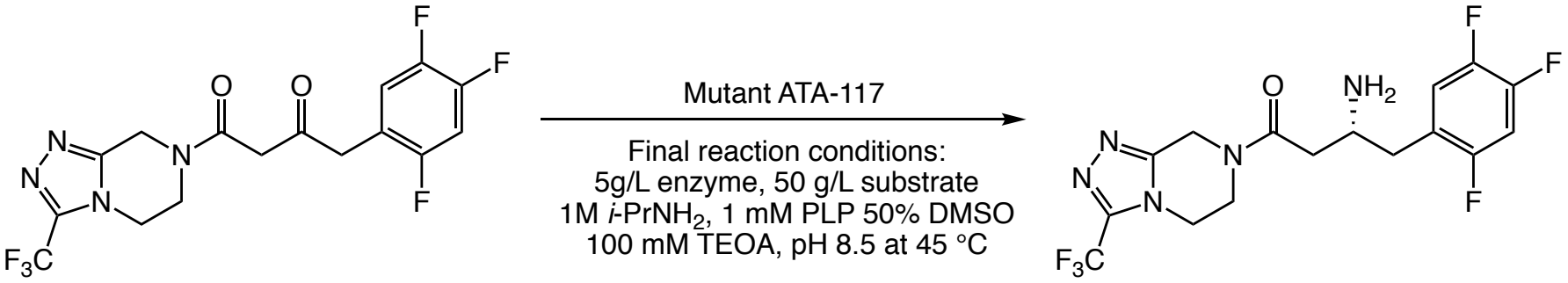
Mutant ATA-117 —(10 generations / evolution)→ 200 g substrate/L, >99.95 ee, 92% yield

– 30000 fold increase in activity
– Types of directed evolution included: structure aided homolgy, sat. mut., and random mut.
– 36,480 mutants screened
– 27 total mutations

Mutation Locations	Number of Mutations
small binding site	5
large binding site	5
peripheral of binding	17

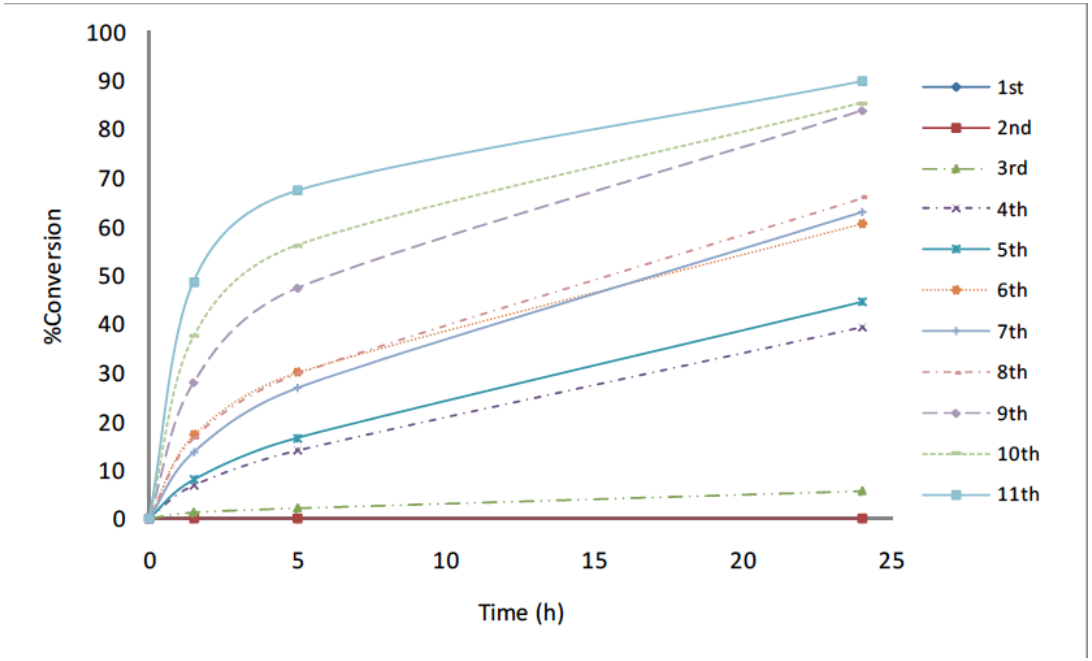
Savile, *Science*, **2010**, *329*, 305–309.

Merck's Sitagliptin Synthesis



Mutant ATA-117 —(10 generations, evolution)→ 200 g substrate/L, >99.95 ee, 92% yield

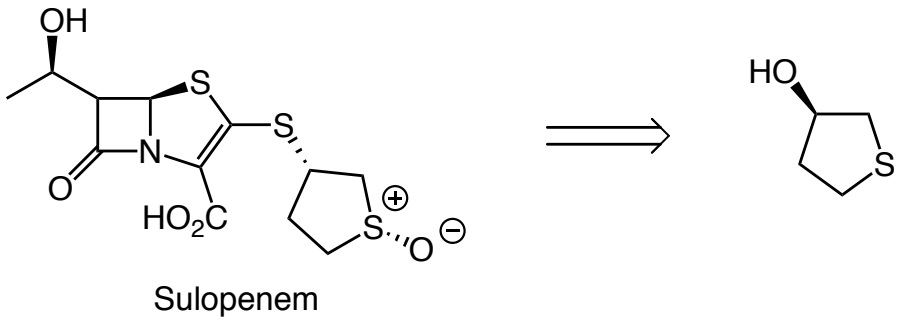
– 30000 fold increase in activity
– Types of directed evolution included: structure aided homolgy, sat. mut., and random mut.
– 36,480 mutants screened
– 27 total mutations



Mutation Locations	Number of Mutations
small binding site	5
large binding site	5
peripheral of binding	17

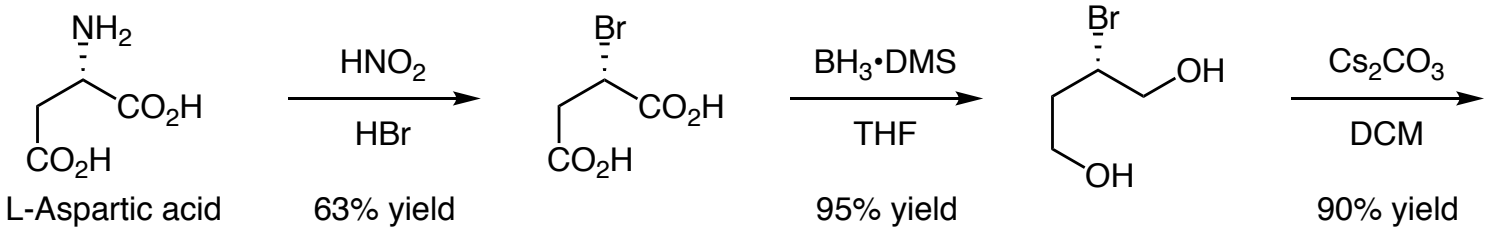
Savile, *Science*, **2010**, *329*, 305–309.

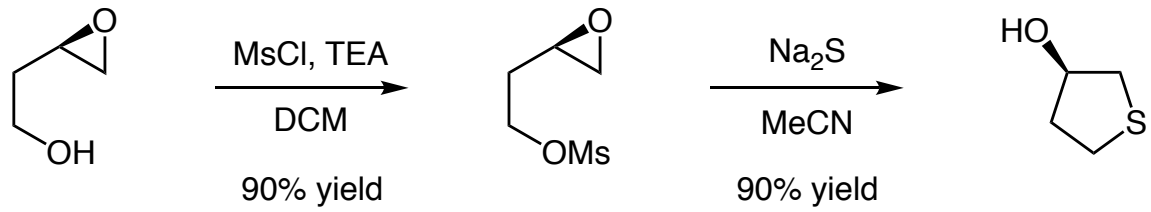
Codexis Enantioselective Ketoreductase



Sulopenem

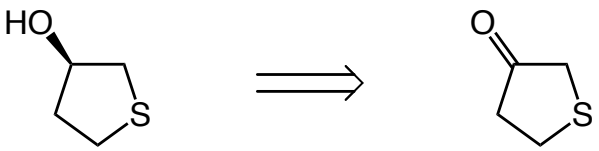
Original synthesis fom the chiral pool



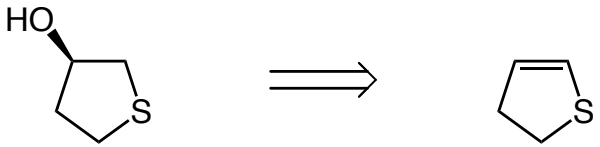


Liang *Org. Process Res. Dev.* **2010**, *14*, 188–192.

Codexis Enantioselective Ketoreductase



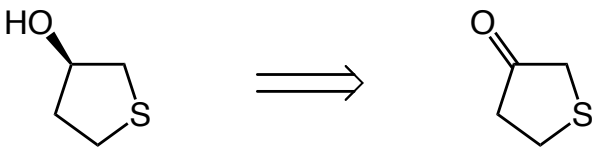
Asymmetric reduction



Asymmetric oxidation

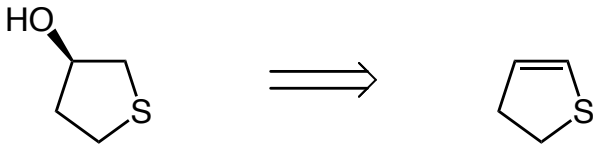
Liang *Org. Process Res. Dev.* **2010**, *14*, 188–192.
Brown, *J. Am. Chem. Soc.* **1986**, *108*, 2049–2054.
Mashima, *J. Am. Chem. Soc.* **1993**, *115*, 3318–3319.

Codexis Enantioselective Ketoreductase



Asymmetric reduction

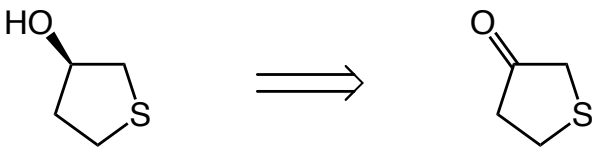
CBS catalyst 23% ee
Ir-BINAP 66–82% ee
horse liver ADH 33% ee
whole-cell fungal 81 – 91% ee



Asymmetric oxidation

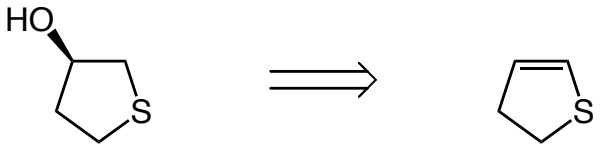
Liang *Org. Process Res. Dev.* **2010**, *14*, 188–192.
Brown, *J. Am. Chem. Soc.* **1986**, *108*, 2049–2054.
Mashima, *J. Am. Chem. Soc.* **1993**, *115*, 3318–3319.

Codexis Enantioselective Ketoreductase



Asymmetric reduction

CBS catalyst 23% ee
Ir-BINAP 66–82% ee
horse liver ADH 33% ee
whole-cell fungal 81 – 91% ee

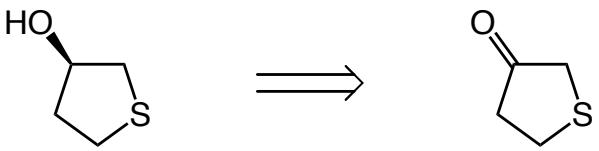


Asymmetric oxidation

Ipc_2BH - 100% ee

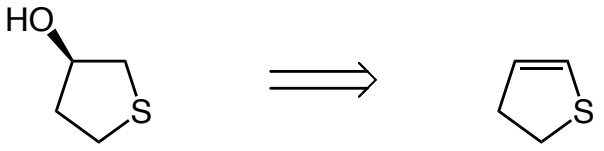
Liang *Org. Process Res. Dev.* **2010**, *14*, 188–192.
Brown, *J. Am. Chem. Soc.* **1986**, *108*, 2049–2054.
Mashima, *J. Am. Chem. Soc.* **1993**, *115*, 3318–3319.

Codexis Enantioselective Ketoreductase



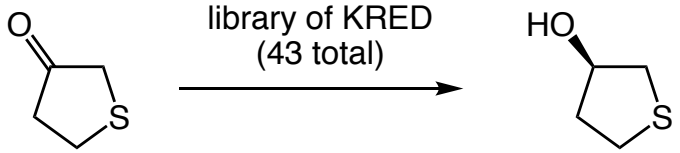
Asymmetric reduction

CBS catalyst 23% ee
Ir-BINAP 66–82% ee
horse liver ADH 33% ee
whole-cell fungal 81 – 91% ee



Asymmetric oxidation

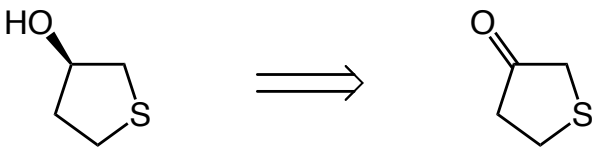
Ipc_2BH - 100% ee



KRED *L. kefir*
63% ee

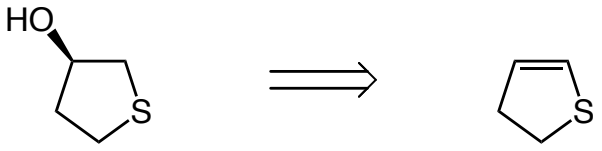
Liang *Org. Process Res. Dev.* **2010**, *14*, 188–192.
Brown, *J. Am. Chem. Soc.* **1986**, *108*, 2049–2054.
Mashima, *J. Am. Chem. Soc.* **1993**, *115*, 3318–3319.

Codexis Enantioselective Ketoreductase



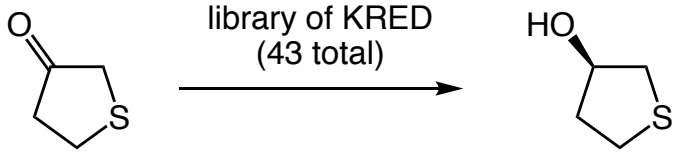
Asymmetric reduction

CBS catalyst 23% ee
Ir-BINAP 66–82% ee
horse liver ADH 33% ee
whole-cell fungal 81 – 91% ee



Asymmetric oxidation

Ipc_2BH - 100% ee



KRED *L. kefir*
63% ee

Parent KRED
63% ee

8 generations
(10 total mutations)

methods used:
random mut.
gene shuffling
semisynthetic shuffling
in silico computation

mutant KRED
88% yield, 99.3% ee

Liang *Org. Process Res. Dev.* **2010**, *14*, 188–192.
Brown, *J. Am. Chem. Soc.* **1986**, *108*, 2049–2054.
Mashima, *J. Am. Chem. Soc.* **1993**, *115*, 3318–3319.

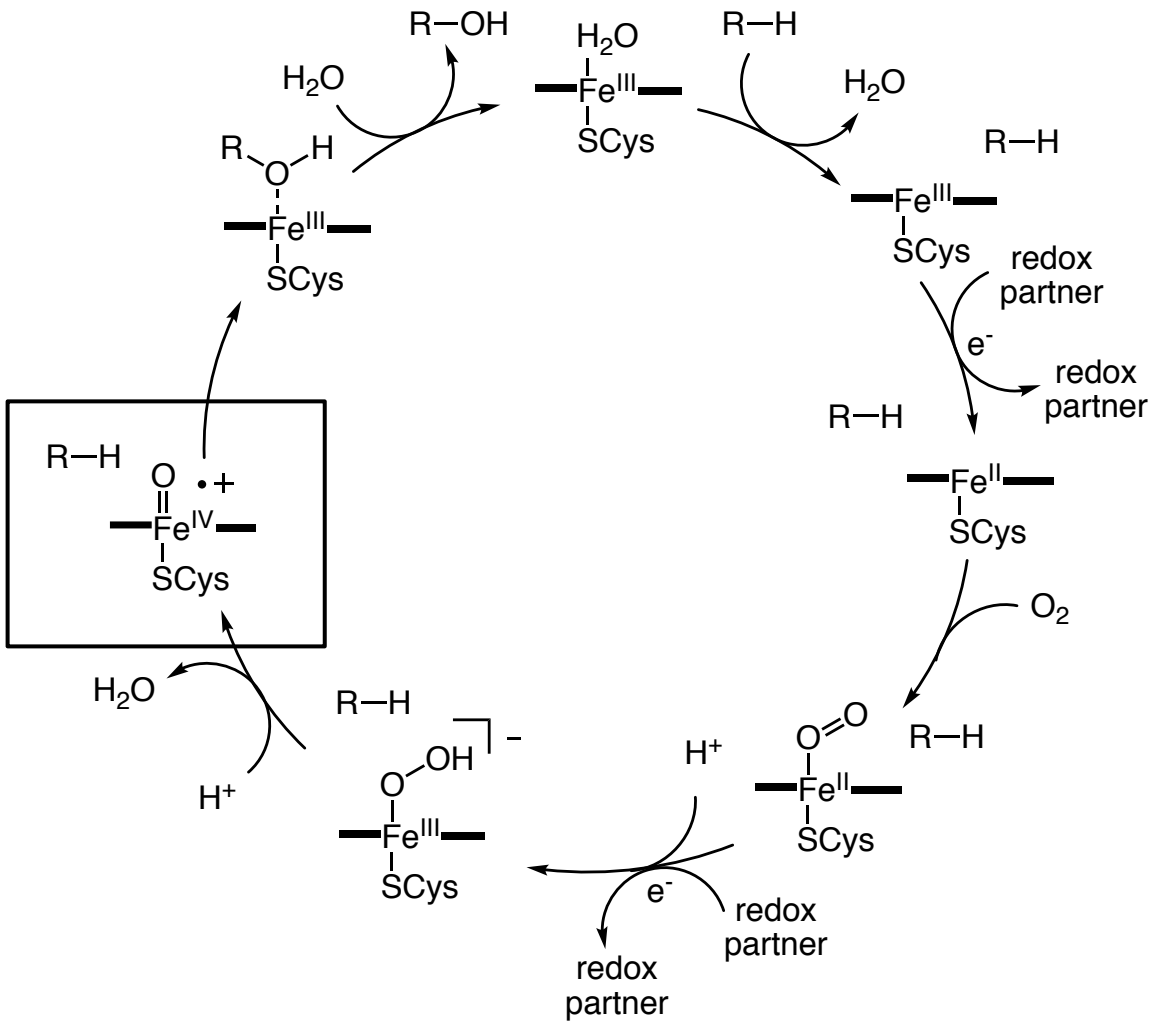
New to Nature Enzymatic Catalysis

3 types of development of new to nature enzymes:

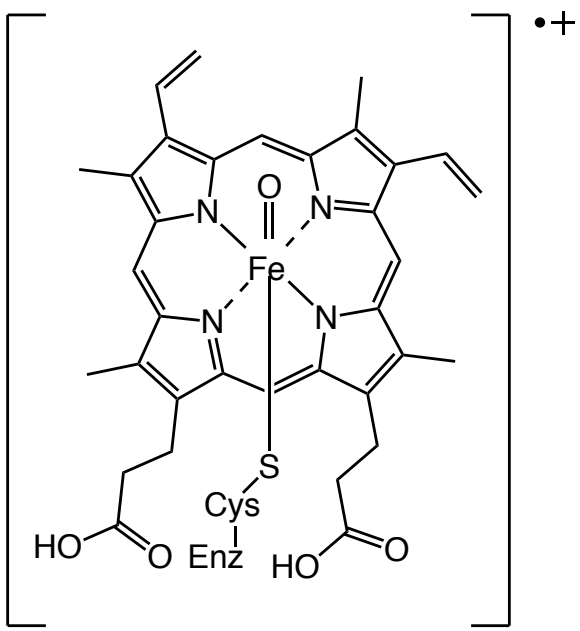
- Repurposing native enzymes
- Redesign/ de novo synthesis
- Artificial cofactors

Ward, *Trends Biotechnol.* **2018**, *36*, 60–72.

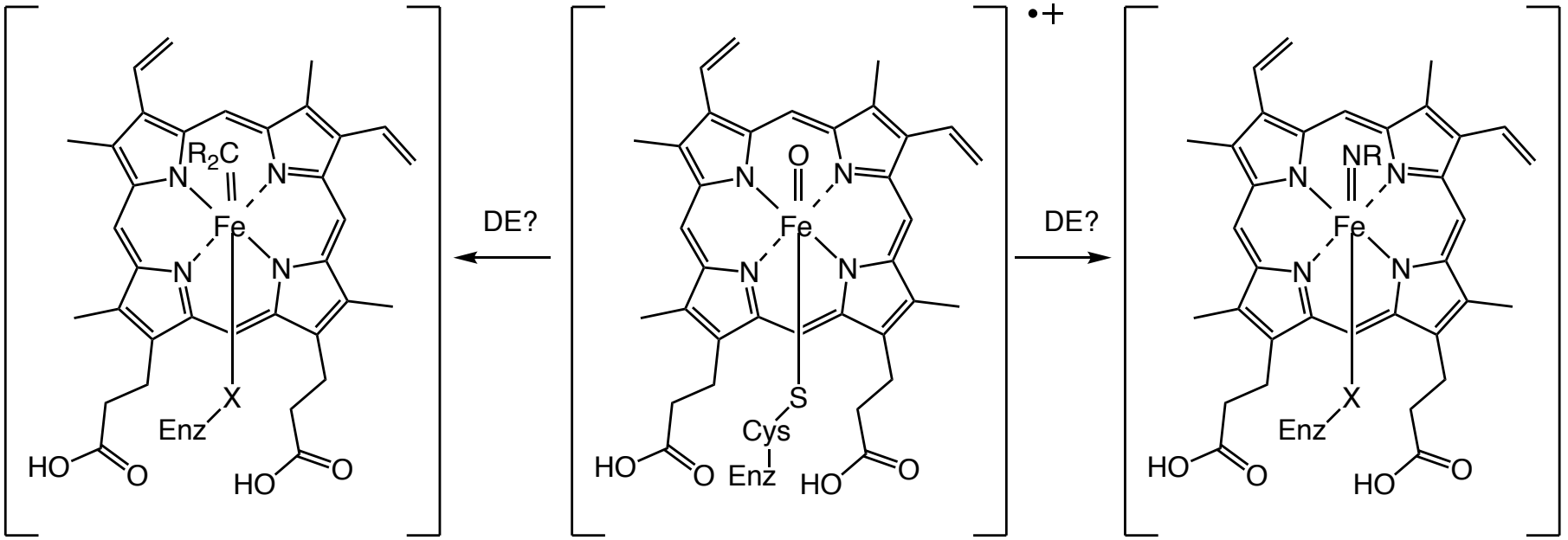
P450 catalytic cycle



Enzymatic Carbenoids and Nitrenoids

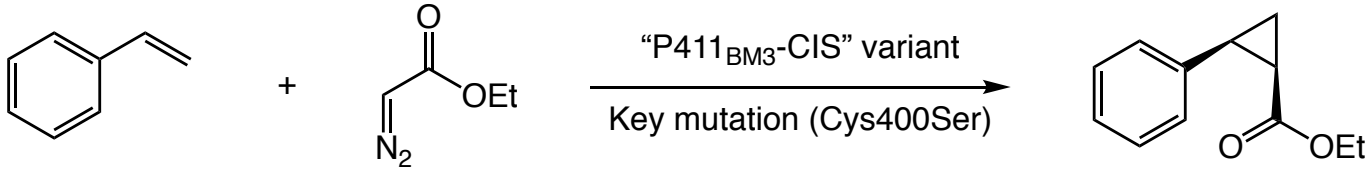


Enzymatic Carbenoids and Nitrenoids



Enzymatic Carbenoids

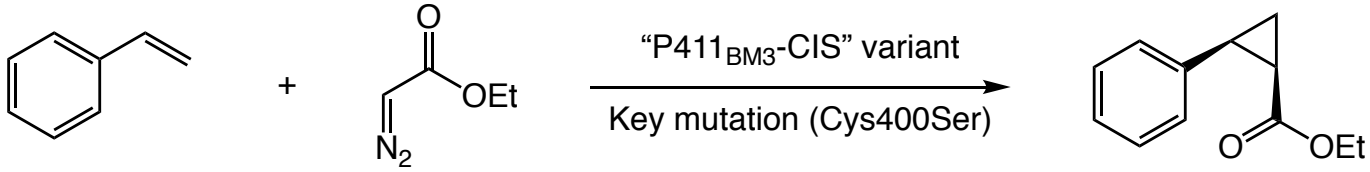
Cyclopropanations (Arnold, 2013)



72% yield, 99% ee
9:1 cis:trans
67,800 TTN

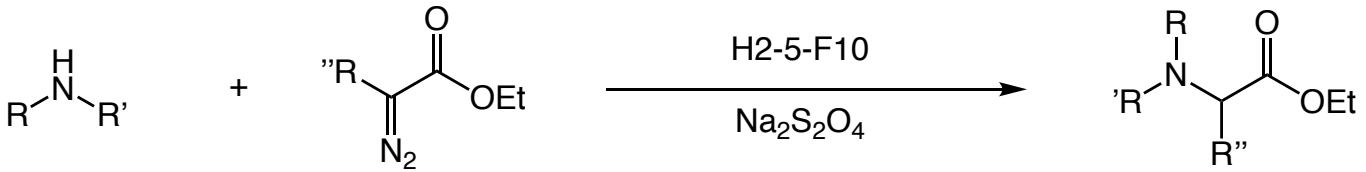
Enzymatic Carbenoids

Cyclopropanations (Arnold, 2013)



72% yield, 99% ee
9:1 cis:trans
67,800 TTN

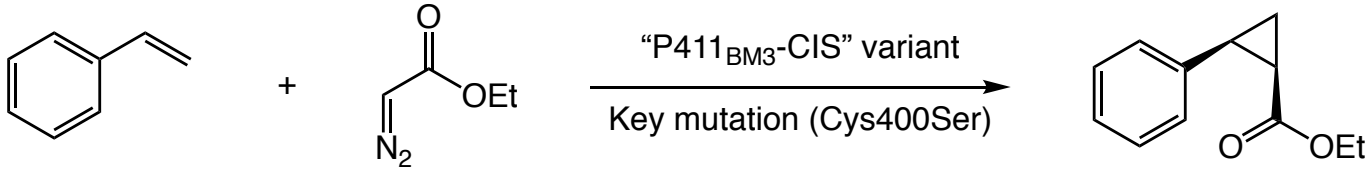
N–H insertions (Arnold, 2014)



37–83% yield
130–354 TTN

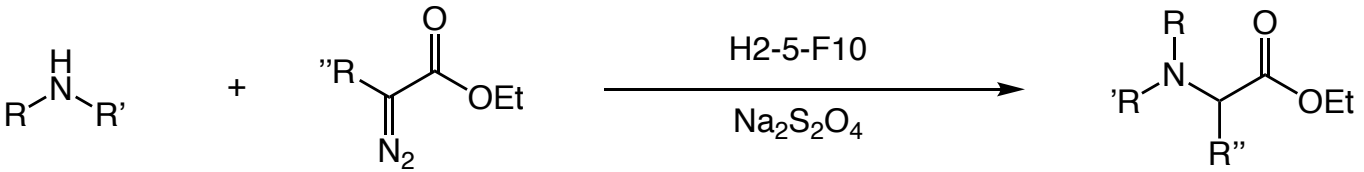
Enzymatic Carbenoids

Cyclopropanations (Arnold, 2013)

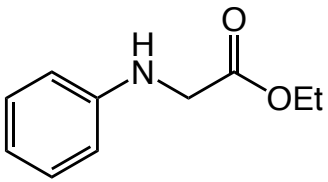


72% yield, 99% ee
9:1 cis:trans
67,800 TTN

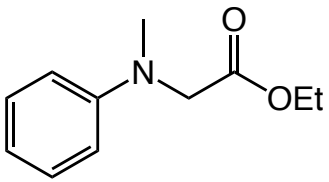
N–H insertions (Arnold, 2014)



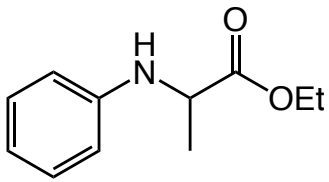
37–83% yield
130–354 TTN



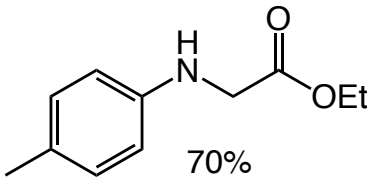
68%



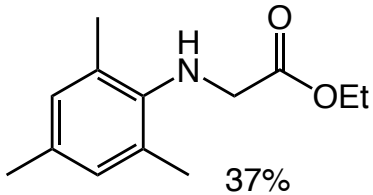
83%



26%



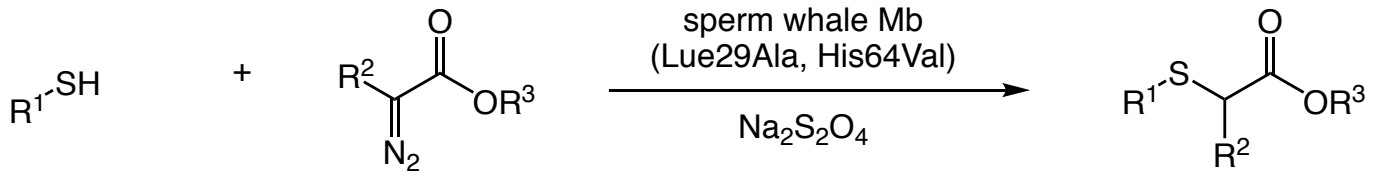
70%



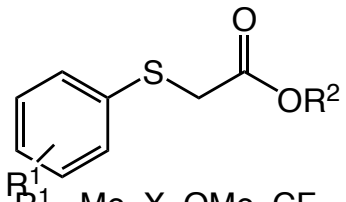
37%

Enzymatic Carbenoids

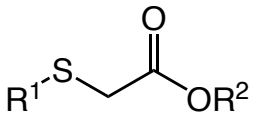
S–H insertions (Fasan, 2015)



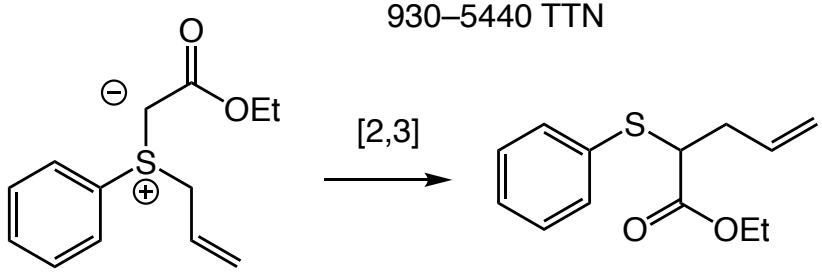
30–99% yield
930–5440 TTN



R^1= Me, X, OMe, CF_3
R^2= Et, *t*-Bu, Cy, Bn
67–99% yield
1500–5440 TTN

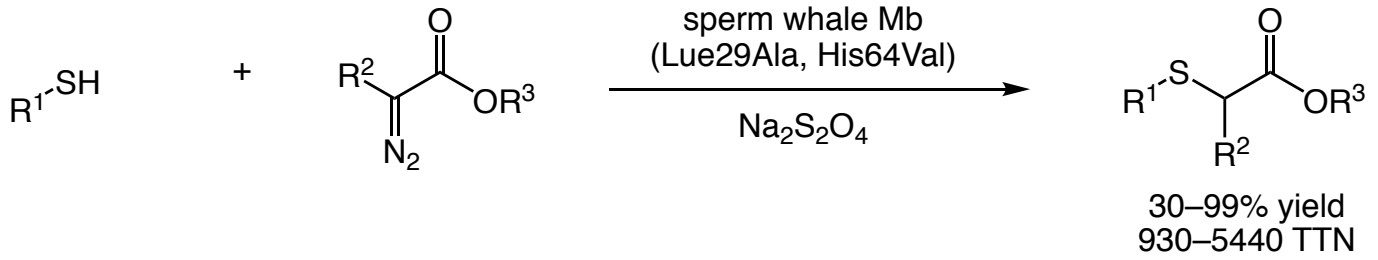


R^1= Bn, Cy, Alkyl
R^2= Et, Bn
30–83% yield
930–2550 TTN



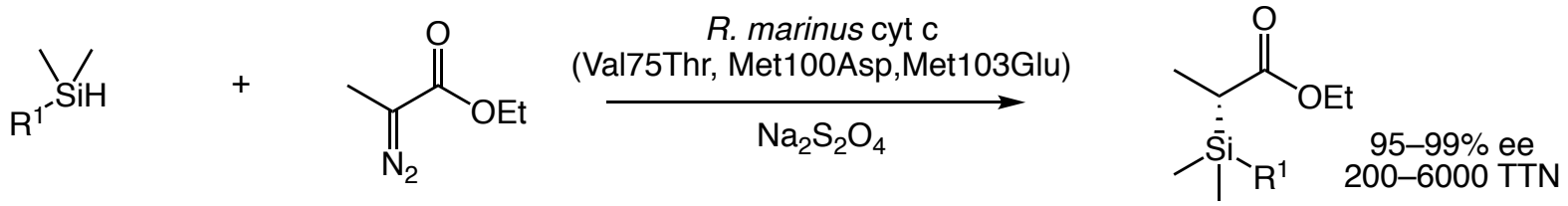
Enzymatic Carbenoids

S–H insertions (Fasan, 2015)



30–99% yield
930–5440 TTN

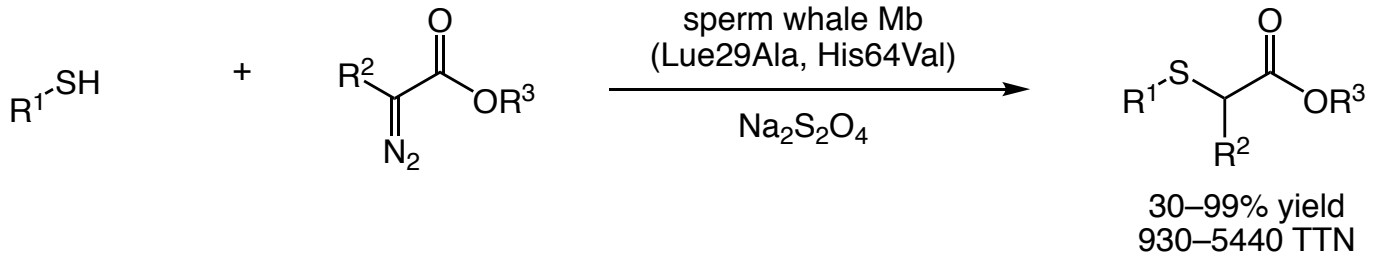
Si–H insertions (Arnold, 2016)



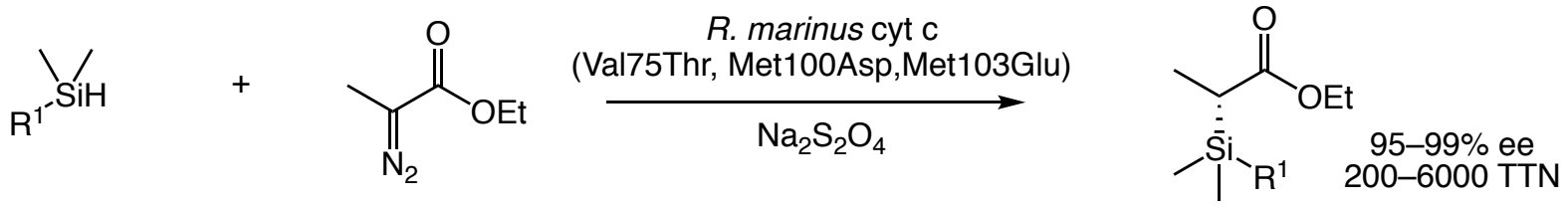
95–99% ee
200–6000 TTN

Enzymatic Carbenoids

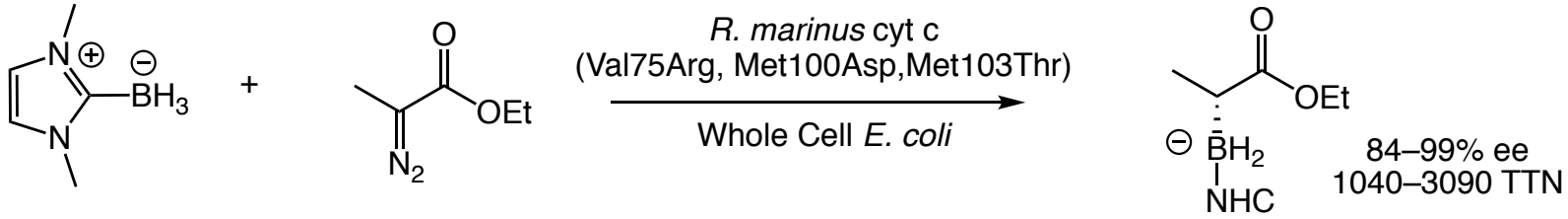
S–H insertions (Fasan, 2015)



Si–H insertions (Arnold, 2016)

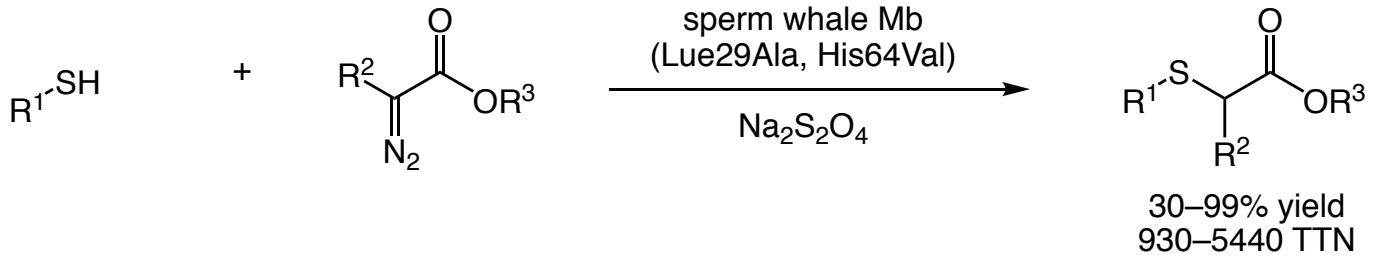


B–H insertions (Arnold, 2017)

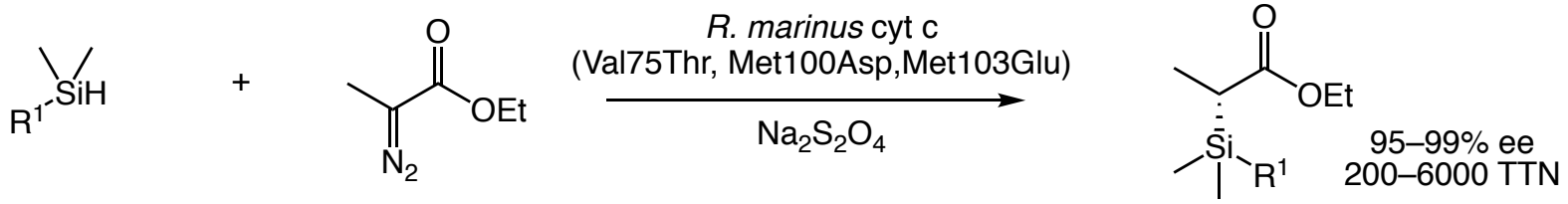


Enzymatic Carbenoids

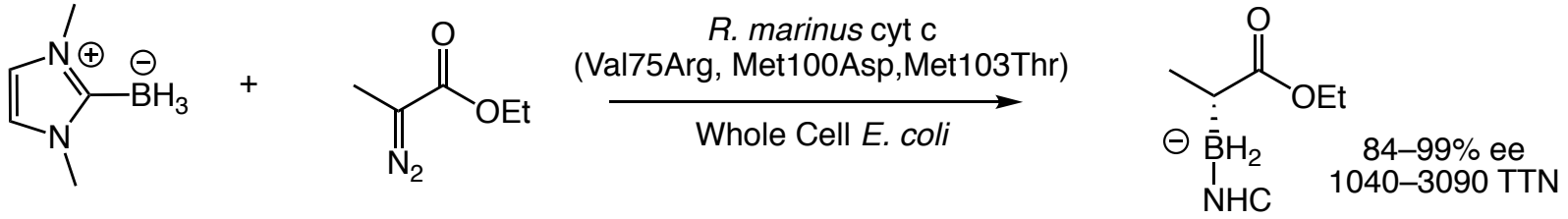
S–H insertions (Fasan, 2015)



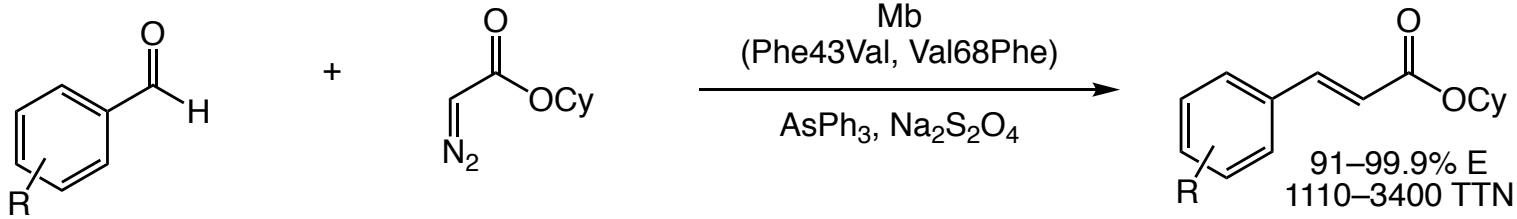
Si–H insertions (Arnold, 2016)



B–H insertions (Arnold, 2017)



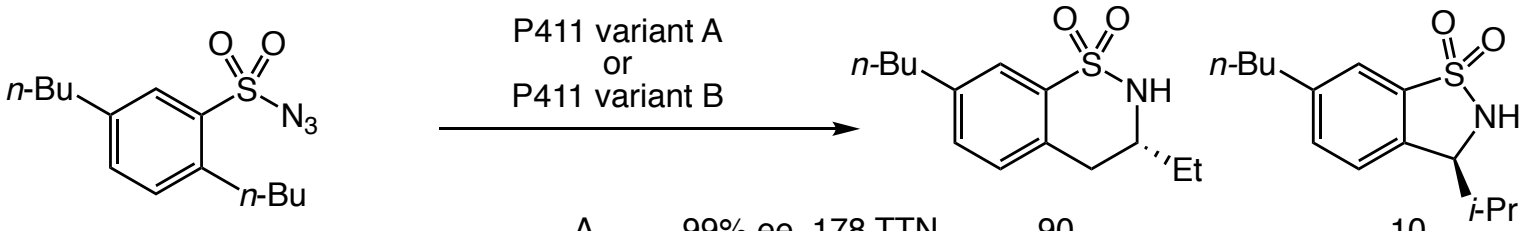
Olefinations (Fasan, 2016)



Fasan; Arnold *Curr. Opin Biotechnol.* **2017**, *47*, 102–111.

Enzymatic Nitrenoid

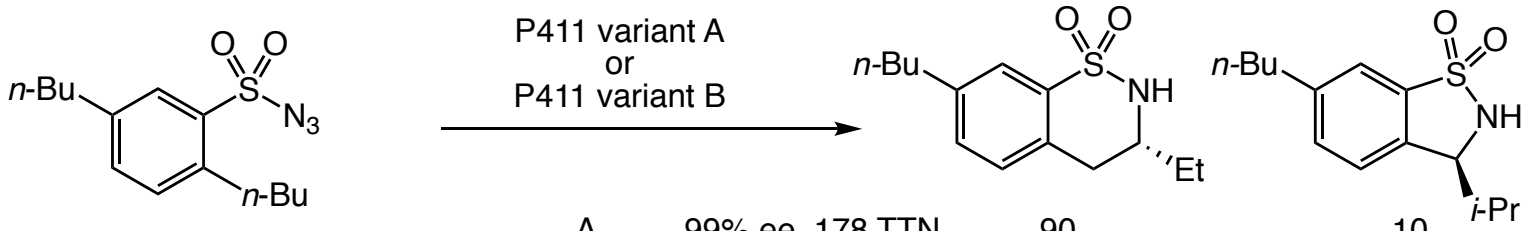
Intramolecular C–H amination (Arnold, 2014)



A	99% ee, 178 TTN	90	10
B	99% ee, 128 TTN	3	97

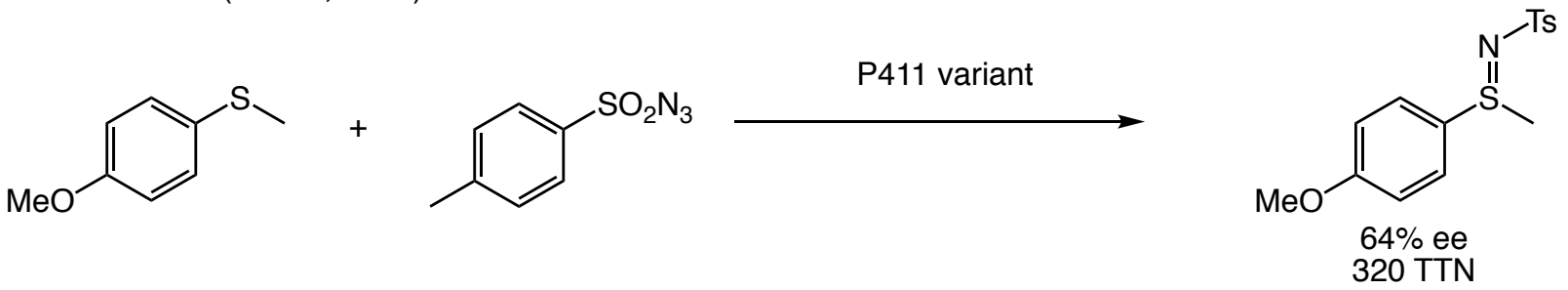
Enzymatic Nitrenoid

Intramolecular C–H amination (Arnold, 2014)



A	99% ee, 178 TTN	90	10
B	99% ee, 128 TTN	3	97

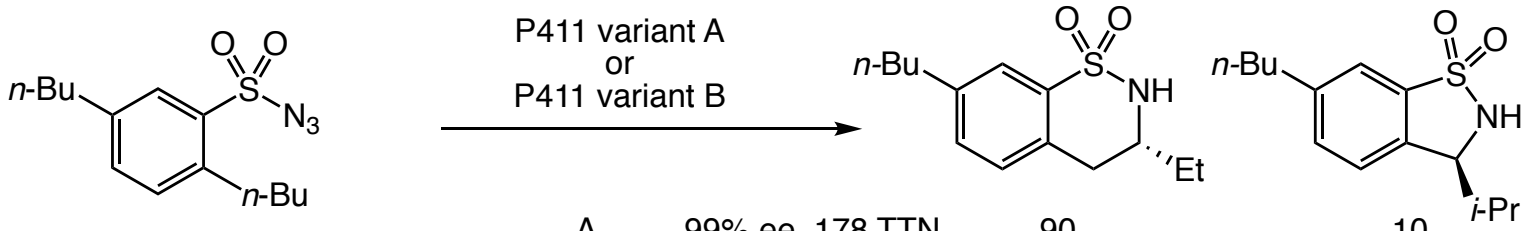
Sulfimidation (Arnold, 2014)



64% ee
320 TTN

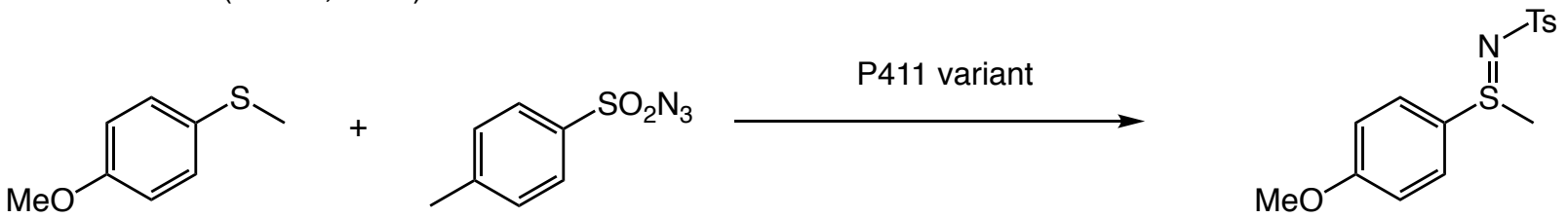
Enzymatic Nitrenoid

Intramolecular C–H amination (Arnold, 2014)



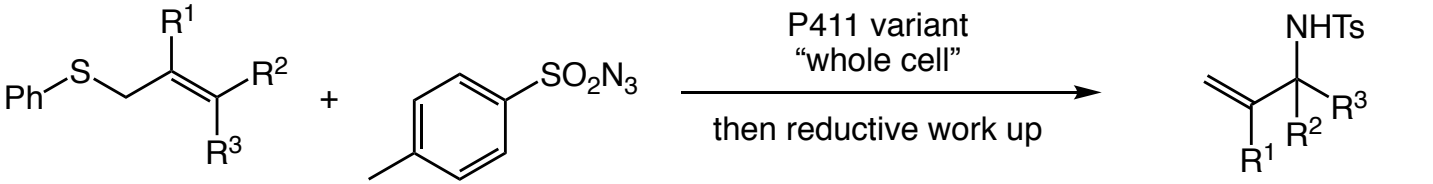
A	99% ee, 178 TTN	90	10
B	99% ee, 128 TTN	3	97

Sulfimidation (Arnold, 2014)



64% ee
320 TTN

Sulfimidation then [2,3] (Arnold, 2016)

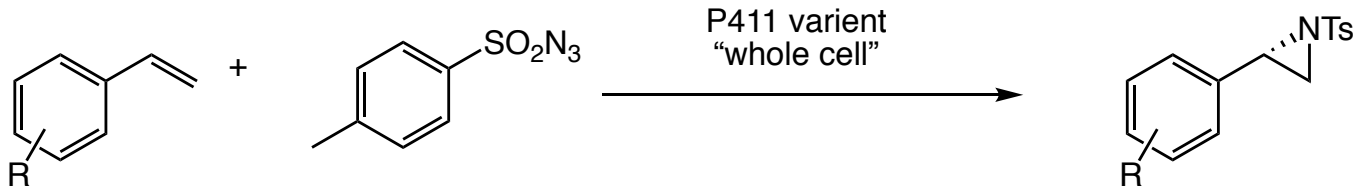


21–77% yield
68–98% ee
590 – 2200 TTN

Fasan; Arnold *Curr. Opin Biotechnol.* **2017**, *47*, 102–111.

Enzymatic Nitrenoid

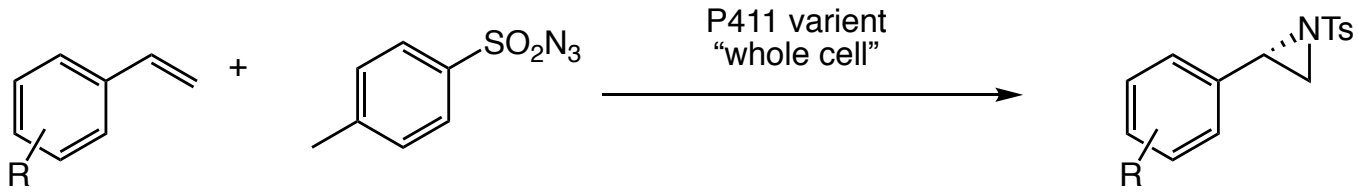
Aziridination (Arnold 2015)



2–70% yield
81–99% ee
21–600 TTN

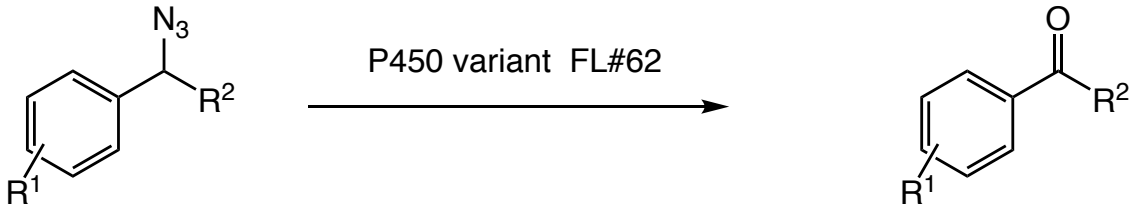
Enzymatic Nitrenoid

Aziridination (Arnold 2015)



2–70% yield
81–99% ee
21–600 TTN

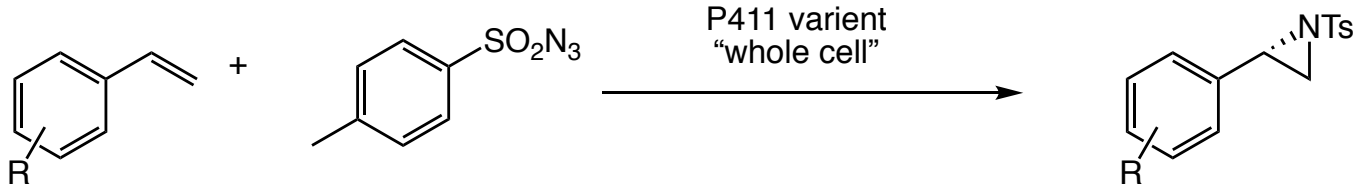
Oxidative deamination (Fasan, 2016)



1–96% yield
60 – 9660 TTN

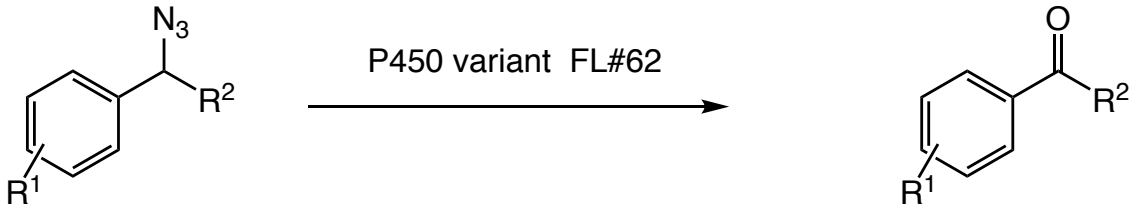
Enzymatic Nitrenoid

Aziridination (Arnold 2015)



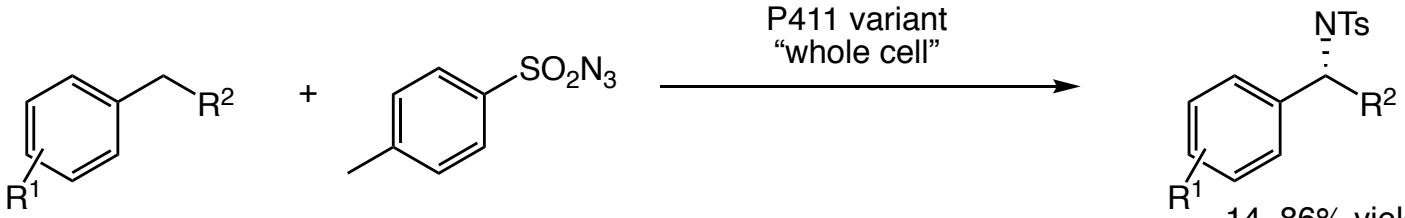
2–70% yield
81–99% ee
21–600 TTN

Oxidative deamination (Fasan, 2016)



1–96% yield
60 – 9660 TTN

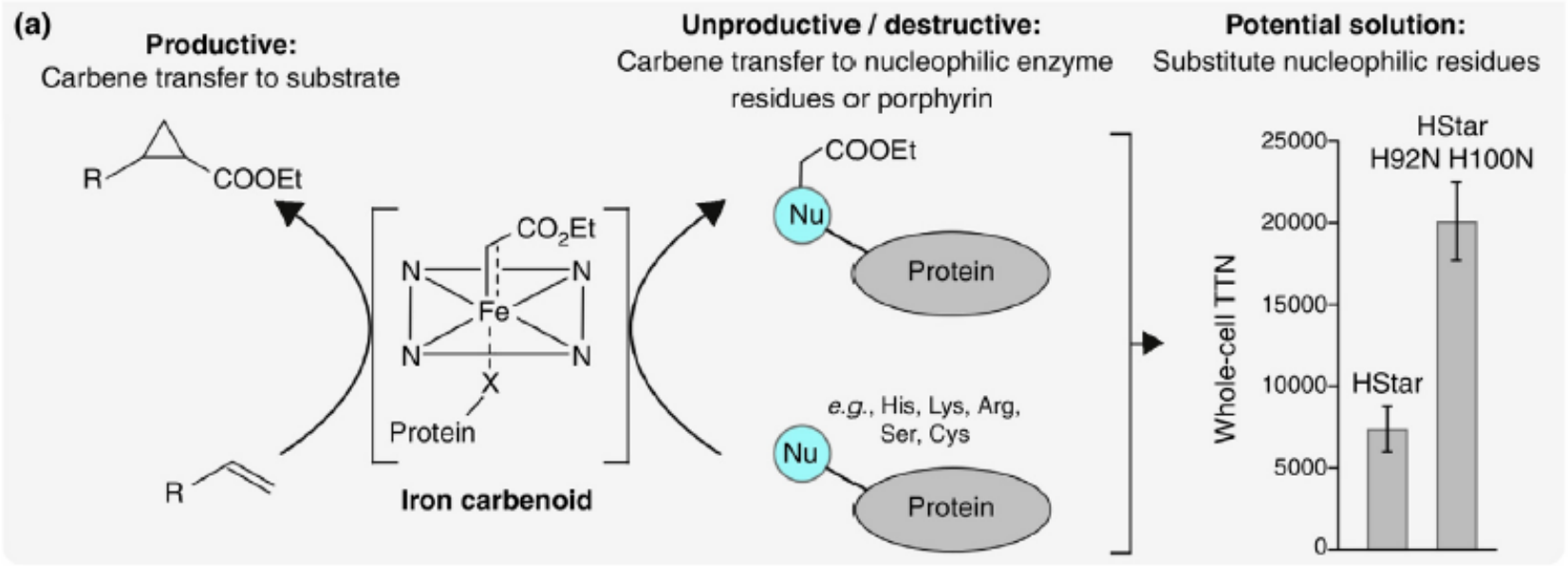
Intermolecular C–H amination (Arnold, 2017)



14–86% yield
87–99% ee
110–670 TTN

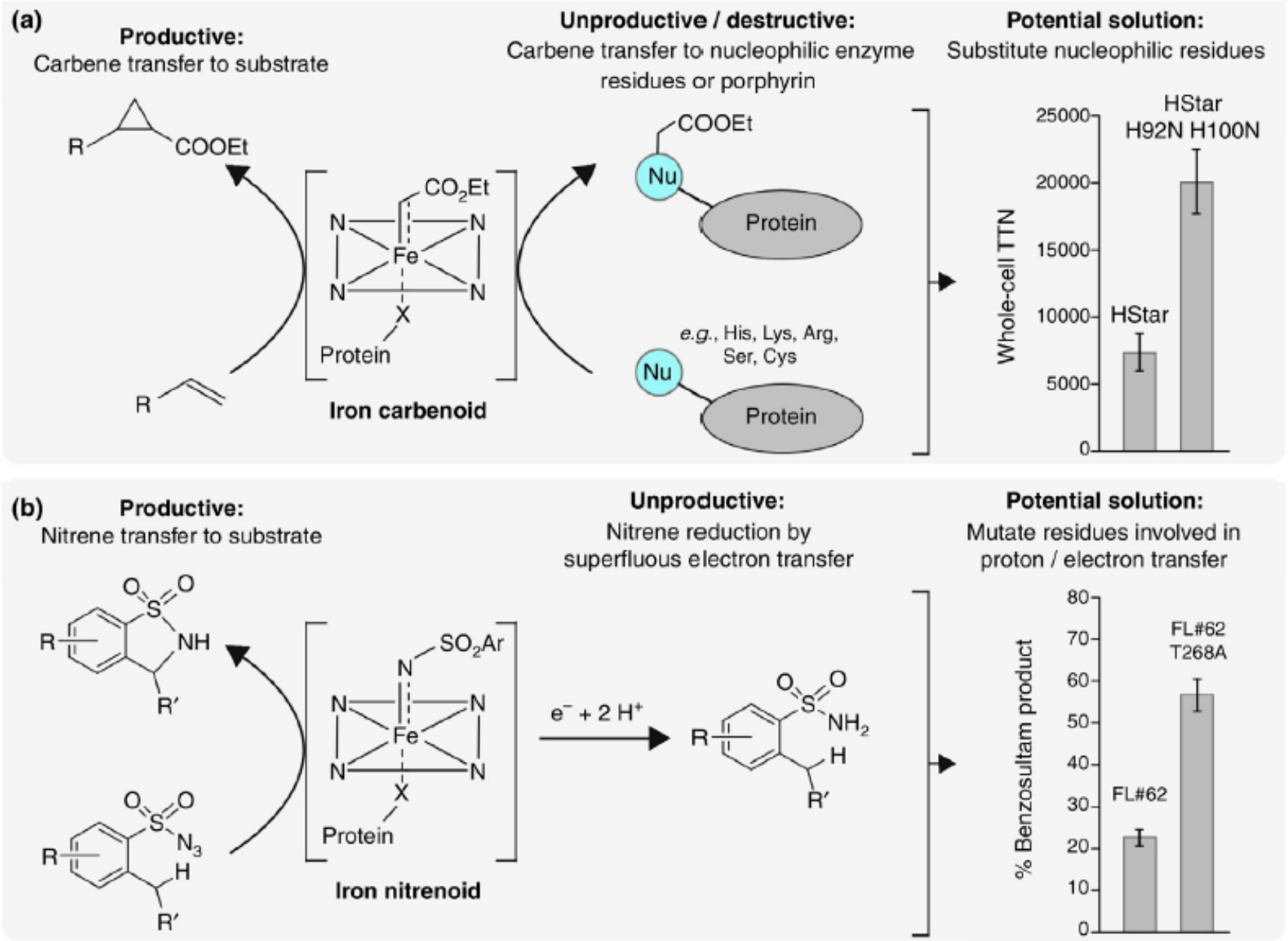
Fasan; Arnold *Curr. Opin Biotechnol.* **2017**, *47*, 102–111.

Potentially Helpful Mutations



Fasan; Arnold *Curr. Opin Biotechnol.* **2017**, *47*, 102–111.

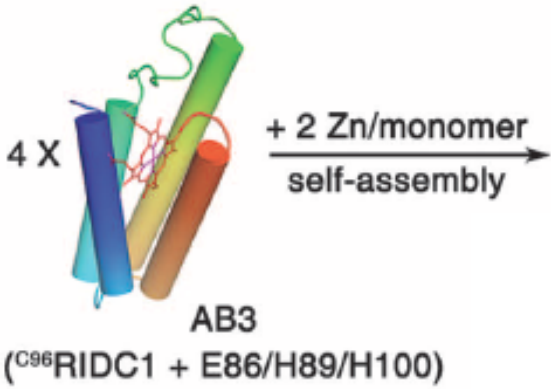
Potentially Helpful Mutations



Fasan; Arnold *Curr. Opin Biotechnol.* **2017**, *47*, 102–111.

Redesigned/ De Novo Synthesis

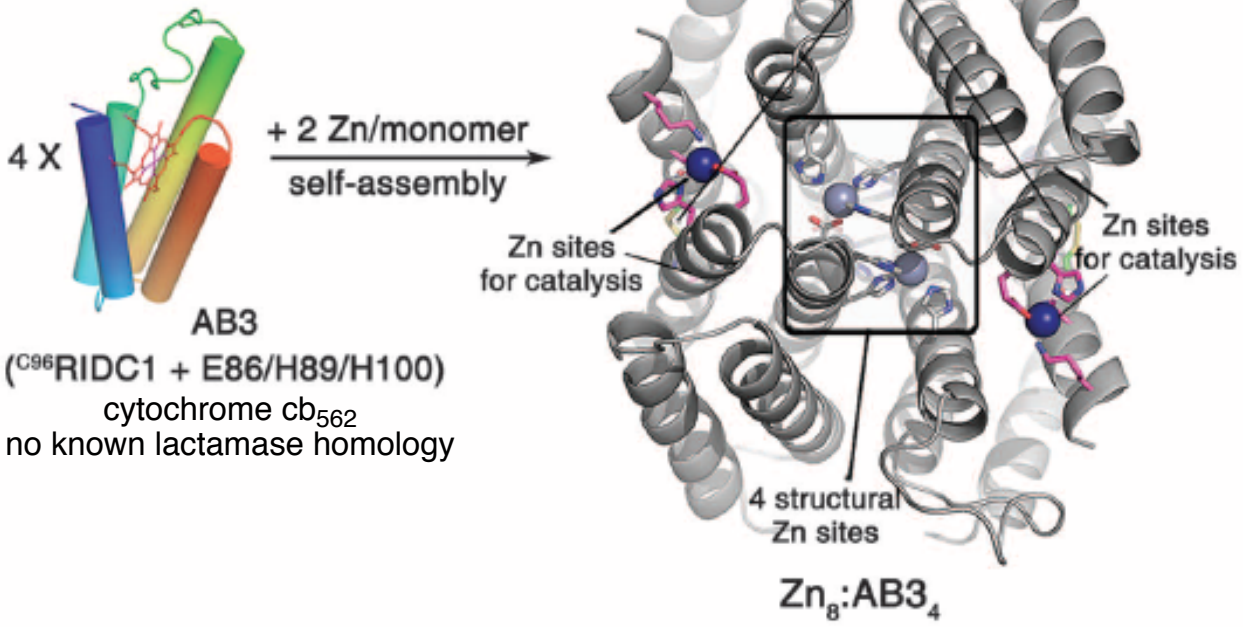
Tezcan (2014)



Tezcan *Science*, **2104**, *346*, 1525–1528.

Redesigned/ De Novo Synthesis

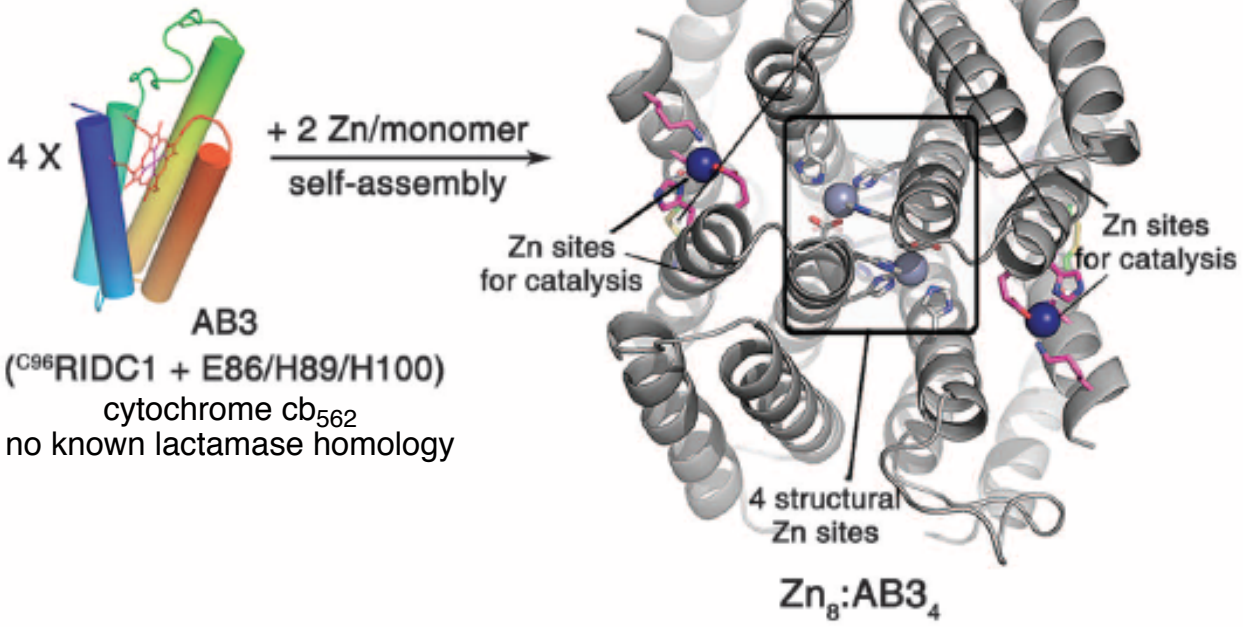
Tezcan (2014)

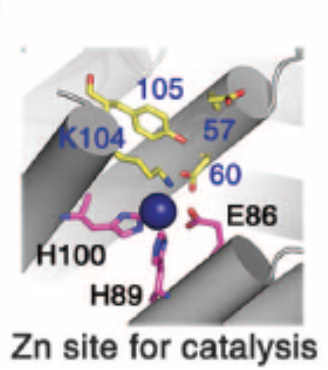


Tezcan *Science*, **2104**, *346*, 1525–1528.

Redesigned/ De Novo Synthesis

Tezcan (2014)

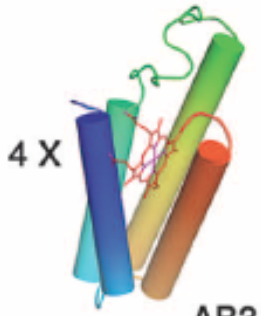




Tezcan *Science*, **2104**, *346*, 1525–1528.

Redesigned/ De Novo Synthesis

Tezcan (2014)

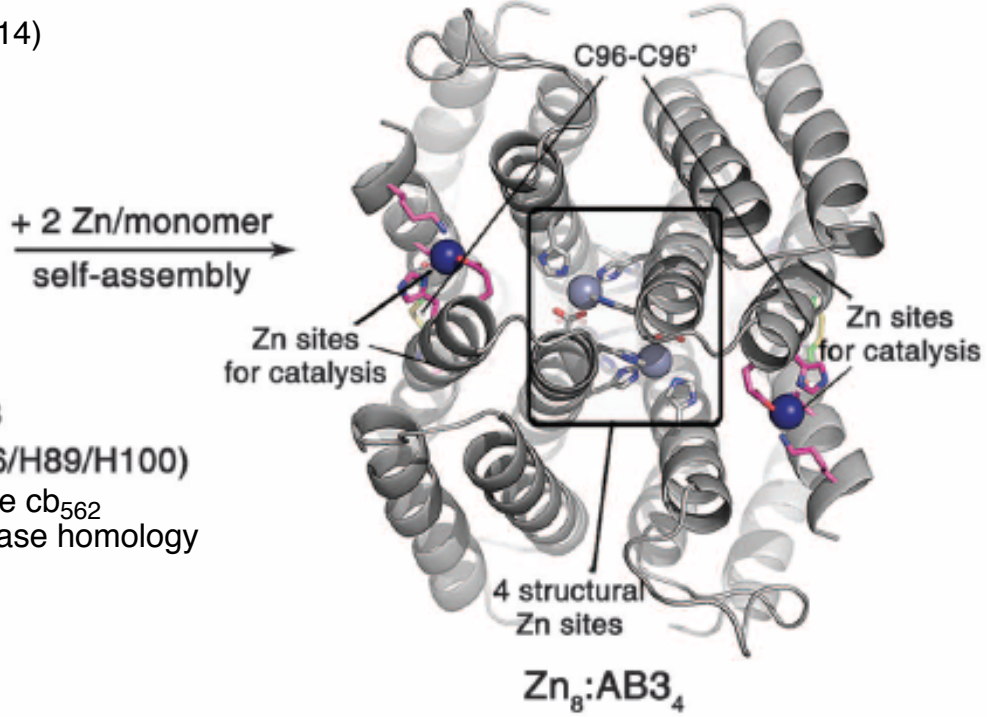


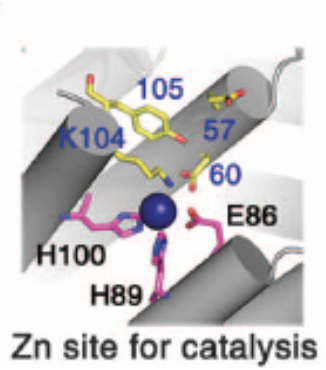
AB3

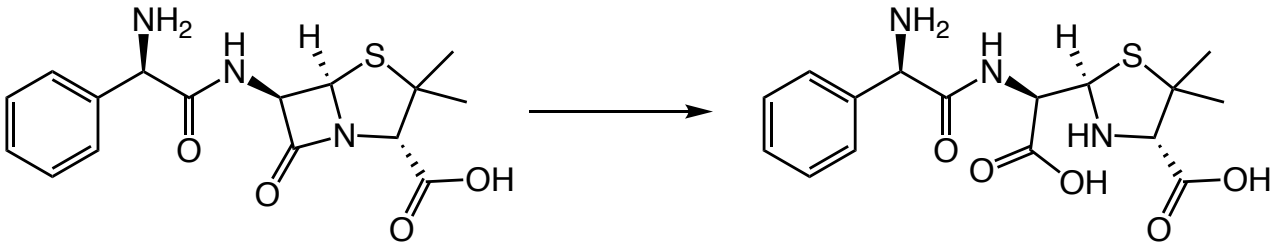
(^{C96}RIDC1 + E86/H89/H100)

cytochrome cb_{562}

no known lactamase homology



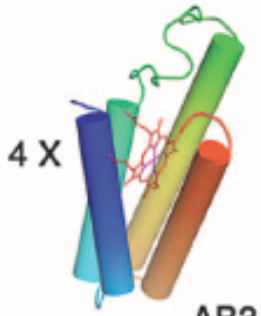




Tezcan *Science*, **2104**, *346*, 1525–1528.

Redesigned/ De Novo Synthesis

Tezcan (2014)

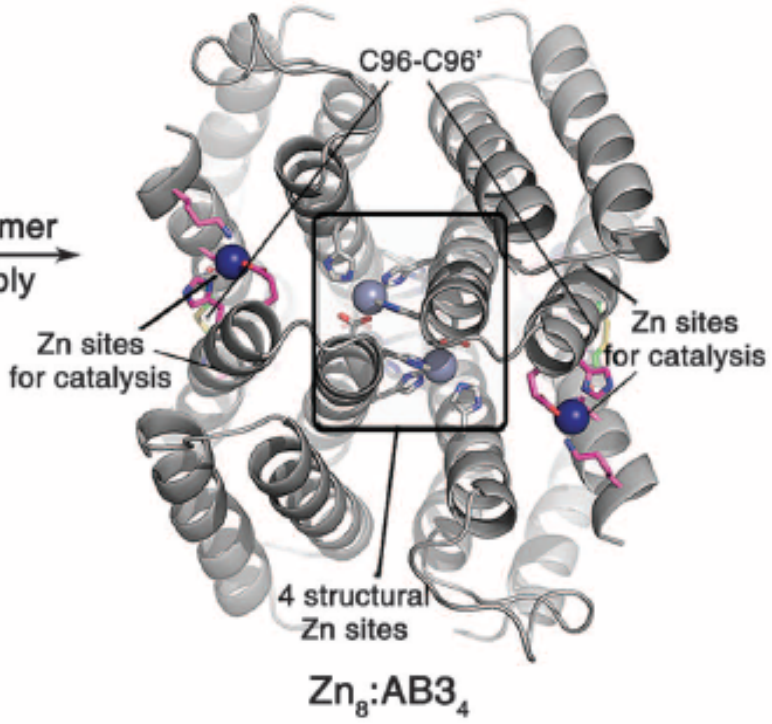


AB3

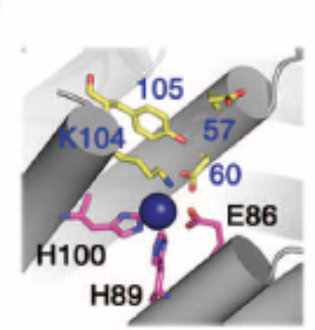
(C96RIDC1 + E86/H89/H100)

cytochrome cb_{562}

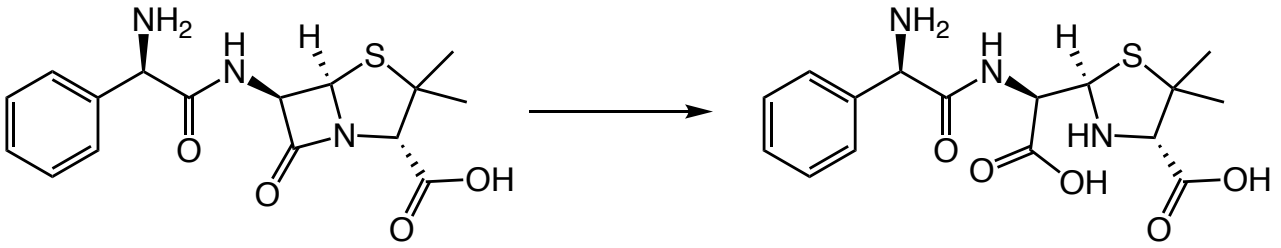
no known lactamase homology



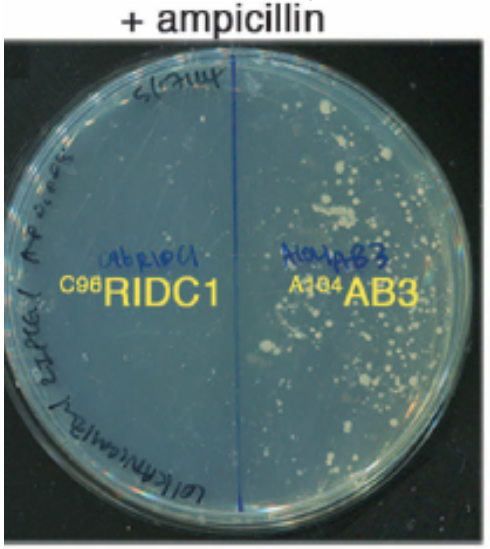
Zn_8:$AB3_4$



Zn site for catalysis



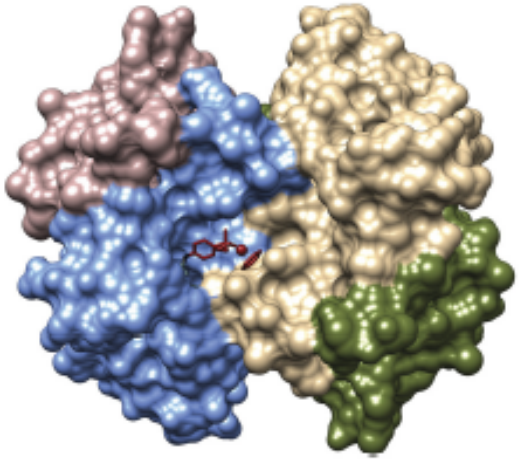
+ ampicillin



Tezcan *Science*, **2104**, *346*, 1525–1528.

Artificial Cofactors

Biotin Linked (Ward, 2016)

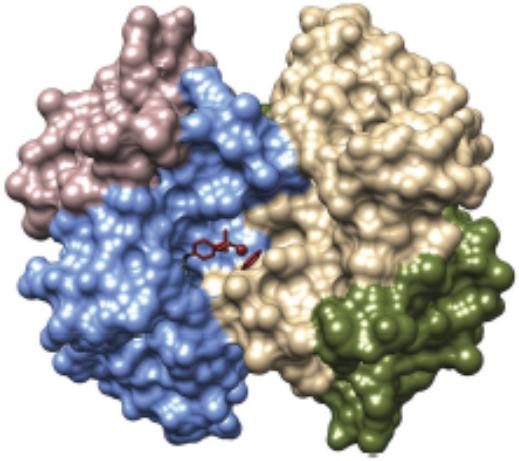


Streptavidin tetramer (SAV)
pM affinity for biotin

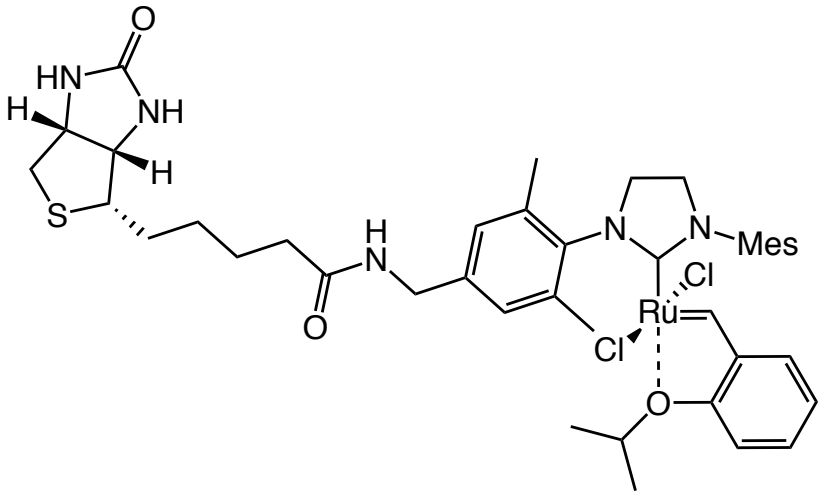
Ward *Narture*, **2016**, *537*, 662–665.

Artificial Cofactors

Biotin Linked (Ward, 2016)



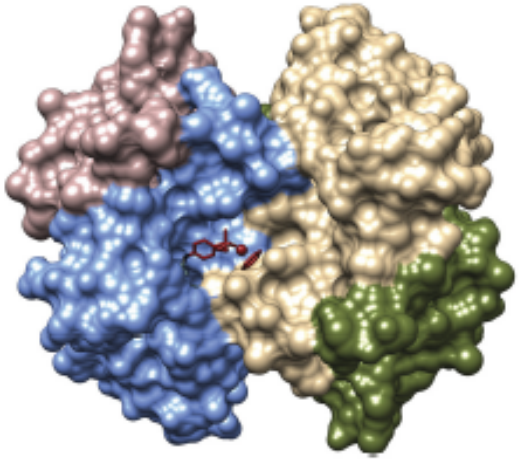
Streptavidin tetramer (SAV)
pM affinity for biotin



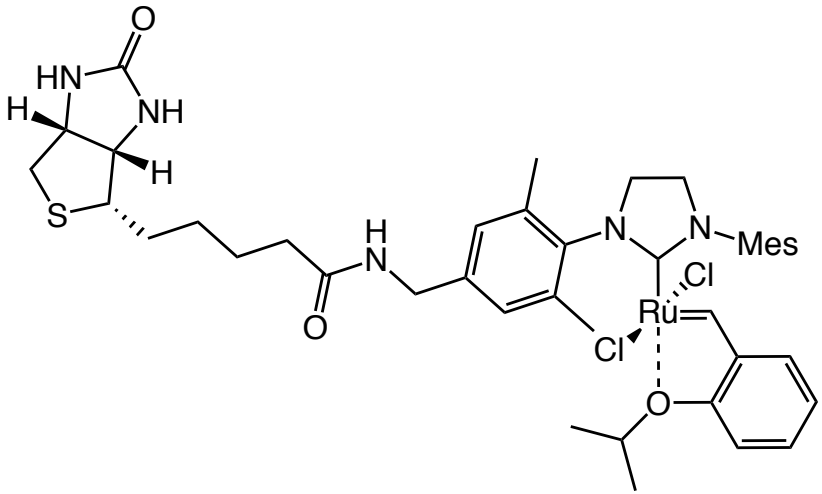
Ward *Narture*, **2016**, *537*, 662–665.

Artificial Cofactors

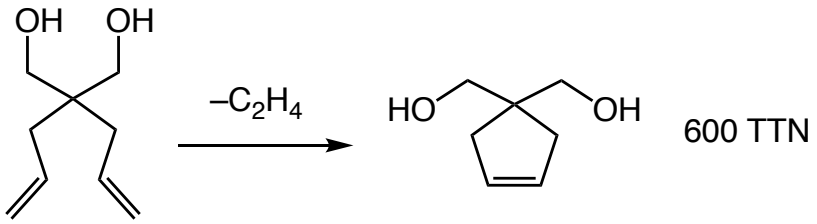
Biotin Linked (Ward, 2016)



Streptavidin tetramer (SAV)
pM affinity for biotin



Wild type SAV —(5 generations, Sat. Mut.)→ SAV^{mut}

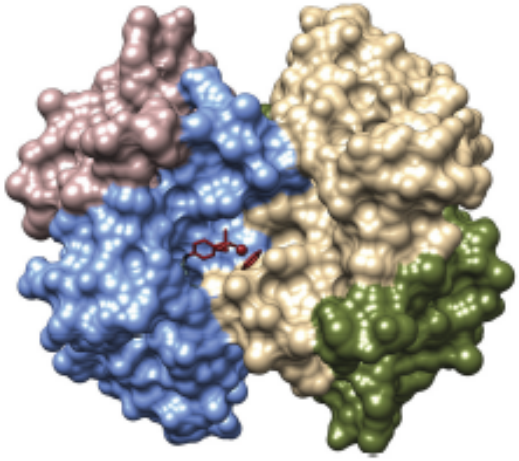


600 TTN

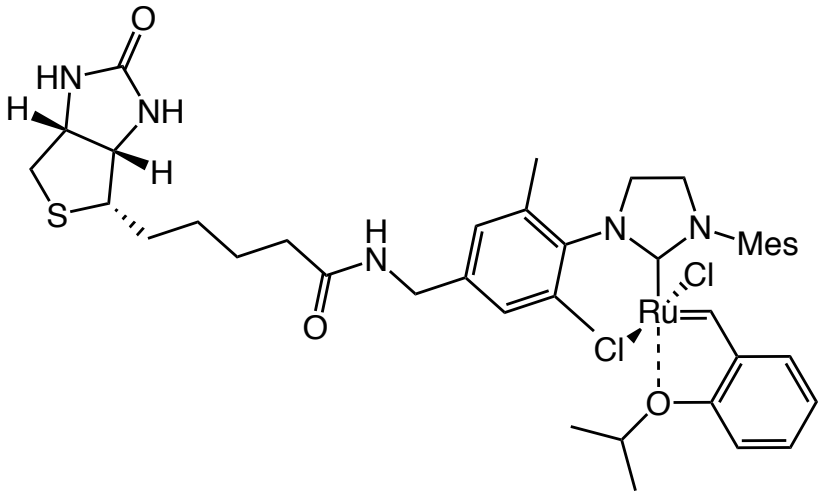
Ward *Narture*, **2016**, *537*, 662–665.

Artificial Cofactors

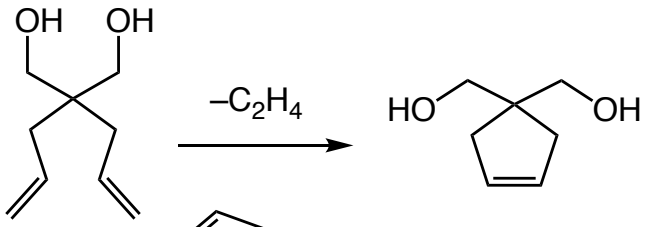
Biotin Linked (Ward, 2016)



Streptavidin tetramer (SAV)
pM affinity for biotin



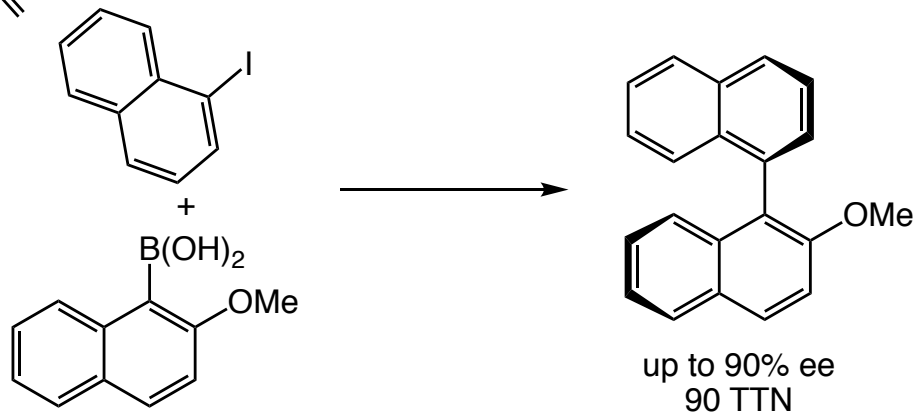
Wild type SAV —(5 generations, Sat. Mut.)→ SAV^{mut}



600 TTN

This system of cofactor anchoring has been utilized in:

Pd-cross coupling (ward, 2016)

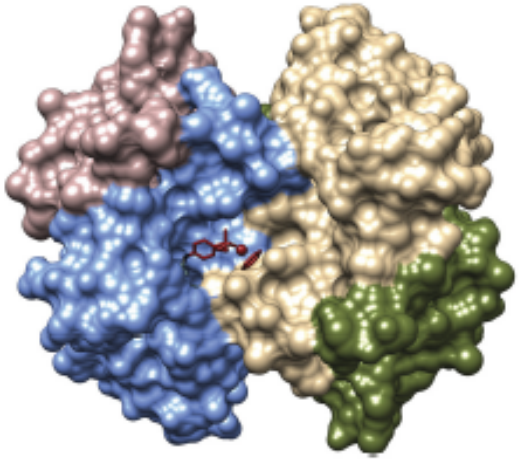


up to 90% ee
90 TTN

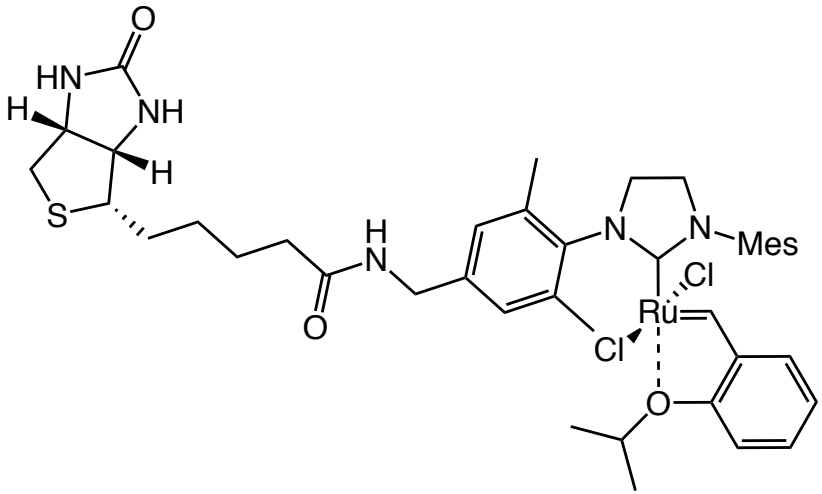
Ward *Narture*, **2016**, *537*, 662–665.

Artificial Cofactors

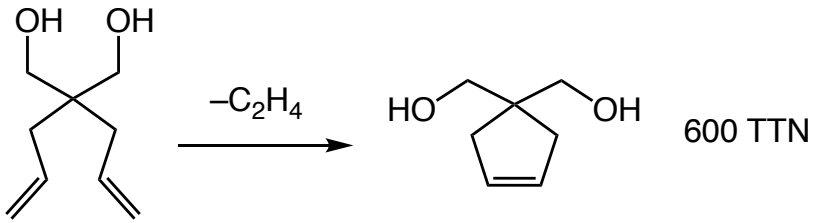
Biotin Linked (Ward, 2016)



Streptavidin tetramer (SAV)
pM affinity for biotin



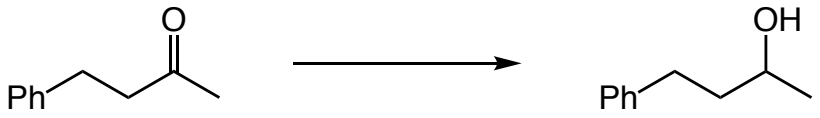
Wild type SAV —(5 generations, Sat. Mut.)→ SAV^{mut}



600 TTN

This system of cofactor anchoring has been utilized in:

Pd-cross coupling	(ward, 2016)
Ru-ketone hydrogenation	(Ward, 2008)

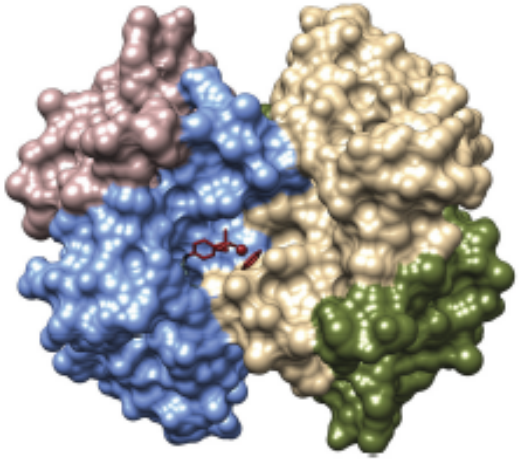


up to 96% ee

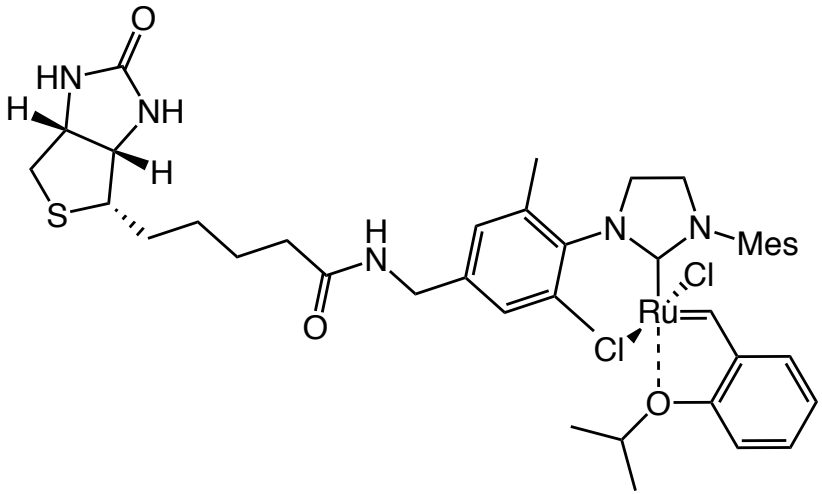
Ward *Narture*, **2016**, *537*, 662–665.

Artificial Cofactors

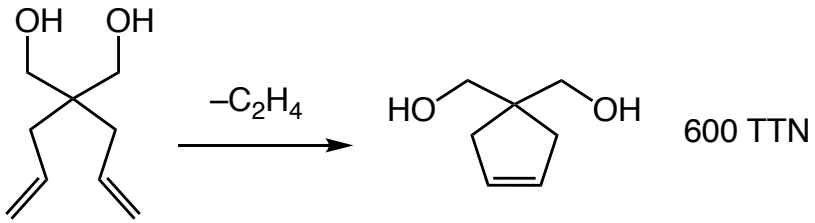
Biotin Linked (Ward, 2016)



Streptavidin tetramer (SAV)
pM affinity for biotin



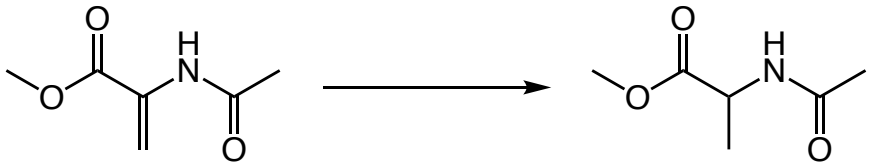
Wild type SAV —(5 generations, Sat. Mut.)→ SAV^{mut}



600 TTN

This system of cofactor anchoring has been utilized in:

Pd-cross coupling	(ward, 2016)
Ru-ketone hydrogenation	(Ward, 2008)
Rh-hydrogenation	(Reetz, 2006)



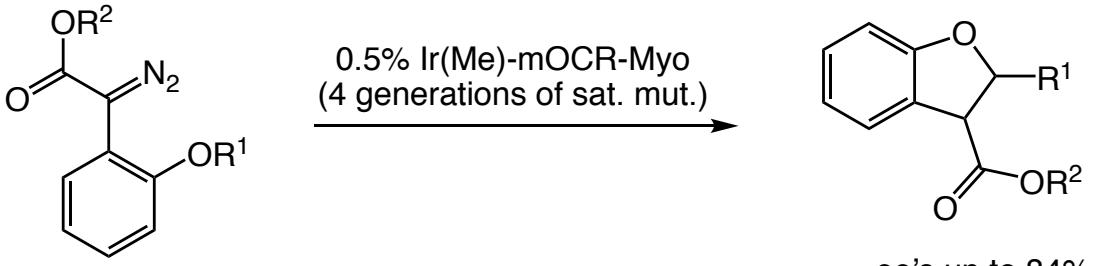
up to 65% ee

Ward *Narture*, **2016**, *537*, 662–665.

Artificial Cofactors

Metal swap of natural cofactors (Hartwig, 2016)

Expressed hemeless P450s and insert [M]-heme complexes after purification



R^1 = Me, Et
R^2 = Me, Et, Bn

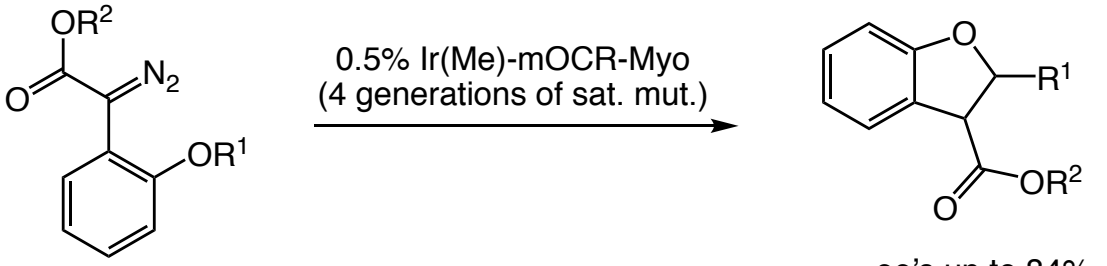
ee's up to 84%
TON up to 7,260

Hartwig *Narture*, **2016**, *534*, 534–537.
Lewis *Nat. Comm.* **2015**, *6*, 1–8.

Artificial Cofactors

Metal swap of natural cofactors (Hartwig, 2016)

Expressed hemeless P450s and insert [M]-heme complexes after purification

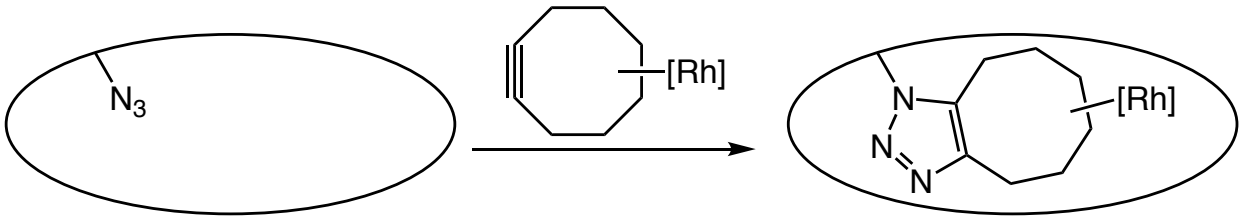


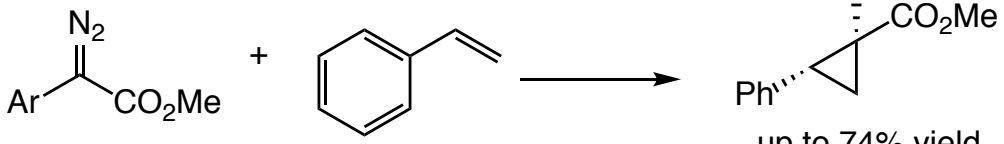
R^1 = Me, Et
R^2 = Me, Et, Bn

ee's up to 84%
TON up to 7,260

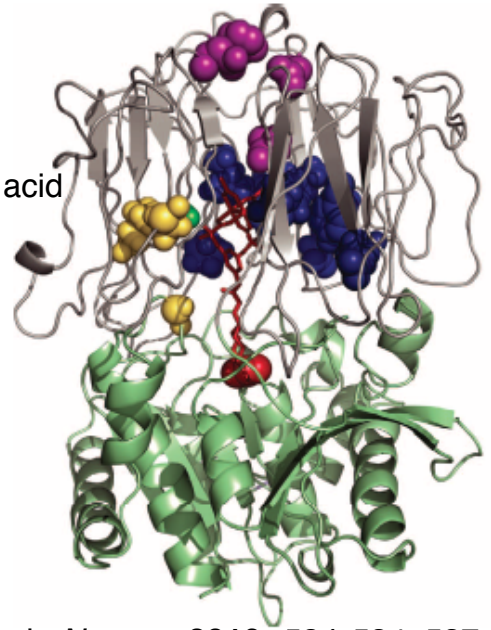
Covalent linkage (Lewis, 2016)

Used the Amber codon (one of 3 stop codons) to introduce a unnatural amino acid





up to 74% yield
92% ee



Hartwig *Narture*, **2016**, *534*, 534–537.
Lewis *Nat. Comm.* **2015**, *6*, 1–8.

Summary

Broaden the scope and selectivity of known Biological reaction

Developed new to nature reactions

- Repurposed enzymes
- Redesigned/synthetic enzymes
- Artifical cofactors

Advantages

- Can in theory evolve tailor-made catalyst for any specific substrate
- Reaction conditions are typlically environmentily friendly
- generally run under mild conditions

Disadvantages/ Limitations

- Significant time required to evolve new enzymes
- Structural information generally required
- Artificial metalloenzymes still need development to be synthetically useful (limited scope and not commercially available)

General Review Articles

Performing Directed Evolution

Liu, D. R. *Nature Rev. Gen.* **2015**, *16*, 379–394.

Reetz, M. T. *Isr. J. Chem.* **2015**, *55,* 51–60.

Reetz, M. T. *Angew. Chem. Int. Ed.* **2001**, *40*, 284–310.

Arnold, F. H. *Curr. Opin. Struct. Biol.* **2005**, *15*, 447–452.

Screening Methodologies

Maillard, N. *Chem. Commun.* **2009***, 1*, 34–46.

Uses in Organic Synthesis

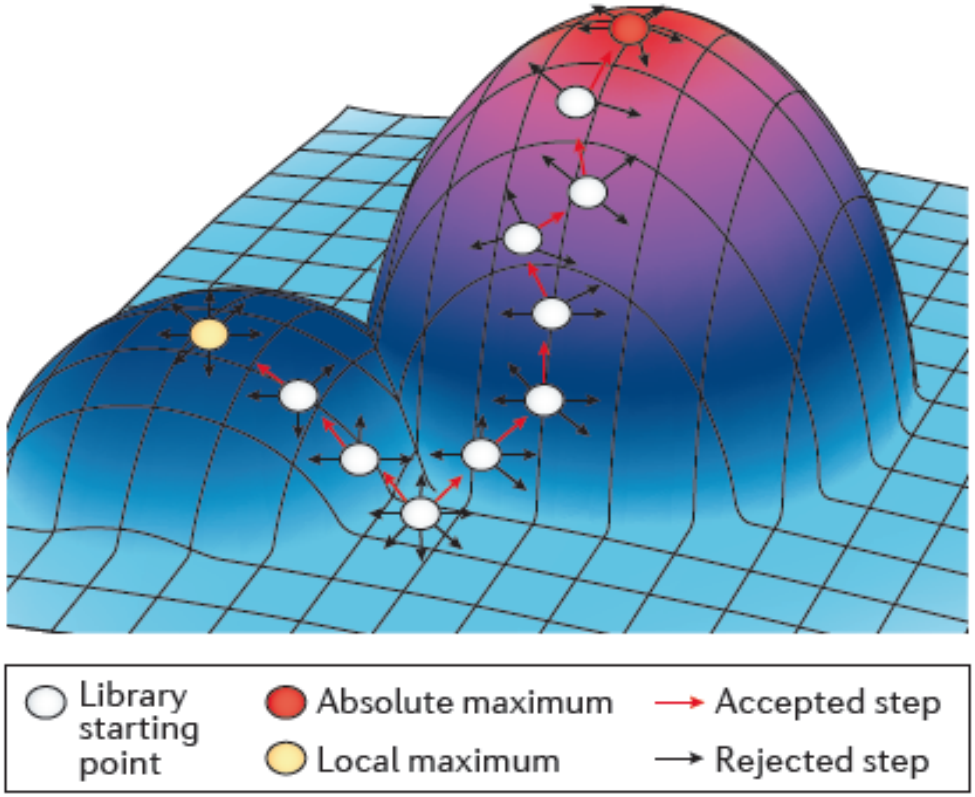
Reetz, M. T. *Bioorg. Med. Chem.* **2017**

Lewis, J. C. *Curr. Opin. Chem. Biol.* **2017**, *37*, 48–55.

Arnold, F. A. *Angew. Chem. Int. Ed.* **2017**, *56*, 2–8.

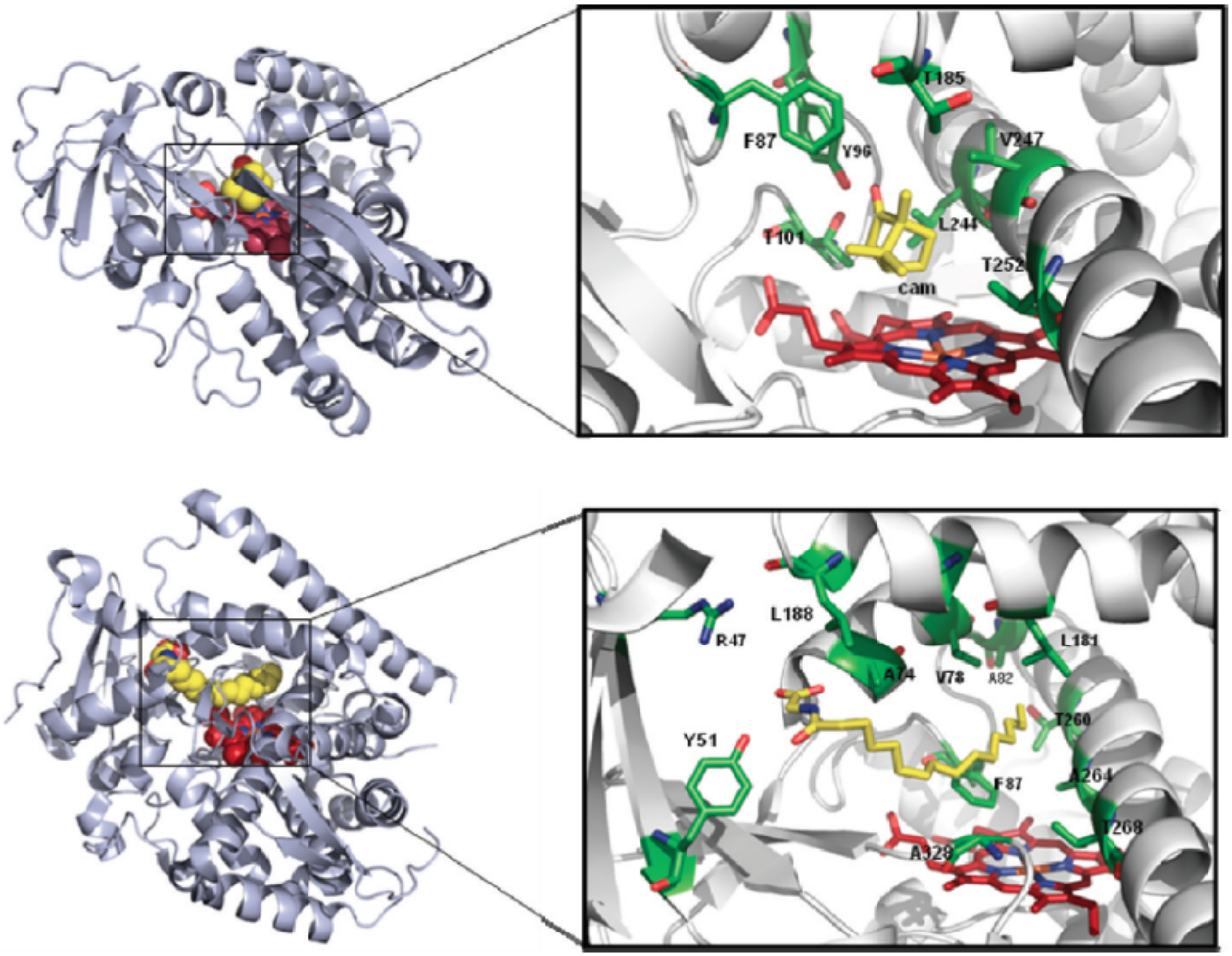
Ward, T. R. *Angew. Chem. Int. Ed.* **2016**, *55*, 7344–7357.

Directed Evolution



Lui, *Nat Rev. Genet.* **2015**, *16*, 379–394.

Background



Fasan, *ACS. Catal.* **2012**, *2*, 647–666.