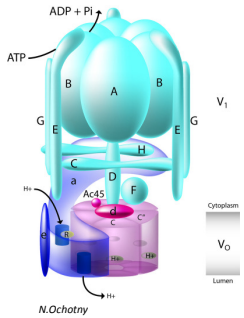
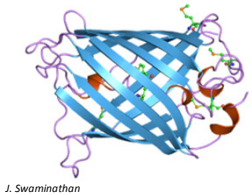
CHAPTER 2



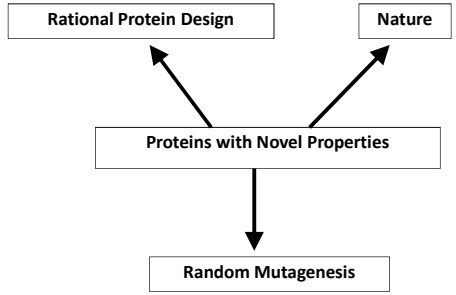


PROTEINS

Part 2 – protein engineering

Protein engineering

- Enhance stability/function under new conditions
  - temperature, pH, organic/aqueous solvent, [salt]
- Alter enzyme substrate specificity
- Enhance enzymatic rate
- Alter epitope binding properties



Protein engineering

RATIONAL DESIGN

-Site directed mutagenesis of one or more residues

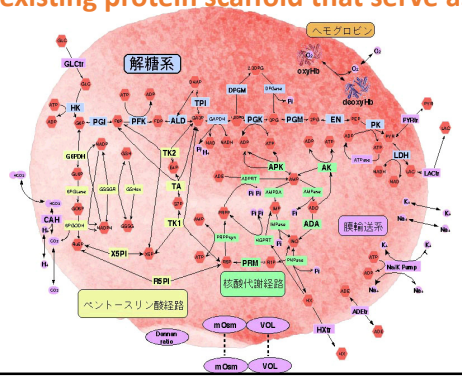
-Fusion of functional domains from different proteins to create chimaeric (Domain swapping)

↓

Functional evaluation

A protein library having the mass of our galaxy could only cover the combinatorial possibilities for a peptide with 50 residues

Therefore even genetic selection approaches for designing novel functional proteins will not generally build on fully random sequences, but will be based on existing protein scaffold that serve as template.



Protein engineering

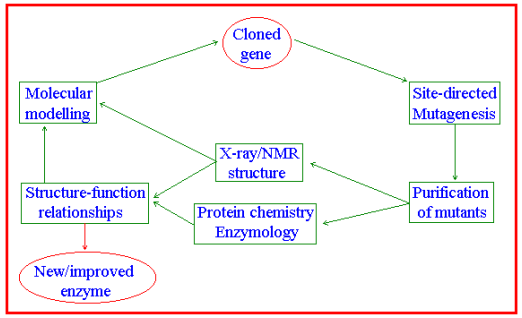
In order to consider the rational design of a target enzyme, one needs to have several pieces of information:

1. A cloned gene coding for the enzyme.
2. The sequence of the gene.
3. Information on the chemistry of the active site, ideally one would know which amino acids in the sequence are involved in activity.
4. Either a crystal/NMR structure for of the enzyme, or the structure of another protein displaying a high degree of structural homology.

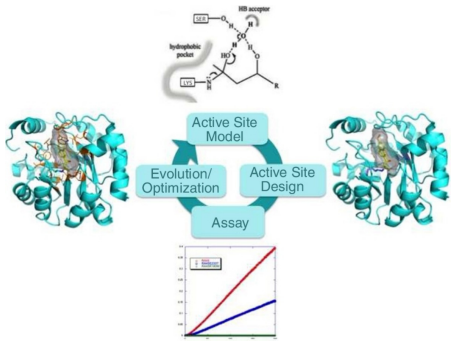
The above information is needed in order to have a clear idea of which amino acids one should mutate to which likely effect.

Protein engineering

Typically, protein engineering is a cyclic activity involving many scientists with different skills:

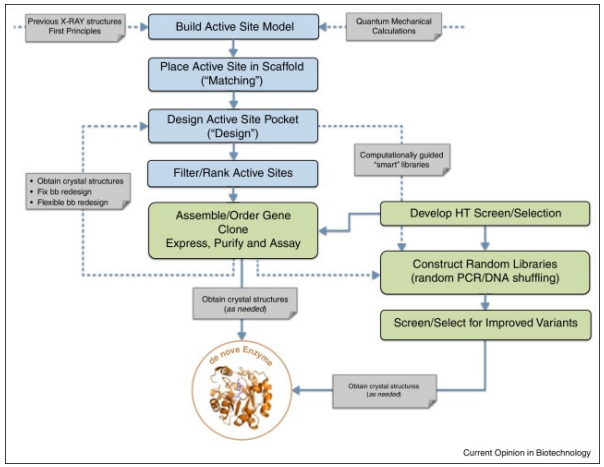


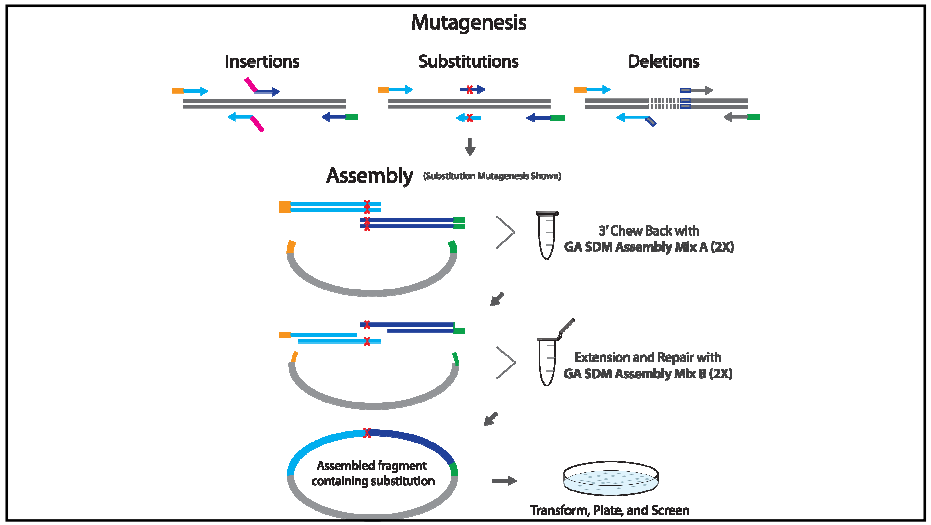
Protein engineering – de novo enzyme design

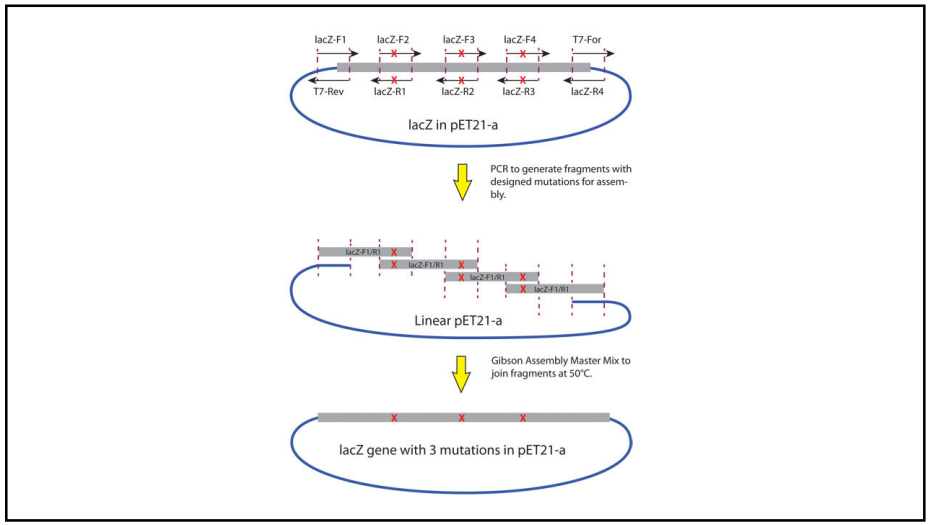


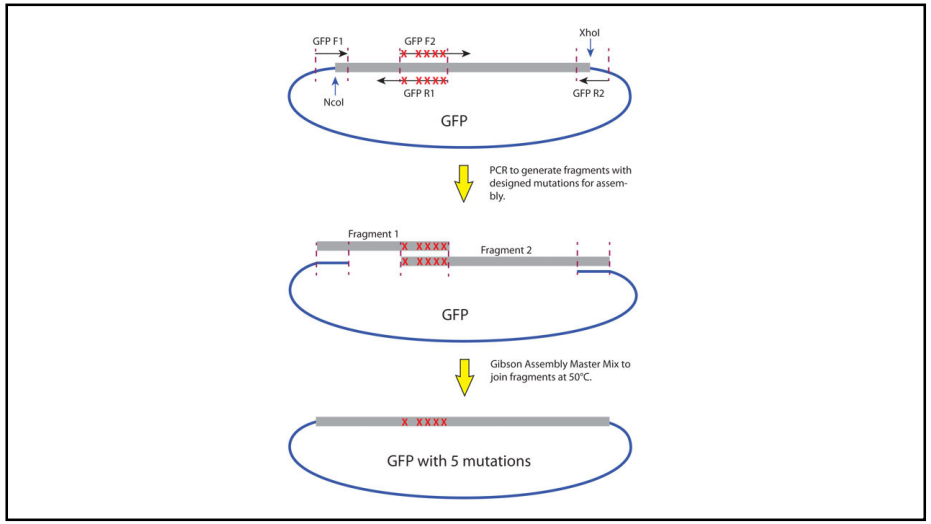
A. Zanghellini *Curr. Opp. Biotechnol.*, **2014**, *29*, 132-138

Protein engineering – de novo enzyme design

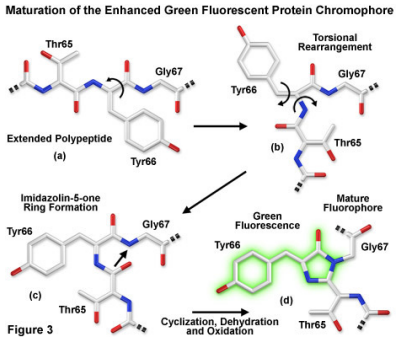








Green Fluorescent Protein



Example – Green Fluorescent Protein

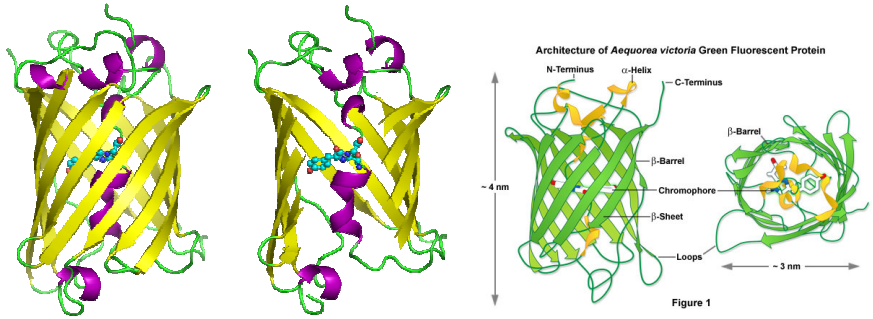
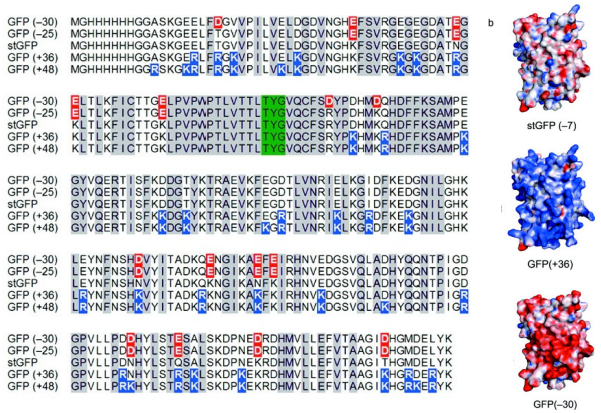
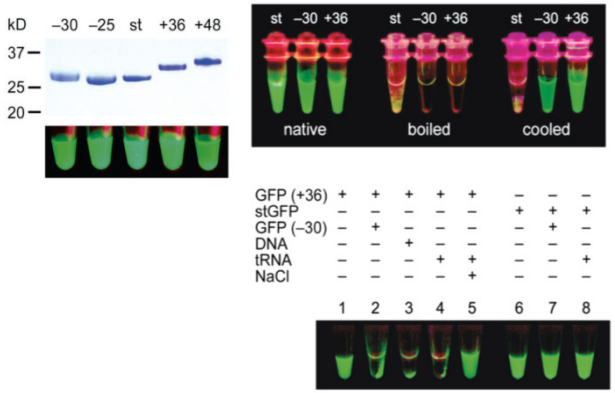


Figure 1

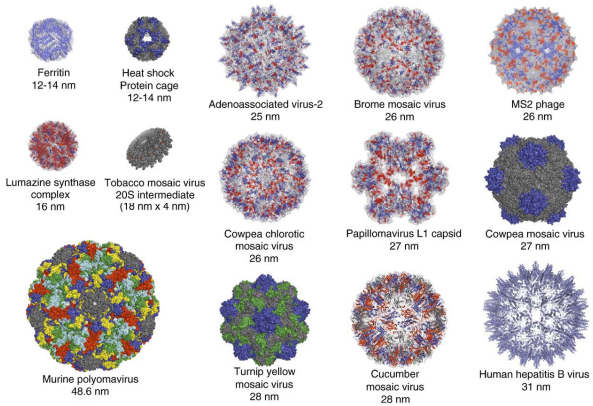
Green Fluorescent Protein



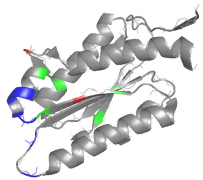
Green Fluorescent Protein



Example – protein capsids



Engineering of lumazine synthase

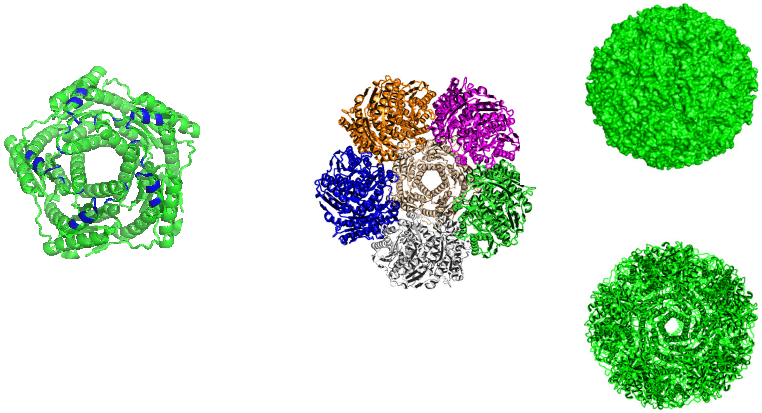


wt AaLS:

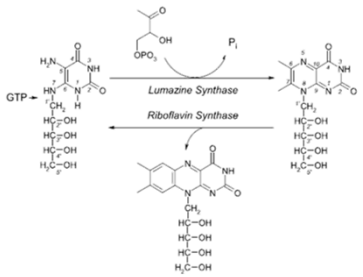
MEIYEGKLTA EGLRFGIVAS RFNHALVDRL VEGAIDCIVR HGGREEDITL VRVPGSWEIP VAAGELARKE DIDAVIAIGV LIRGATPHFD YIASEVSKGL ANLSLELRKP ITFGVITADT LEQAIERAGT KHGNKGWEAA LSAIEMANLF KSLRLEHHHH HH***

162 Aminoacids

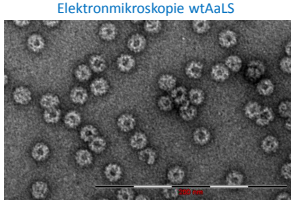
Engineering of lumazine synthase

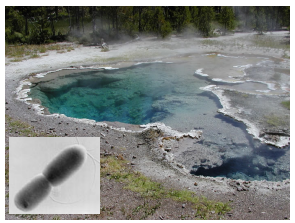


The A. Aeolicus ***lumazine synthase – origin and mode of action***



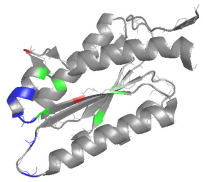
Elektronmikroskopie wtAaLS





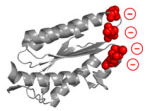
A. Aeolicus - warme Quellen(80-95 °C)
(Aae: K.O. Stetter, R. Rachel, University of Regensburg; springs: Octopus Spring, Yellowstone National Park, Allan Treiman, LPI)

Engineering of lumazine synthase

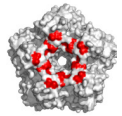


wt AaLS:

MEIYEGKLTA EGLRFGIVAS RFNHALVDRL VEGAIDCIVR HGGREEDITL VRVPGSWEIP VAAGELARKE DIDAVIAIGV LIRGATPHFD YIASEVSKGL ANLSLELRKP ITFGVITADT LEQAIERAGT KHGNKGWEAA LSAIEMANLF KSLRLEHHHH HH***

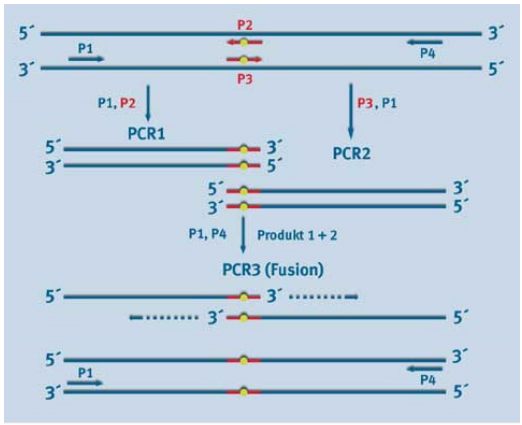


Arg83Glu/Thr86Glu/Thr120Glu/Gln123Glu
300 extra charges per 60-meric capsid



F. P. Seebeck, K. J. Woycechowsky, W. Zhuang, J. P. Rabe, D. Hilvert *J. Am. Chem. Soc.* 2006, *128*, 4516

The Overlap-Extension PCR



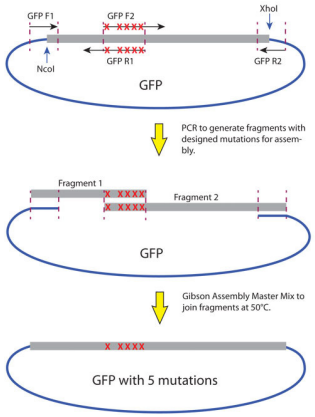
Sieht komplizierter aus, als es ist: Schema der Overlap-Extension-PCR.
Mutagene Primer: P2 u. P3, flankierende Primer: P1 u. P4

Recombinant lumazine synthase from *Aquifex aeolicus*. The lumazine synthase gene from *A. aeolicus* was amplified from plasmid pNCO-AA-ribH. by PCR using the primers 200304AQS (GATATACCATGGAAATCTACGAAGGTAAACTA) and 200304AQA (GATATACTCGAGTCGGAGAGACTTGAATAAGT). The restriction sites are underlined.
The NcoI/XhoI double digested fragment was ligated into a modified version of the pMG209 vector.

Mutations R83E, T86E, Q123E and T120E were introduced by standard overlap-extension PCR using Pfu-polymerase and the following additional primers:

Arg83Glu/Thr86Glu
200304NEG1 (GTTCTCATC**GAA**GGGGCA**GAG**CCACATTTC),
200304NEG2 (GAAATGTGG**CTC**TGCCCC**TTC**GATGAGAAC),

Thr120Glu/Gln123Glu
203004NEG3(CAGCTGAC**GAA**TTGGAA**GAG**GCTATCGAG),
203004NEG4 (CTCGATAGC**CTC**TTCCAA**TTC**GTCAGCTG).
The altered codons are shown in bold.
All coding portions of the constructed plasmids were confirmed by DNA sequencing.

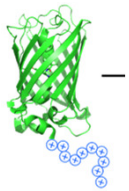


F. P. Seebeck, K. J. Woycechowsky, W. Zhuang, J. P. Rabe, D. Hilvert *J. Am. Chem. Soc.* 2006, *128*, 4516

Recombinant GFP-R$_{10}$. The plasmid pMG-GFP$_3$ encodes GFP from *Aequorea victoria* with an inframe, *C*-terminal His-tag. We replaced the His-tag with a deca-arginine tag by ligating a duplex of the primers 220304polyRs (TCGAGCGTAGACGACGCCGTCGGCGACGTCGACGTTAA) and 220304polyRa (CTAGTTAACGTCGACGTCGCCGACGGCGTCGTCTACGC) into pMG-GFP that was linearized by double digestion with XhoI and SpeI.

The resulting low copy number plasmid (pAC4C-GFP-R) encodes GFP-R$_{10}$ (under control of the T7 and salicylate promoters) and also chloramphenicol resistance.

This plasmid served as a template for site-directed mutagenesis in order to generate a construct (pAC4C-GFP) encoding GFP that lacks any *C*-terminal tag. Mutations (in bold) were introduced using the QuickChange™ site-directed mutagenesis kit (Stratagene) with the primers 050804killargS (CTCGAGCGT**TAA**CGACGCCGT) and 050804killargA (ACGGCGTCG**TTA**ACGCTCGAG)

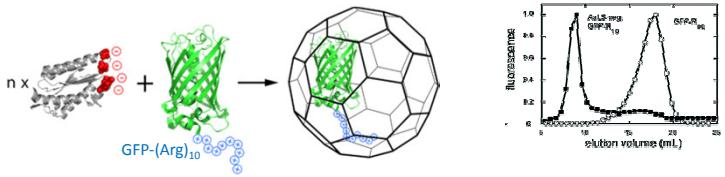


UAA – Stop (Ochre)

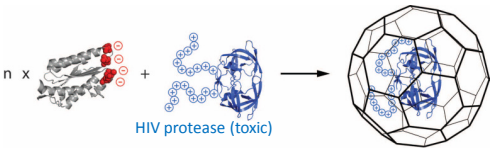
UGA	Stop (*Opal*)	A
UGG	(Trp/W) Tryptophan	G
CGU	(Arg/R) Arginine	U
CGC		C
CGA		A
CGG		G
AGU	(Ser/S) Serine	U
AGC		C
AGA	(Arg/R) Arginine	A
AGG		G

F. P. Seebeck, K. J. Woycechowsky, W. Zhuang, J. P. Rabe, D. Hilvert *J. Am. Chem. Soc.* 2006, *128*, 4516

Engineering of lumazine synthase



F. P. Seebeck, K. J. Woycechowsky, W. Zhuang, J. P. Rabe, D. Hilvert *J. Am. Chem. Soc.* 2006, *128*, 4516



B. Woersdoerfer, K.J.Woycechowsky, D.Hilvert *Science* **2011**, *331*, 589-592

Error prone PCR

Taq DNA polymerase lacks 3' to 5' exonuclease activity → error rate of 0.001-0.002% per nucleotide per replication.

- choosing the gene, or the area within a gene, one wishes to mutate.
- the extent of error required is calculated based upon the type and extent of activity one wishes to generate.
- This extent of error determines the error prone PCR strategy to be employed.
- Following PCR, the genes are cloned into a plasmid and introduced to competent cell systems.
- These cells are then screened for desired traits.
- Plasmids are then isolated for colonies which show improved traits, and are then used as templates the next round of mutagenesis.

Rates of error in PCR can be increased in the following ways:

Increase concentration of magnesium chloride, which stabilizes non complementary base pairing.
Add manganese chloride to reduce base pair specificity.
Increased and unbalanced addition of dNTPs.
Addition of base analogs like dITP, 8 oxo-dGTP, and dPTP.
Increase concentration of Taq polymerase.
Increase extension time.
Increase cycle time.
Use less accurate Taq polymerase.

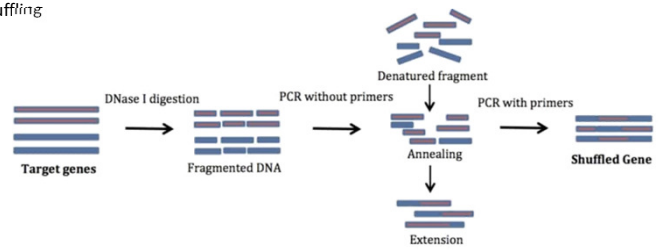
Error-prone PCR library construction. the plasmid pMG-AaLS-neg - template in an error-prone PCR (epPCR). The primers AQs (GATATACCATGGAAATCTACGAAGGTAAACTA) and AQa (GATATACTCGAGTCGGAGAGACTTGAATAAGT), flanking the coding region of the capsid gene, were employed for amplification.

The resulting epPCR product (482 bp) was purified by agarose gel electrophoresis. The epPCR product and the acceptor vector pMG-AaLS-neg were digested with *NcoI* and *XhoI*. The fragments were ligated with T4 DNA ligase to give a plasmid library containing the mutant capsid genes.

B. Woersdoerfer, K.J.Woycechowsky, D.Hilvert *Science* **2011,** *331,* 589-592

DNA Shuffling

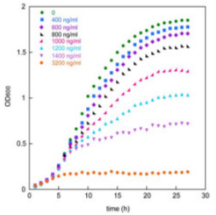
DNA shuffling is a way to rapidly propagate beneficial mutations in a directed evolution experiment. It is used to rapidly increase DNA library size.
First, DNase is used to fragment a set of parent genes into pieces of 50-100 bp in length. This is then followed by a polymerase chain reaction (PCR) without primers- DNA fragments with sufficient overlapping homologous sequence will anneal to each other and are then extended by DNA polymerase.
Several rounds of this PCR extension are allowed to occur, after some of the DNA molecules reach the size of the parental genes. These genes can then be amplified with another PCR, this time with the addition of primers that are designed to complement the ends of the strands.
It is possible to recombine portions of these genes to generate hybrids or chimeric forms with unique properties, hence the term DNA shuffling



DNA shuffling library construction. For the second and third rounds of evolution, genes encoding capsid variants selected in the previous round were subjected to DNA shuffling. Capsid genes were amplified (PCR, primers AQs and AQa), then digested with DNaseI (6 ng/µl) to give 50 bp to 100 bp fragments.The fragments were purified and assembled in a PCR-like process, but without primers. The reassembled genes were amplified in a final PCR using primers AQs and AQa, digested with NcoI and XhoI, and ligated into the pMG vector, desalted and concentrated to afford plasmid libraries containing mutant capsid genes.

B. Woersdoerfer, K.J.Woycechowsky, D.Hilvert *Science* **2011,** *331,* 589-592

Selection: Capsid mutants produced under conditions of high protease content
AaLS plasmids + the plasmid pACYC-tetHIV-R_{10} encoding the R_{10}-tagged HIV protease → electrocompetent *E. coli* XL1-Blue. Growth in presence of ampicillin (150 µg/ml) and chloramphenicol (30 µg/ml).
Capsid production was induced by salicylate (100 µM).
The HIV protease-R_{10} gene was induced with 400 to 1600 ng/ml tetracycline.



Internal standards / negative controls: plasmids of AaLS-wt or AaLS-neg, + pACYC-tetHIV-R10
After the 4th round of evolution, 24 surviving capsid variants were isolated, sequenced, and characterized.

For **detailed biophysical characterization**, AaLS variants were overproduced using the T7 promoter system in *E. coli* KA13 cells that had been transformed with the appropriate plasmid (pMG-AaLS-wt, pMG-AaLS-neg, or pMG-AaLS- 13). For coproduction with HIV protease, the cells were also transformed with either pACYCtetHIV or pACYC-tetHIV-R10.

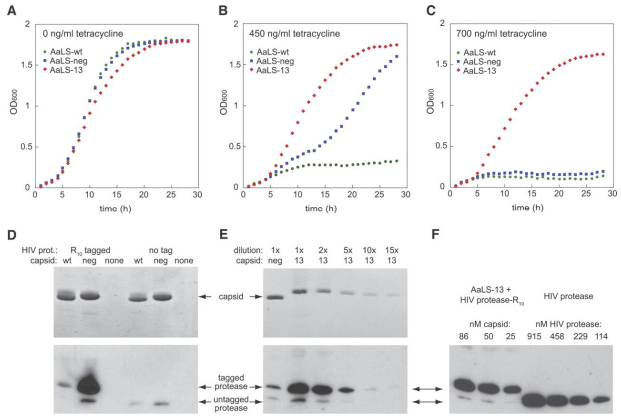
	Round I	Round II	Round III	Round IV
Library transformants	$5x10^6$	$3x10^6$	$3x10^7$	$8x10^6$
HIV protease-R_{10} transformants	$1x10^7$	$4x10^7$	$4x10^6$	$2x10^6$

DNA level:	unselected	selected	unselected	selected	unselected	selected	unselected	selected
# sequences considered	21	28	8	18	4	23	6	24
mutations per gene (range)	1-7	1-6	4-9	4-10	6-13	3-14	8-10	6-14
mutations per gene (average)	3 = 0.6 %	2.8	6.3	6.8	8.5	8.2	7.7	9.3
standard deviation	1.5	1.3	1.9	2.0	3.1	2.5	3.9	2.2
frame shifts	4	0	2	6	3	7	1	9

Amino acid level:	selected	common mutations	selected	common mutations	selected	common mutations	selected	common mutations
# sequences considered	27		12		16		15	
mutation per protein (range)	0-5	0-2	3-8	1-3	1-8	1-6	4-9	3-7
mutation per protein (average)	1.9	0.8	5.2	1.9	3.8	2.1	7.1	4.4
standard deviation	1.4	0.8	1.5	0.8	1.9	1.5	1.8	1.2

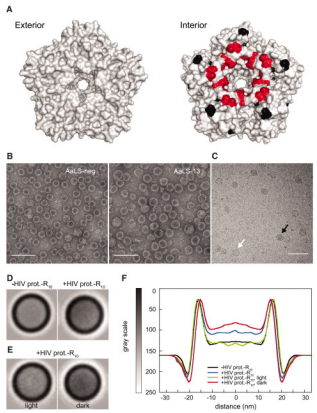
B. Woersdoerfer, K.J.Woycechowsky, D.Hilvert
Science **2011,** *331,* 589-592

Engineering of lumazine synthase



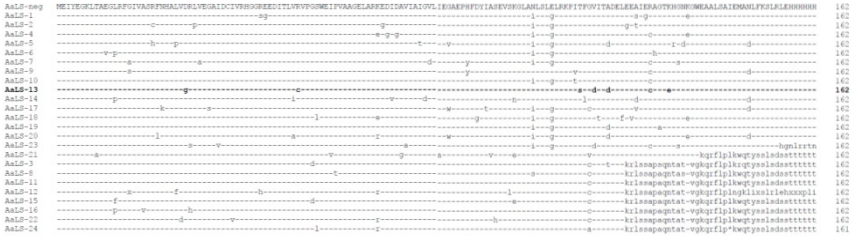
B. Woersdoerfer, K.J.Woycechowsky, D.Hilvert *Science* **2011,** *331,* 589-592

Engineering of lumazine synthase



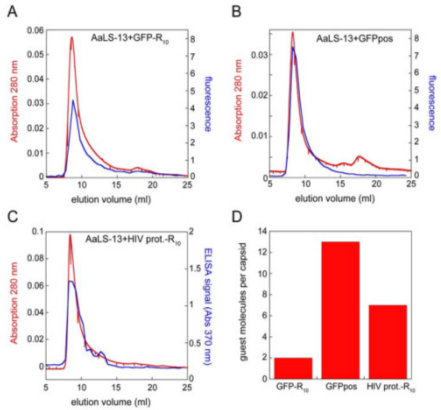
B. Woersdoerfer, K.J.Woycechowsky, D.Hilvert *Science* **2011,** *331,* 589-592

Engineering of lumazine synthase



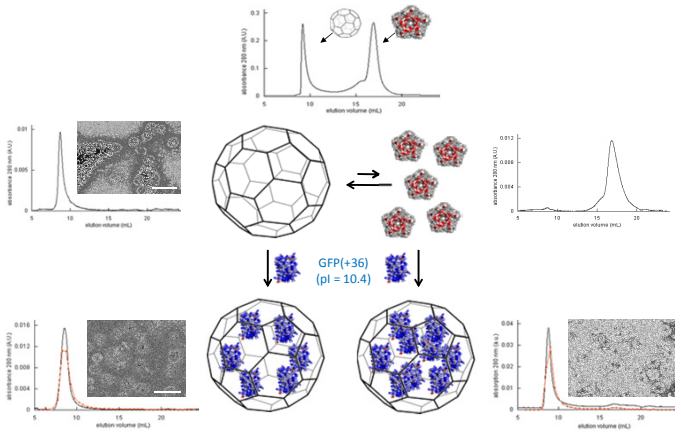
B. Woersdoerfer, K.J.Woycechowsky, D.Hilvert *Science* **2011,** *331,* 589-592

Engineering of lumazine synthase



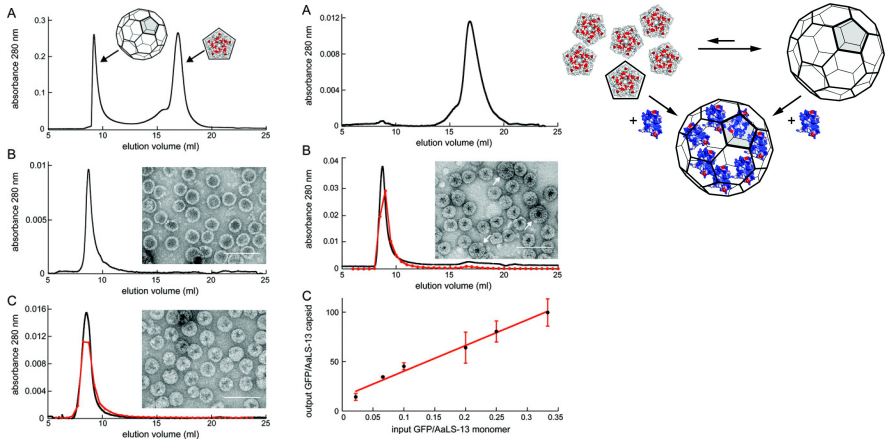
B. Woersdoerfer, K.J.Woycechowsky, D.Hilvert *Science* **2011,** *331,* 589-592

Engineering of lumazine synthase



B. Woersdoerfer, Z. Pianowski, D.Hilvert *J. Am. Chem. Soc.* **2012,** *134,* 909-911

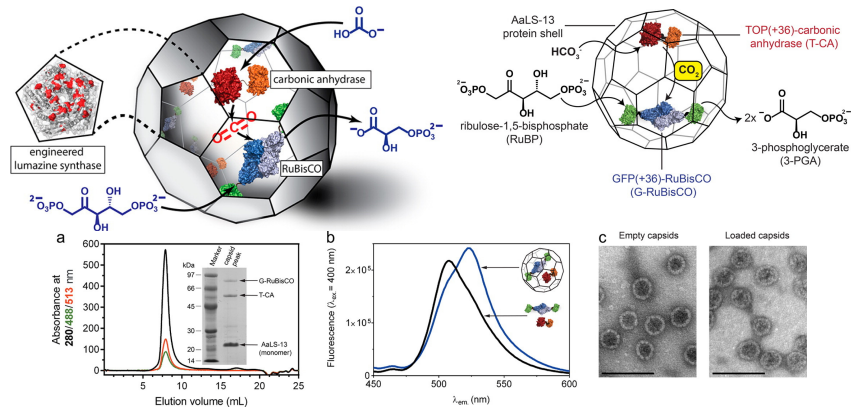
Engineering of lumazine synthase



B. Woersdoerfer, Z. Pianowski, D.Hilvert *J. Am. Chem. Soc.* **2012,** *134,* 909-911

Engineering of lumazine synthase

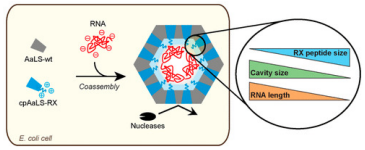
Bottom-up Construction of a Primordial Carboxysome Mimic



R. Frey, S. Mantri, M. Rocca, D.Hilvert *J. Am. Chem. Soc.*, **2016**, *138 (32),* pp 10072–10075

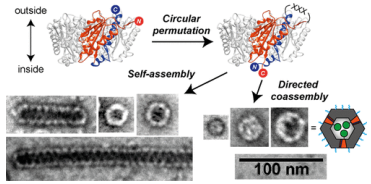
Engineering of lumazine synthase

Modular Protein Cages for Size-Selective RNA Packaging in Vivo



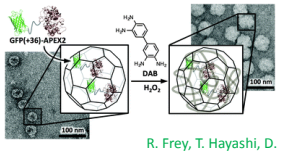
Y. Azuma, T. Edwardson, N. Terasaka, D. Hilvert
J. Am. Chem. Soc., 2018, 140 (2), pp 566–569

Diversification of Protein Cage Structure Using Circularly Permuted Subunits



Y. Azuma, M. Herger, D. Hilvert
J. Am. Chem. Soc., 2018, 140 (2), pp 558–561

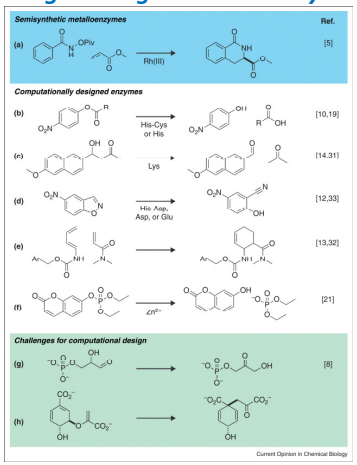
Enzyme-mediated polymerization inside engineered protein cages



Oxidation of 3,3-diaminobenzidine (DAB) by the engineered ascorbate peroxidase APEX2 encapsulated in AaLS capsids resulted in templated formation of polyDAB–capsid nanoparticles of homogeneous size and shape.

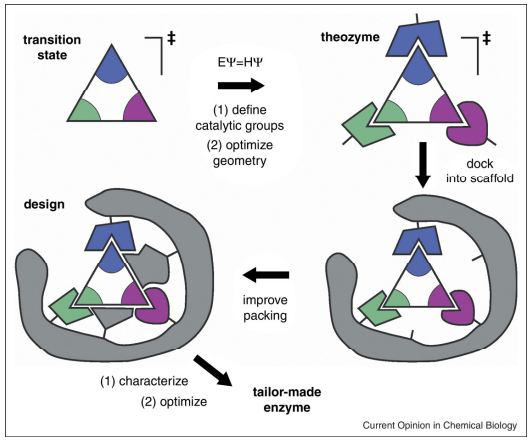
R. Frey, T. Hayashi, D. Hilvert,
Chem. Commun., **2016**, *52*, 10423-10426

Protein engineering – de novo enzyme design



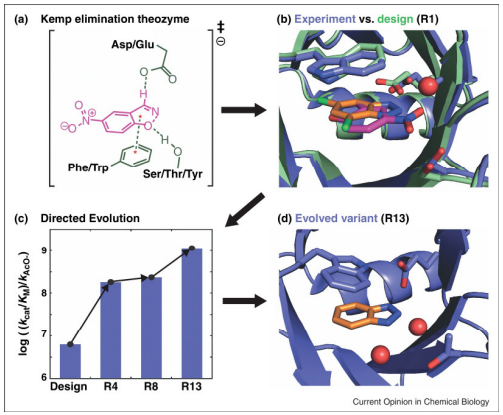
H. Kries, R. Blomberg, D. Hilvert *Curr. Opp. Chem. Biol.*, **2013**, *17*, 1-8

Protein engineering – de novo enzyme design



H. Kries, R. Blomberg, D. Hilvert *Curr. Opp. Chem. Biol.*, **2013**, *17*, 1-8

Protein engineering – de novo enzyme design



H. Kries, R. Blomberg, D. Hilvert *Curr. Opp. Chem. Biol.*, **2013**, *17*, 1-8

Computational Design of Catalytic Dyads and Oxyanion Holes for Ester Hydrolysis

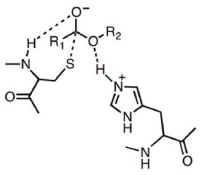
Nucleophilic catalysis is a general strategy for accelerating ester and amide hydrolysis.

In natural active sites, nucleophilic elements such as catalytic dyads and triads are usually paired with oxyanion holes for substrate activation.
The evolutionary origin is difficult to track back.

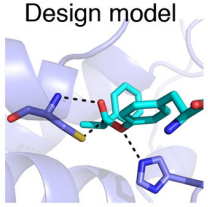
Minimal requirements for esterase activity have been explored by computationally designing artificial catalysts using catalytic dyads and oxyanion holes.

Four active designs in different scaffolds have been obtained by combining the ***oxyanion hole*** motif with a ***Cys-His dyad***. Rapid acylation of active site cysteines followed by slow hydrolysis of the acyl-enzyme intermediate limits overall catalytic efficiency.

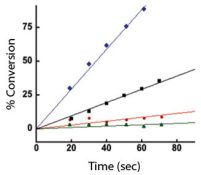
Theozyme



Design model



Catalytic activity

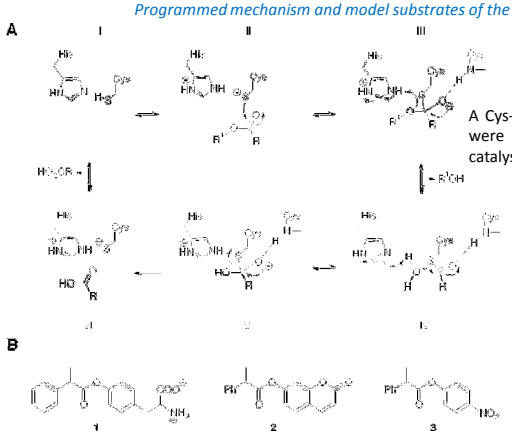


Most active variant:
k_{cat}/K_M of 400 M^{-1} s^{-1}
(*p*-nitrophenyl ester cleavage).

F. Richter, R. Blomberg, S. D. Khare, G. Kiss, A. P. Kuzin, A. J. T. Smith, J. Gallaher, Z. Pianowski, R. C. Helgeson, A. Grjasnow, R. Xiao, J. Seetharaman, M. Su, S. Vorobiev, S. Lew, F. Forouhar, G. J. Kornhaber, J. F. Hunt, G. T. Montelione, L. Tong, K. N. Houk, D. Hilvert, and D. Baker
J. Am. Chem. Soc., **2012**, *134* (39), pp 16197–16206

Computational Design of Catalytic Dyads and Oxyanion Holes for Ester Hydrolysis

Programmed mechanism and model substrates of the de novo designed esterases.



A Cys-His dyad, in combination with an oxyanion binder, were used to hydrolyze activated esters via covalent catalysis.

The tyrosyl ester **1** served as the target substrate for computational design; the fluorogenic coumarin ester **2** and the chromogenic *p*-nitrophenyl ester **3** were used for screening purposes.

F. Richter, *et al. J. Am. Chem. Soc.*, **2012**, *134* (39), pp 16197–16206

Computational Design of Catalytic Dyads and Oxyanion Holes for Ester Hydrolysis

Snapshots of the computational design process.

Representation of the calculated theozyme of the ester substrate framed by the catalytic dyad (Cys-His) and the backbone NH-oxyanion contact.

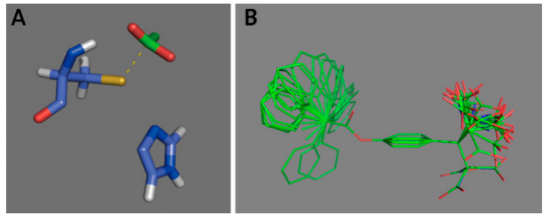
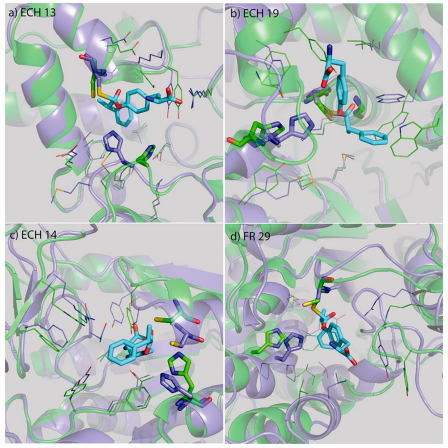


Image of the theoretical conformer ensemble of tyrosyl ester **1**.

To increase the number of matches, both the histidine sidechain and the substrate could rotate with respect to the cysteine (not shown). Note that in this case, the backbone NH contact is made by the cysteine itself.

F. Richter, *et al. J. Am. Chem. Soc.*, **2012**, *134* (39), pp 16197–16206

Computational Design of Catalytic Dyads and Oxyanion Holes for Ester Hydrolysis

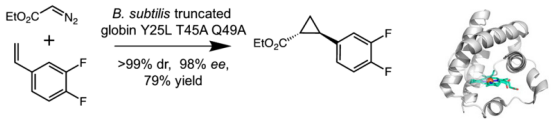


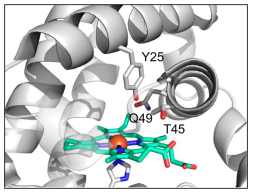
Crystal structures of the four active designs. In each case, the design model is shown in purple with the ligand in cyan, and the crystal structure in green. The theozyme residues and the ligand are shown in stick representation, and selected other active site residues in line representation.

F. Richter, *et al. J. Am. Chem. Soc.*, **2012**, *134* (39), pp 16197–16206

Directed evolution – bringing new chemistry to life

A *B. subtilis* globin variant, engineered by directed evolution, catalyzes the cyclopropanation of 3,4-difluorostyrene to make the desired stereoisomer of a ticagrelor precursor with high selectivity and yield



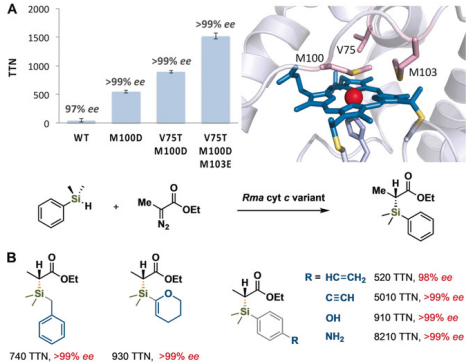


Positions of the Y25, T45, and Q49 residues near the heme iron in the B. subtilis wild-type protein (PDB ID: 1UX8)

K. E. Hernandez *et al. ACS Catal.*, **2016**, *6 (11)*, pp 7810–7813

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

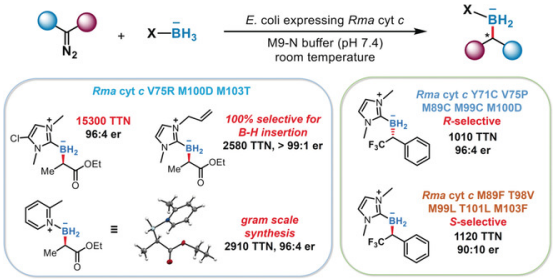
Directed evolution – bringing new chemistry to life



A) Chiral Si–C bond formation catalyzed by a laboratory-evolved variant of *Rhodothermus marinus* cytochrome *c*. The three amino acid residues that were mutated to increase this abiological activity include the methionine axial ligand (M100). B) The enzyme catalyzes formation of different organosilane products with high enantiomeric excess from silane and diazo substrates.

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

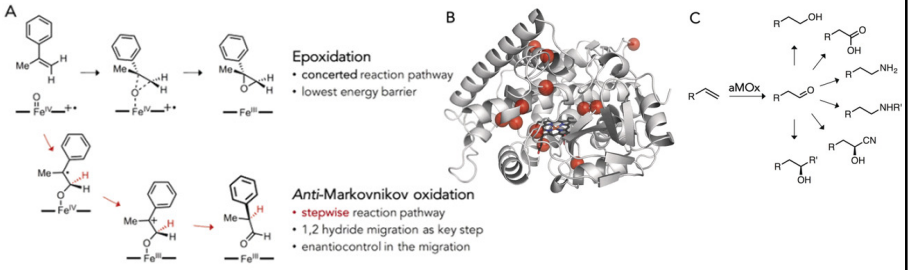
Directed evolution – bringing new chemistry to life



Production of chiral organoboranes by *E. coli* expressing *Rhodothermus marinus* cytochrome *c*. The bacterial catalyst uses borane-Lewis base complexes and diazo reagents to construct boron-containing carbon stereocenters efficiently and selectively in cells by carbene B–H insertion. The bioconversion can be conducted readily on gram scale, and the enantio-preference of borylation was switched to give either enantiomer of the organoborane products.

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

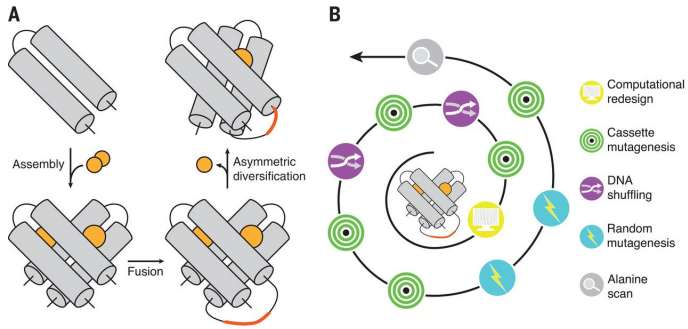
Directed evolution – bringing new chemistry to life



A cytochrome P450 anti-Markovnikov oxygenase. A) Competing reaction pathways for P450-catalyzed oxo transfer to alkenes. The concerted epoxidation pathway is favored over the stepwise anti-Markovnikov oxidation consisting of oxo transfer followed by an (enantioselective) 1,2-hydride migration. B) Ten rounds of directed evolution accumulated 12 amino acid mutations, many of which are distant from the active site. C) aMOx can be combined with established (bio)catalysts for various challenging anti-Markovnikov alkene functionalization reactions.

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

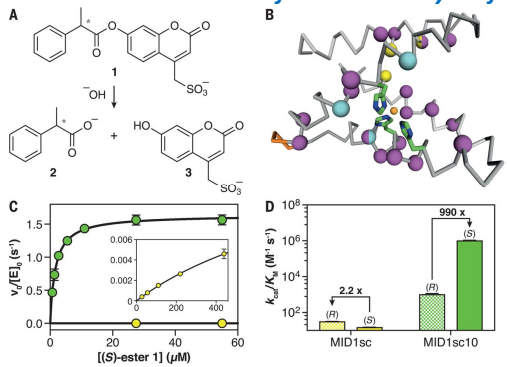
Evolution of a metalloenzyme from short peptides



Zinc-mediated assembly of helix-turn-helix fragments, followed by fusion and asymmetric diversification, afforded MID1sc10, an efficient metalloesterase.

S. Studer, D.A. Hansen, Z. Pianowski, P.R.E. Mittl, A. Debon, S.L. Guffy, B.S. Der, B. Kuhlman, D. Hilvert *Science,* **2018**, *362,* 1285-1288

Evolution of a metalloenzyme from short peptides



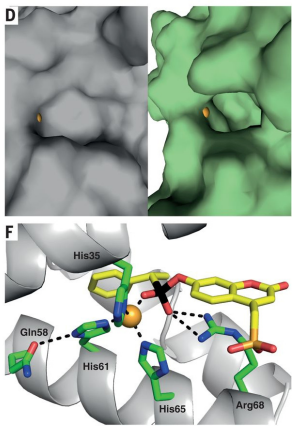
Crystal structure of MID1sc10
zinc ion - orange sphere,
coordinating histidines - green sticks
linkage of two polypeptides – orange sticks
beneficial mutations - magenta spheres,
residues replaced to prevent competitive zinc binding modes - cyan spheres).

The evolved variant MID1sc10 is highly enantioselective as a consequence of a 2200-fold specificity switch from the modestly (R)-selective starting catalyst MID1sc

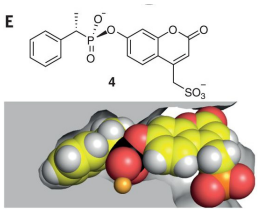
Michaelis-Menten plots for **MID1sc** (yellow and inset) and **MID1sc10** (green) show a 70,000-fold improvement in hydrolysis efficiency for (S)-configured **1** after optimization.

S. Studer, D.A. Hansen, Z. Pianowski, P.R.E. Mittl, A. Debon, S.L. Guffy, B.S. Der, B. Kuhlman, D. Hilvert *Science,* **2018**, *362,* 1285-1288

Evolution of a metalloenzyme from short peptides



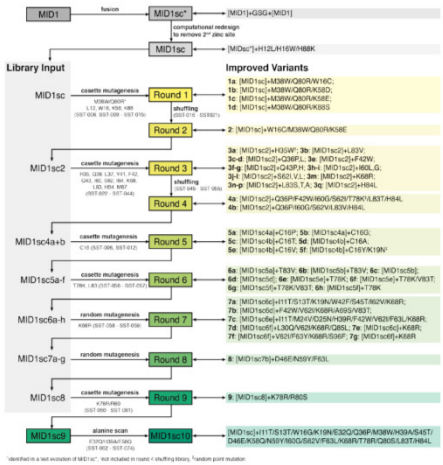
The observed structural changes transformed the shallow binding site of MID1 (gray) into a deep, hydrophobic pocket in MID1sc10 (green).



Cut-away view of the active site, showing the snug fit of phosphonate **4** in the binding pocket. (zinc ion – orange)

View of the MID1sc10 active site with phosphonate **4** (yellow) coordinating to the Zn(II)His_3 complex (orange sphere and green sticks). Arg68 and Gln58 form mechanistically relevant hydrogen bonds to phosphonate **4** and the backside nitrogen of His61, respectively

S. Studer, D.A. Hansen, Z. Pianowski, P.R.E. Mittl, A. Debon, S.L. Guffy, B.S. Der, B. Kuhlman, D. Hilvert *Science,* **2018**, *362,* 1285-1288



S. Studer, D.A. Hansen, Z. Pianowski, P.R.E. Mittl, A. Debon, S.L. Guffy, B.S. Der, B. Kuhlman, D. Hilvert *Science,* **2018**, *362,* 1285-1288