Synthetic life

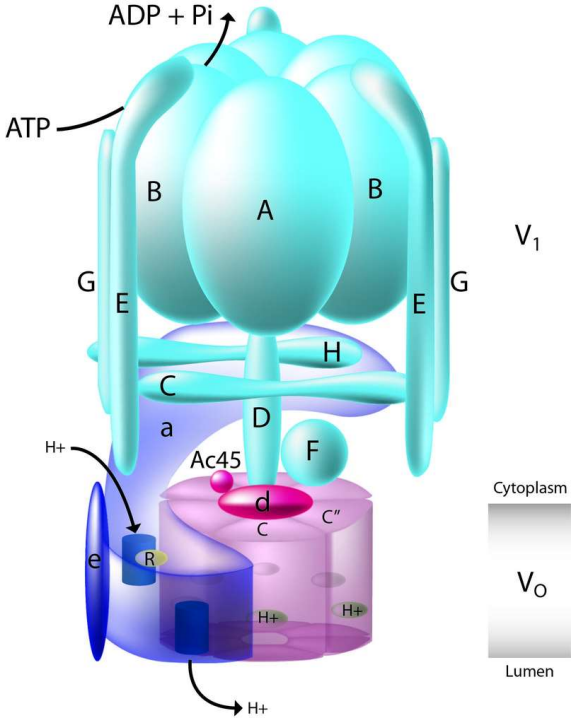
(continuation of „The molecular origins of life" SoSe 2020)

NaturalNews.com

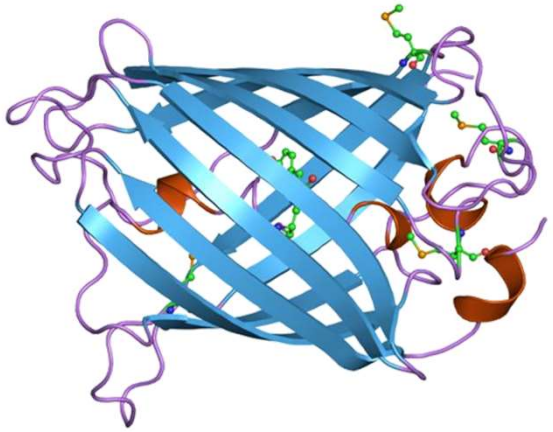
WiSe 2020/21

Zbigniew Pianowski

CHAPTER 2



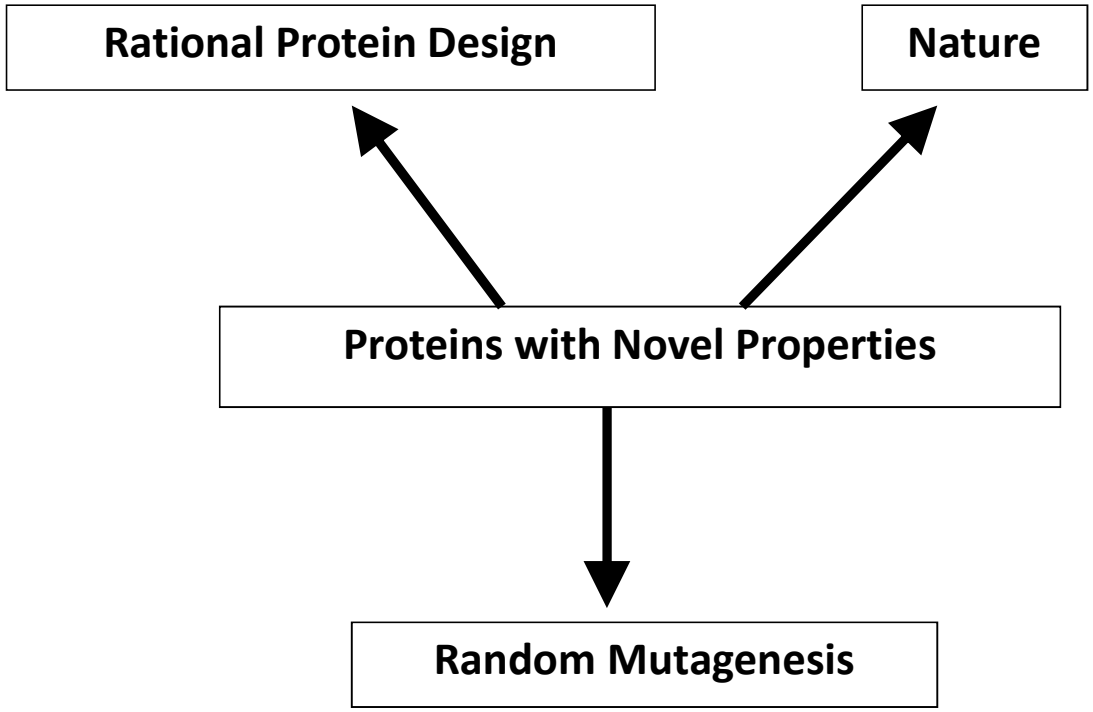
N.Ochotny

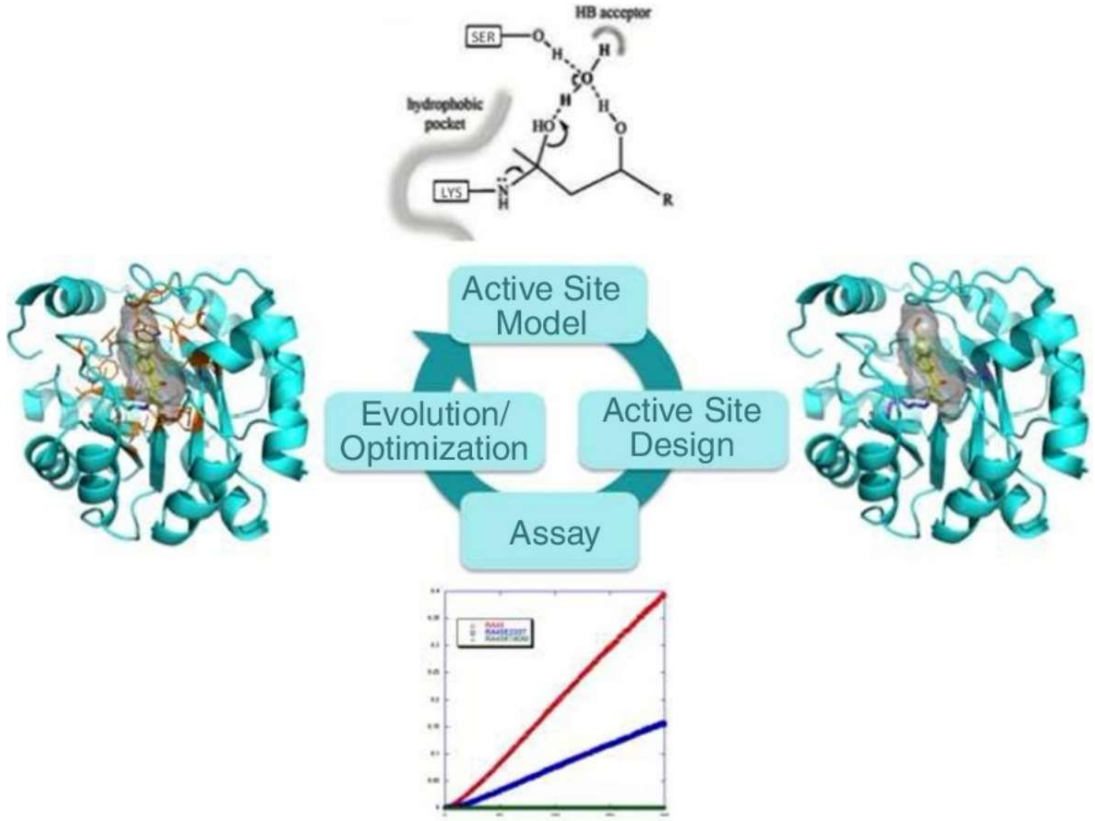


J. Swaminathan

PROTEINS

Part 2 – 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**

RATIONAL DESIGN

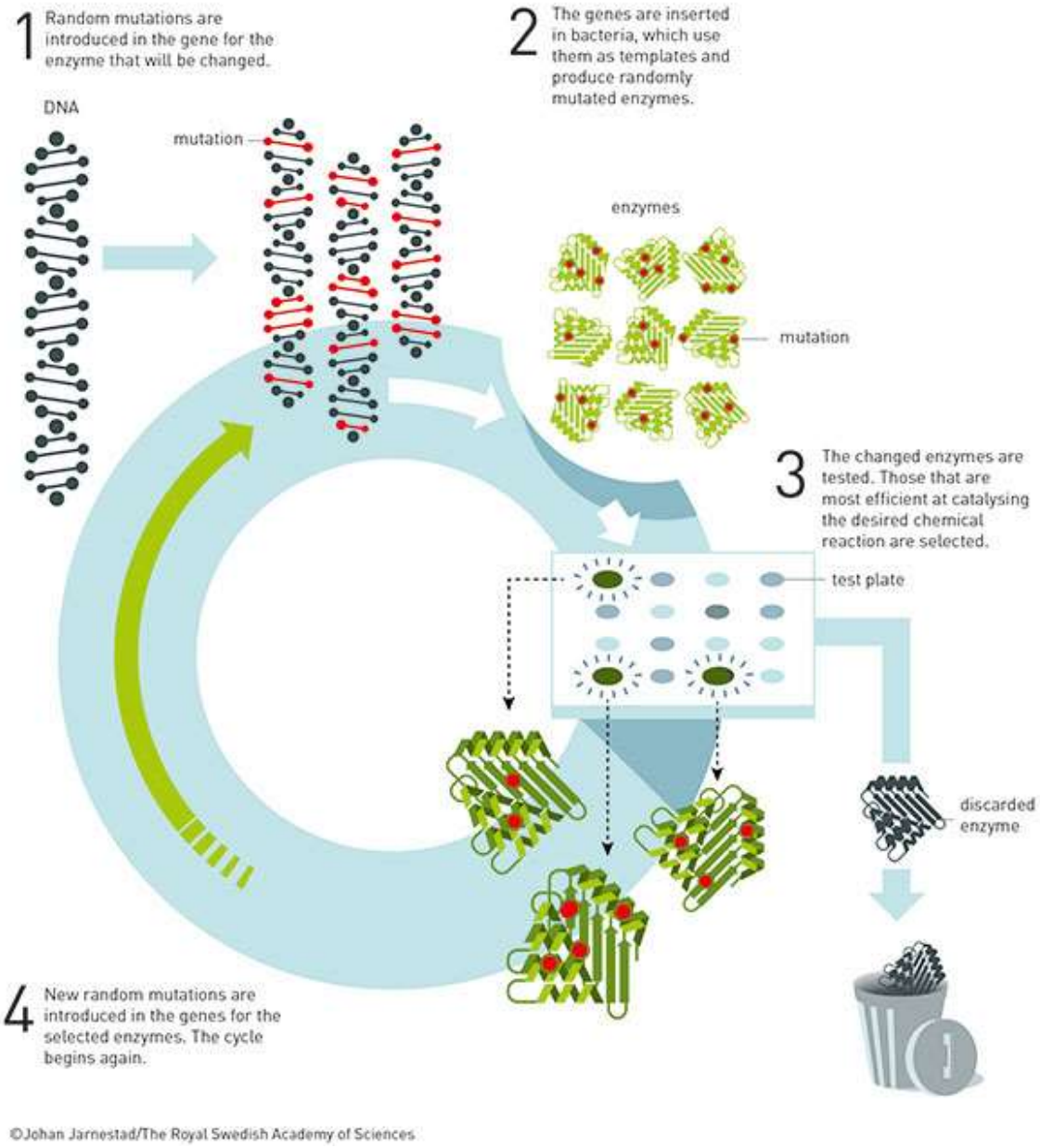
- Site directed mutagenesis of one or more residues
- Fusion of functional domains from different proteins to create chimaeric (Domain swapping)
- Functional evaluation

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

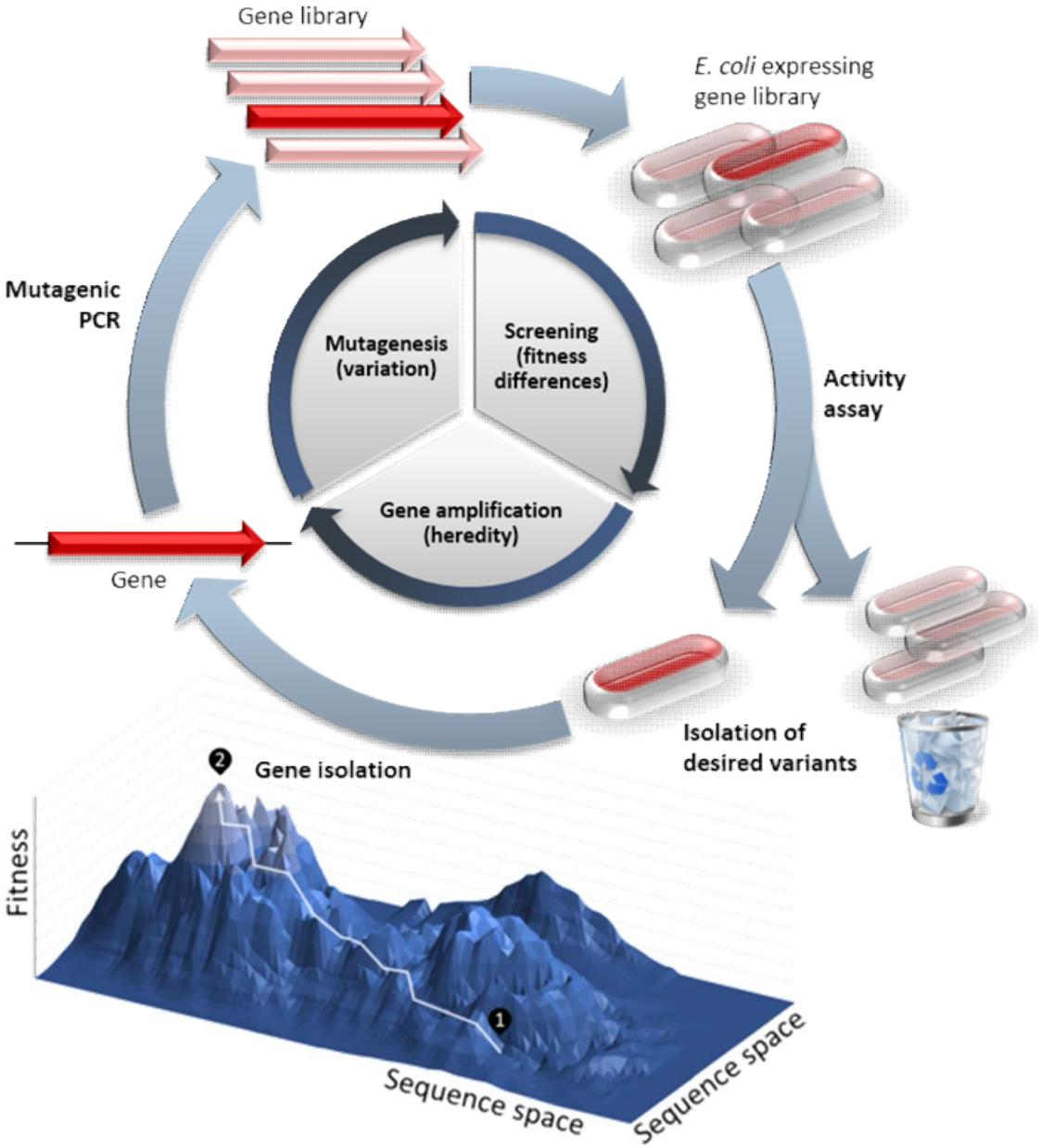
Protein engineering: rational design and directed evolution

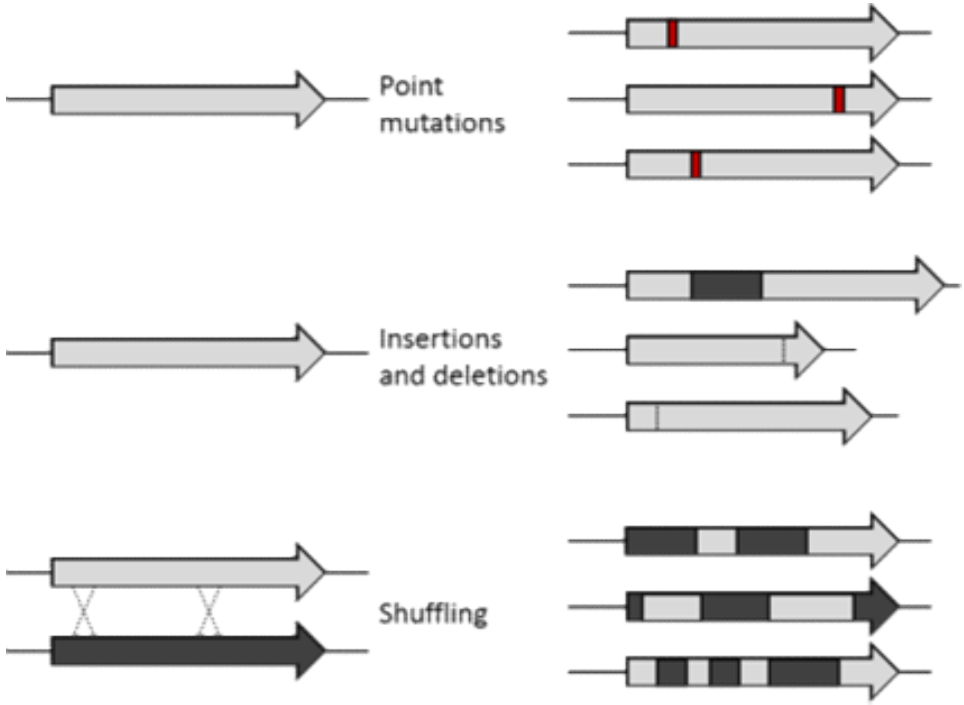
Nobel prize chemistry 2018
Frances Arnold

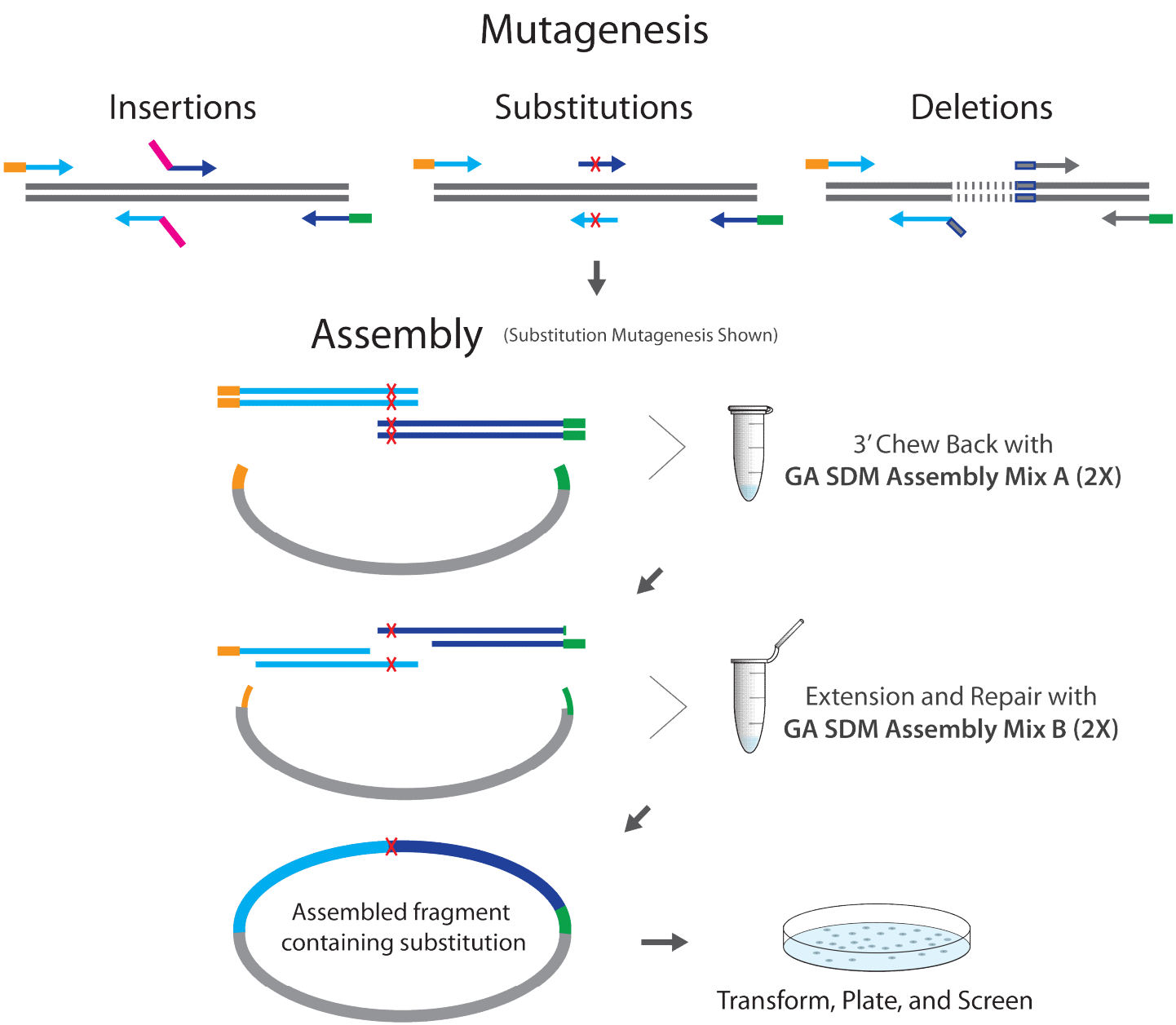
„for pioneering the use of directed evolution to engineer enzymes"



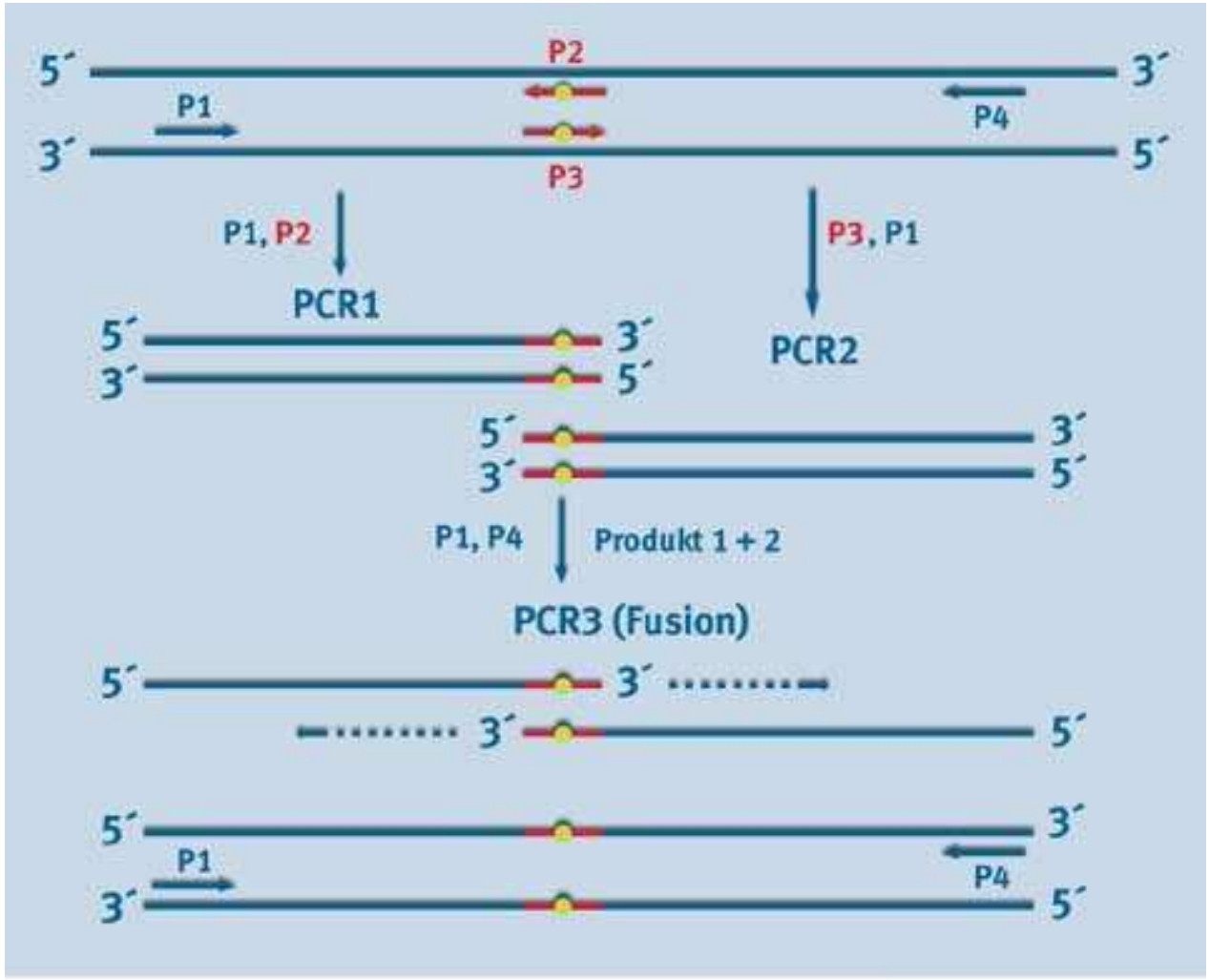
Directed evolution





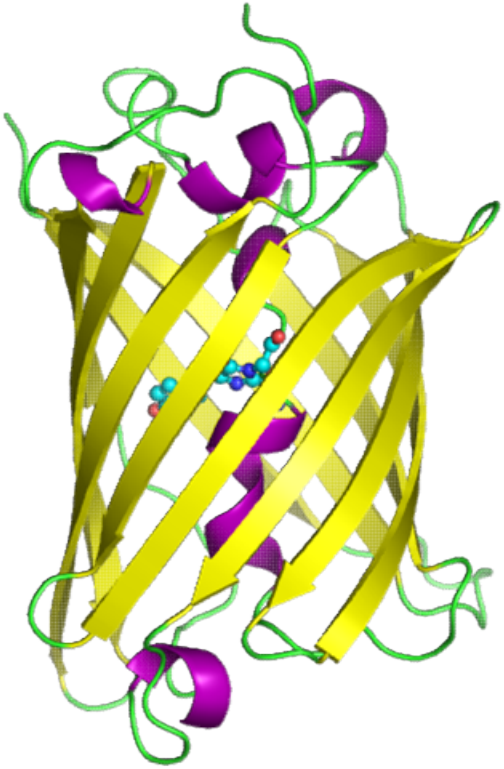

Mutagenesis
Insertions
Substitutions
Deletions
Assembly
(Substitution Mutagenesis Shown)
3' Chew Back with
GA SDM Assembly Mix A (2X)
Extension and Repair with
GA SDM Assembly Mix B (2X)
Assembled fragment
containing substitution
Transform, Plate, and Screen

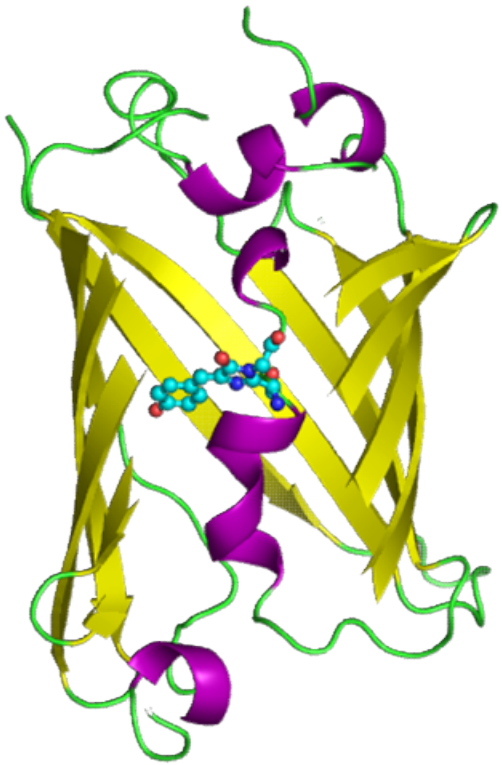
The Overlap-Extension PCR



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

Evolution of Green Fluorescent Protein





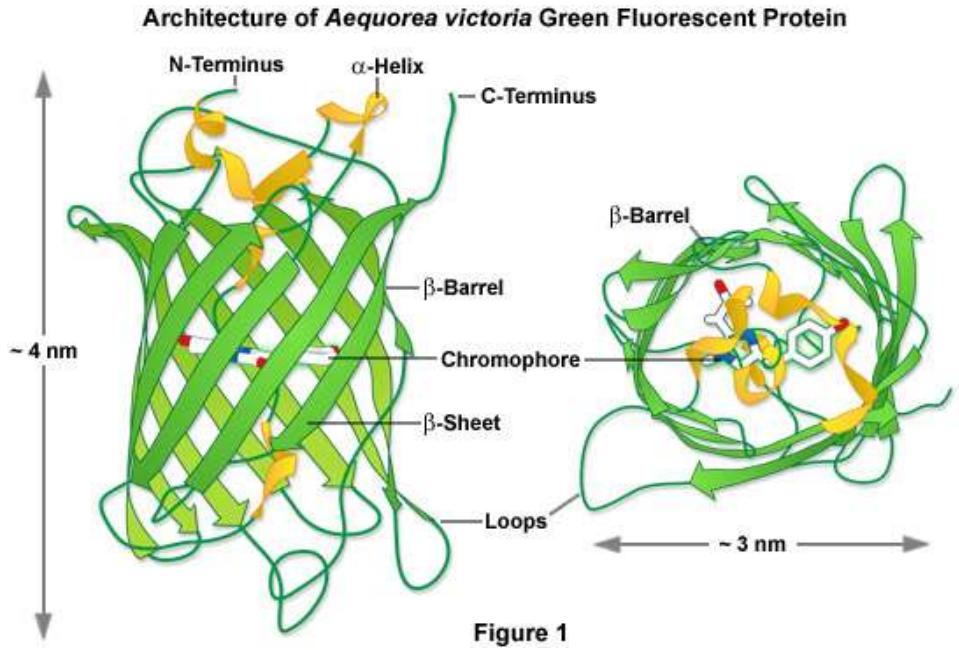
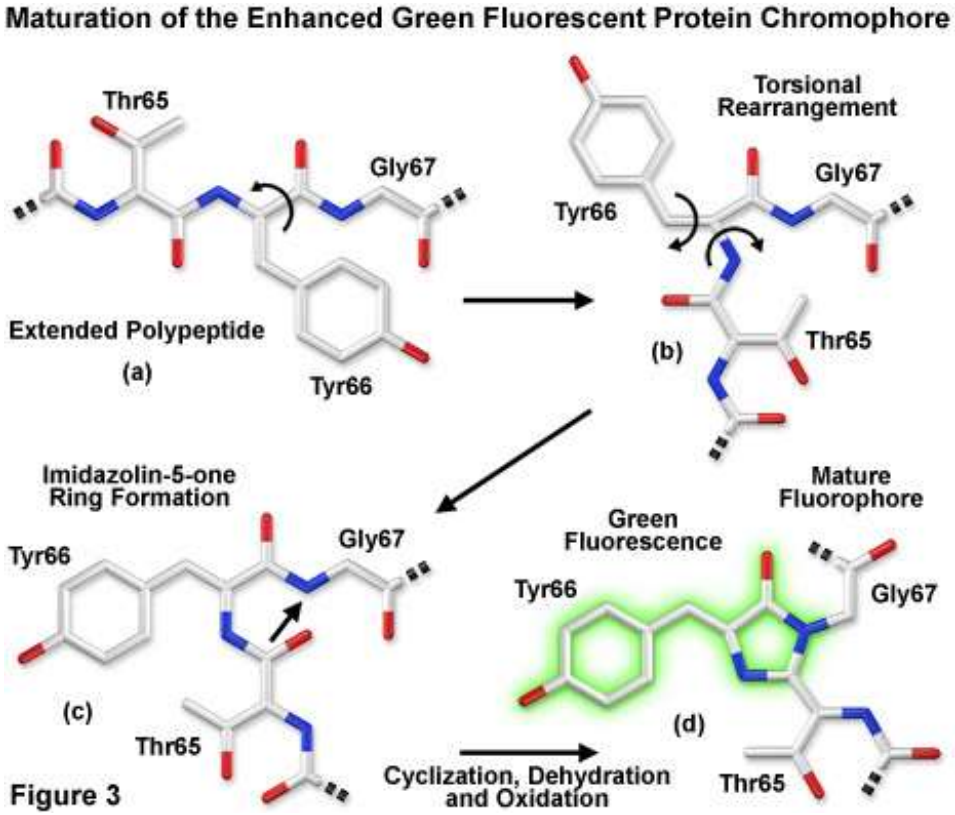
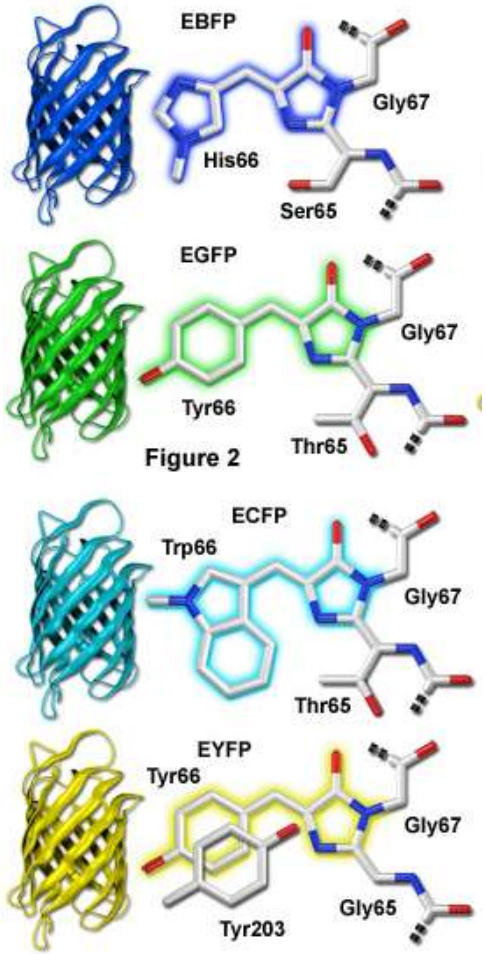
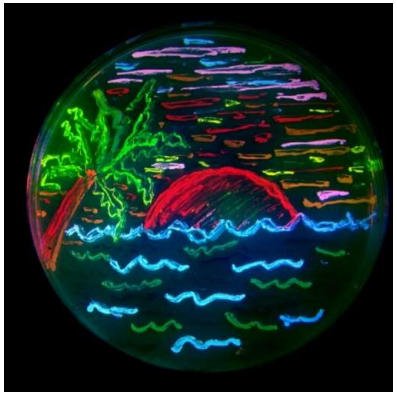


Figure 1

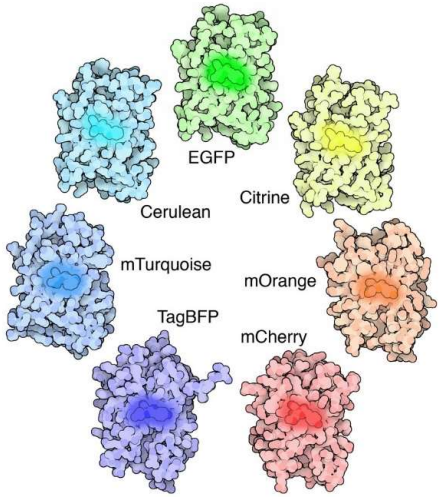
Green Fluorescent Protein







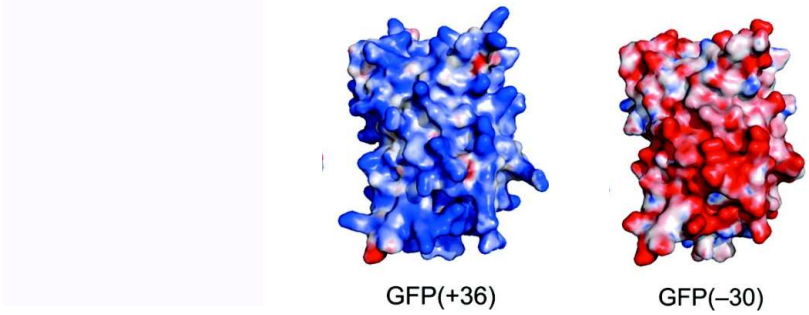
N. Shaner

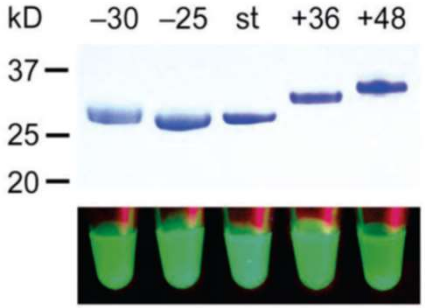


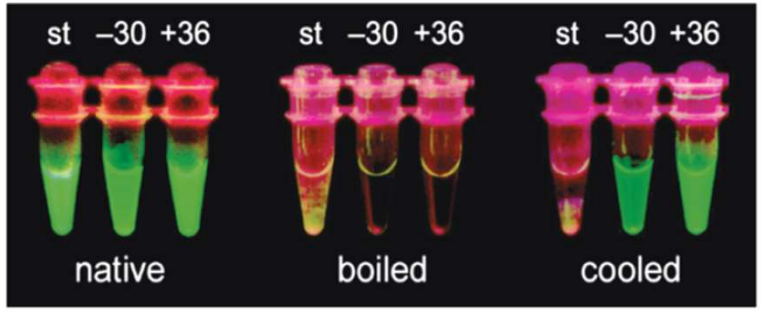
D. Goodsell

Green Fluorescent Protein

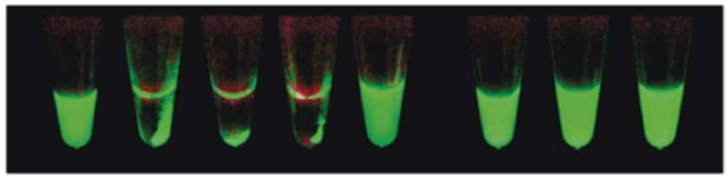
```
GFP (–30)  MGHHHHHHGGASKGEELFDGVVPILVELDGDVNGHEFSVRGEGEGDATEG
GFP (–25)  MGHHHHHHGGASKGEELFTGVVPILVELDGDVNGHEFSVRGEGEGDATEG
stGFP      MGHHHHHHGGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNG
GFP (+36)  MGHHHHHHGGASKGERLFRGKVPILVELKGDVNGHKFSVRGKGKGDATRG
GFP (+48)  MGHHHHHHGGRSKGKRLFRGKVPILVKLKGDVNGHKFSVRGKGKGDATRG

GFP (–30)  ELTLKFICTTGELPVPWPTLVTTLTYGVQCFSDYPDHMDQHDFFKSAMPE
GFP (–25)  ELTLKFICTTGELPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPE
stGFP      KLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPE
GFP (+36)  KLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPKHMKRHDFFKSAMPK
GFP (+48)  KLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPKHMKRHDFFKSAMPK

GFP (–30)  GYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK
GFP (–25)  GYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK
stGFP      GYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK
GFP (+36)  GYVQERTISFKKDGKYKTRAEVKFEGRTLVNRIKLKGRDFKEKGNILGHK
GFP (+48)  GYVQERTISFKKDGKYKTRAEVKFKGRTLVNRIKLKGRDFKEKGNILGHK

GFP (–30)  LEYNFNSHDVYITADKQENGIKAEFEIRHNVEDGSVQLADHYQQNTPIGD
GFP (–25)  LEYNFNSHDVYITADKQENGIKAEFEIRHNVEDGSVQLADHYQQNTPIGD
stGFP      LEYNFNSHNVYITADKQKNGIKANFKIRHNVEDGSVQLADHYQQNTPIGD
GFP (+36)  LRYNFNSHKVYITADKRKNGIKAKFKIRHNVKDGSVQLADHYQQNTPIGR
GFP (+48)  LRYNFNSHKVYITADKRKNGIKAKFKIRHNVKDGSVQLAKHYQQNTPIGR

GFP (–30)  GPVLLPDDHYLSTESALSKDPNEDRDHMVLLEFVTAAGIDHGMDELYK
GFP (–25)  GPVLLPDDHYLSTESALSKDPNEDRDHMVLLEFVTAAGIDHGMDELYK
stGFP      GPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK
GFP (+36)  GPVLLPRNHYLSTRSKLSKDPKEKRDHMVLLEFVTAAGIKHGRDERYK
GFP (+48)  GPVLLPRKHYLSTRSKLSKDPKEKRDHMVLLEFVTAAGIKHGRKERYK
```



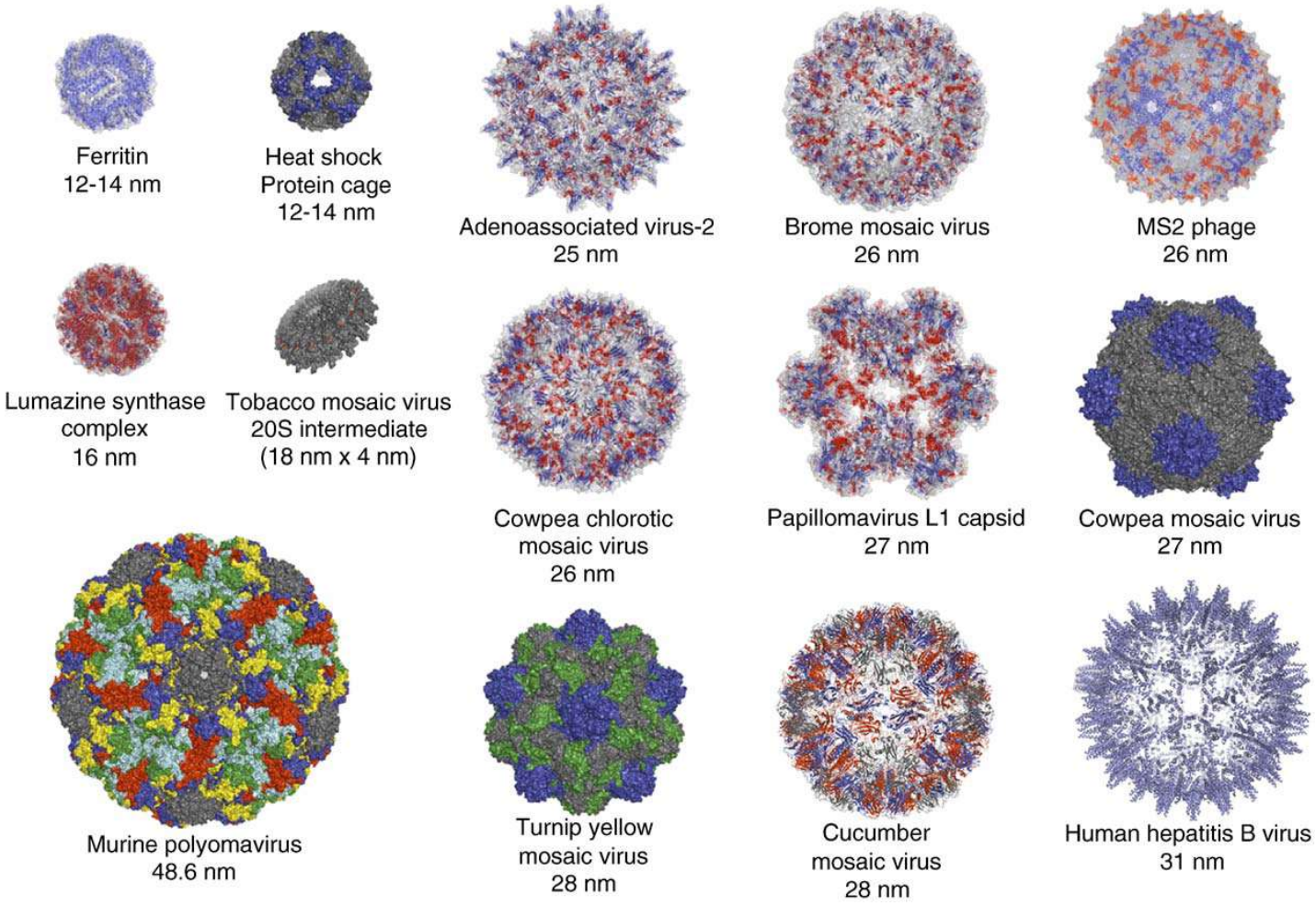




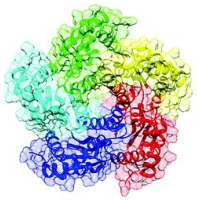
	1	2	3	4	5	6	7	8
GFP (+36)	+	+	+	+	+	–	–	–
stGFP	–	–	–	–	–	+	+	+
GFP (–30)	–	+	–	–	–	–	+	–
DNA	–	–	+	–	–	–	–	–
tRNA	–	–	–	+	+	–	–	+
NaCl	–	–	–	–	+	–	–	–



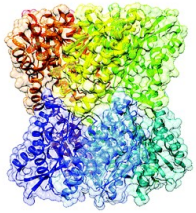
Example – protein capsids



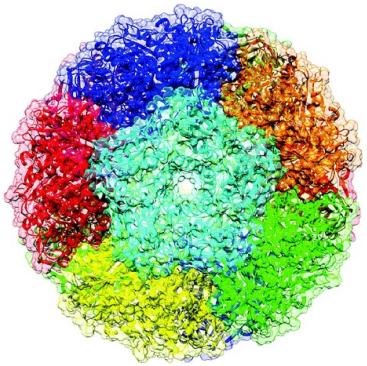
Lumazine synthase



pentamer
Schizosaccharomyces pombe

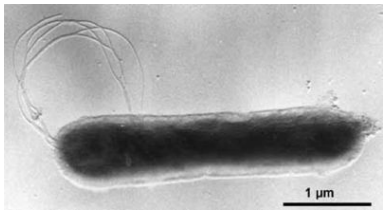


dimer of pentamers
Brucella abortus

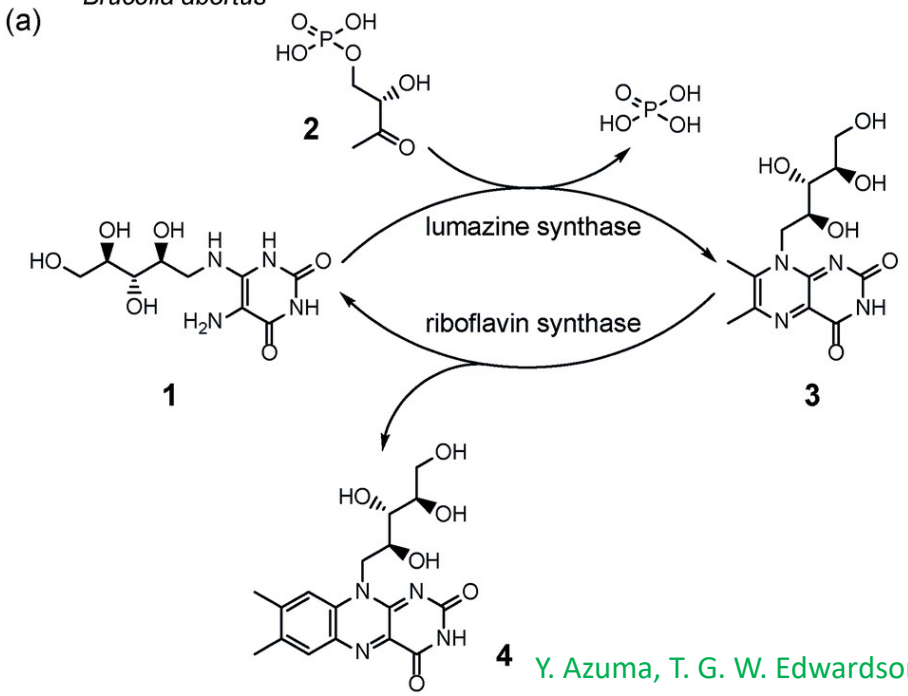


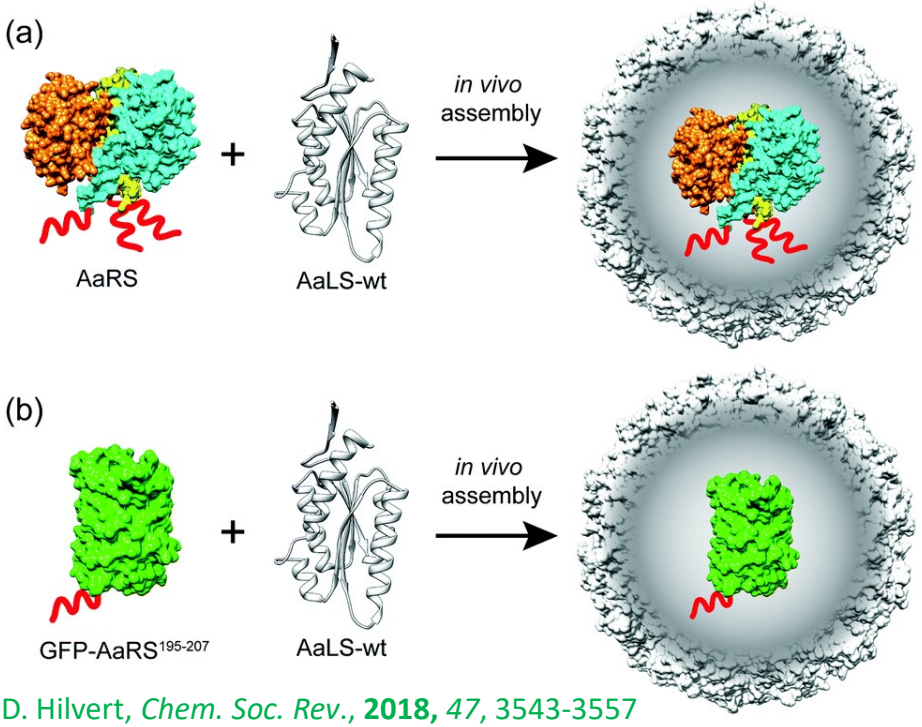
dodecahedron
formed from 12 pentamers
Aquifex aeolicus

Evolvability of proteins from thermophiles



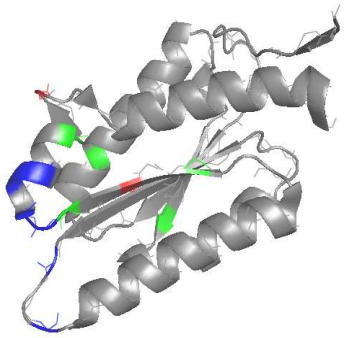
Aquifex aeolicus (Aa) – thermophilic bacteria

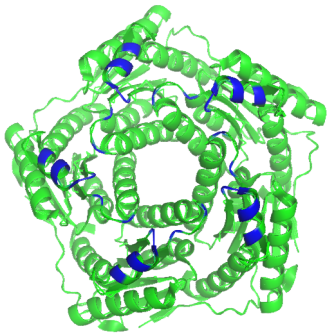


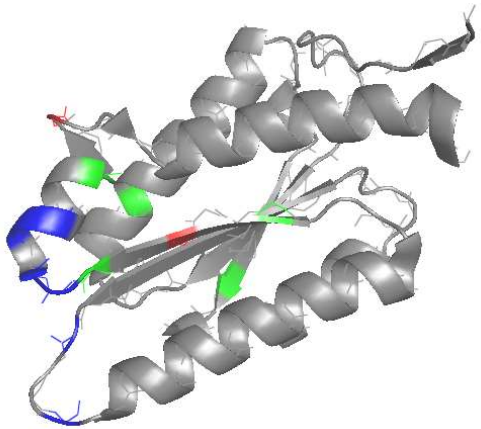


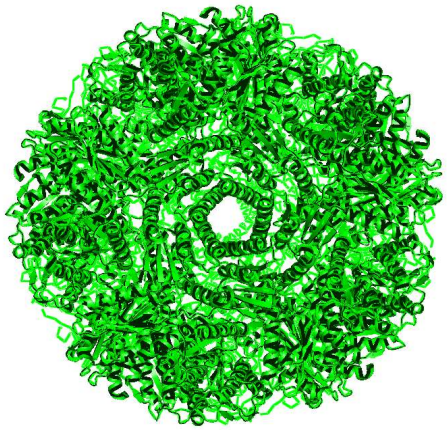
Y. Azuma, T. G. W. Edwardson and D. Hilvert, *Chem. Soc. Rev.*, **2018,** *47*, 3543-3557

Protein capsids - lumazine synthase







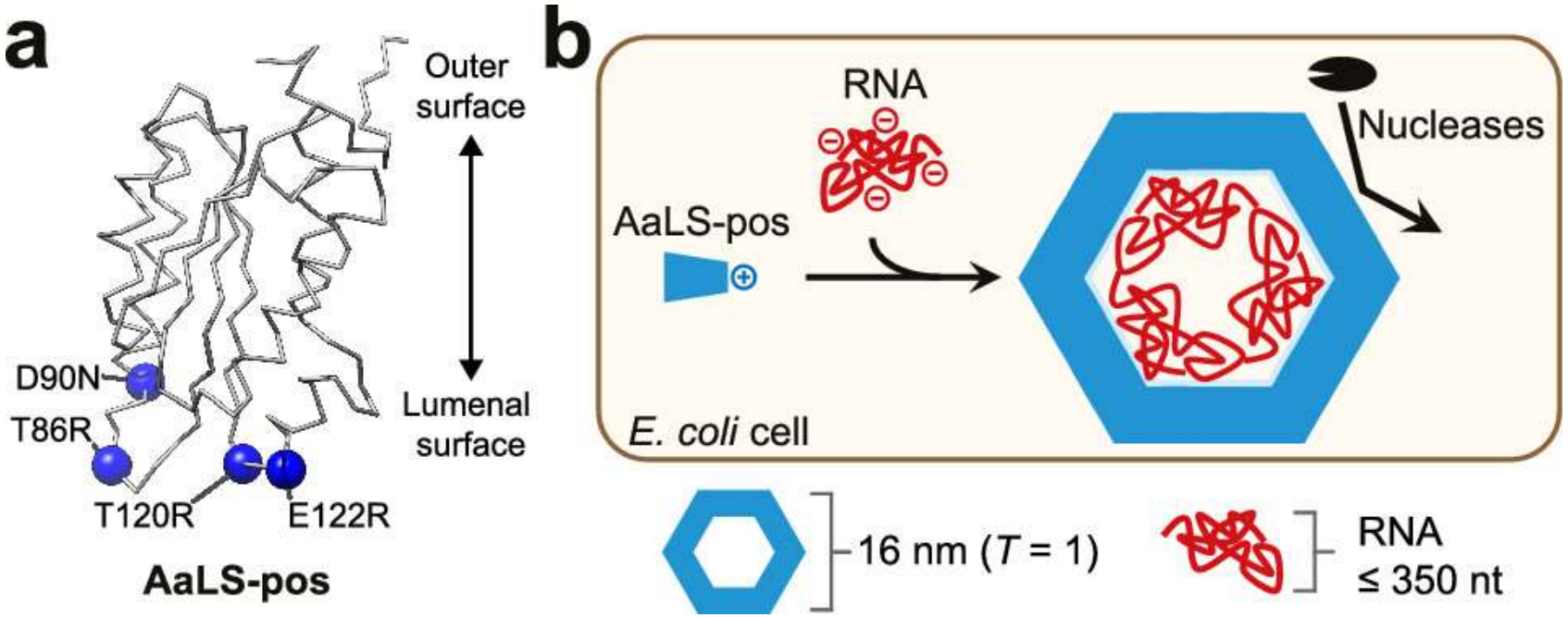


wt AaLS:

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

162 Aminoacids

Engineering of lumazine synthase

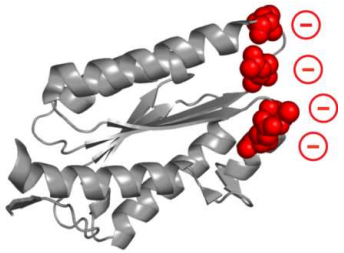


AaLS-pos co-assembles with RNA *in vivo* and protects its cargo from nuclease digestion.

The capsid retains the same 16 nm diameter as its parent AaLS-wt and encapsulates RNAs of up to 350 nt in length.

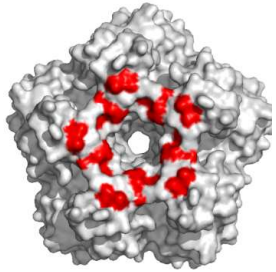
T. G. W. Edwardson and D. Hilvert, *J. Am. Chem. Soc.*, **2019**, *141*, 9432-9443

Engineering of lumazine synthase

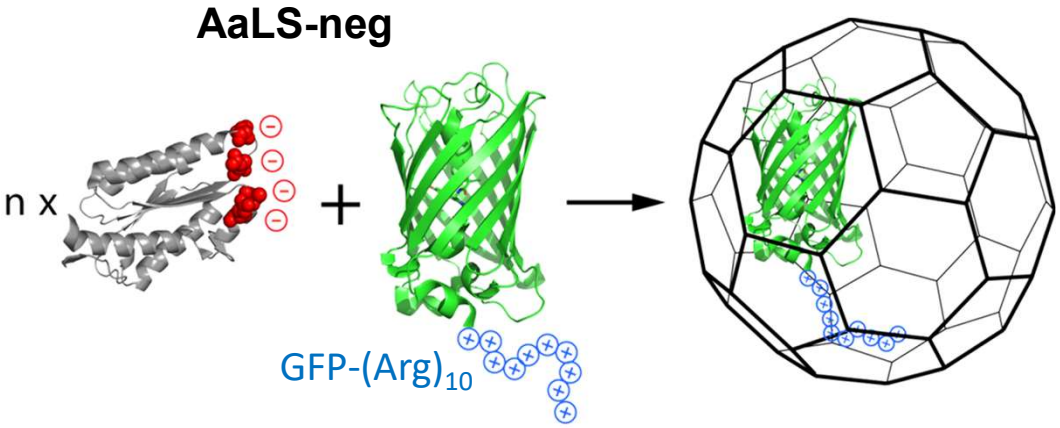


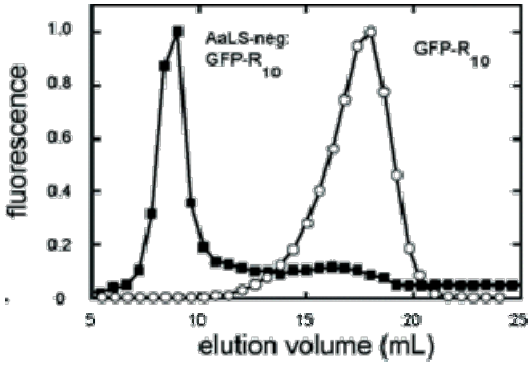
Arg83Glu/Thr86Glu/Thr120Glu/Gln123Glu

300 extra charges per 60-meric capsid

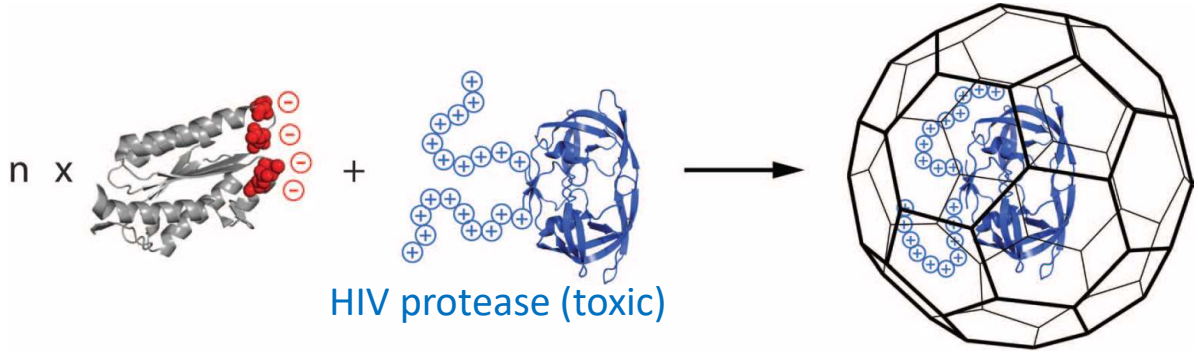


AaLS-neg





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

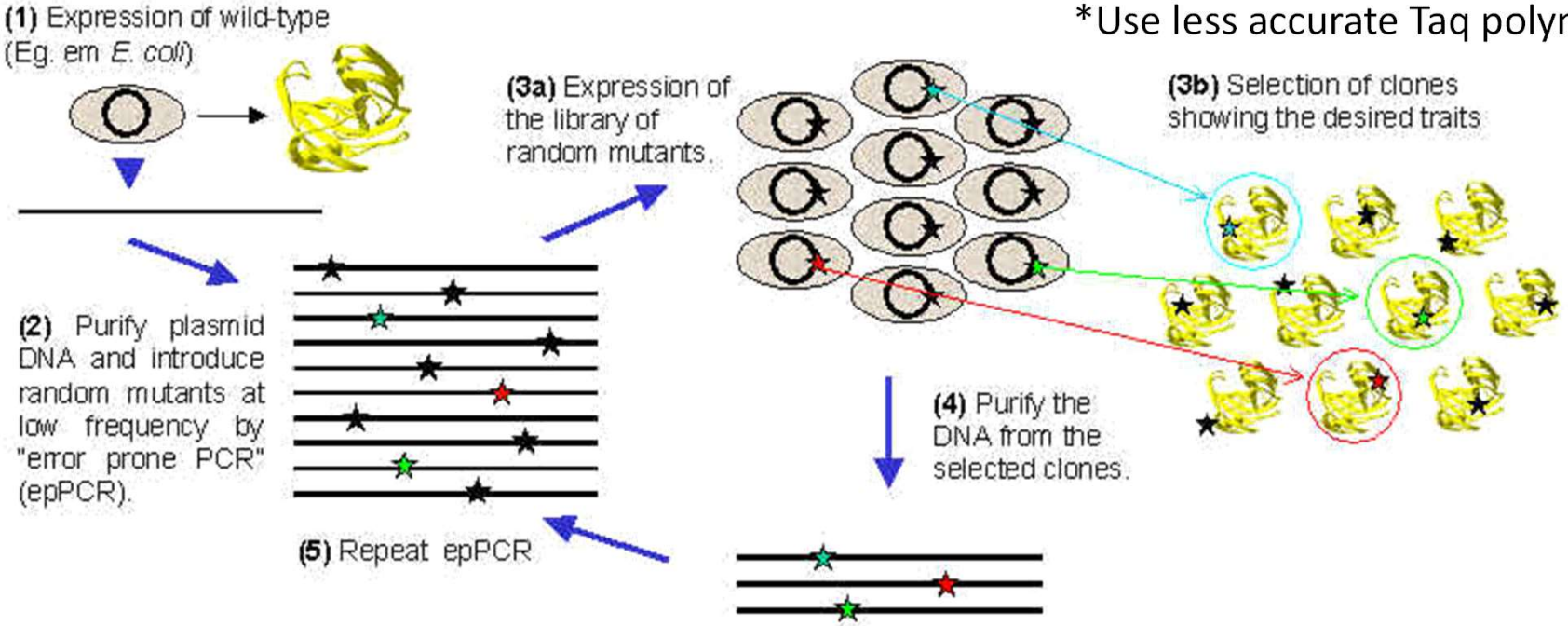


B. Woersdoerfer, K.J.Woycechowsky, D.Hilvert *Science* **2011,** *331,* 589-592

Error prone PCR library construction

Taq DNA polymerase - Rates of error in PCR can be increased in the following ways:

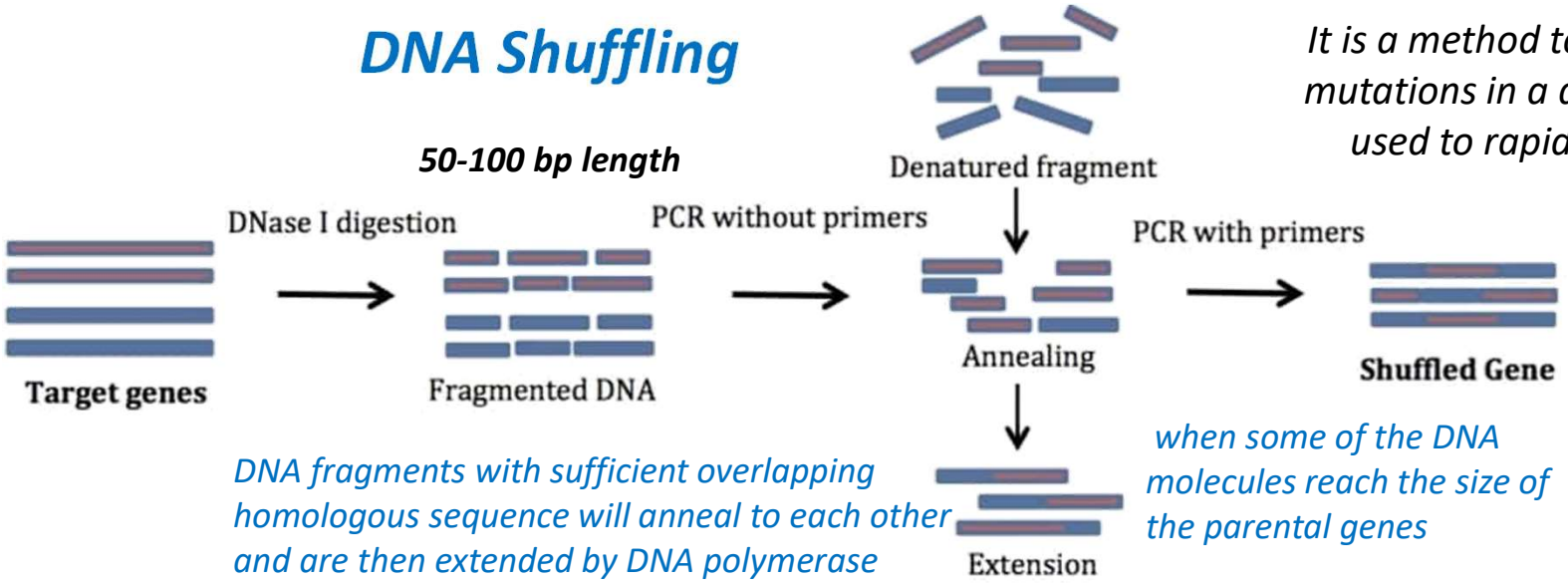
*Increase concentration of magnesium chloride, which stabilizes non complementary base pairing.
*Add $MnCl_2$ to reduce base pair specificity.
*Increased and unbalanced addition of dNTPs.
*Addition of base analogs dITP, 8 oxo-dGTP,dPTP.
*Increase concentration of Taq polymerase, extension time, cycle time.
*Use less accurate Taq polymerase.



DNA Shuffling

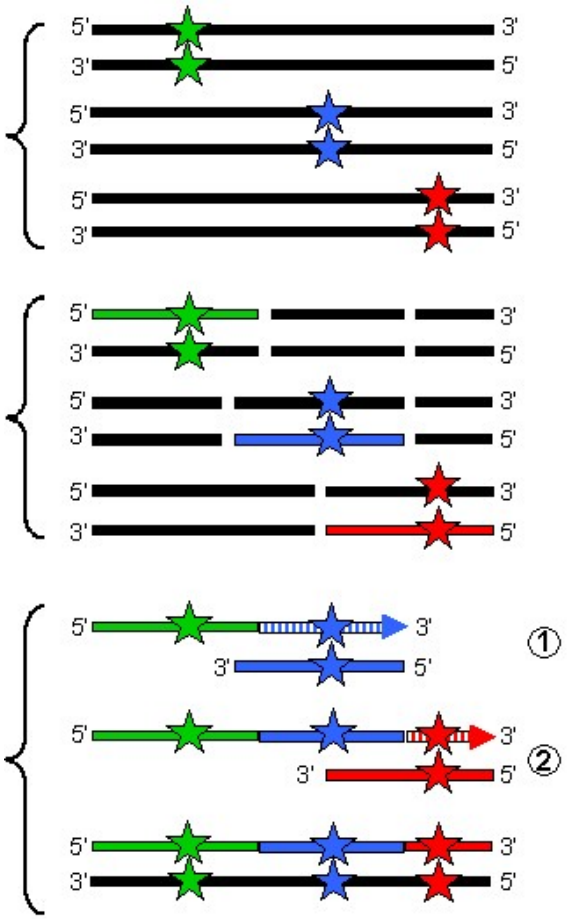
It is a method to rapidly propagate beneficial mutations in a directed evolution experiment, used to rapidly increase DNA library size

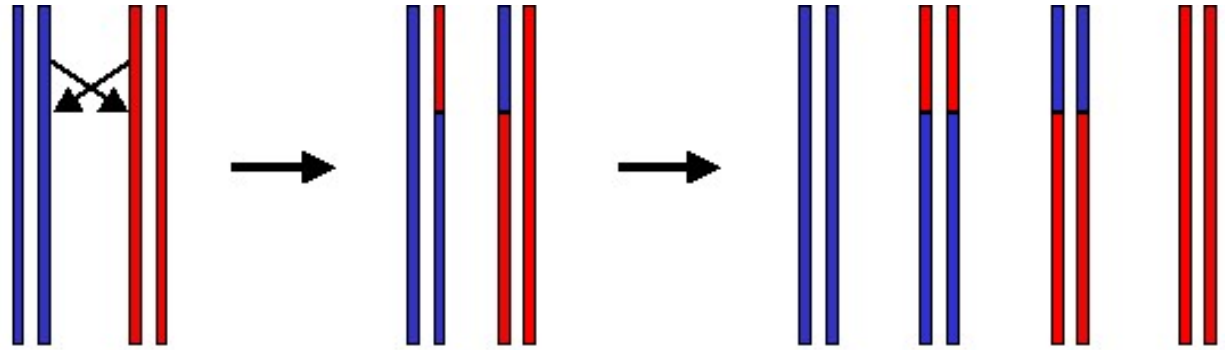
50-100 bp length



DNA fragments with sufficient overlapping homologous sequence will anneal to each other and are then extended by DNA polymerase

when some of the DNA molecules reach the size of the parental genes



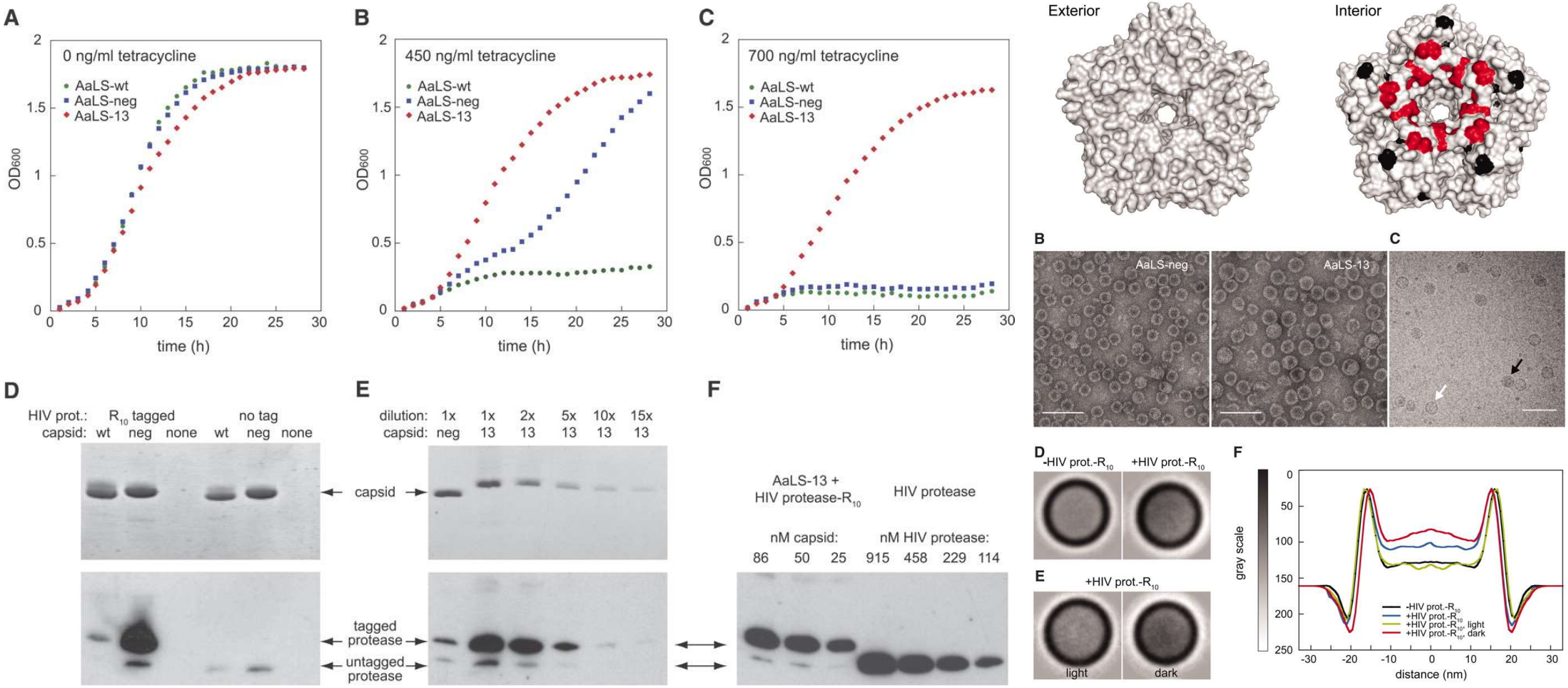


(a) Two DNA double helices sharing high sequence similarity approach each other

(b) One strand from each helix swaps with a strand fron the other helix

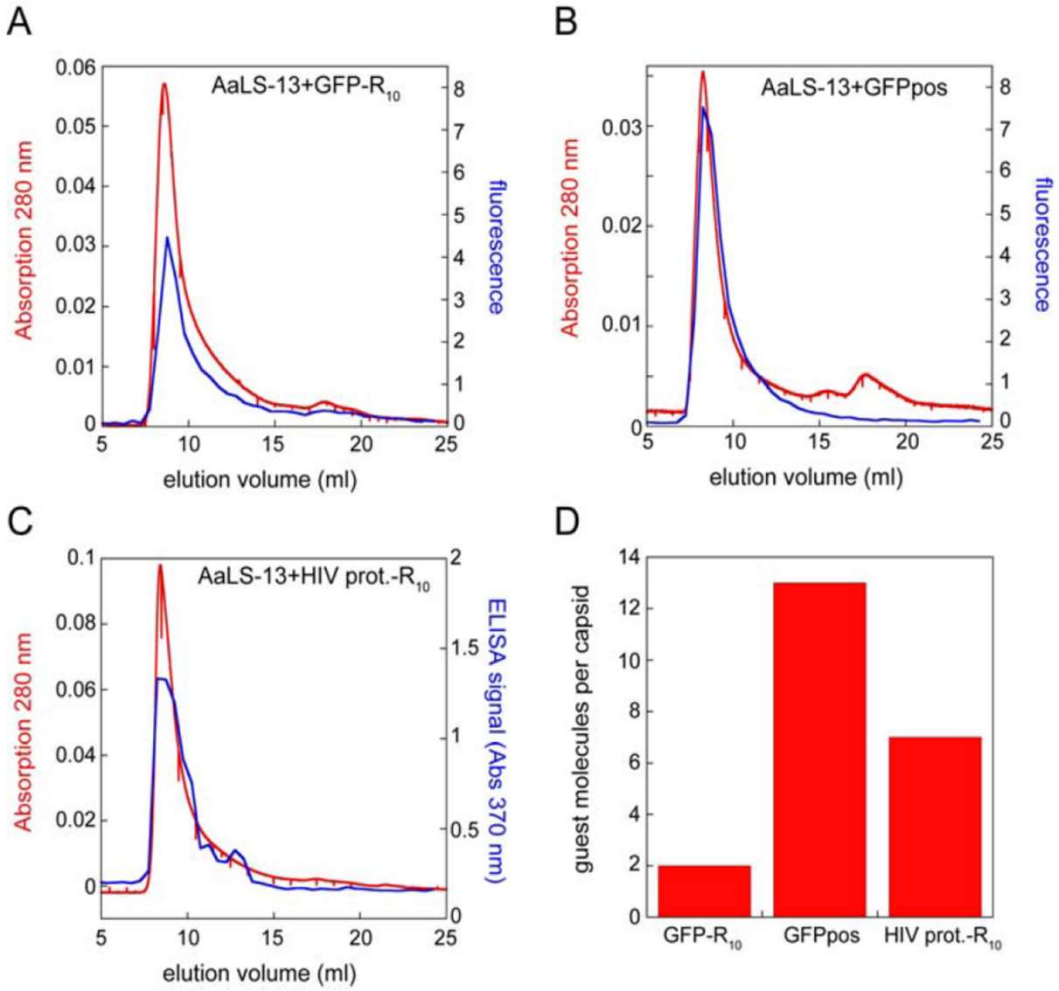
(c) All strands go through one round of normal semi-conservative DNA replication, resulting in the regeneration of the two original double helices, together with two new hybrid DNA double helices.

Engineering of lumazine synthase



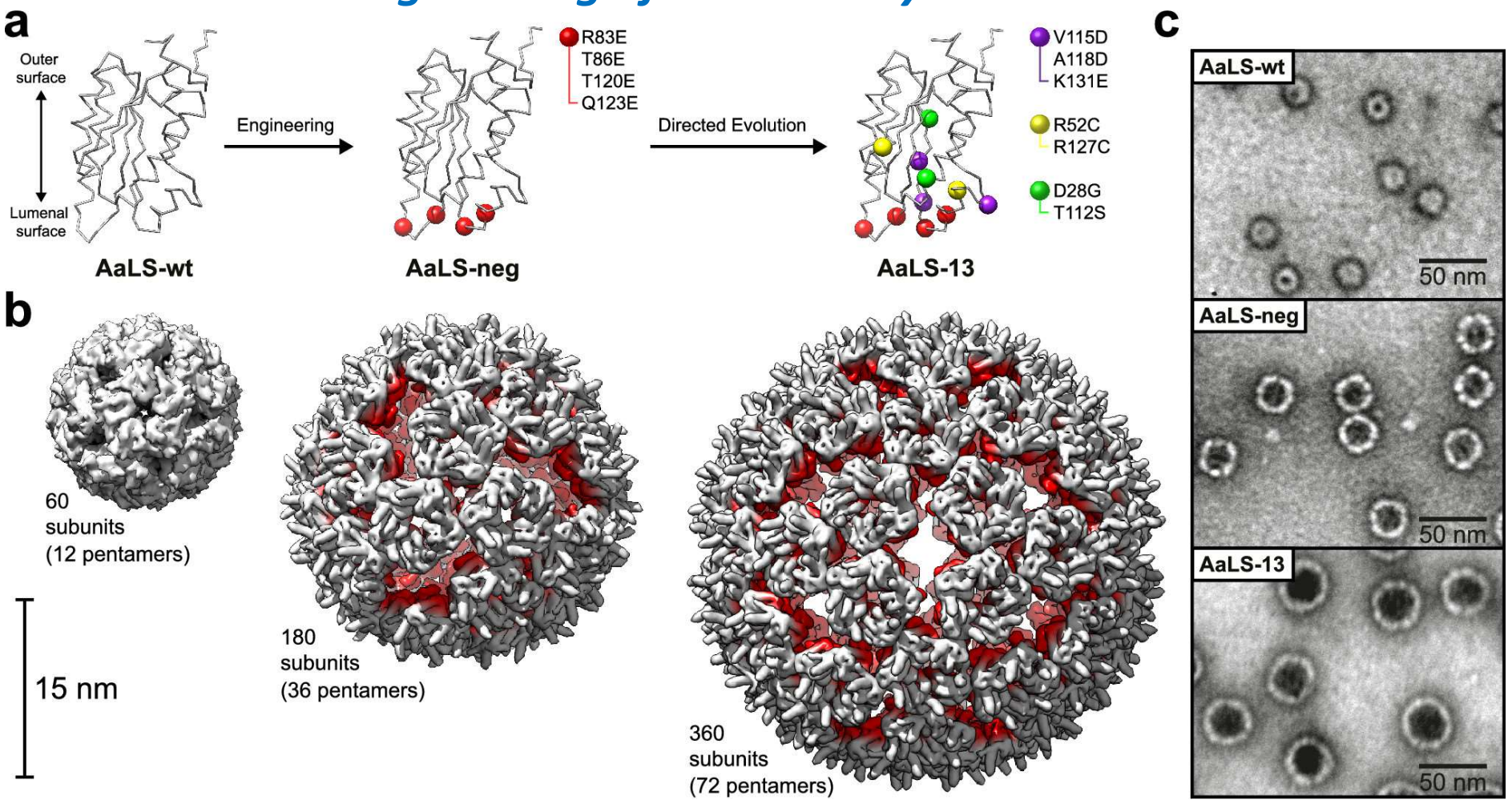
B. Woersdoerfer, K.J.Woycechowsky, D.Hilvert *Science* **2011,** *331,* 589-592

Engineering of lumazine synthase



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

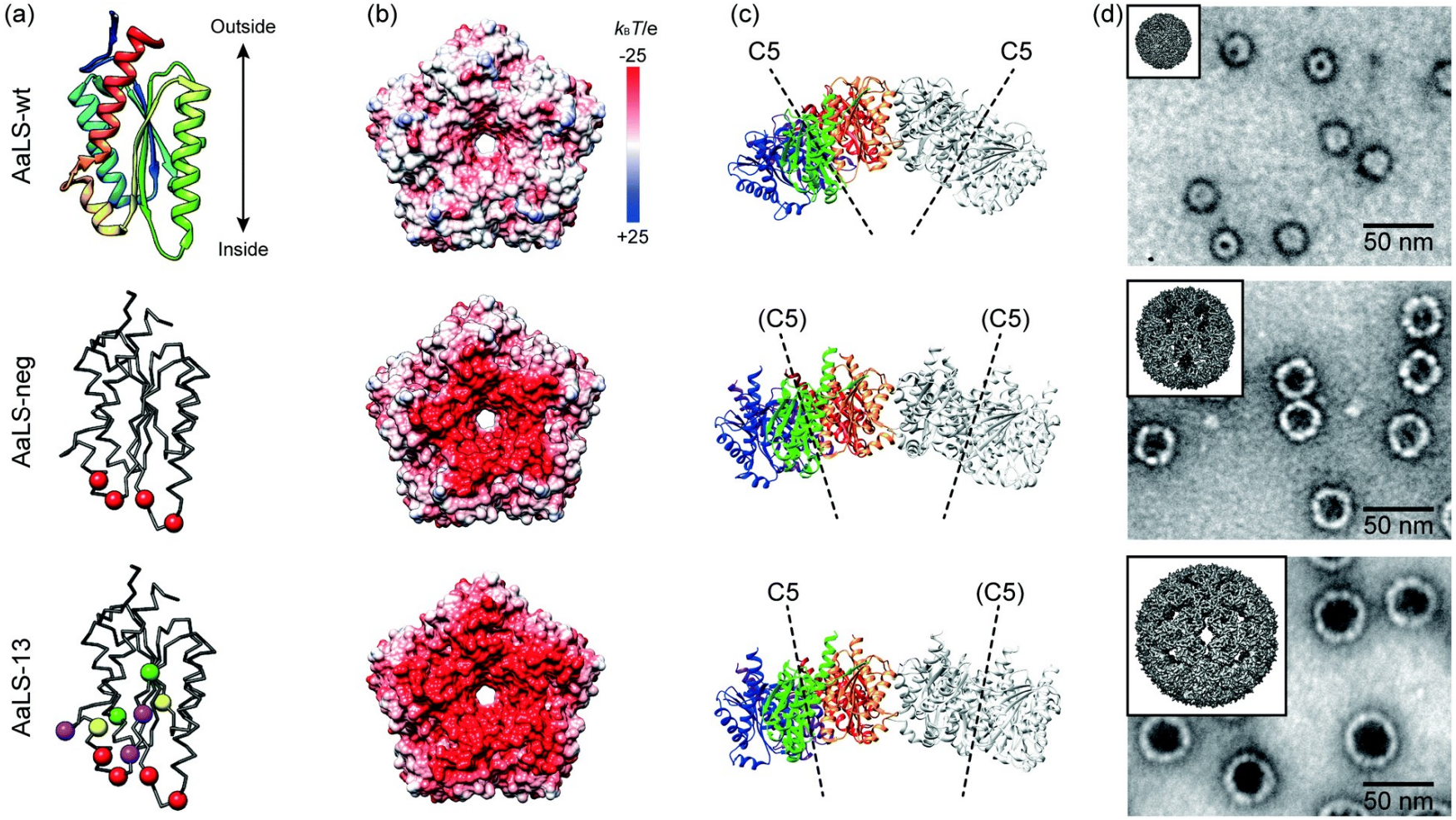
Engineering of lumazine synthase



PDB IDs: 1HQK (AaLS-wt), 5MQ3 (AaLS-neg), and 5MQ7 (AaLS-13)

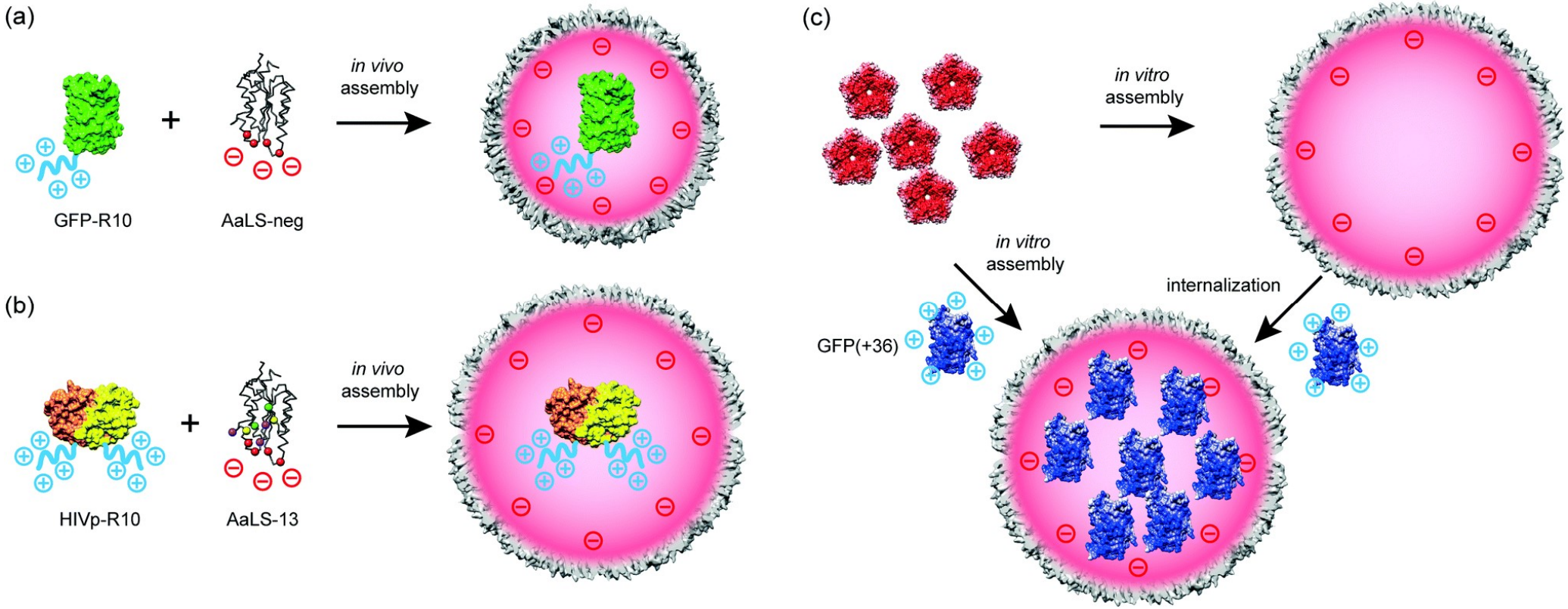
T. G. W. Edwardson and D. Hilvert, *J. Am. Chem. Soc.*, **2019**, *141*, 9432-9443

Engineering of lumazine synthase



B. Wörsdörfer, K. J. Woycechowsky and D. Hilvert, *Science*, **2011**, *331*, 589–592

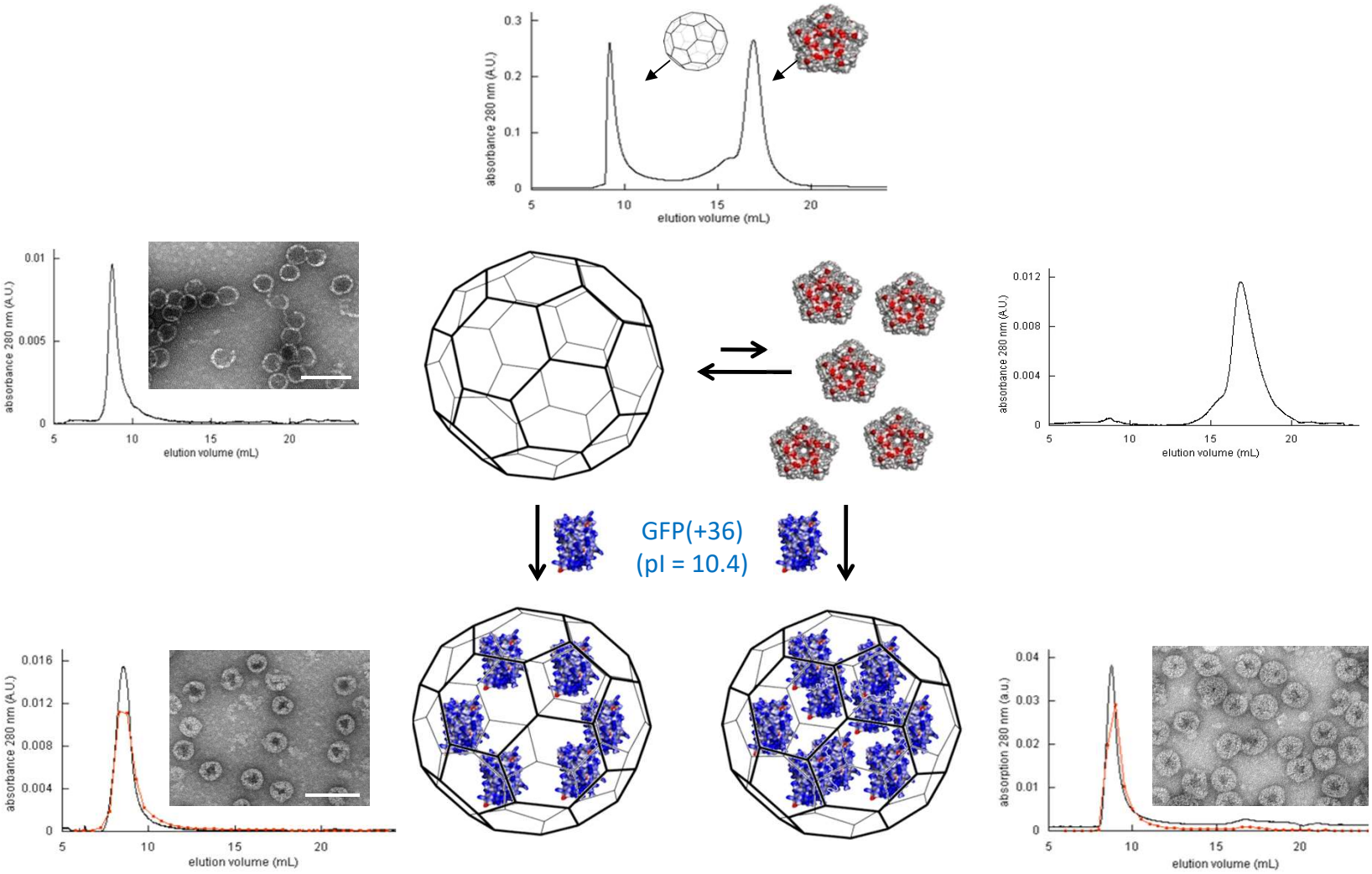
Engineering of lumazine synthase



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

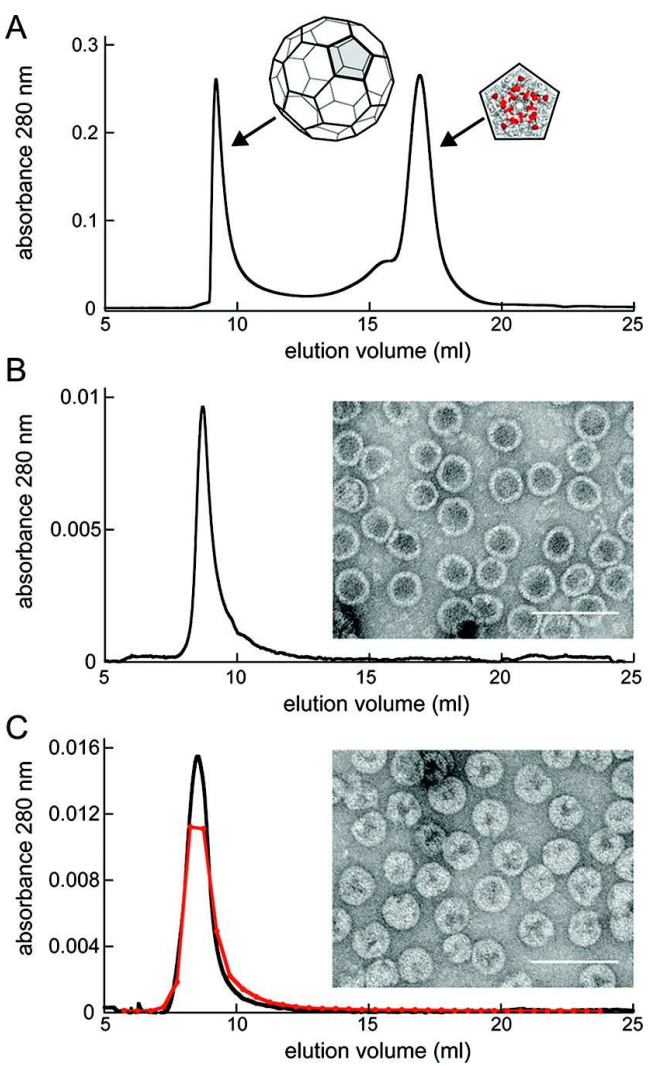
B. Wörsdörfer, Z. Pianowski and D. Hilvert, *J. Am. Chem. Soc.*, **2012**, *134*, 909–911

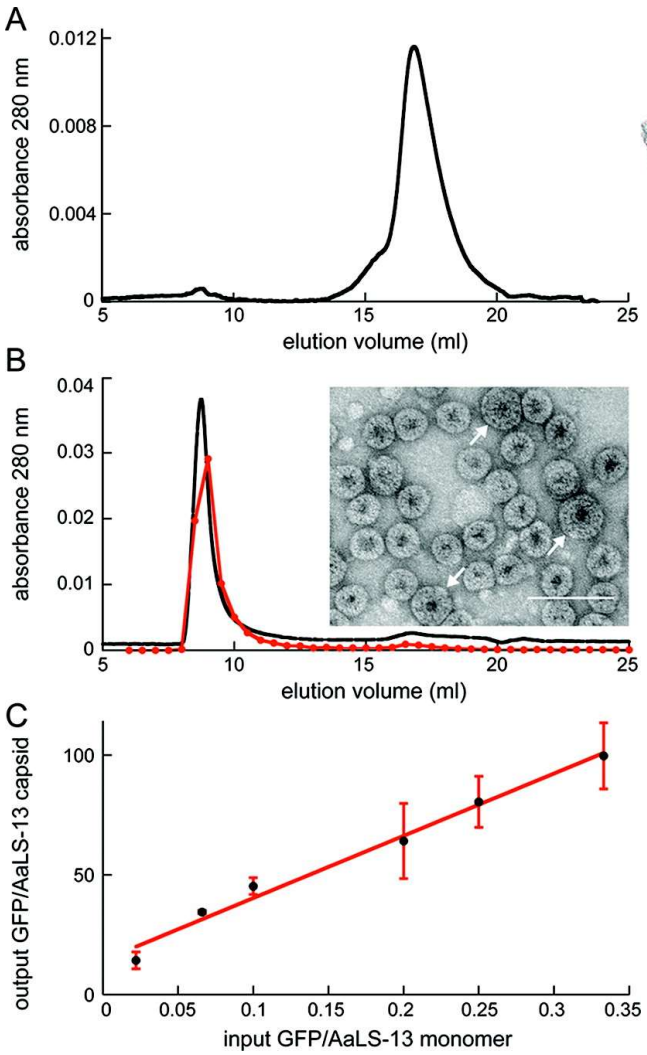
Engineering of lumazine synthase

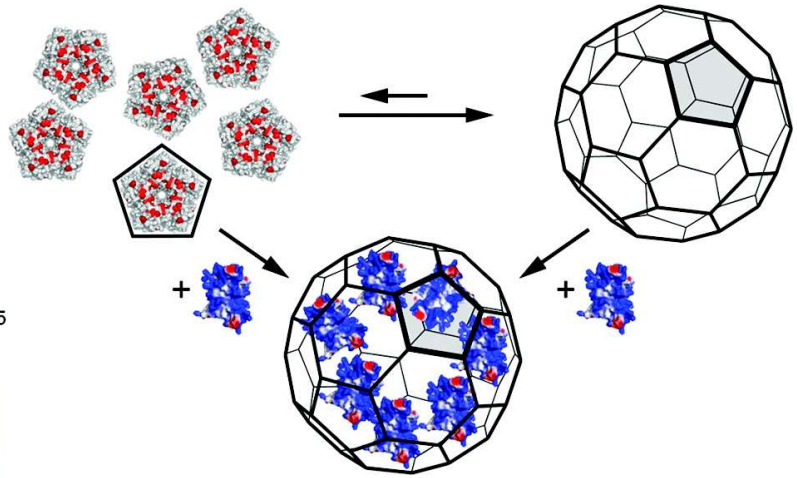


B. Woersdoerfer, Z. Pianowski, D.Hilvert *J. Am. Chem. Soc.* **2012,** *134,* 909-911

Engineering of lumazine synthase





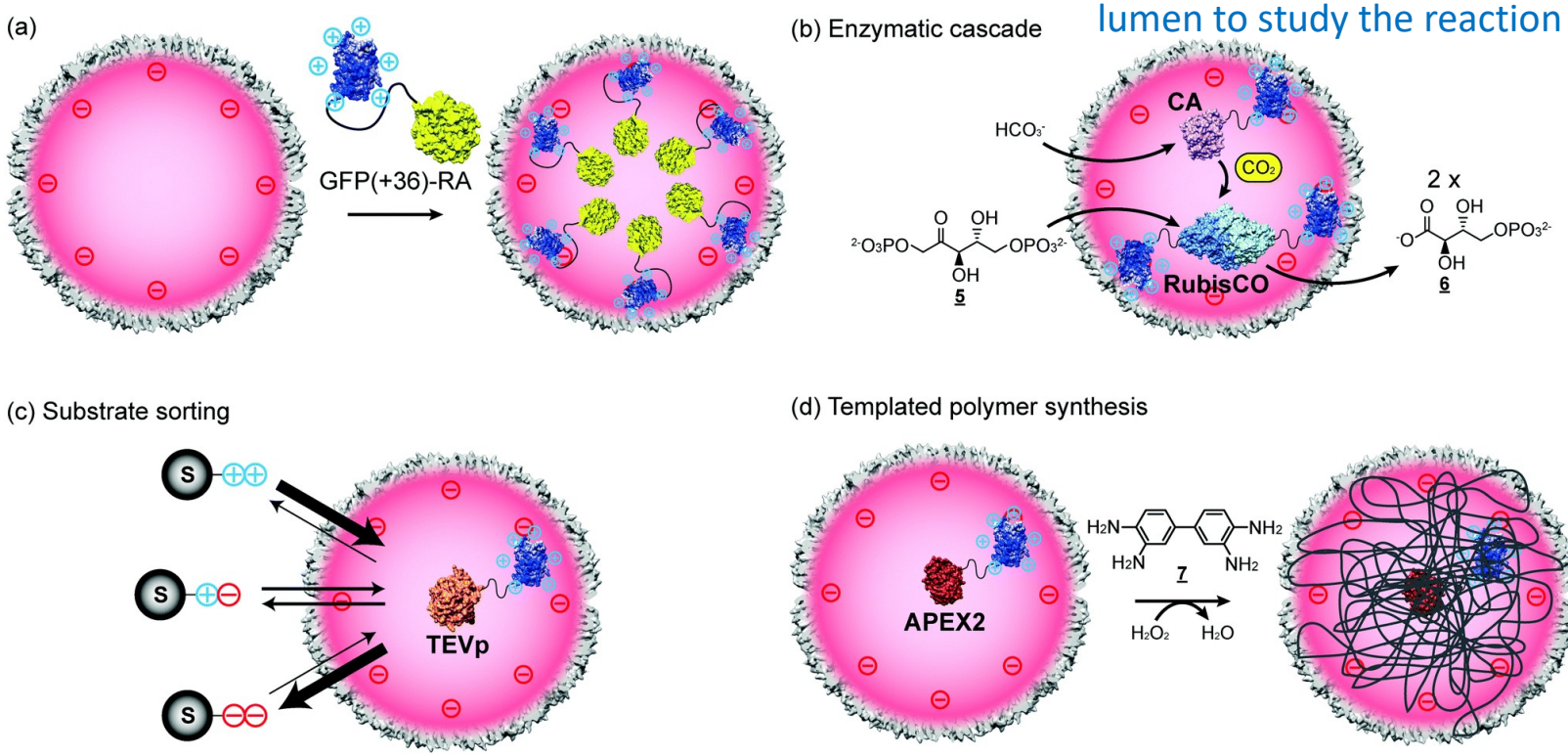


B. Woersdoerfer, Z. Pianowski, D.Hilvert *J. Am. Chem. Soc.* **2012,** *134,* 909-911

Lumazine synthase-based nanoreactors

Rapid and quantitative encapsulation of GFP(+36)-fusion enzymes by empty AaLS-13 cages in vitro (RA, retroaldolase)

A synthetic mimic of the carboxysome: ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) and carbonic anhydrase (CA), were co-packaged in the lumen to study the reaction cascade.

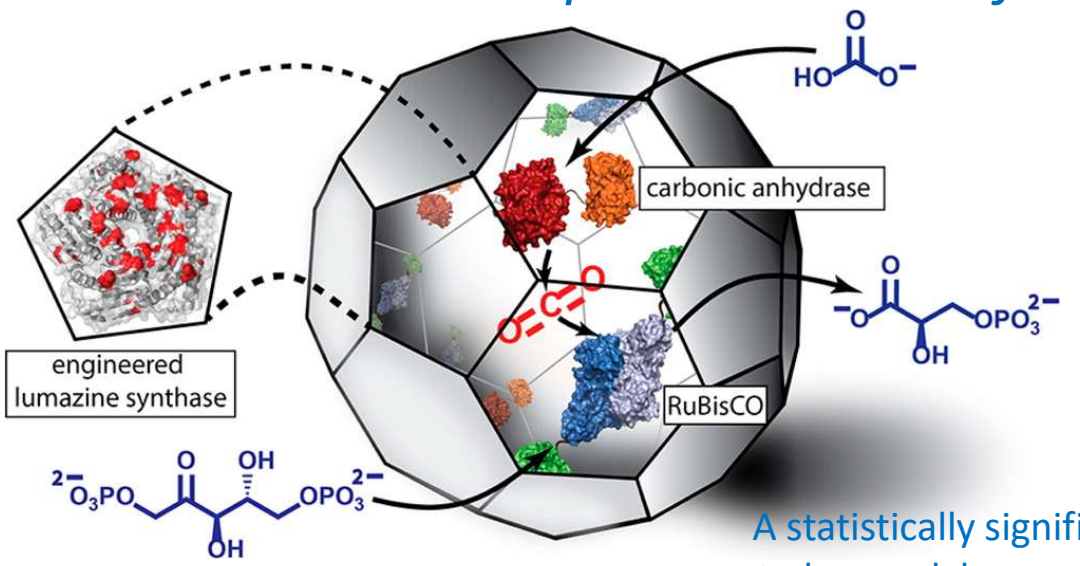


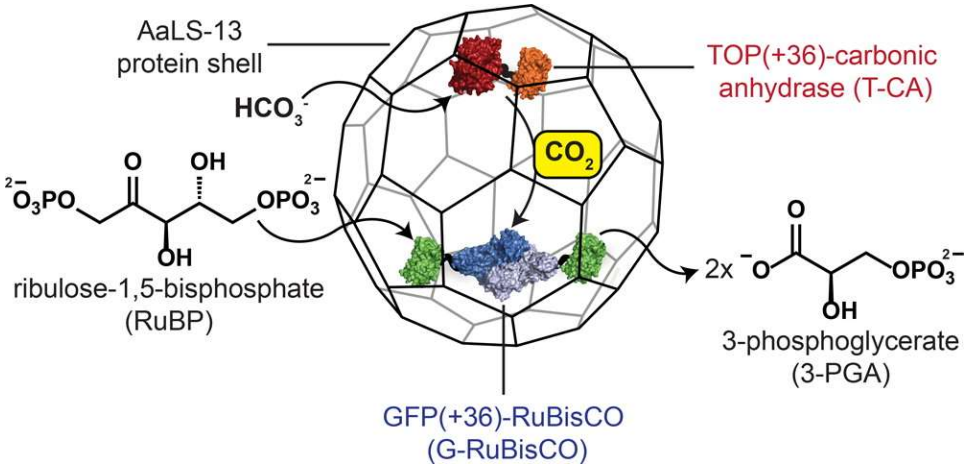
Substrate sorting for encapsulated tobacco etch virus protease (TEVp) based on electrostatic interactions with the lumenal surface of the AaLS-13 cage

AaLS-13 as templates for monodisperse polymer synthesis. An engineered variant of ascorbate peroxidase (APEX2) enables confined polymerization of compound (7) in the lumen of the cage.

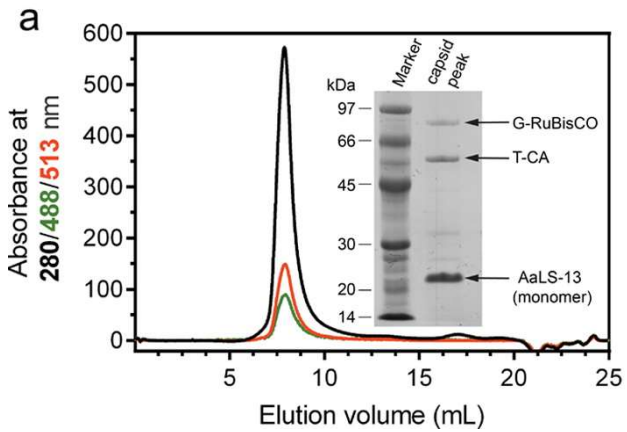
Engineering of lumazine synthase

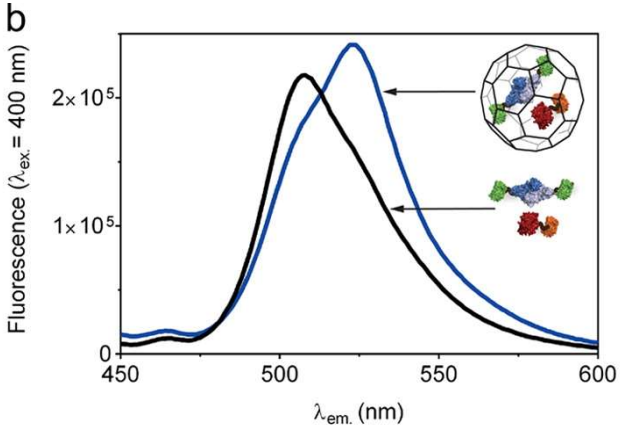
Bottom-up Construction of a Primordial Carboxysome Mimic

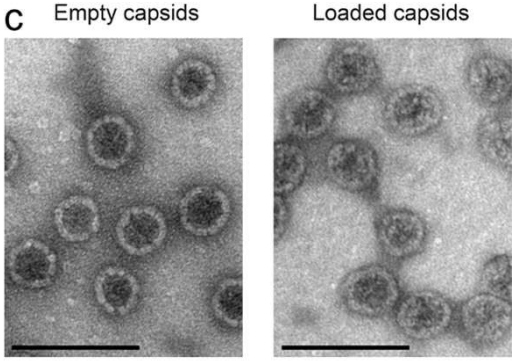




A statistically significant kinetic effect of coencapsulated CA on RuBisCO activity was not observed, however the capsid shell protected the enzyme from proteolytic damage.



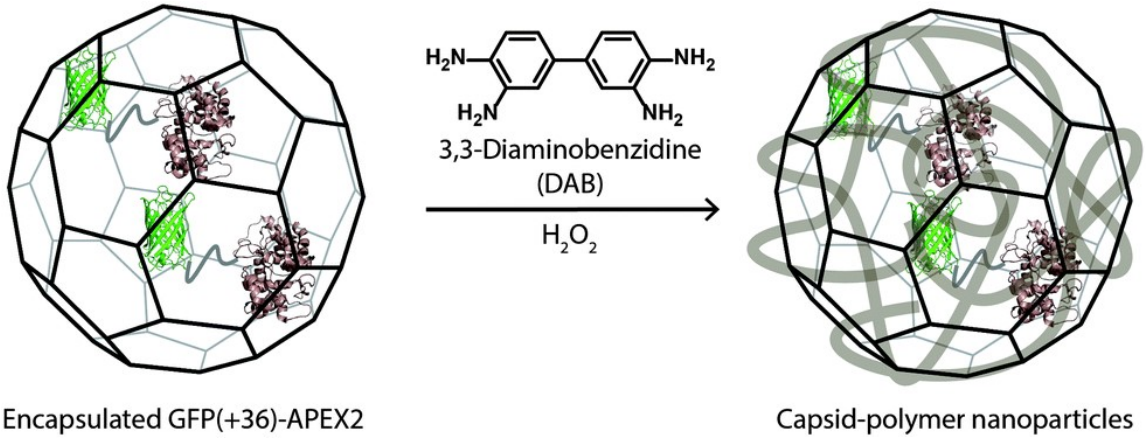




R. Frey, S. Mantri, M. Rocca, D.Hilvert *J. Am. Chem. Soc.*, **2016**, *138 (32),* pp 10072–10075

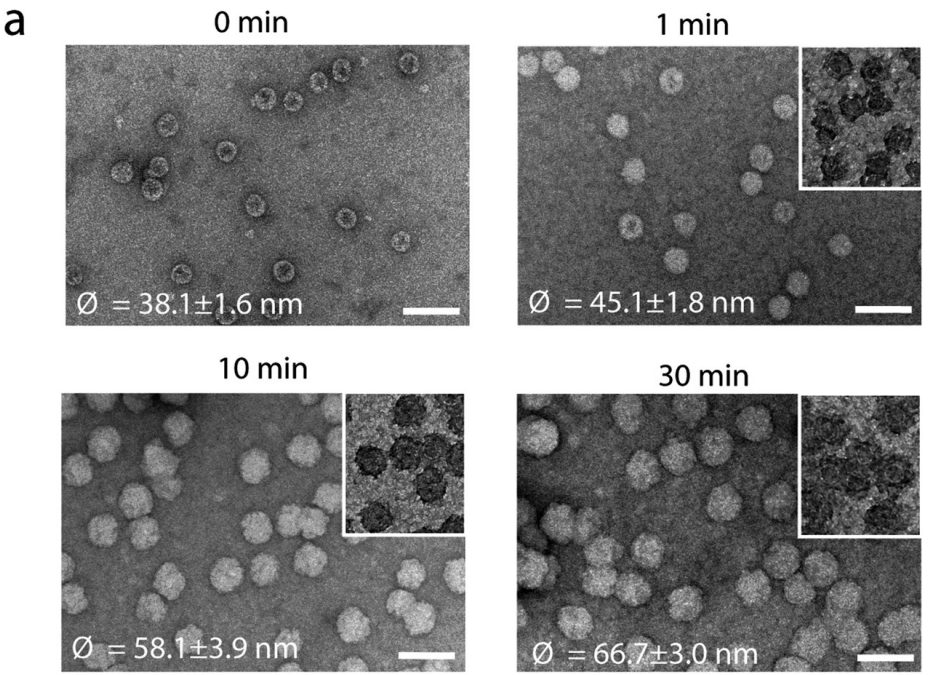
Engineering of lumazine synthase

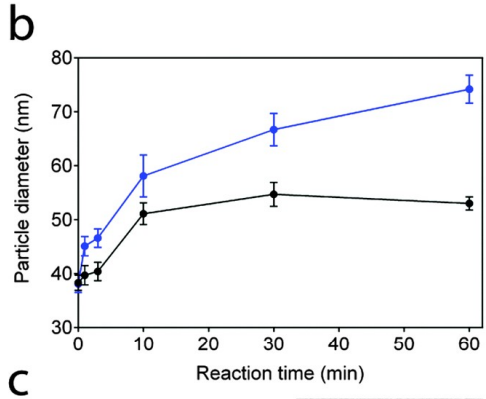
Enzyme-mediated polymerization inside engineered protein cages

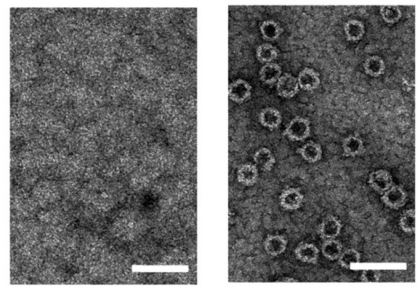


Oxidation of 3,3-diaminobenzidine (DAB) by the engineered ascorbate peroxidase APEX2 encapsulated in AaLS capsids resulted in templated formation of polyDAB–capsid nanoparticles of homogeneous size and shape.

Progress of the polymerization reaction monitored by electron microscopy



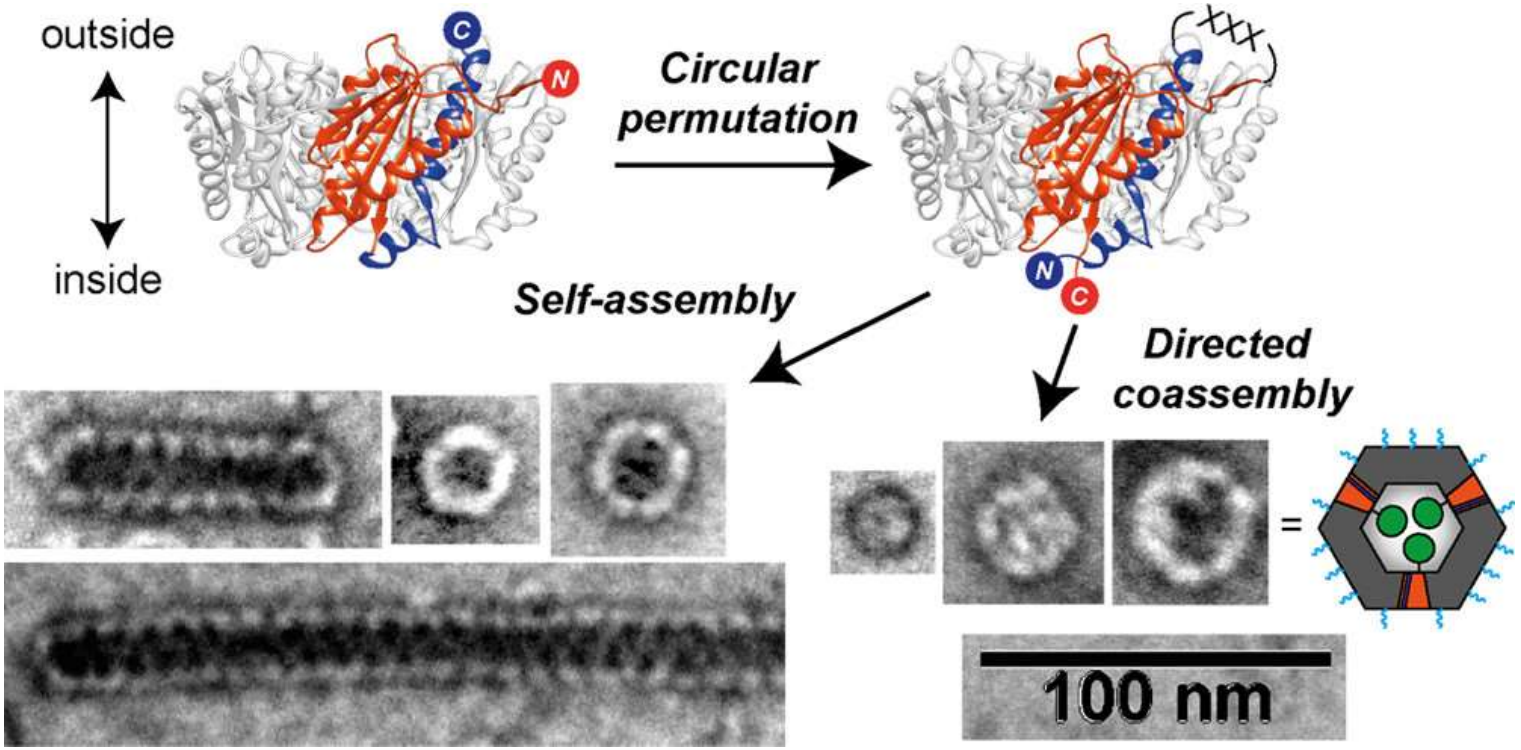




R. Frey, T. Hayashi, D. Hilvert,
Chem. Commun., **2016**, *52*, 10423-10426

Engineering of lumazine synthase

Diversification of Protein Cage Structure Using Circularly Permuted Subunits



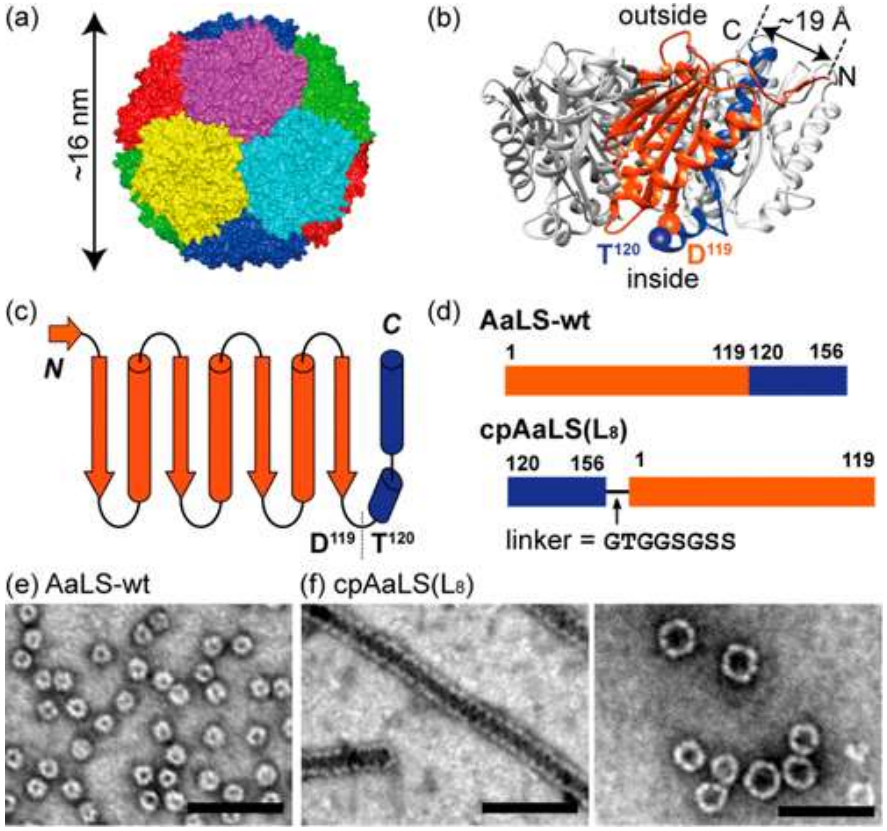
Design of circularly permuted AaLS (cpAaLS). The left structure shows a pentameric capsomer of AaLS-wt, where one monomer unit is colored: residues 1–119, orange; residues 120–156, blue. Upon circular permutation, the native termini are connected with a flexible peptide linker, and new sequence termini are generated between residues 119 and 120 in a loop region facing the interior of the assembly.

Y. Azuma, M. Herger, D. Hilvert

J. Am. Chem. Soc., 2018, 140 (2), pp 558–561

Engineering of lumazine synthase

Diversification of Protein Cage Structure Using Circularly Permuted Subunits

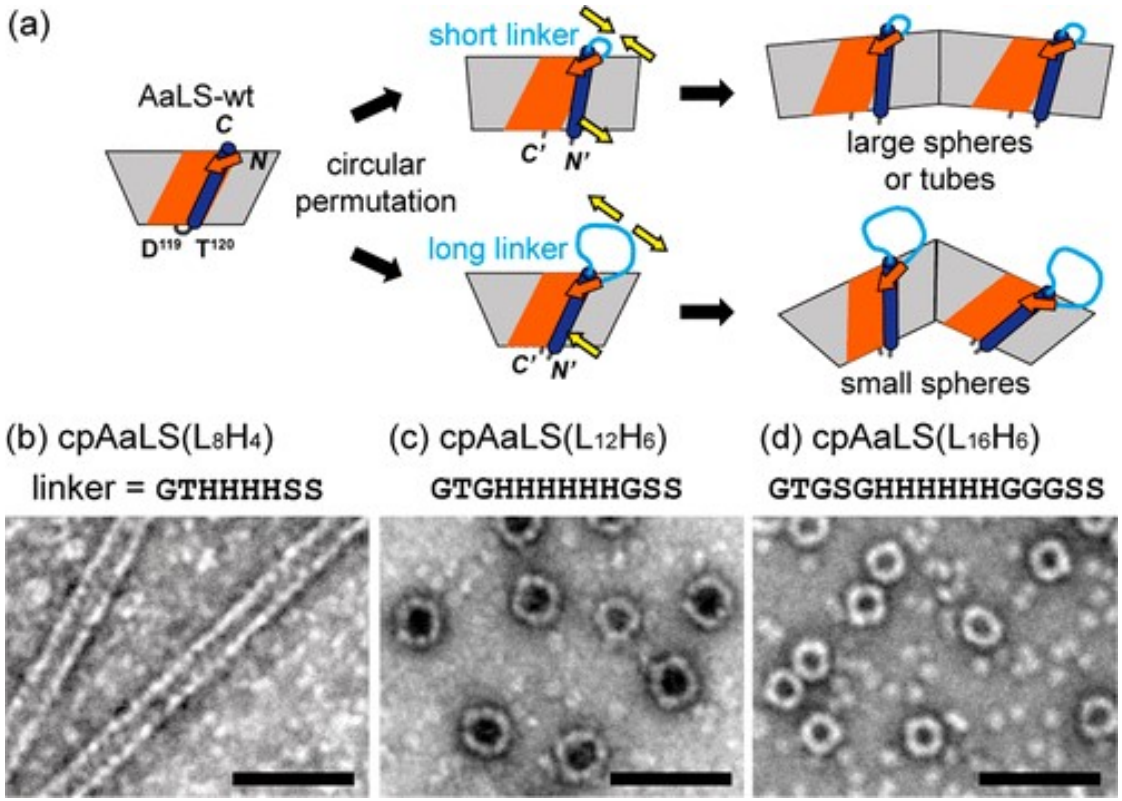


Design of circularly permuted AaLS (cpAaLS).
A pentameric capsomer of AaLS-wt, where one monomer unit is colored: residues 1–119, orange; residues 120–156, blue. Upon circular permutation, the native termini are connected with a flexible peptide linker, and new sequence termini are generated between residues 119 and 120 in a loop region facing the interior of the assembly.

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

Engineering of lumazine synthase

Diversification of Protein Cage Structure Using Circularly Permuted Subunits



Linker length controls the assembly state of circularly permuted AaLS.

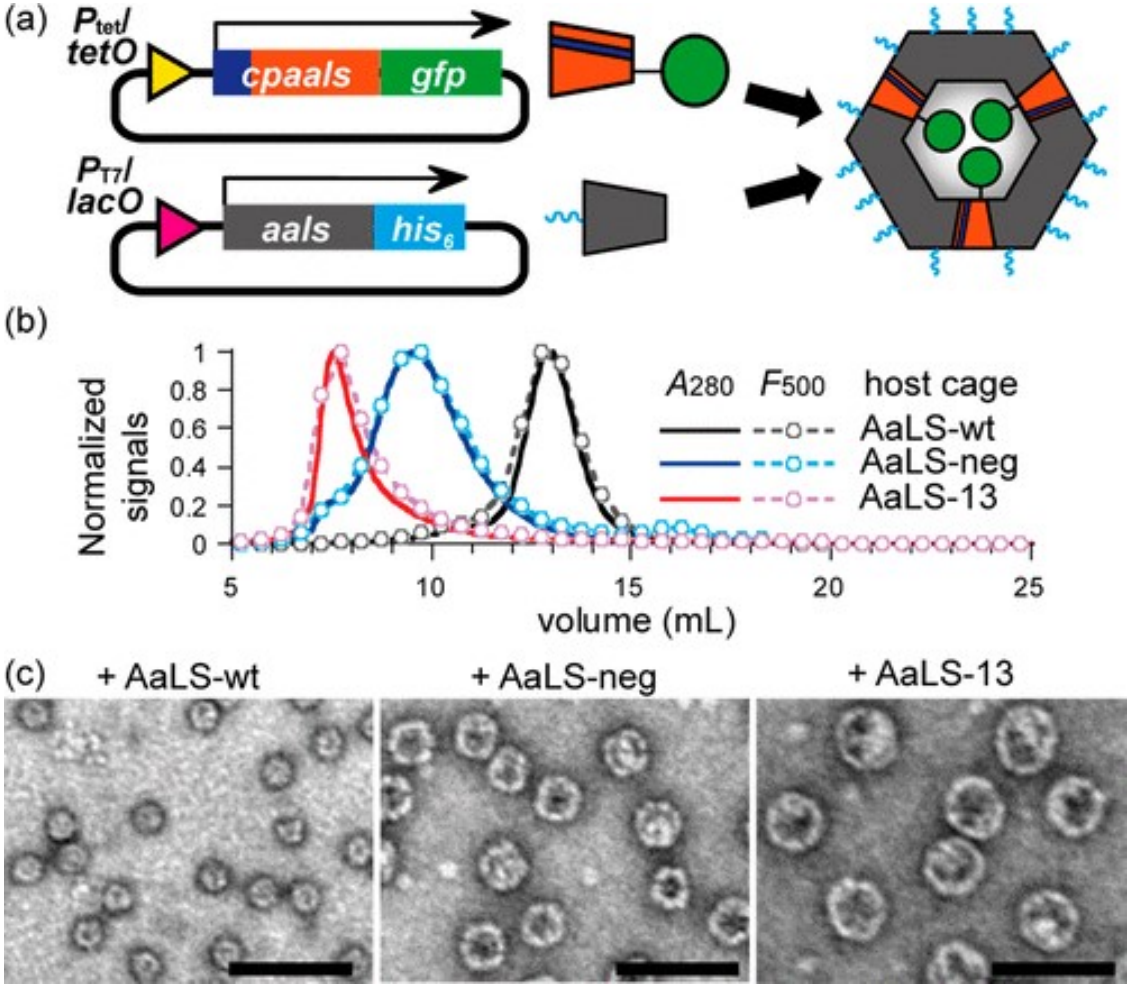
(a) Scheme illustrating the hypothetical relationship between linker length, capsomer shape, and assembly of higher-order structures.

(b–d) TEM images of the assemblies produced by different cpAaLS(LxHy) variants. The rod-shaped cpAaLS(L8H4) structures were obtained by self-assembly of isolated capsid fragments.
Scale bar = 100 nm

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

Engineering of lumazine synthase

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

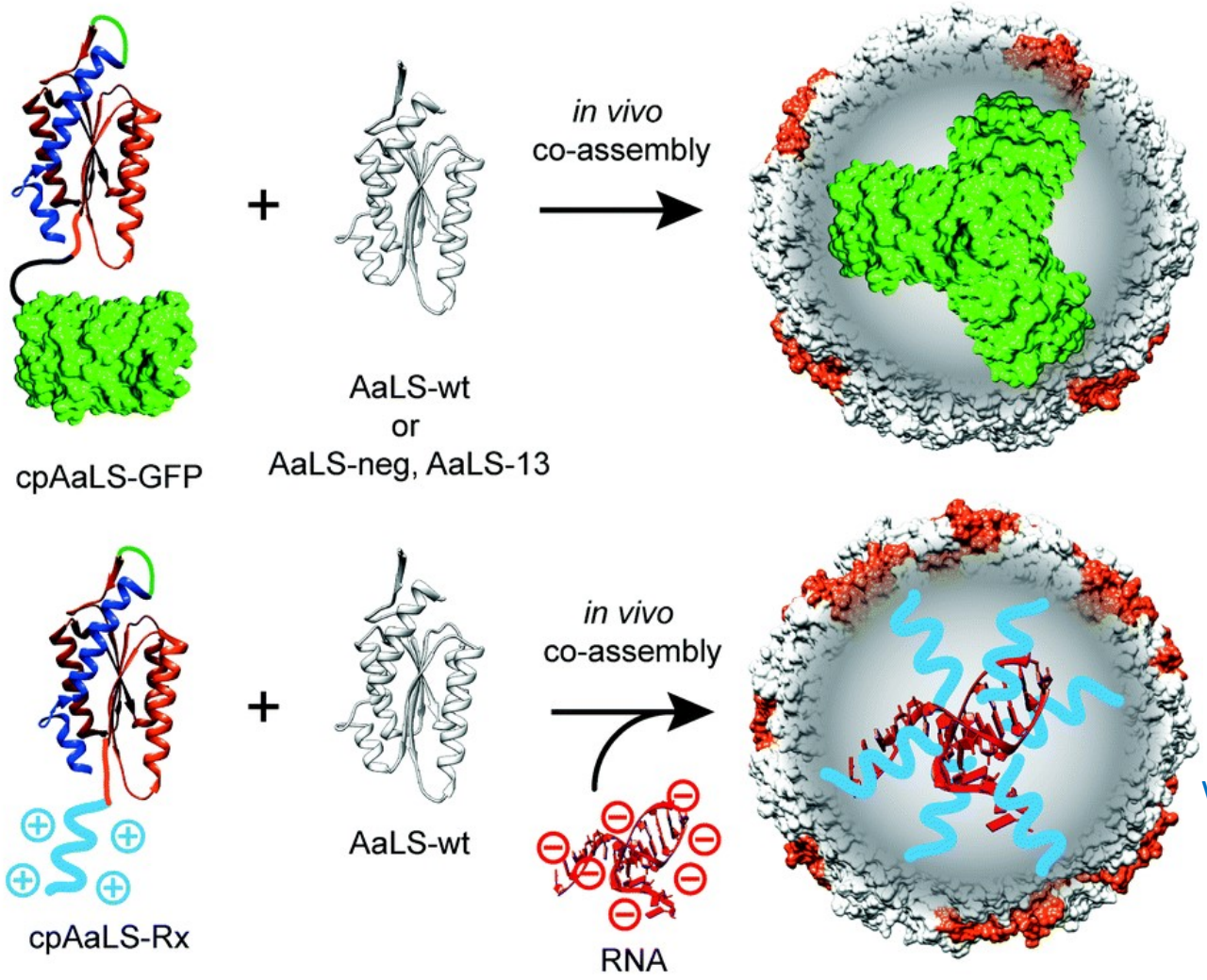
Incorporation of cpAaLS into other AaLS assemblies.

(a) Scheme illustrating the formation of patchwork assemblies in E. coli.
P_{tet}, tetracycline promoter;
tetO, tetracycline operon;
P_{T7}, T7 promoter;
lacO, lactose operon.

(b) Size-exclusion chromatogram of cpAaLS(L8)-GFP coassembled with AaLS-wt (black), AaLS-neg (blue), and AaLS-13 (red). Continuous and dashed lines respectively indicate absorbance at 280 nm (A280) and fluorescence (F500) (ex, 470 nm; em, 500 nm) for each fraction.

(c) TEM images of cpAaLS(L8)-GFP coassembled with AaLS-wt, AaLS-neg, and AaLS-13. Scale bar = 100 nm.

Diversification of Protein Cage Structure Using Circularly Permuted Subunits

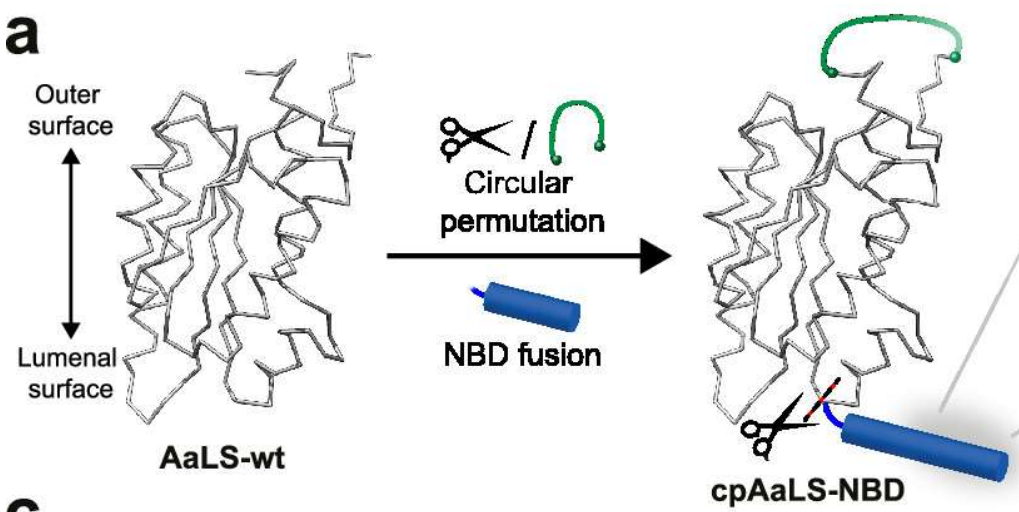


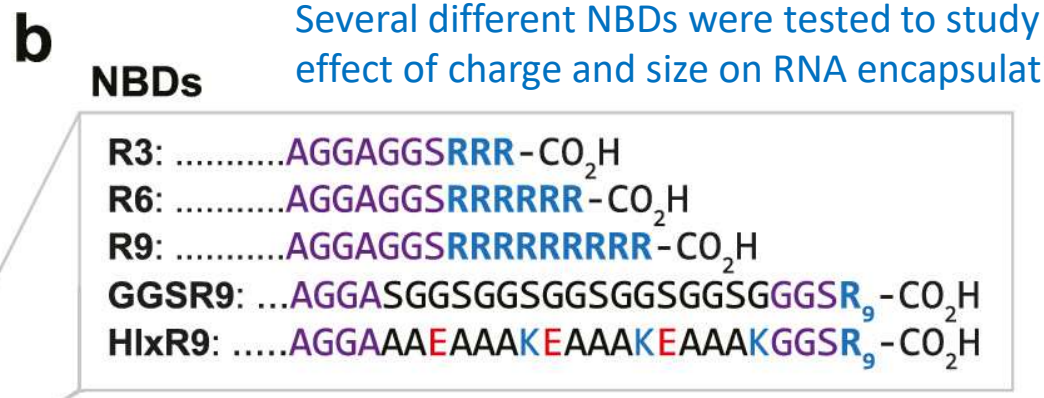
Encapsulation of a protein guest by fusion to cpAaLS and patchwork formation with other AaLS variants.

Capture of small RNAs in AaLS patchwork assemblies using cpAaLS variants with an appended oligoarginine tag (cpAaLS-RX).

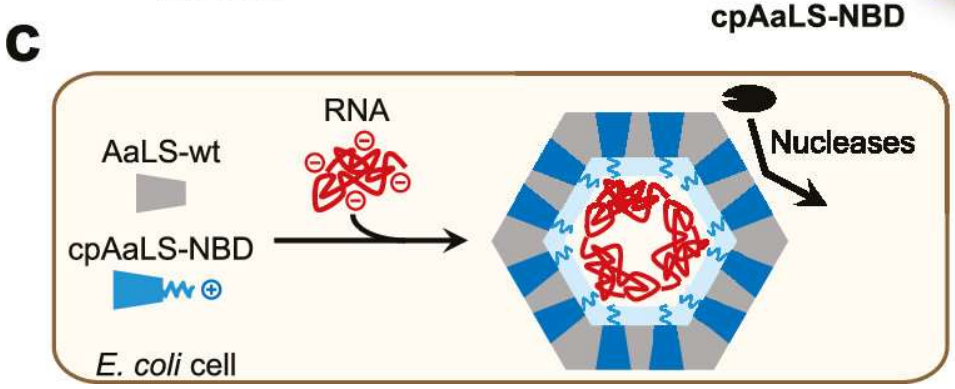
Y. Azuma, T. G. W. Edwardson, N. Terasaka, D. Hilvert *J. Am. Chem. Soc.*, **2018**, *140 (2)*, pp 566-569

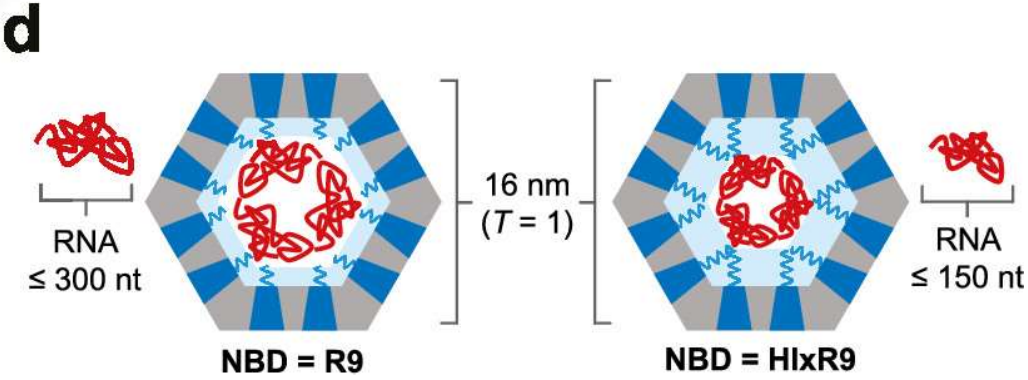
Diversification of Protein Cage Structure Using Circularly Permuted Subunits





Several different NBDs were tested to study the effect of charge and size on RNA encapsulation.





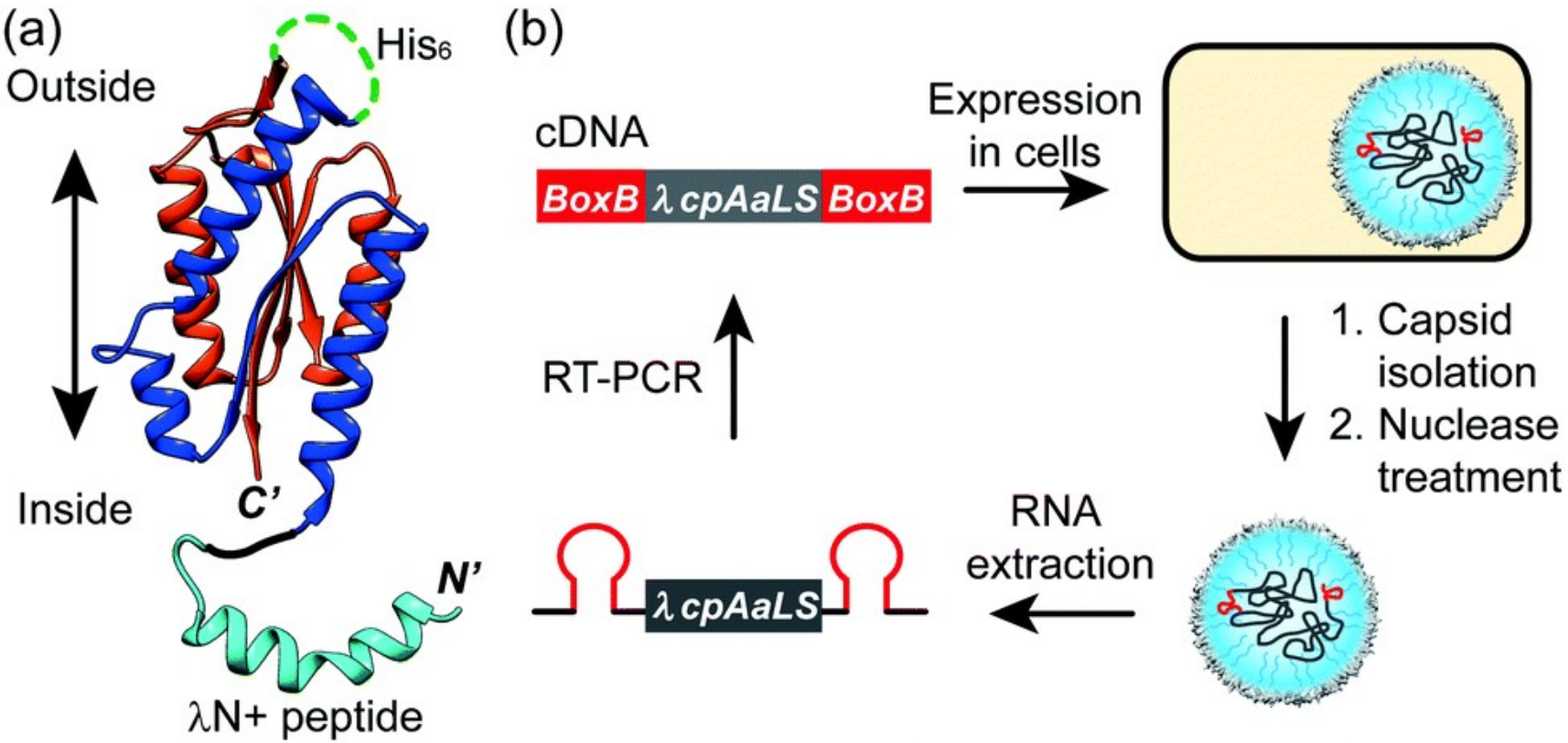
Co-expression of cpAaLS-NBD and AaLS-wt proteins provides patchwork capsids that encapsidate RNA upon assembly in vivo.

Patchwork capsids with 1:1 ratios of the two proteins have the same 16 nm external diameter. However, increasing the steric bulk of the NBDs decreases the capsid cavity volume and favors the packaging of shorter RNAs.

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

An artificial virus

Design of λcpAaLS for specific RNA recognition

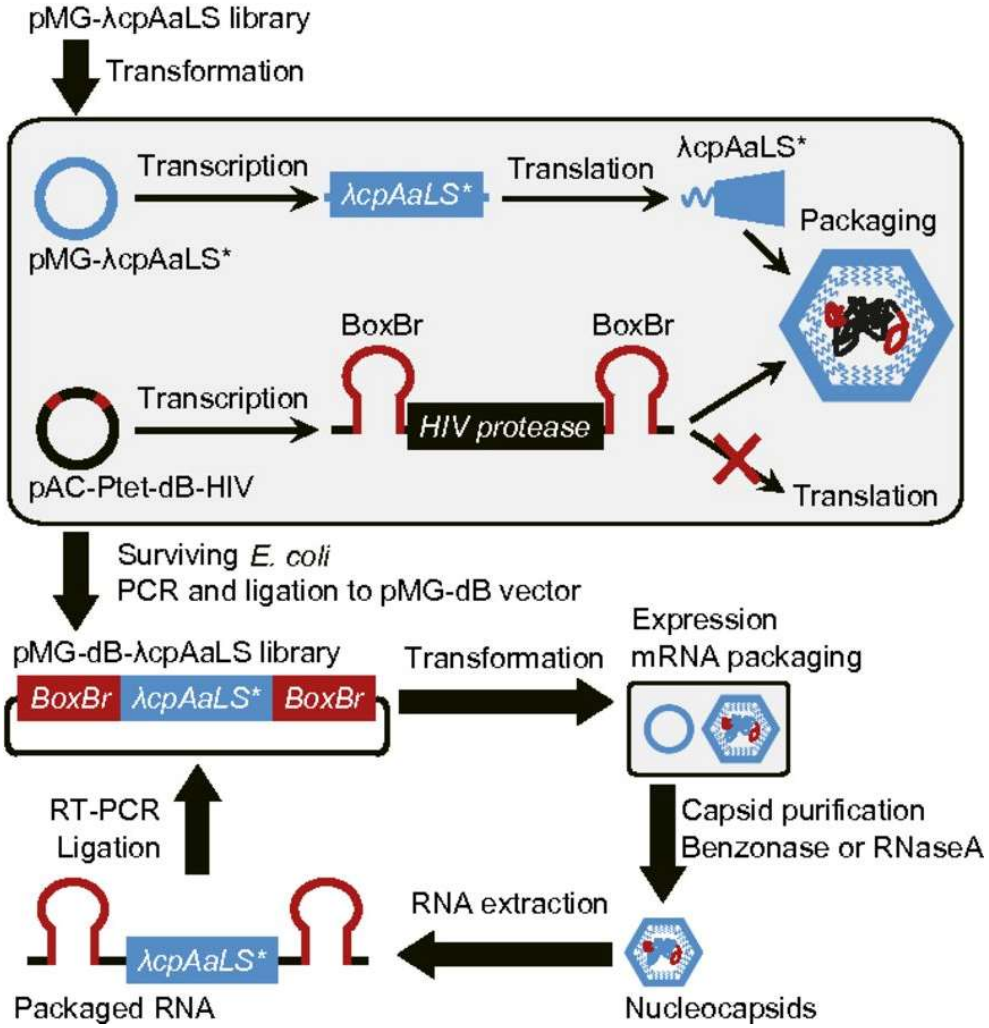


The λN+ peptide, which recognizes the BoxB RNA sequence, was appended to the N terminus and a hexahistidine tag inserted in the exterior loop for purification.

Scheme for the directed evolution of nucleocapsids encapsulating their own genome

N. Terasaka, Y. Azuma and D. Hilvert, Proc. Natl. Acad. Sci. U.S.A., **2018**, DOI: 10.1073/pnas.1800527115

Laboratory evolution of virus-like nucleocapsids



Directed evolution strategy.

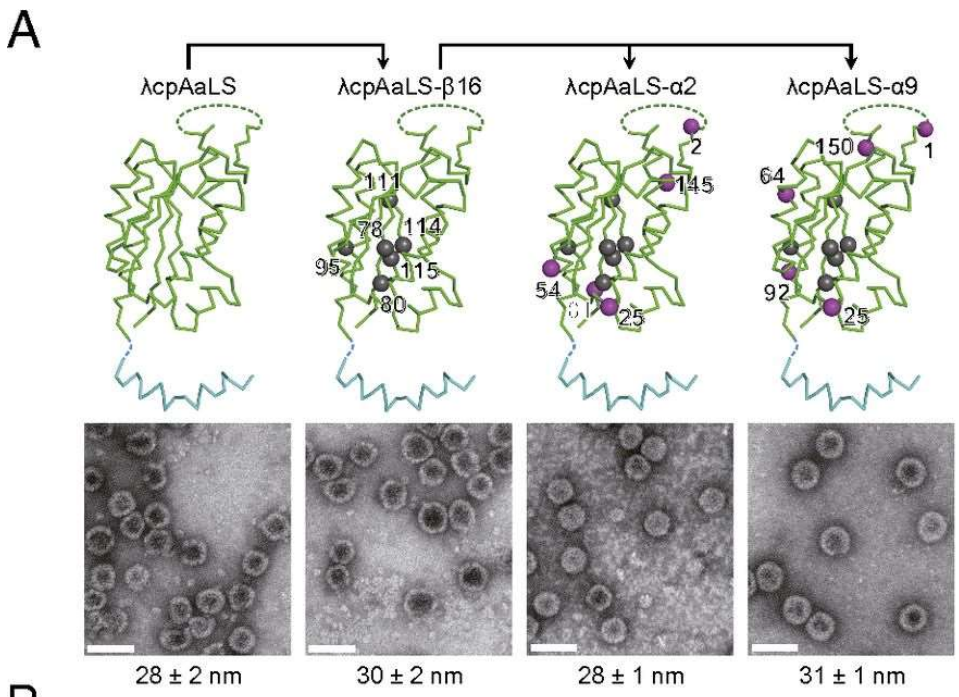
In the first evolutionary round, BoxBr tags that bind to λN+ peptides were introduced in the 5'- and 3'-UTRs of HIV protease mRNA. Capsid variants were initially selected from large gene libraries for their ability to inhibit production of the toxic protease by sequestration of its BoxBr-tagged mRNA (*Upper*).

For further optimization (*Lower*), the BoxBr tags were transferred to the UTRs of mRNA encoding the selected λcpAaLS variants, and the encoded cage proteins were produced in *E. coli*.

After capsid purification and nuclease treatment, encapsulated mRNAs were extracted and converted to cDNA by RT-PCR. The best variant, λcpAaLS-β16, was further diversified by epPCR, and the resulting library was subjected to the procedures shown.

N. Terasaka, Y. Azuma and D. Hilvert, Proc. Natl. Acad. Sci. U.S.A., **2018**, *115 (21)* , 5432-5437 DOI: 10.1073/pnas.1800527115

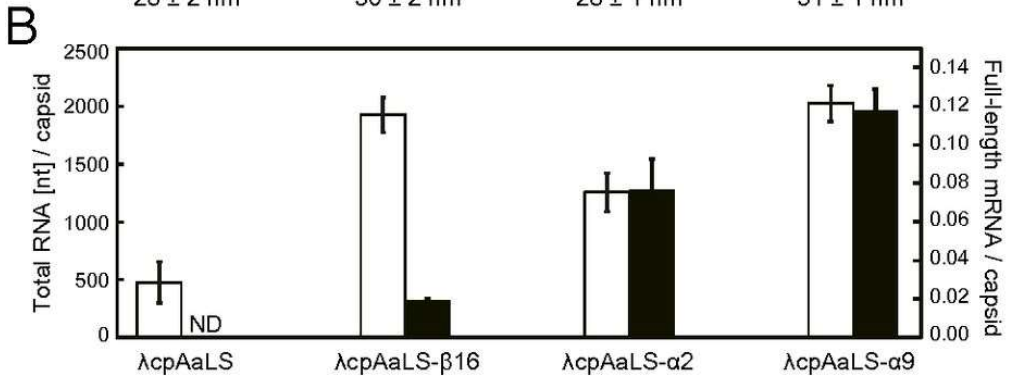
Laboratory evolution of virus-like nucleocapsids

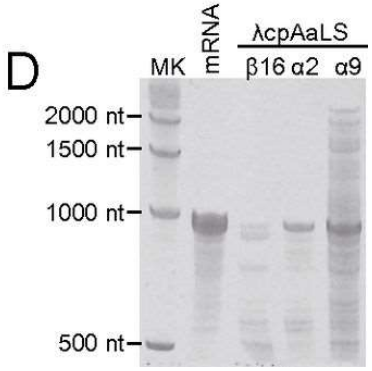


Characterization of evolved λcpAaLS variants.

(A) Backbone traces of capsid subunits (Upper) and negatively stained EM images of assembled cages (Lower) are shown for λcpAaLS, λcpAaLS-β16 (I78F, V80D, E95V, I111V, G114D, and V115I); , λcpAaLS-α2 (E2G, A25T, P54T, L81P, and E145G) , and λcpAaLS-α9 (M1I, A25T, G64D, I92N, and F150Y) .

(B) The amounts of total RNA packaged in λcpAaLS capsids expressed from mRNAs possessing two BoxBr tags (white bars) were determined from the 260/280-nm absorbance ratio, whereas full-length capsid-encoding mRNA (black bars) was determined by quantitative RT-PCR. (D) denaturing PAGE. “mRNA” indicates in vitro-transcribed full-length mRNA (863 nt).

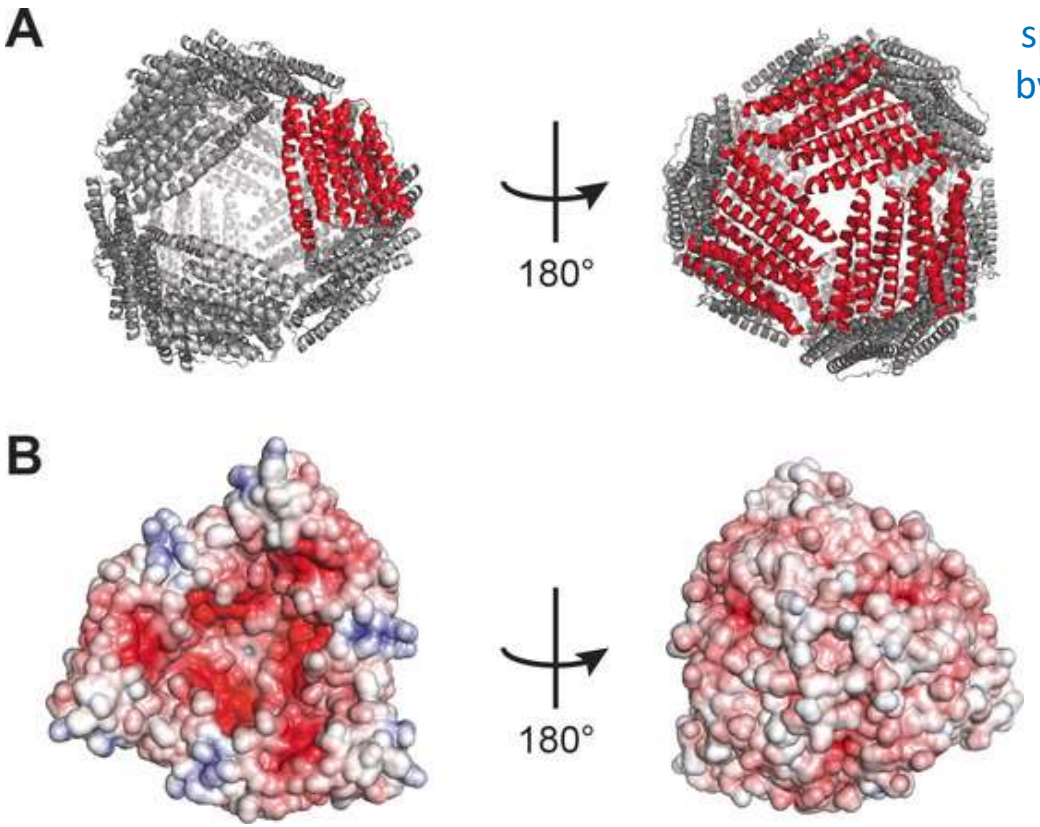




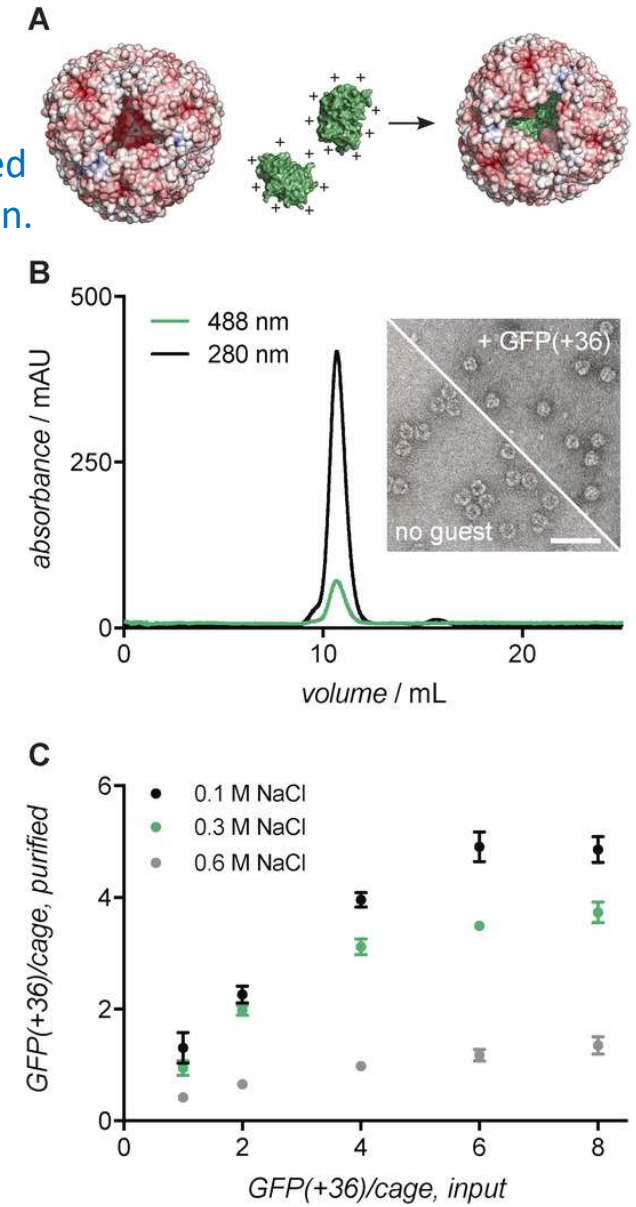
N. Terasaka, Y. Azuma and D. Hilvert, Proc. Natl. Acad. Sci. U.S.A., **2018**, *115 (21)* , 5432-5437 DOI: 10.1073/pnas.1800527115

Enzyme Encapsulation by a Ferritin Cage

Structural and electrostatic properties of
Archaeoglobus fulgidus ferritin (AfFtn, PDB ID: 1SQ3).



GFP(+36) (green) is spontaneously encapsulated by the AfFtn cage in solution.



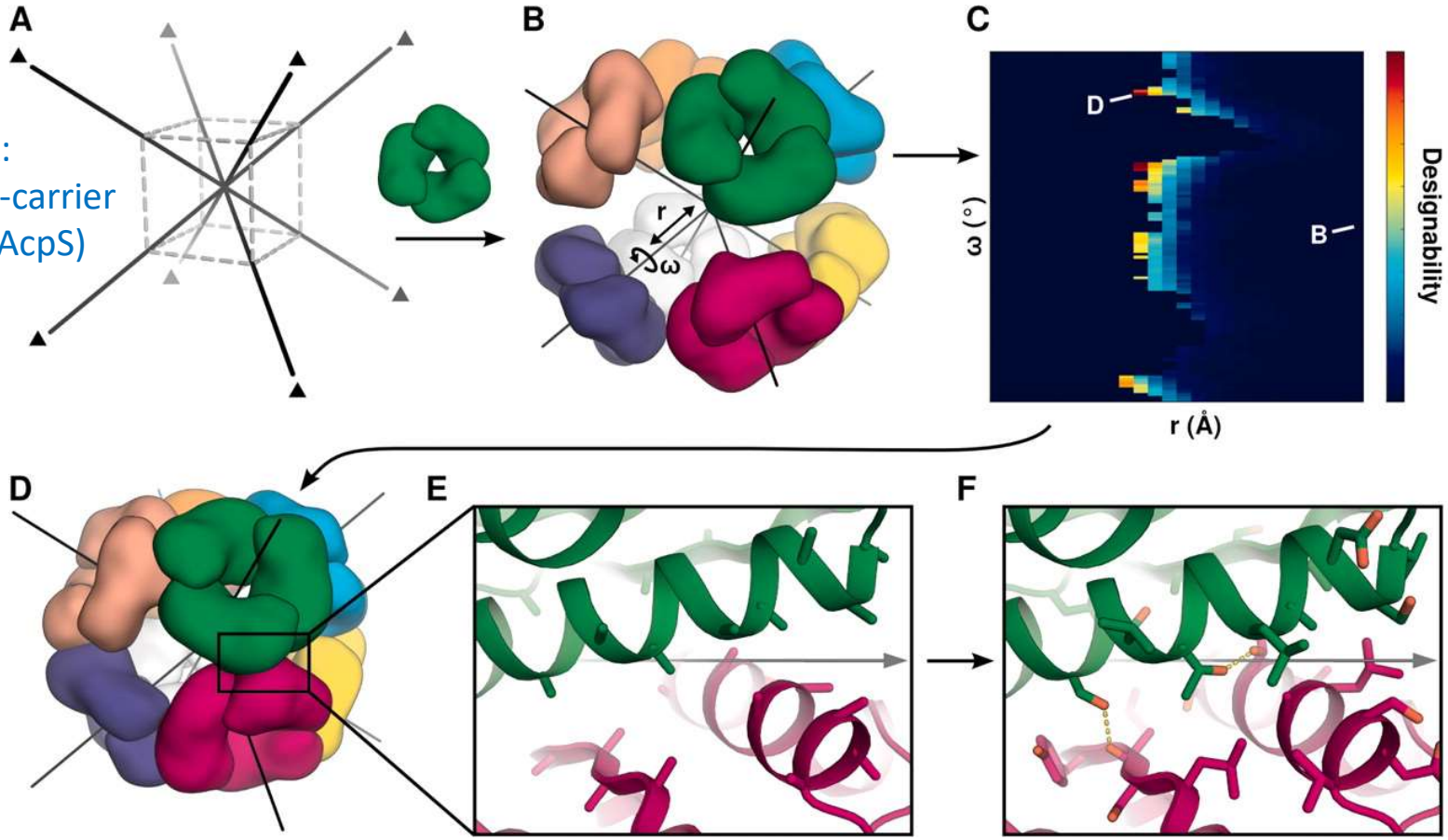
Mapping the electrostatic potential onto the solvent-accessible surface of an AfFtn hexamer33 highlights the elevated negative charge of the lumenal (left) compared to the exterior (right) surface of the cage

S. Tetter and D. Hilvert, *Angew. Chem. Int. Ed.*, **2017**, *56,* 14933-14936

Computational design of self-assembling protein nanomaterials

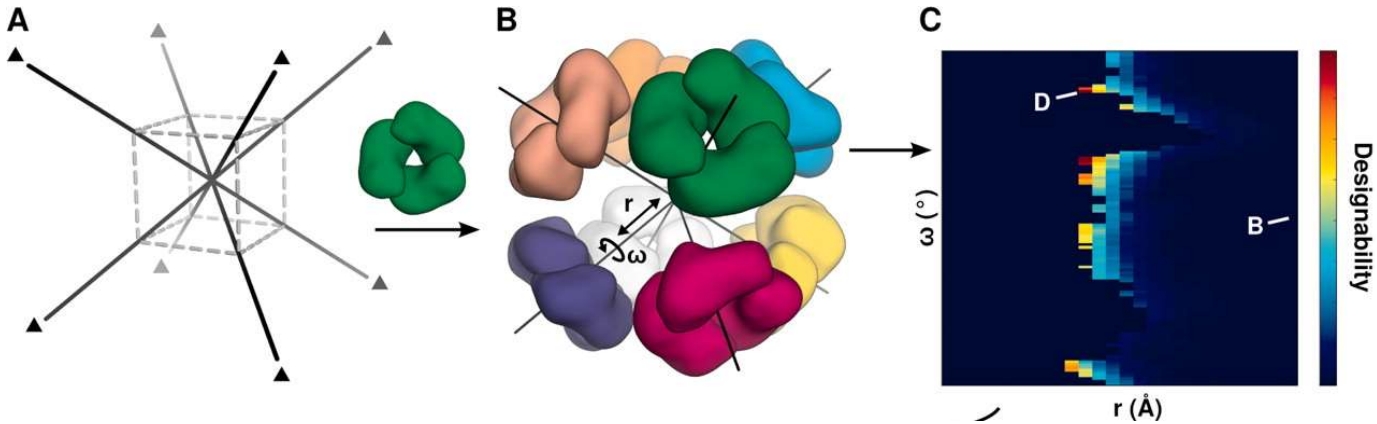
Protein building blocks are docked together symmetrically to
identify complementary packing arrangements, and low-energy protein-protein interfaces
are then designed between the building blocks in order to drive self-assembly.

The building block:
trimeric *E.coli* acyl-carrier
protein synthase (AcpS)



King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; André, I.; Gonen, T.; Yeates, T. O.; Baker, D. Science **2012**, *336*, 1171

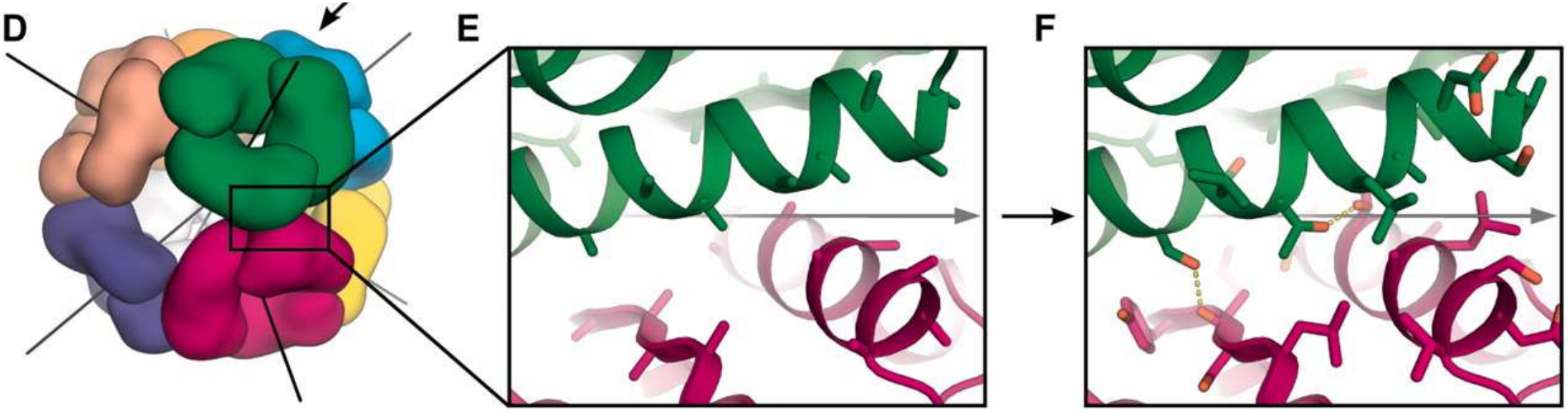
Computational design of self-assembling protein nanomaterials



(A) First, a target symmetric architecture is chosen. Octahedral point group symmetry is used in this example; the threefold rotational axes are marked here by triangles and shown as black lines throughout. The dashed cube is shown to orient the viewer. A symmetric oligomer which shares an element of symmetry with the target architecture, here a C3 symmetric trimer (green), is selected as a building block.
(B) Multiple copies of the building block are symmetrically arranged in the target architecture by aligning their shared symmetry axes. The preexisting organization of the oligomeric building block fixes several (in this case four) rigid-body degrees of freedom (DOFs). The two remaining DOFs, radial displacement (r) and axial rotation (w), are indicated.
(C) Symmetrical docking is performed by systematically varying the two DOFs (moves are applied symmetrically to all subunits) and computing the suitability of each configuration for interface design (red: more suitable; blue: less suitable). Points corresponding to the docked configurations in (B), in which the building blocks are not in contact, and (D), a highly complementary interface, are indicated.

King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; André, I.; Gonen, T.; Yeates, T. O.; Baker, D. Science **2012**, *336*, 1171

Computational design of self-assembling protein nanomaterials



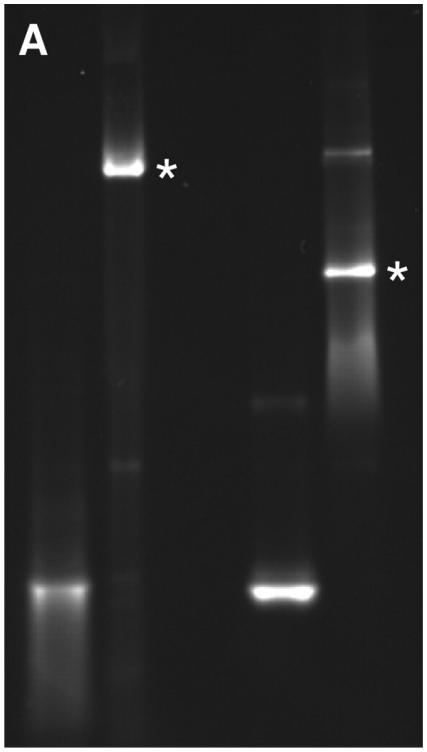
(E) Closer view of the interface in (D). The interface lies on an octahedral two-fold symmetry axis shown as a gray line. In all steps before interface design, only backbone (shown in cartoon) and carbon beta (shown in sticks) atoms are considered. (F) Sequence design calculations are used to create low-energy protein-protein interfaces that drive self-assembly of the desired material. Designed hydrogen bonds across the interface are indicated by dashed lines.

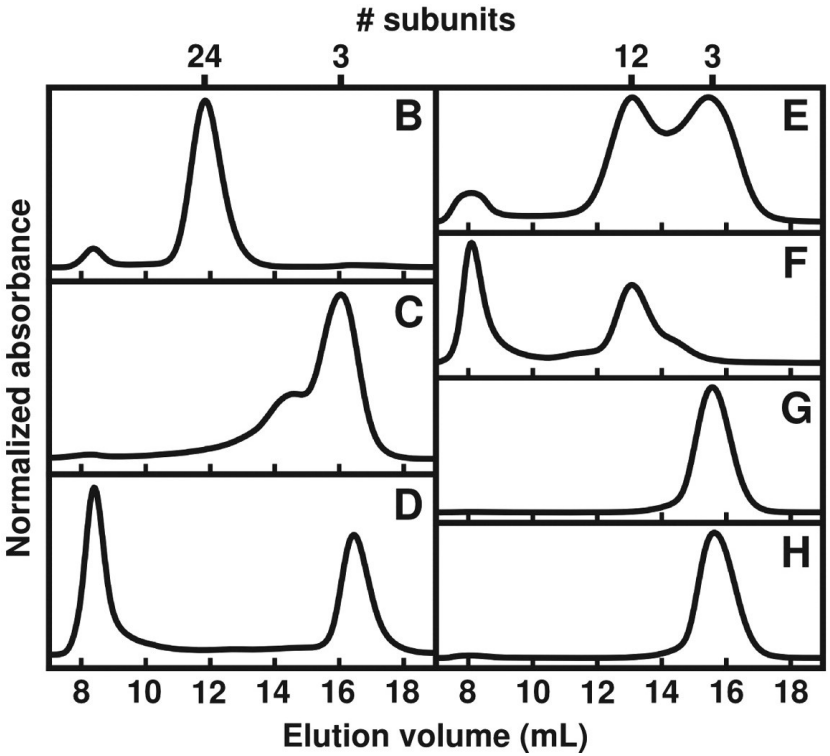
King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; André, I.; Gonen, T.; Yeates, T. O.; Baker, D. Science **2012**, *336*, 1171

Computational design of self-assembling protein nanomaterials

Experimental characterization of O3-33, T3-08, and T3-10.

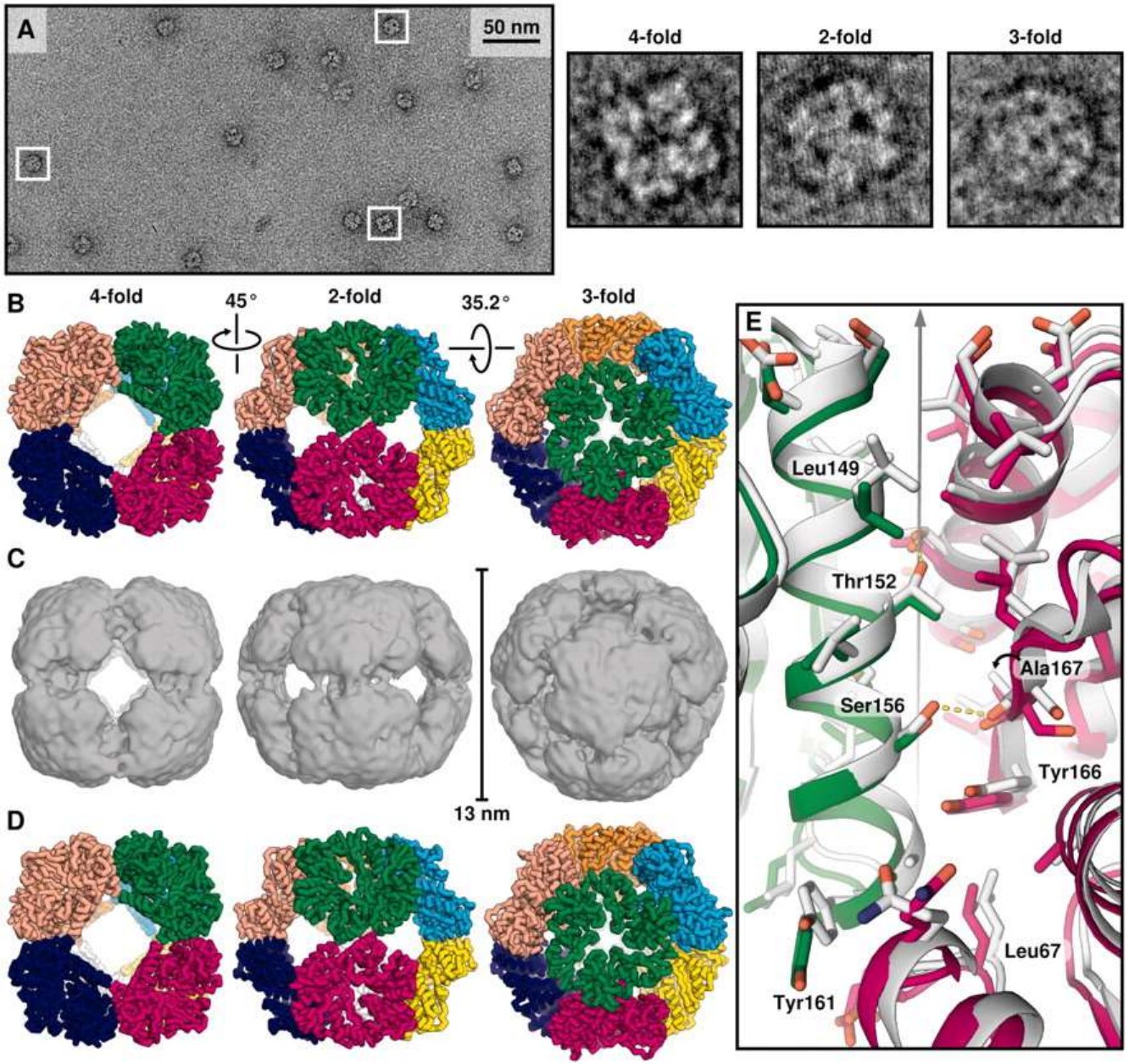
(A) Native PAGE of fluorescently labeled (from left) 3n79-wt, O3-33, 3fttwt, and T3-08 in lysates. Bands corresponding to the designed octahedral (O3-33) and tetrahedral (T3-08) assemblies are indicated with asterisks.





SEC chromatograms of nickel-purified (B) O3-33, (C) 3n79-wt, (D) O3-33(Ala167Arg), (E) T3-08, (F) T3-10, (G) 3ftt-wt, and (H) T3-08(Ala52Gln) collectively demonstrate that the assembly of the designed proteins is a result of the designed interfaces

King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; André, I.; Gonen, T.; Yeates, T. O.; Baker, D. Science **2012**, *336*, 1171



Structural characterization of O3-33.

(A) negative-stain electron micrographs of O3-33.

(B) The O3-33 design model, depicted in ribbon format. Each trimeric building block is shown in a different color.

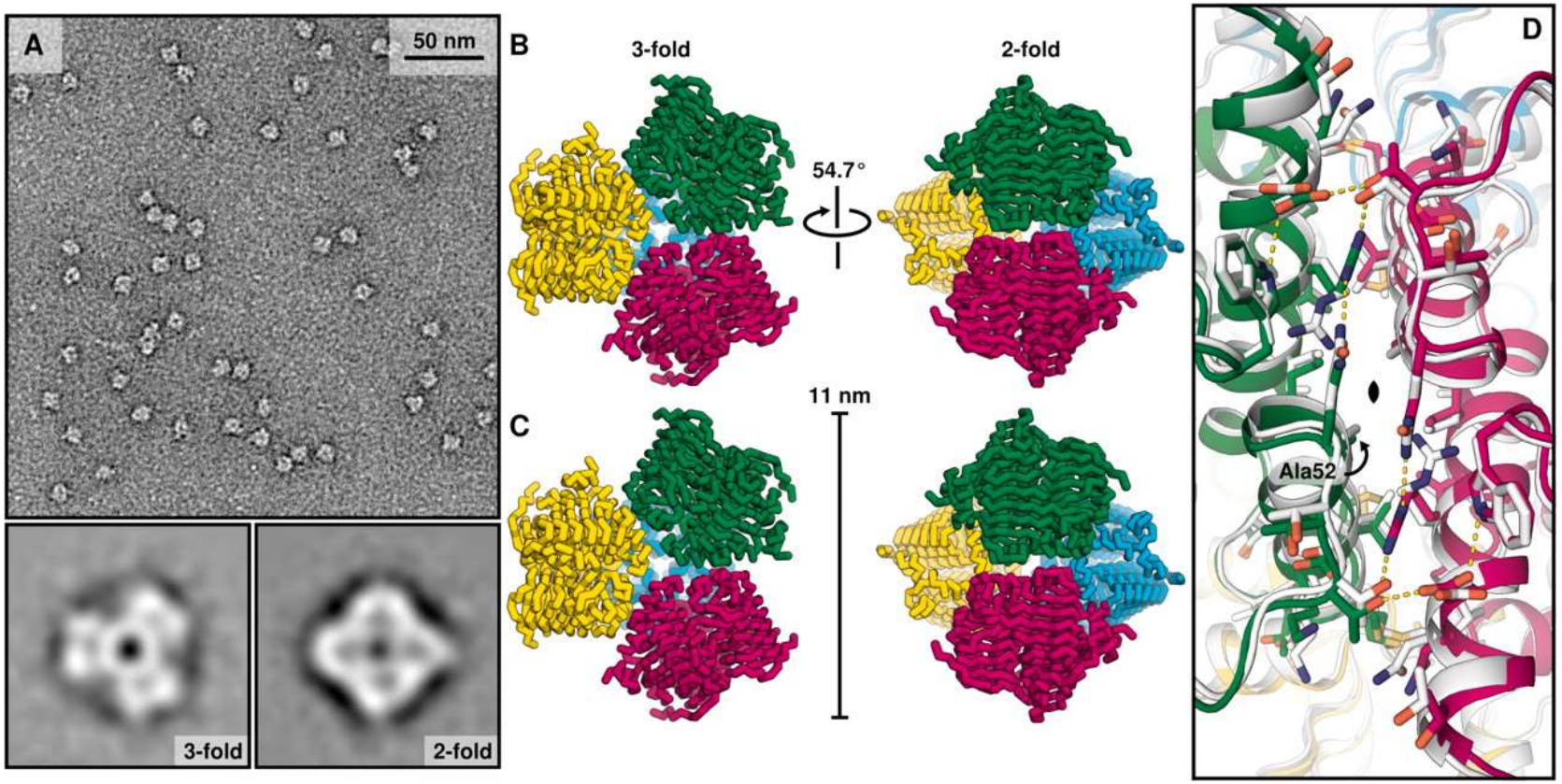
(C) The density map from a 20 Å resolution cryo-EM reconstruction of O3-33.

(D) The crystal structure of O3-33 (R32 crystal form).

(E) The designed interface in O3-33, highlighting the close agreement between the crystal structure (green and magenta) and the design model (white).

Science **2012**, *336*, 1171

Computational design of self-assembling protein nanomaterials

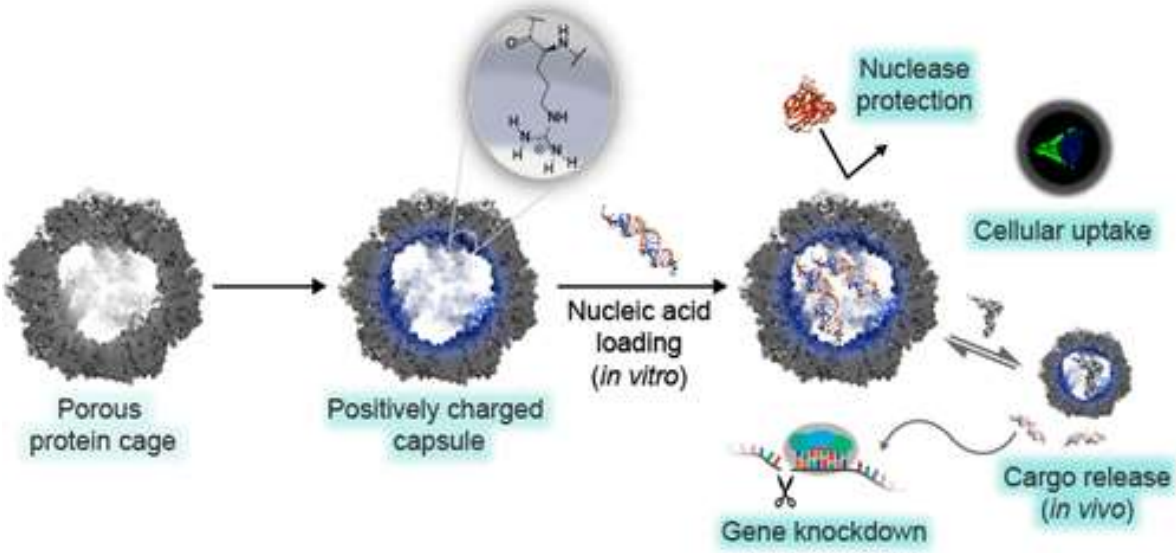


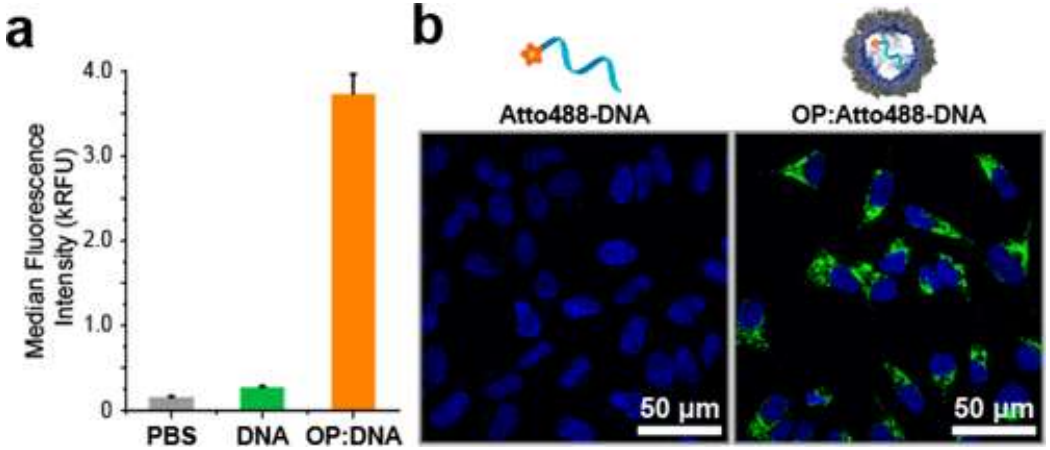
Structural characterization of T3-10.

(A) A representative negative stain electron micrograph of T3-10.
(B) Backbone representation T3-08–T3-10 design model.
(C) The T3-10 crystal structure.
(D) The designed interface in T3-10, revealing the close agreement of the crystal structure (green and magenta) to the design model (white). Alanine 52 is labeled; when mutated to glutamine in T3-08, it disrupts assembly of the designed material.

Science **2012**, *336*, 1171

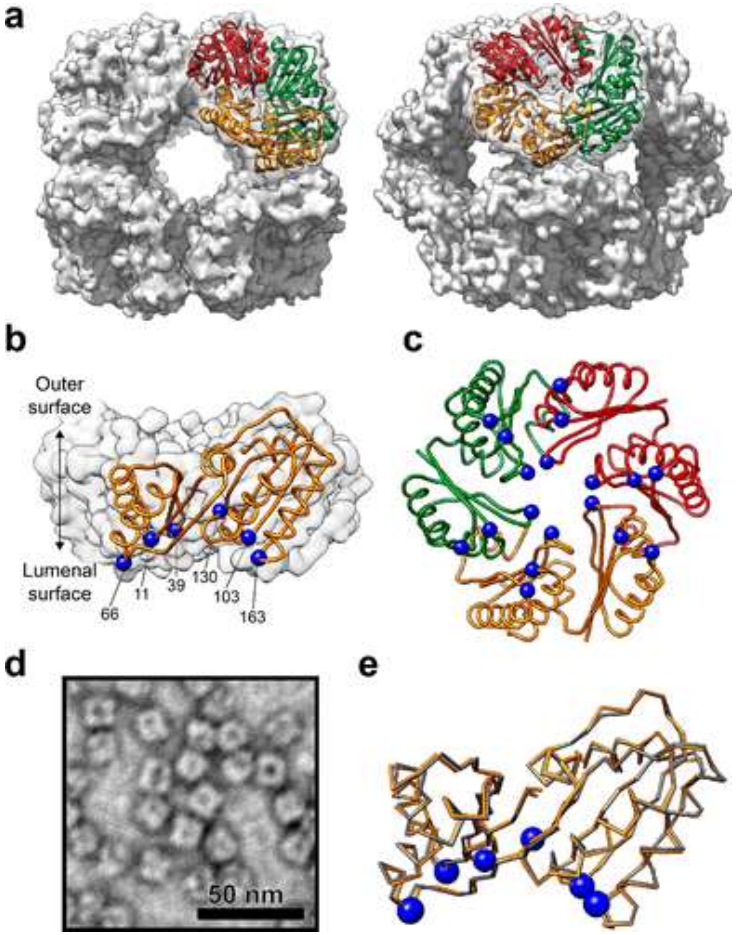
Rational Engineering of a Designed Protein Cage for siRNA Delivery





HeLa cells treated with Atto488-DNA or OP:Atto488-DNA.
Blue, Nuclei (Hoechst stained); Green, Atto488.

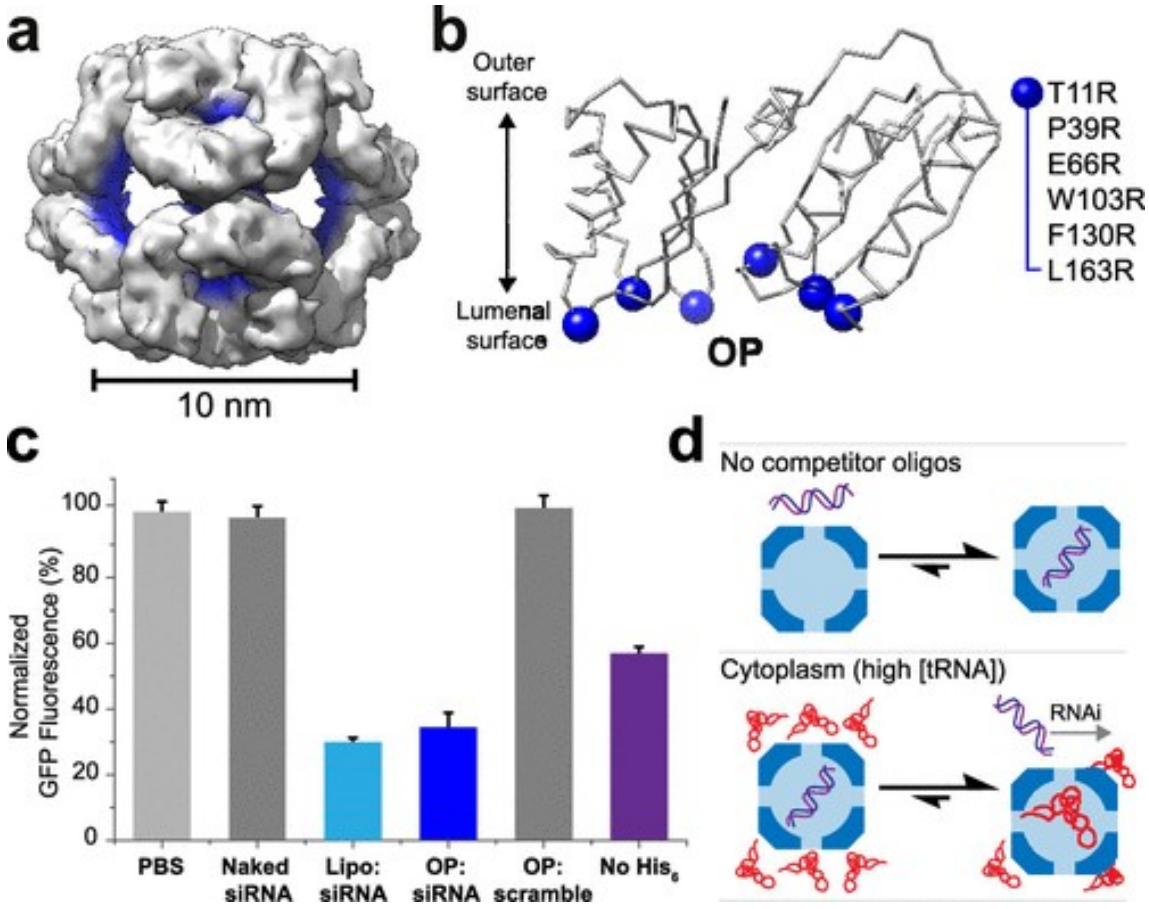
Surface models of the OP cage –
one trimer shown as a colored ribbon.



Superimposed monomers of O3-33 (gray) and OP (orange)

T. G. W. Edwardson, T. Mori, D. Hilvert *J. Am. Chem. Soc.* **2018**, *140*, 10439-10442

Rational Engineering of a Designed Protein Cage for siRNA Delivery



(a) Surface structure of the OP cage assembly, with blue highlighting the positively charged lumen.

(b) OP monomer showing the six arginine mutations introduced with respect to the original O3-33 scaffold.

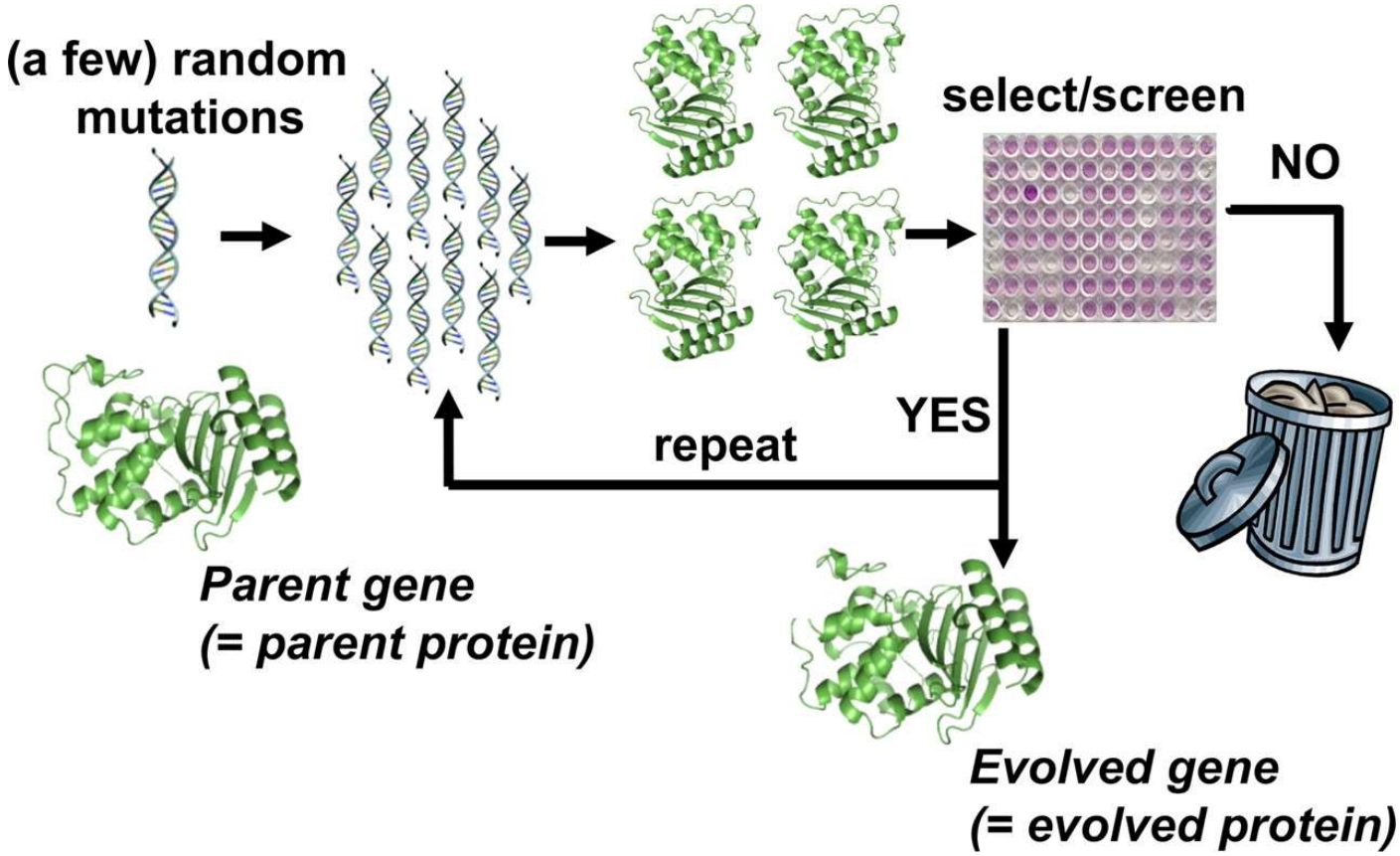
(c) OP cages carrying GFP-targeting siRNA (OP:siRNA) knock down protein expression to the same degree as the commercial reagent lipofectamine (Lipo:siRNA). PBS buffer, naked siRNA, and OP carrying a scrambled sequence siRNA (OP:scramble) have no effect on GFP signal. OP cages lacking His6 tags (No His6) are less active.

(d) While stable in the absence of other oligonucleotides, high concentrations of tRNA can displace the siRNA cargo from OP cages, freeing them to induce RNA interference.

T. G. W. Edwardson, T. Mori, D. Hilvert *J. Am. Chem. Soc.* **2018**, *140*, 10439-10442

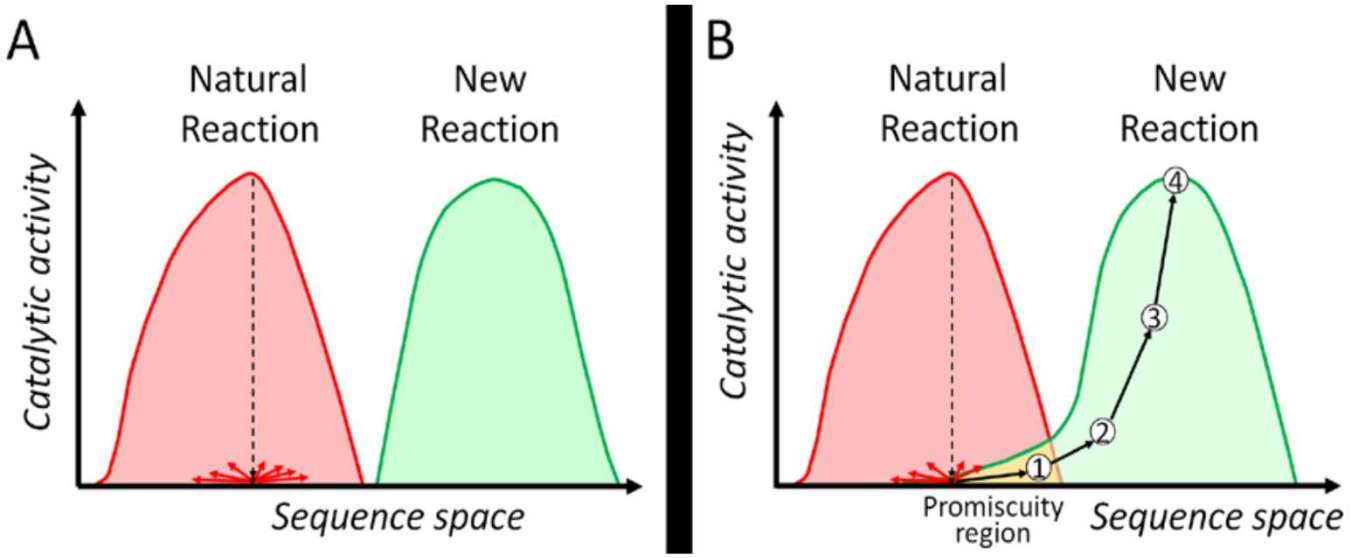
PROTEINS

Enzyme engineering



J. D. Bloom, F. H. Arnold *Proc. Natl. Acad. Sci. USA.*, **2009**, *106 (Suppl. 1),* 9995-10000

Enzyme engineering – engineering of promiscuous enzymes

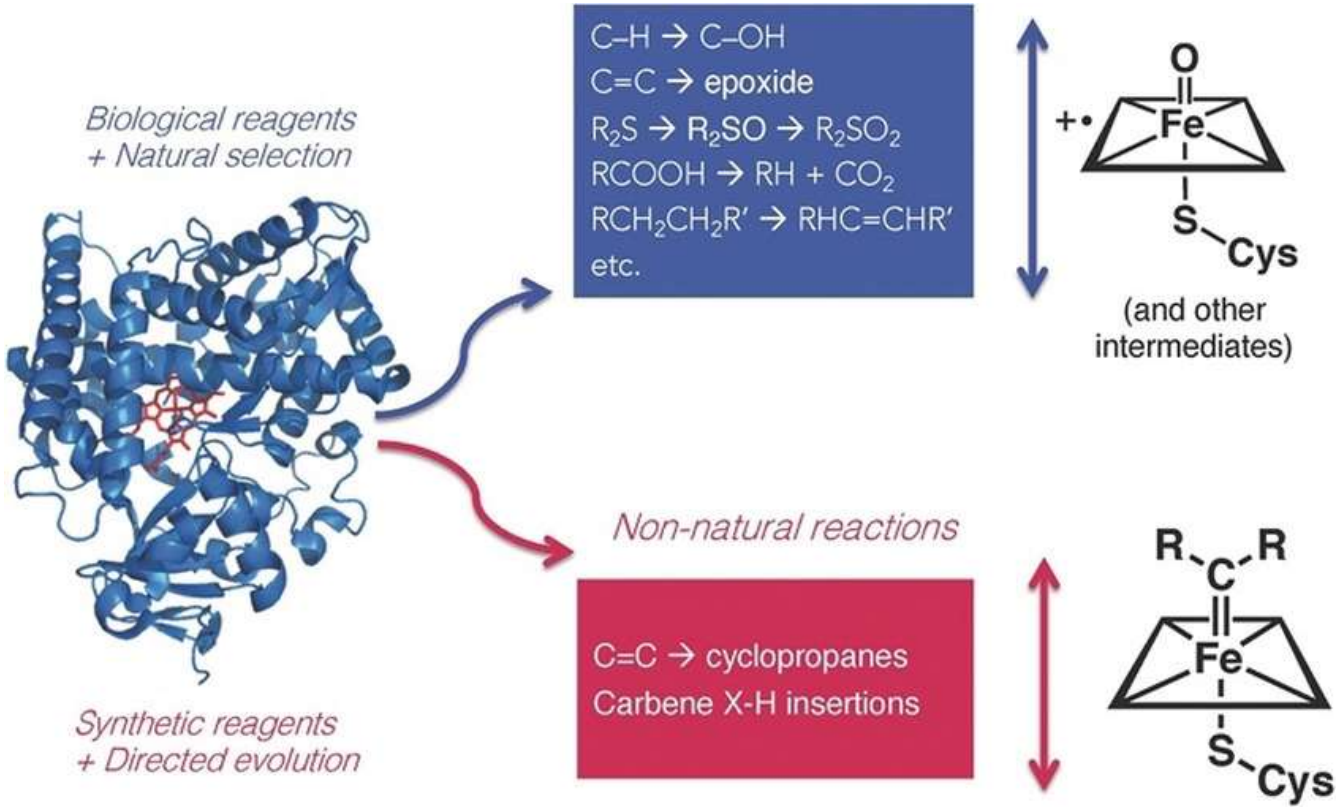


A: A starting point with no activity for the intended reaction is useless since no sequence variations (red arrows) create the new reactivity.

B: A promiscuous enzyme with at least low activity for the intended reaction is a suitable staring point. Some combinations of random mutations may improve the new reactivity (black arrow). The first variant (1) serves as a starting state for sequential rounds of variation and screening →(2)→(3)→(4) for improved variants. Only a small number of cycles and are typically needed to boost up the new reactivity

F. Arnold *Nobel lecture 2018; Angew.Chem.Int.Ed.* ***2019****, 58,* 14420–14426

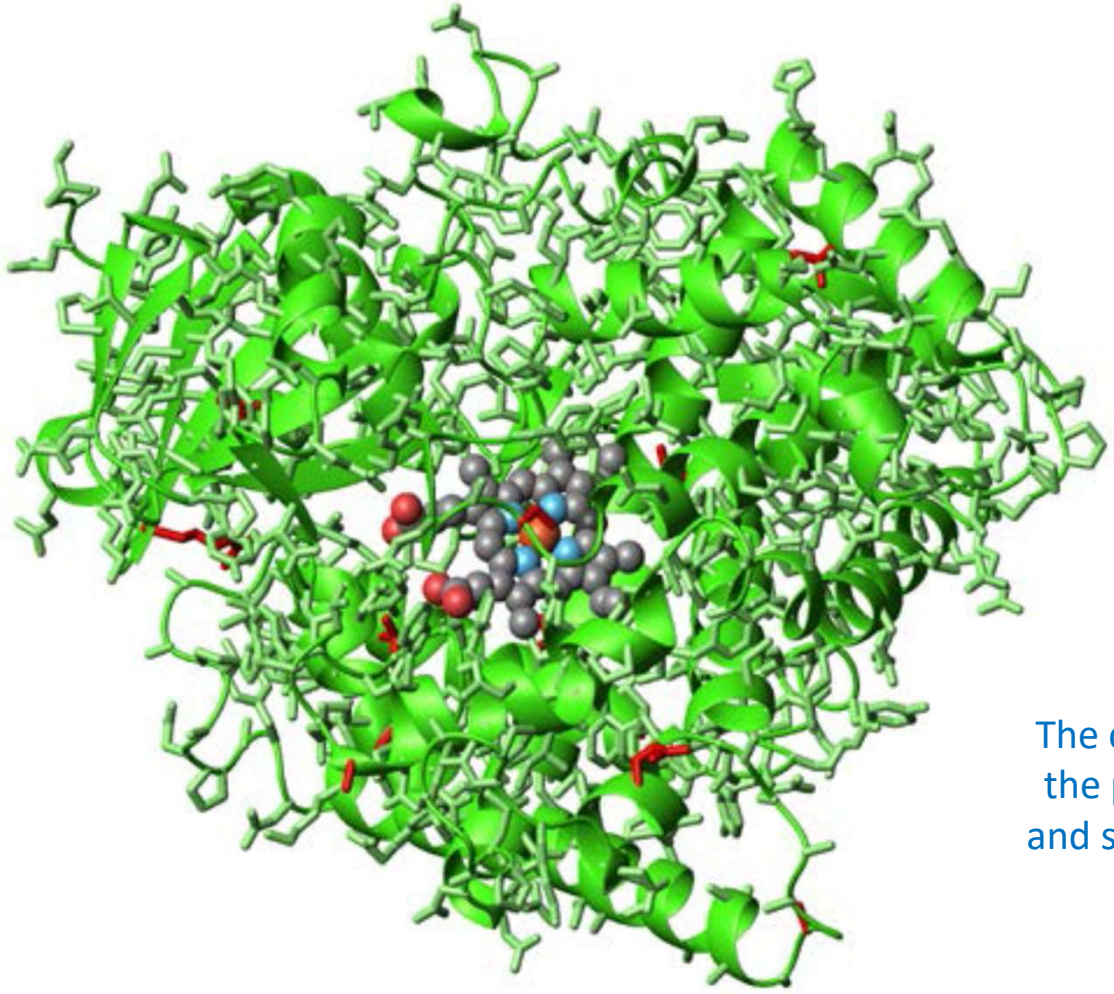
Expanding the scope of P450 chemistry.



The cytochrome P450 family, whose members were presumably created by gene duplication and natural selection of promiscuous functions, comprises enzymes that use reactive oxygen intermediates to catalyze a wide range of reactions. We reasoned that we could expand the scope of P450 chemistry by using synthetic carbene and nitrene precursors to drive formation of new reactive intermediates. Directed evolution would be used to mold the enzyme, controlling and enhancing new-to-nature activities..

F. Arnold *Angew.Chem.Int.Ed.* ***2019****, 58,* 14420–14426

Expanding the scope of P450 chemistry.



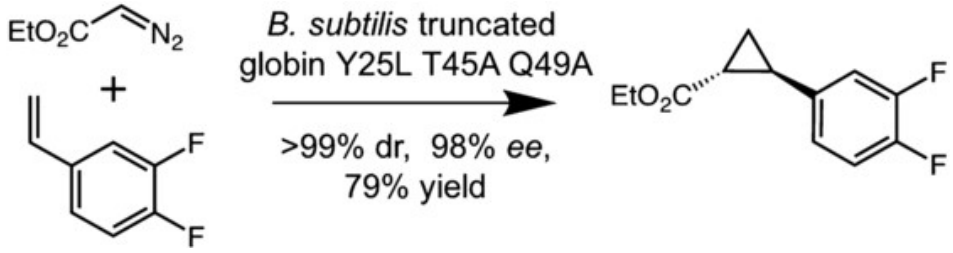
An evolved biocatalyst for cyclopropanation.

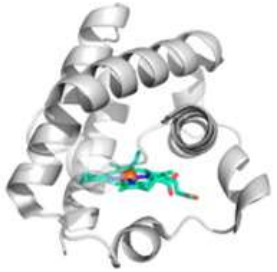
The cytochrome P411 variant of cytochrome P450 with the protein backbone shown as ribbon representation and side-chains as sticks. Side-chains that were mutated in engineered variants are shown in red.

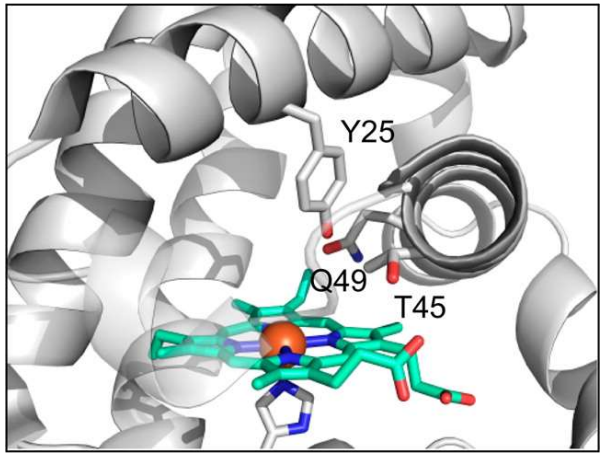
F. Arnold *Angew.Chem.Int.Ed.* ***2019****, 58,* 14420–14426

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





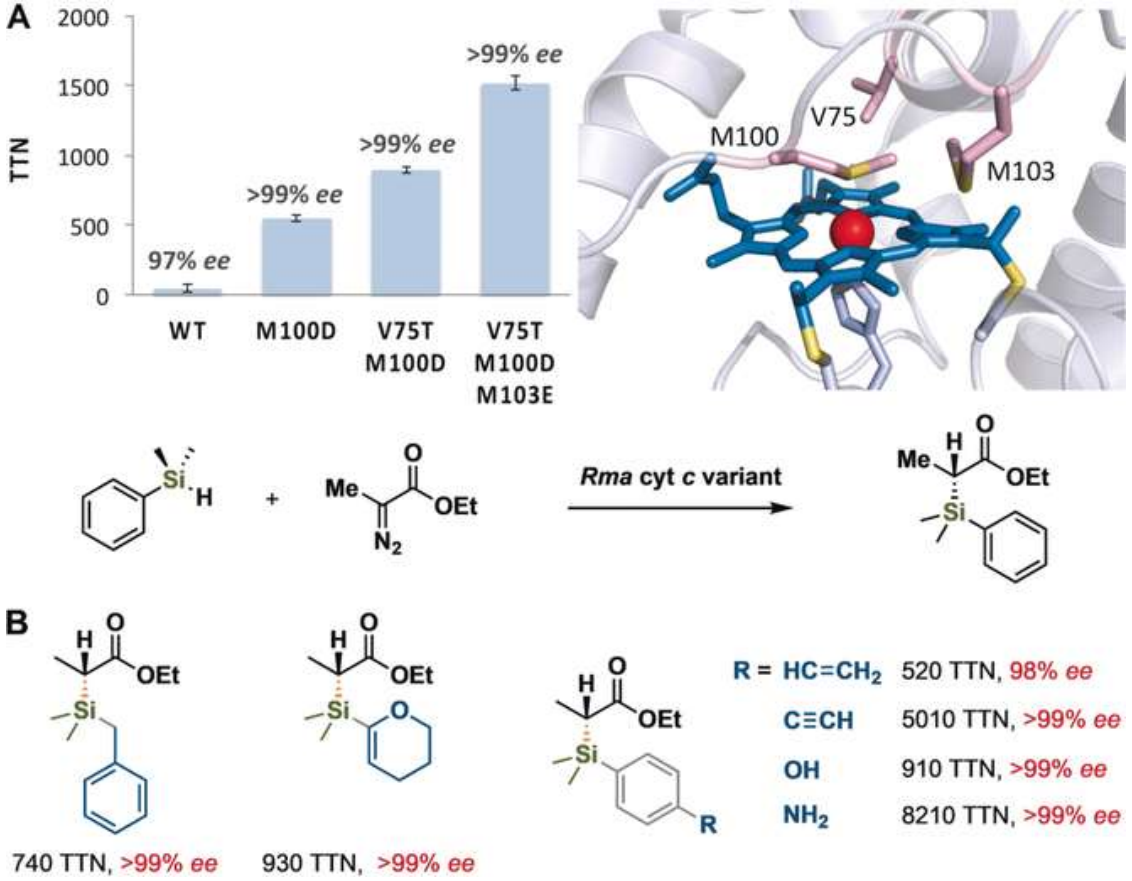


Positions of the Y25, T45, and Q49 residues near the heme iron in the B. subtilis wild-type protein (PDB ID: 1UX8)

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

F. Arnold *Angew. Chem. Int. Ed.*, **2018**, *57,* 4143-4148

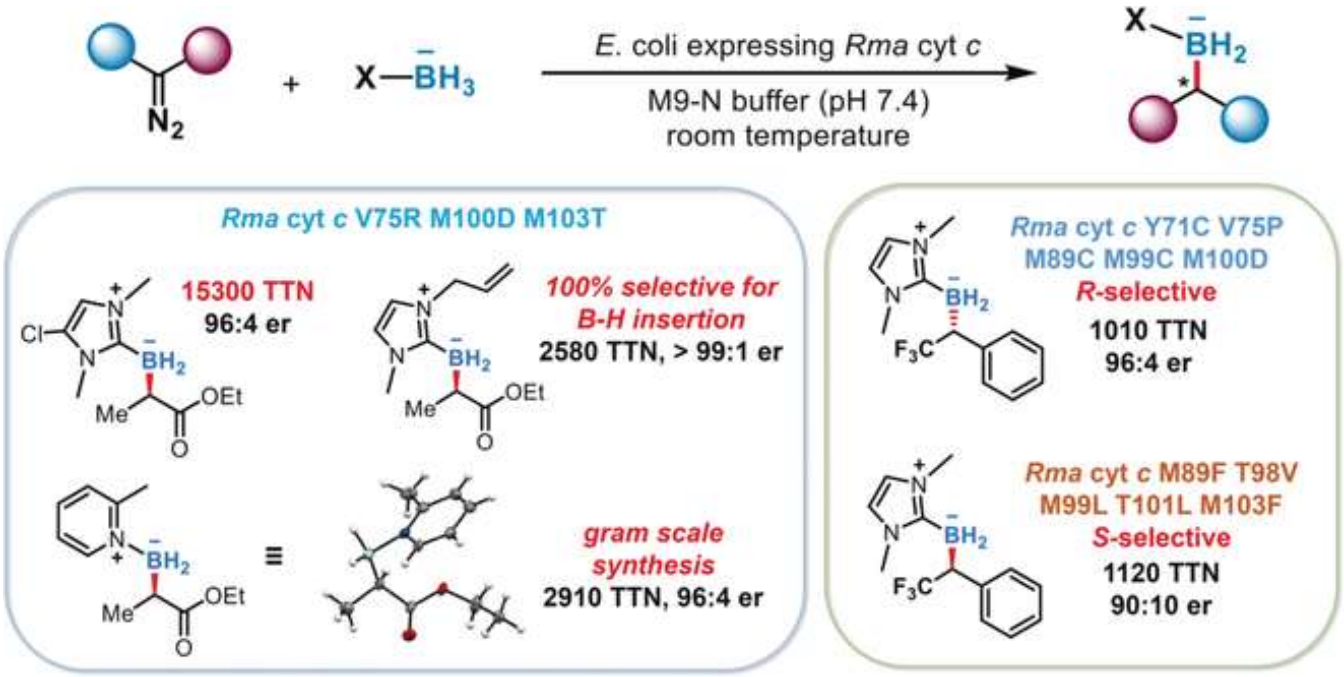
Directed evolution – bringing new chemistry to life



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

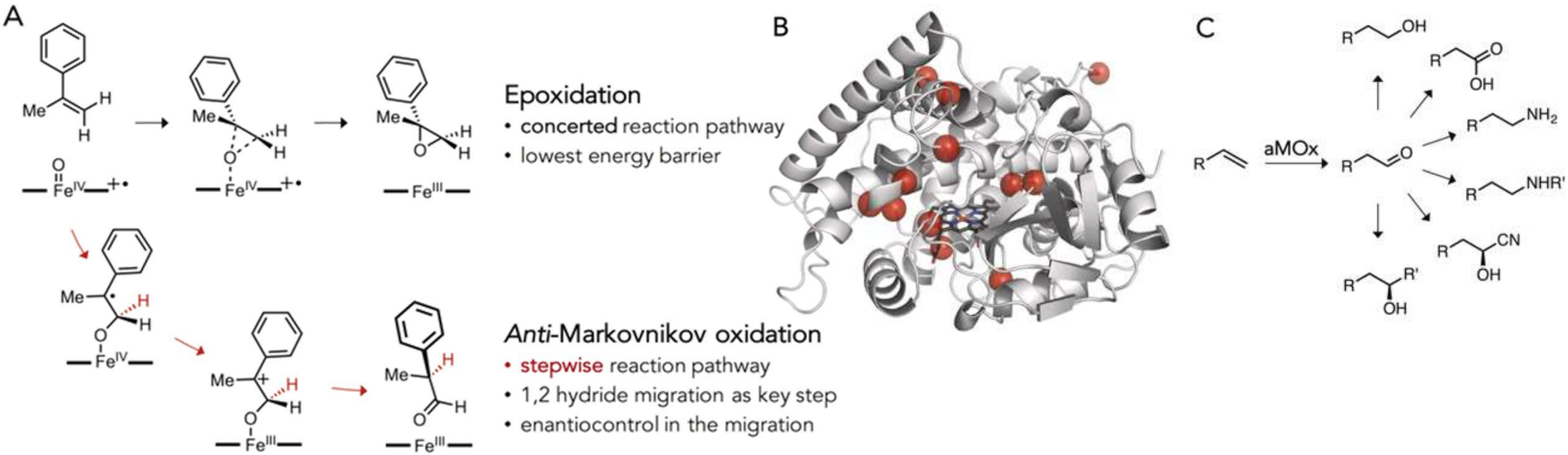
Directed evolution – bringing new chemistry to life



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

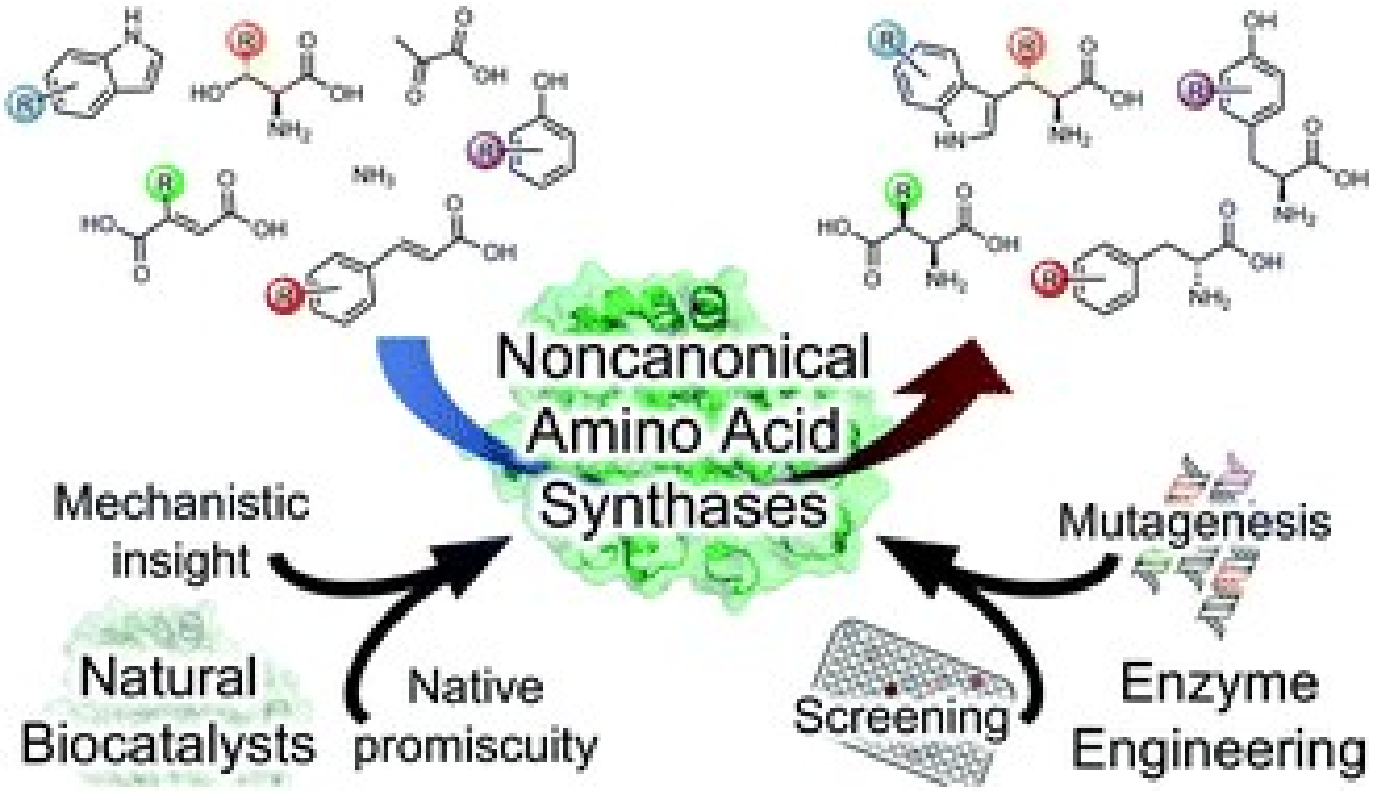
Directed evolution – bringing new chemistry to life



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.

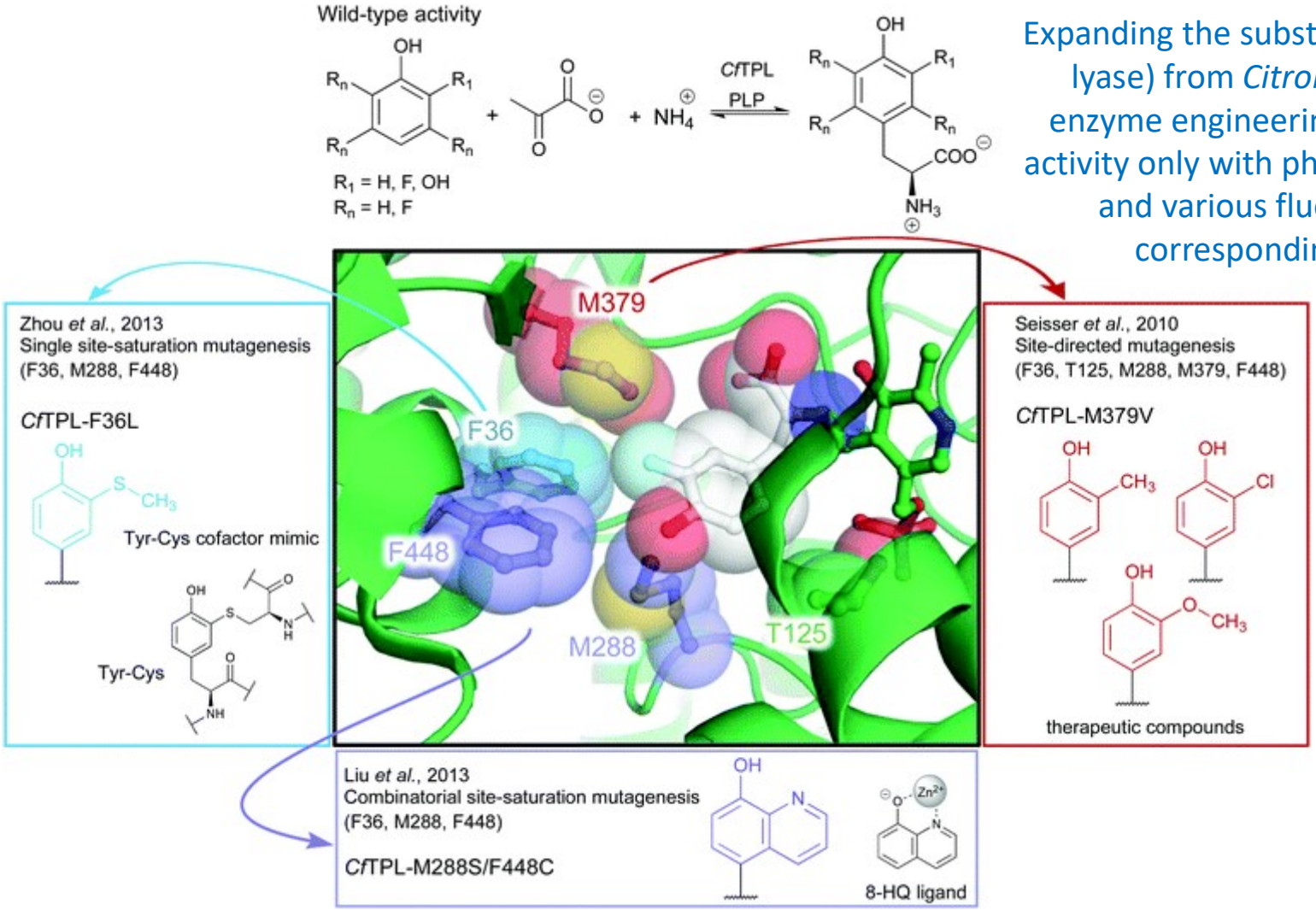
F. Arnold *Angew. Chem. Int. Ed.*, **2017**, *56,* 2-8

Engineering enzymes for noncanonical amino acid synthesis



P. J. Almhjell, C. E. Boville, F. Arnold *Chem. Soc. Rev.,* **2018,** *47*, 8980-8997

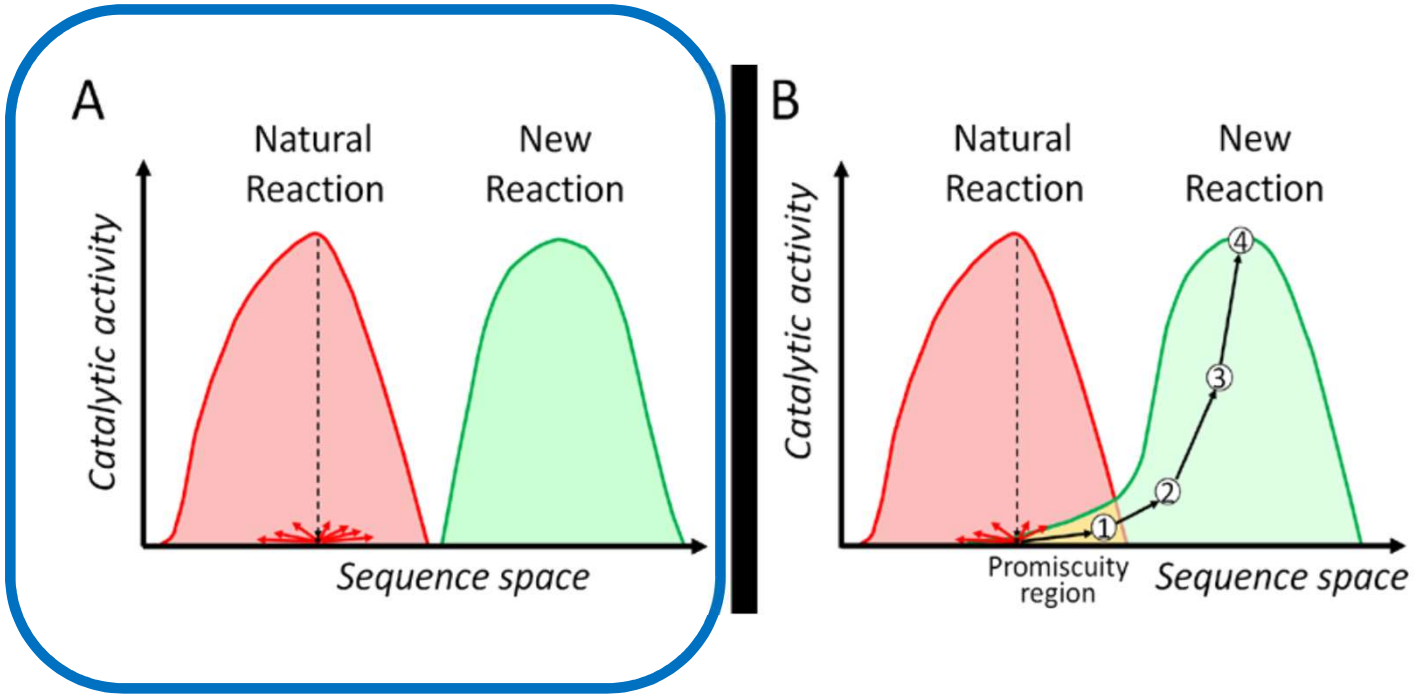
Engineering enzymes for noncanonical amino acid synthesis



Expanding the substrate scope of TPL (tyrosine phenyl lyase) from *Citrobacter freundii* (*Cf*TPL) through enzyme engineering. The wild-type enzyme shows activity only with phenol, catechol (3-hydroxyphenol), and various fluorophenols to synthesize the corresponding tyrosine (Tyr) analogue.

P. J. Almhjell, C. E. Boville, F. Arnold *Chem. Soc. Rev.,* **2018,** *47*, 8980-8997

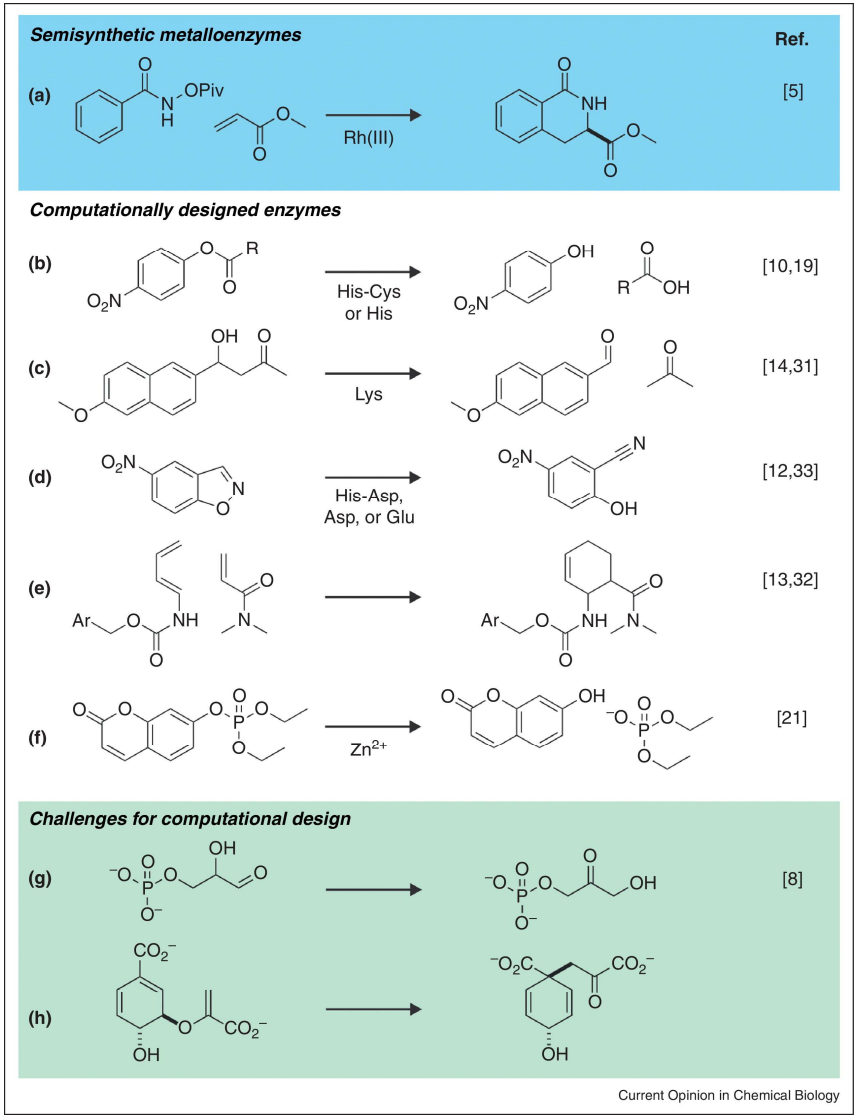
De novo enzyme engineering



A: no sequence variations (red arrows) create the new reactivity.

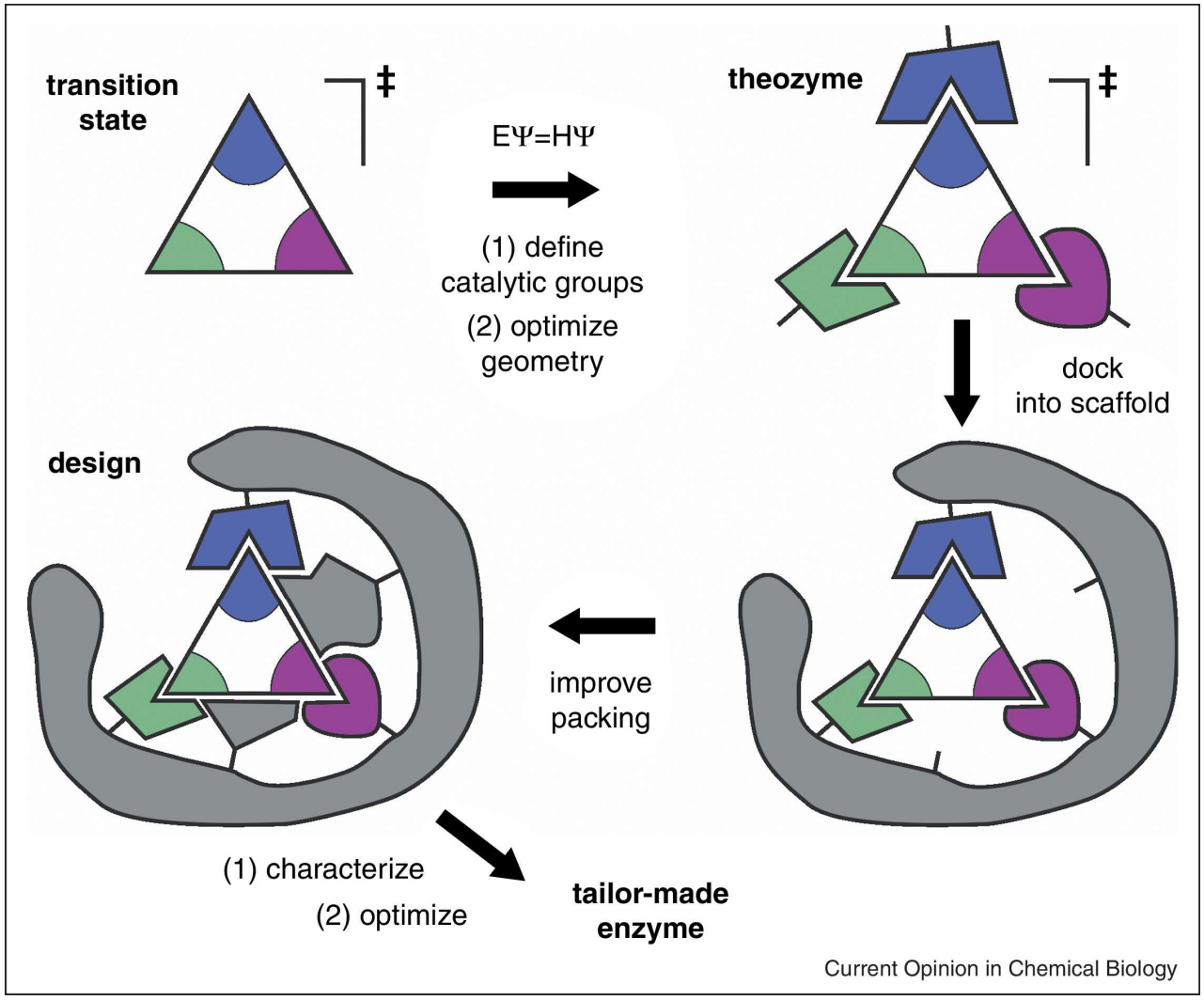
F. Arnold *Nobel lecture 2018; Angew.Chem.Int.Ed.* ***2019****, 58,* 14420–14426

Protein engineering – de novo enzyme design



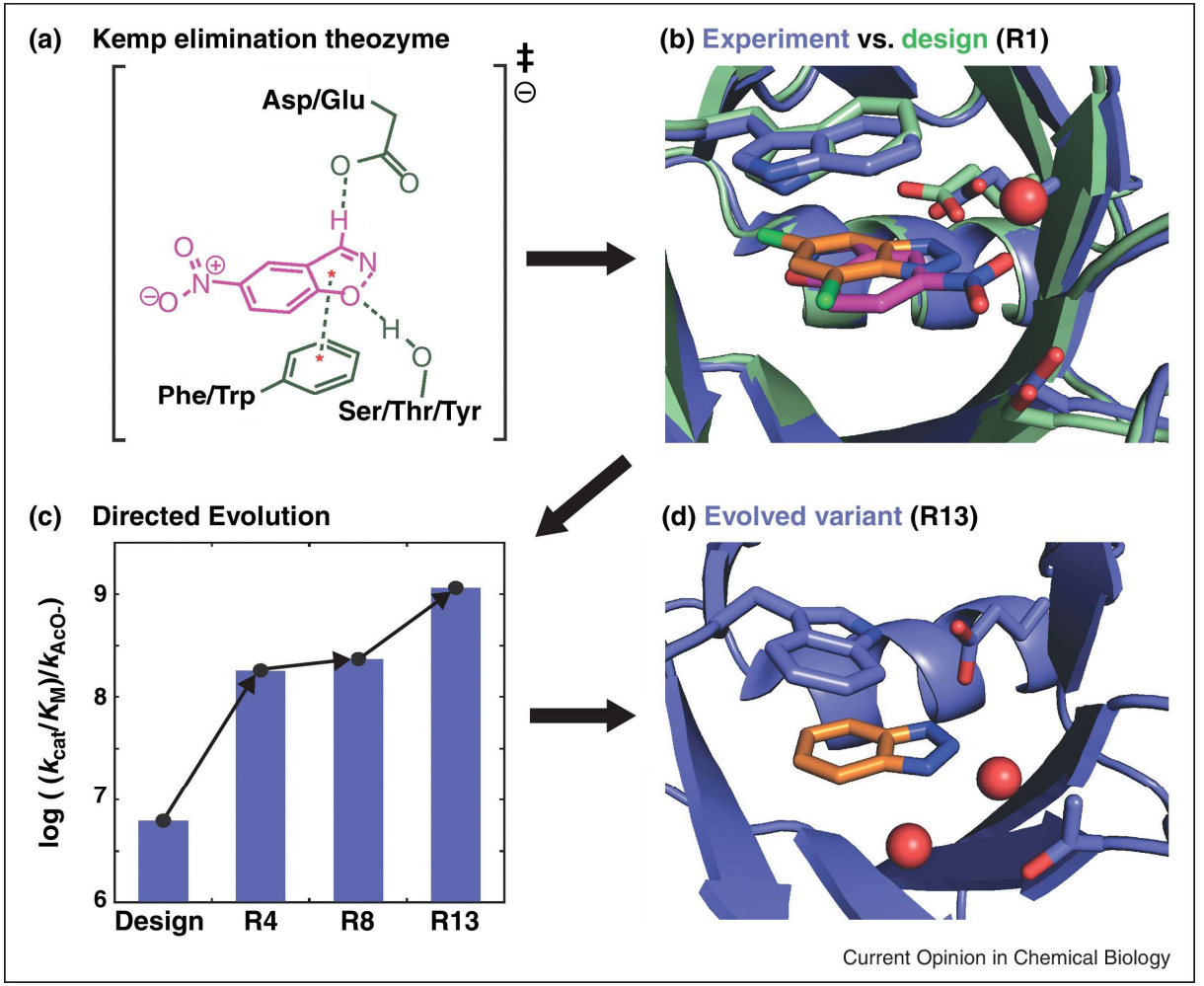
H. Kries, R. Blomberg, D. Hilvert *Curr. Opp. Chem. Biol.*, **2013**, *17,* 1-8

Protein engineering – de novo enzyme design

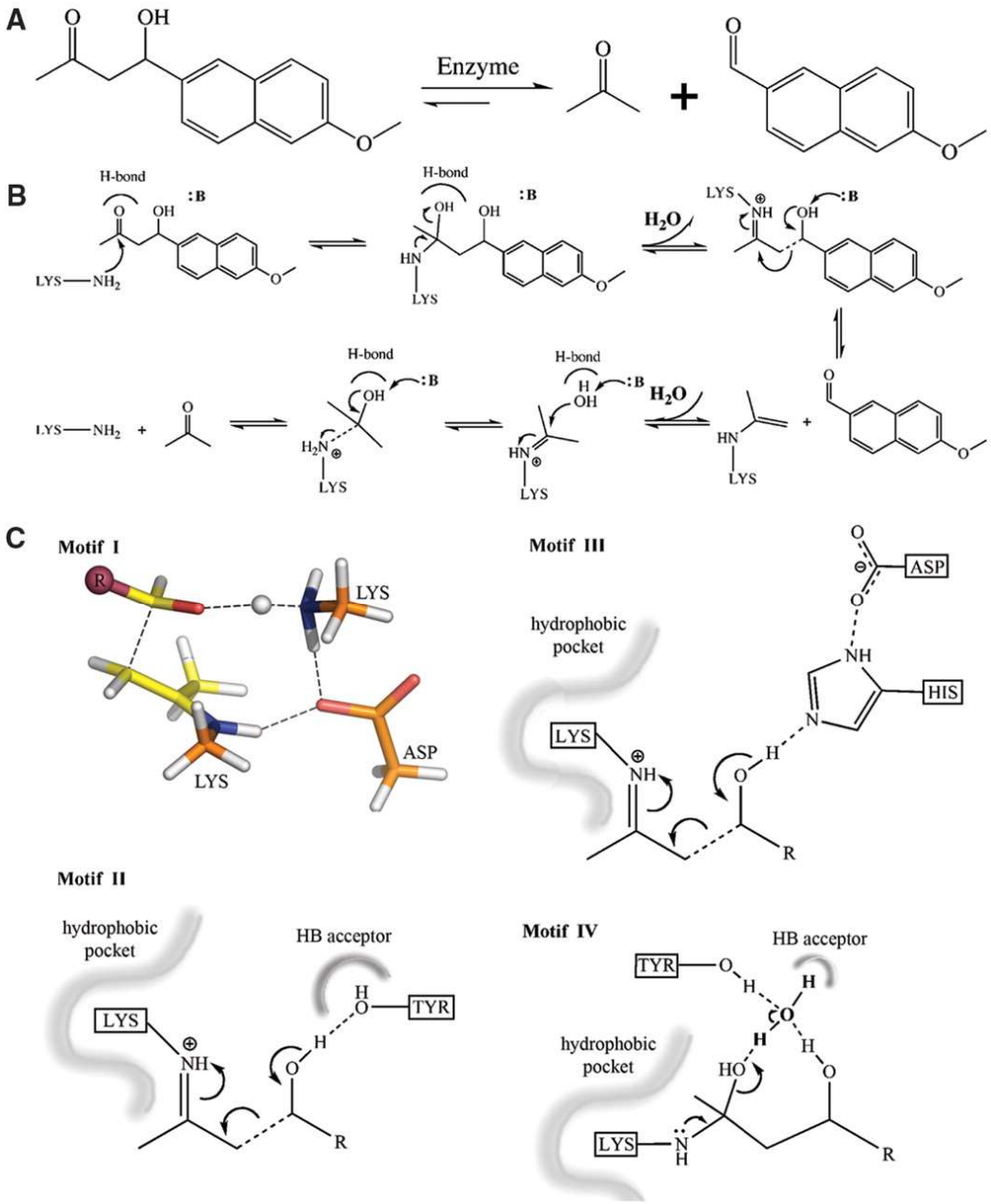


H. Kries, R. Blomberg, D. Hilvert *Curr. Opp. Chem. Biol.*, **2013**, *17,* 1-8

Protein engineering – de novo enzyme design



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



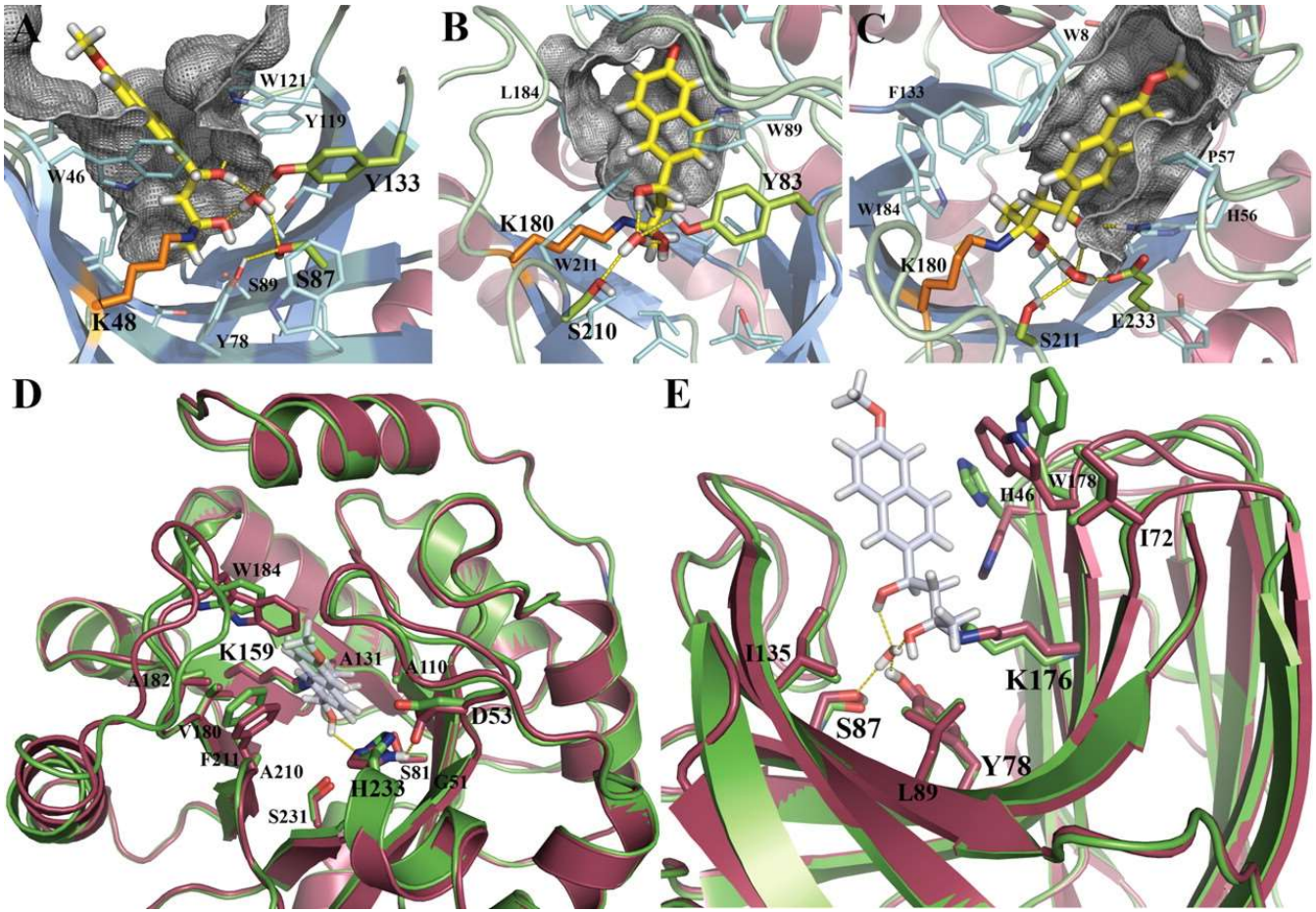
Retro-aldolase - de novo enzyme design

De novo retro-aldolases that use four different catalytic motifs to catalyze the breaking of a carbon-carbon bond in a nonnatural substrate were designed. Of the 72 designs that were experimentally characterized, 32, spanning a range of protein folds, had detectable retro-aldolase activity. Designs that used an explicit water molecule to mediate proton shuffling were significantly more successful, with rate accelerations of up to four orders of magnitude and multiple turnovers, than those involving charged side-chain networks. The atomic accuracy of the design process was confirmed by the x-ray crystal structure of active designs embedded in two protein scaffolds, both of which were nearly superimposable on the design model.

•Carlos F. Barbas III
•Donald Hilvert
•Kendall N. Houk
•Barry L. Stoddard
•David Baker

L. Jiang, E. A. Althoff *et al. Science,* **2008**, Vol. 319, Issue 5868, pp. 1387-1391 DOI: 10.1126/science.1152692

Retro-aldolase - de novo enzyme design

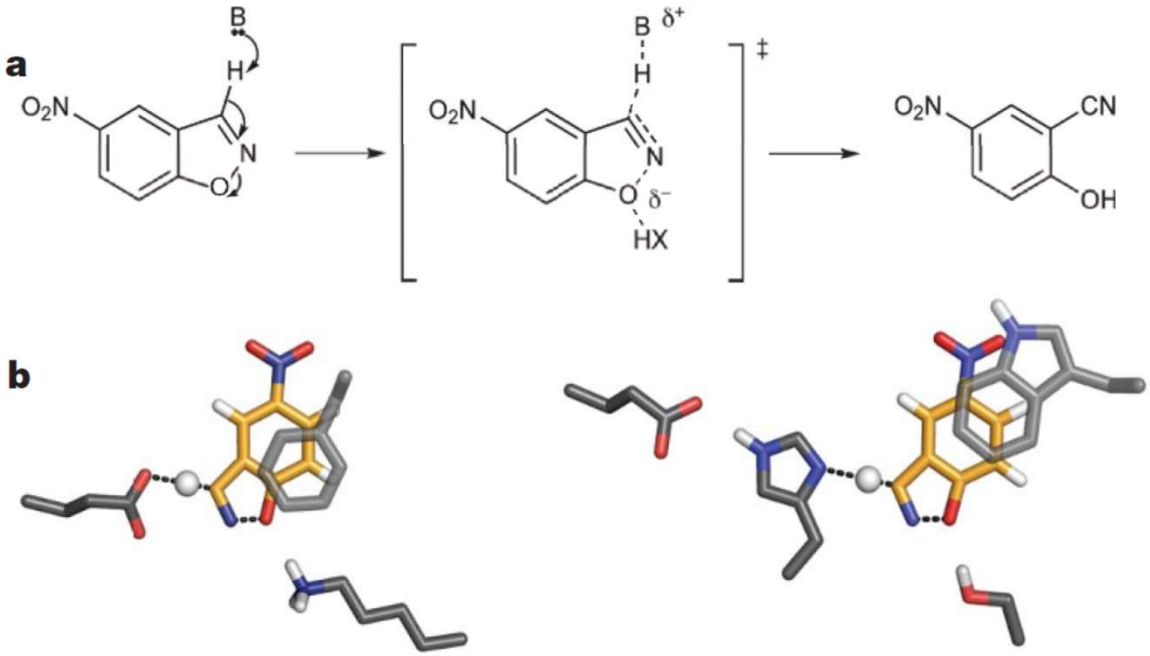


Examples of design models for active designs highlighting groups important for catalysis. The nucleophilic imine-forming lysine is in orange, the TS model is in yellow, the hydrogen-bonding groups are in light green, and the catalytic water is shown explicitly. The designed hydrophobic binding site for the aromatic portion of the TS model is indicated by the gray mesh

L. Jiang, E. A. Althoff *et al. Science,* **2008**, Vol. 319, Issue 5868, pp. 1387-1391 DOI: 10.1126/science.1152692

Kemp eliminase - de novo enzyme design

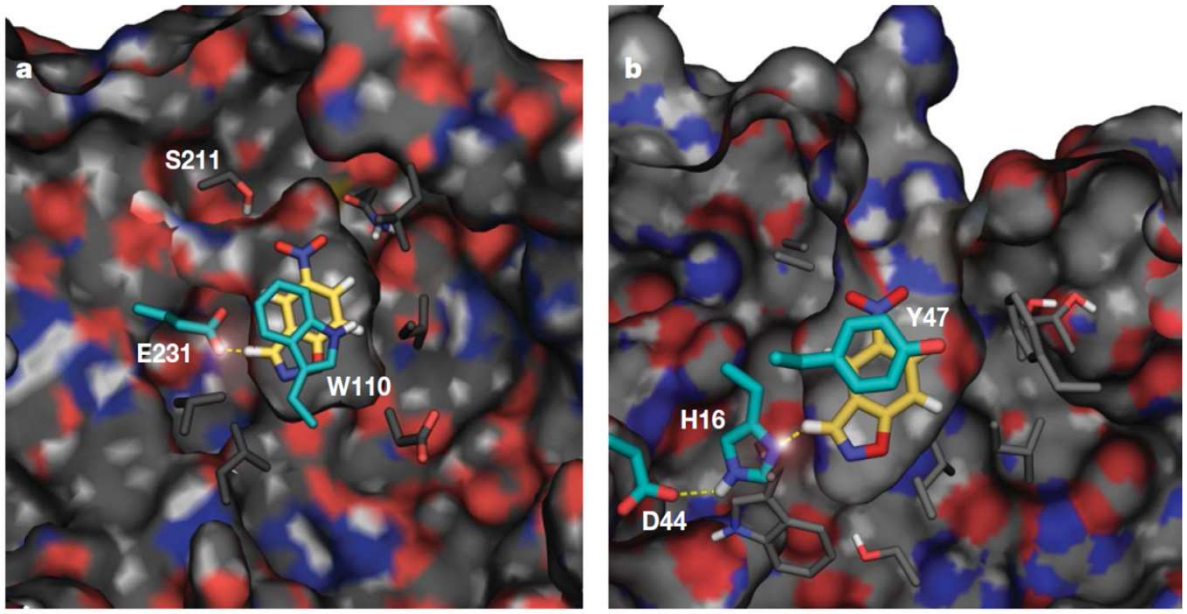
The Kemp elimination proceeds by means of a single transition state, which can be stabilized by a base deprotonating the carbon and the dispersion of the resulting negative charge; a hydrogen bond donor can also be used to stabilize the partial negative charge on the phenolic oxygen.



Examples of active site motifs highlighting the two choices for the catalytic base (a carboxylate (left) or a His–Asp dyad (right)) used for deprotonation, and a p-stacking aromatic residue for transition state stabilization. For each catalytic base, all combinations of hydrogen bond donor groups (Lys, Arg, Ser, Tyr, His, water or none) and p-stacking interactions (Phe, Tyr, Trp) were input as active site motifs into RosettaMatch

Daniela Röthlisberger, *et al., Nature,* **2008**, 453, 190-195

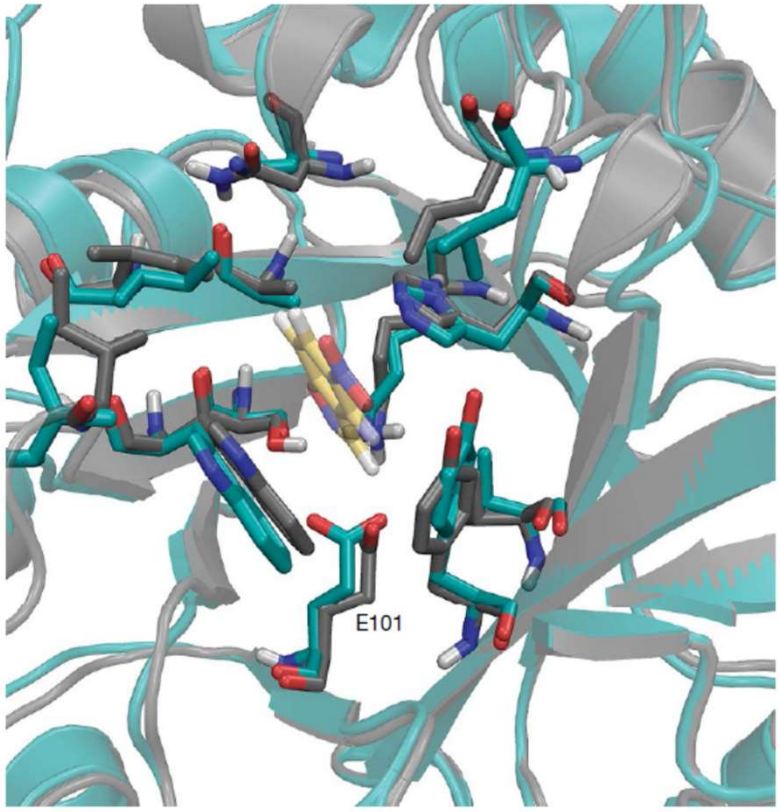
Kemp eliminase - de novo enzyme design



Computational design models of the two most active catalysts.
a, KE59 uses indole-3-glycerolphosphate synthase from Sulfolobus solfataricus as a scaffold. The transition state model is almost completely buried, with loops covering the active site. The mostly hydrophobic residues in the active site pocket pack the transition state model tightly, providing
high shape complementarity. The polar residue Ser 211 interacts with the nitro group of the transition state to promote binding. The key catalytic residues (Glu 231 and Trp 110) are depicted in cyan.
b, The deoxyribose-phosphate aldolase from E. coli is the scaffold for KE70. The shorter loops leave the active-site pocket freely accessible for the substrate. The transition state is surrounded by hydrophobic residues that provide high shape complementarity. His 16 and Asp 44 (in cyan) constitute the catalytic dyad whereas Tyr 47 (in cyan) provides p-stacking interactions.

Daniela Röthlisberger, *et al., Nature,* **2008**, 453, 190-195

Kemp eliminase - de novo enzyme design



Comparison of the designed model of KE07 and the crystal structure.

The crystal structure (cyan) was solved in the unbound state and shows only modest rearrangement of active site side chains compared to the designed structure (grey) modelled in the presence of the transition state (yellow, transparent). KE07 contains 13 mutations compared to the starting template scaffold (PDB code 1thf).

Daniela Röthlisberger, *et al., Nature,* **2008**, 453, 190-195

Computational Design of Catalytic Dyads and Oxyanion Holes for Ester Hydrolysis

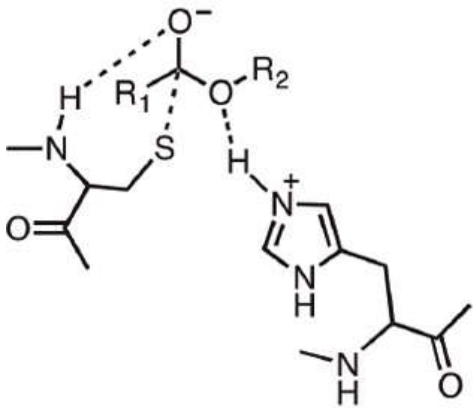
Nucleophilic catalysis is a general strategy for accelerating ester and amide hydrolysis.

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.

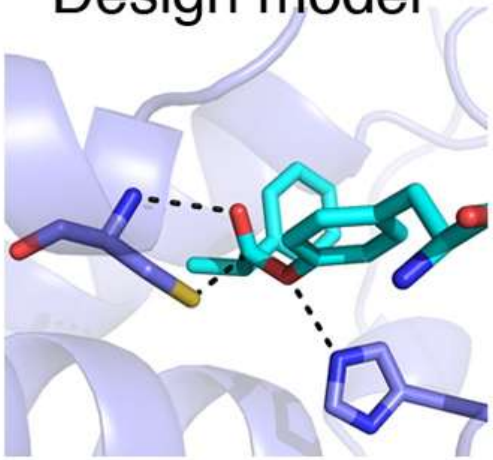
Minimal requirements for esterase activity have been explored by computationally designing artificial catalysts using catalytic dyads and oxyanion holes.

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.

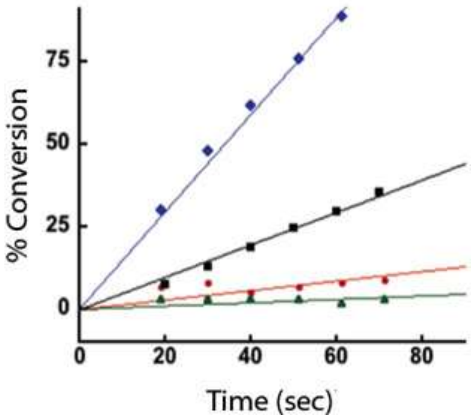
Theozyme



Design model



Catalytic activity



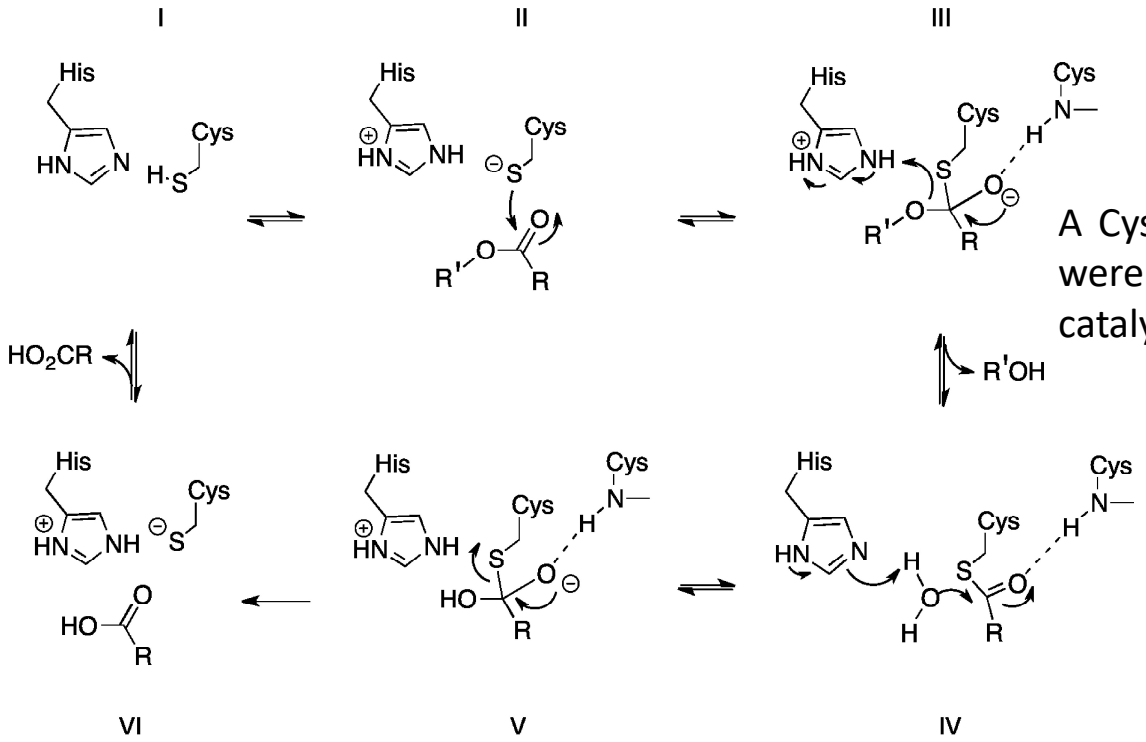
Most active variant:
k_{cat}/K_M of 400 M^{-1} s^{-1}
(*p*-nitrophenyl ester cleavage).

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

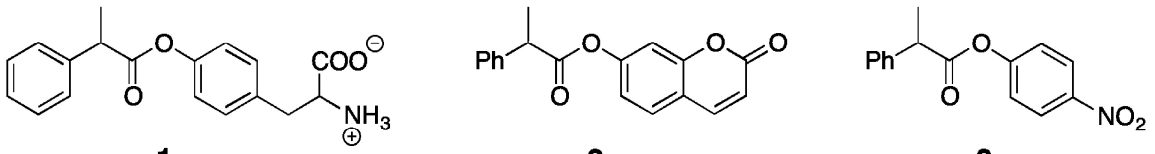
Programmed mechanism and model substrates of the de novo designed esterases.

A



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

B



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.

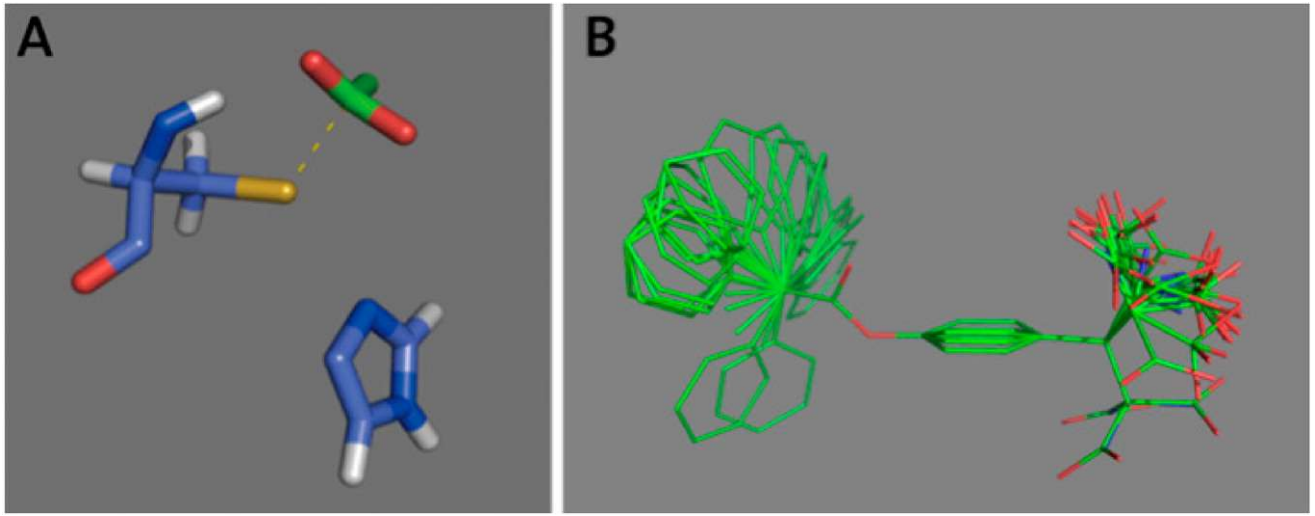
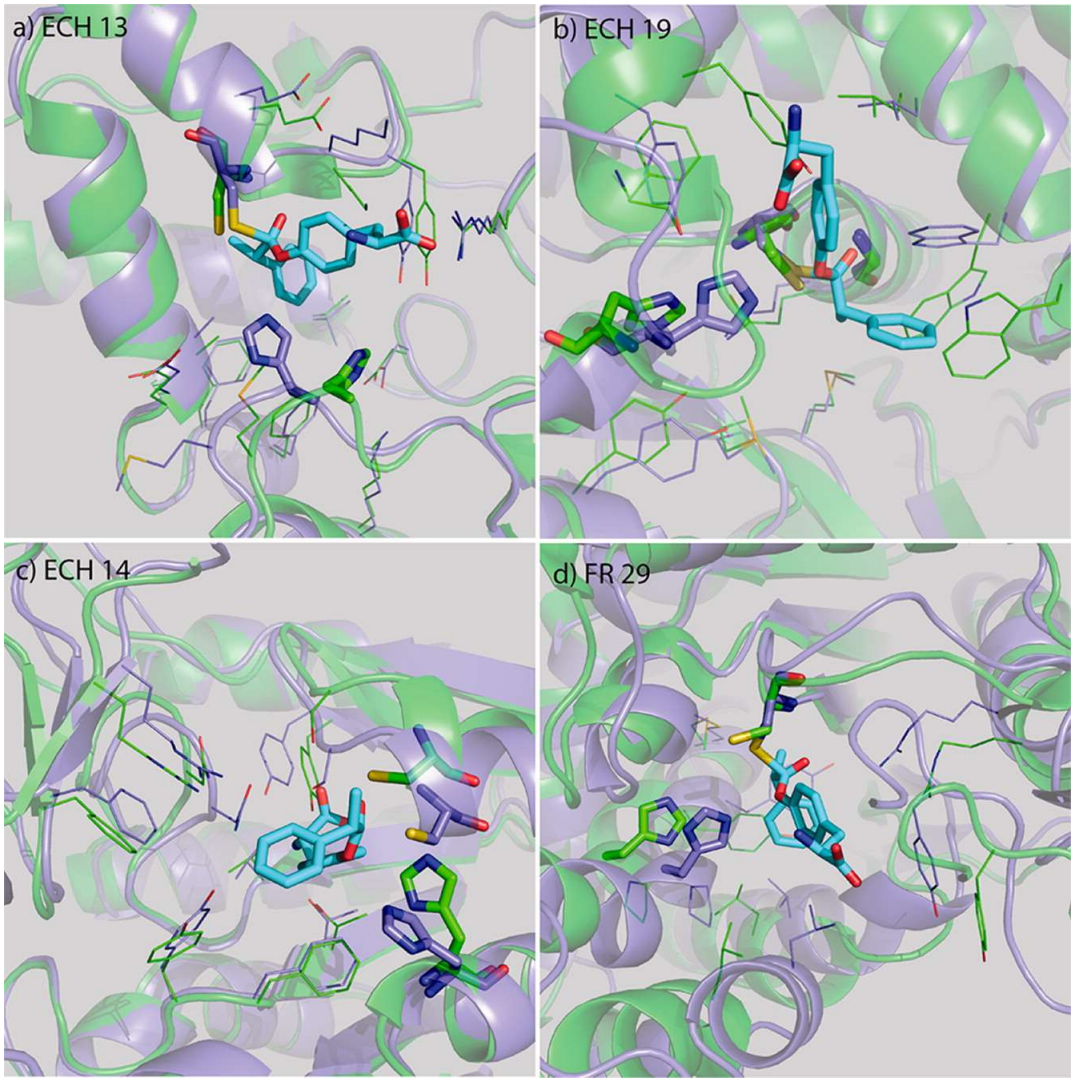


Image of the theoretical conformer ensemble of tyrosyl ester **1**.

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

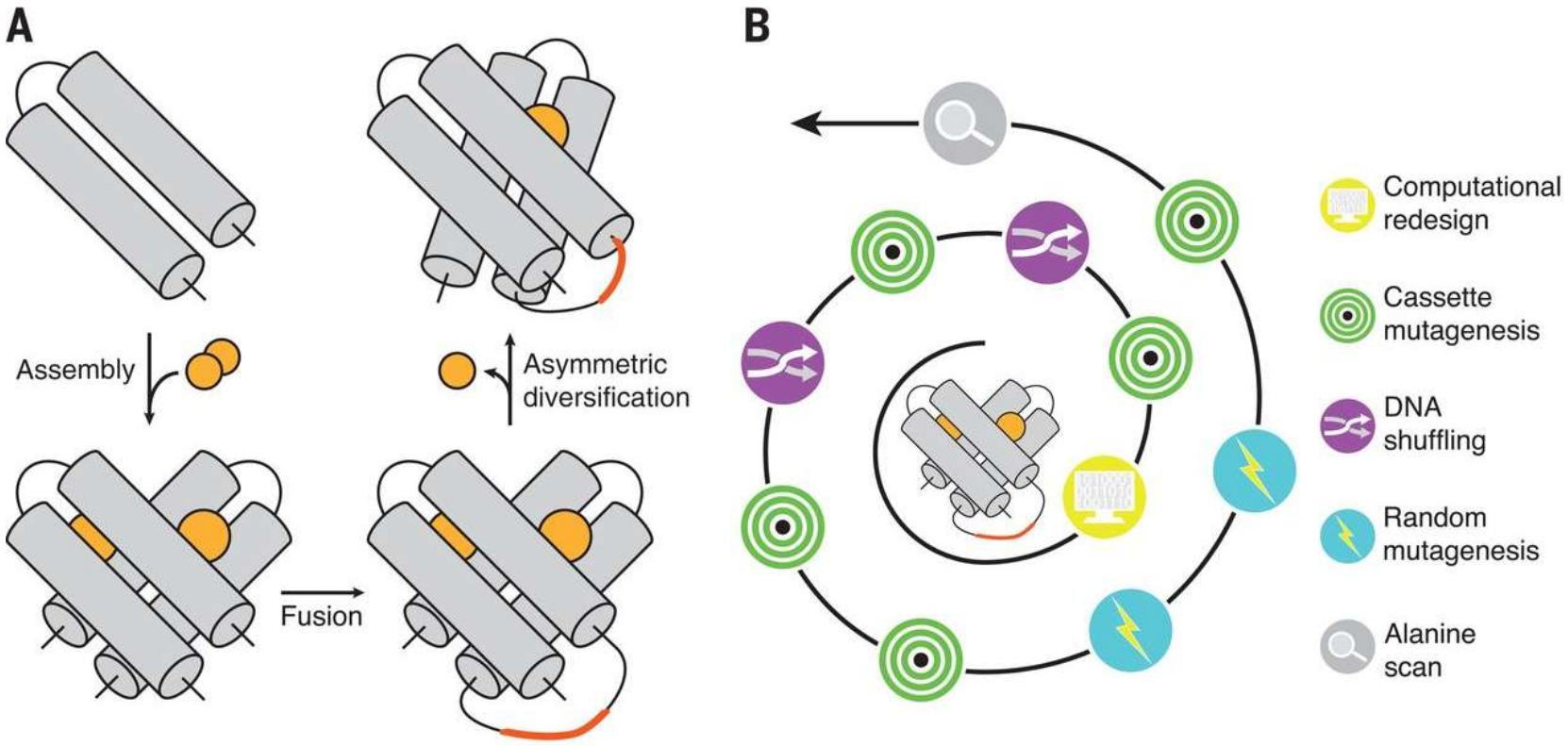
Computational Design of Catalytic Dyads and Oxyanion Holes for Ester Hydrolysis



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

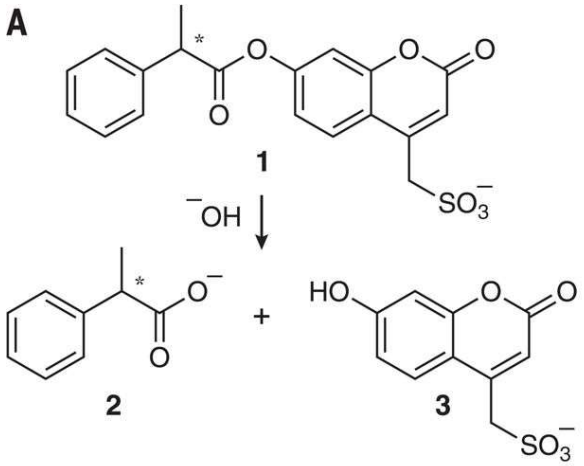
Evolution of a metalloenzyme from short peptides

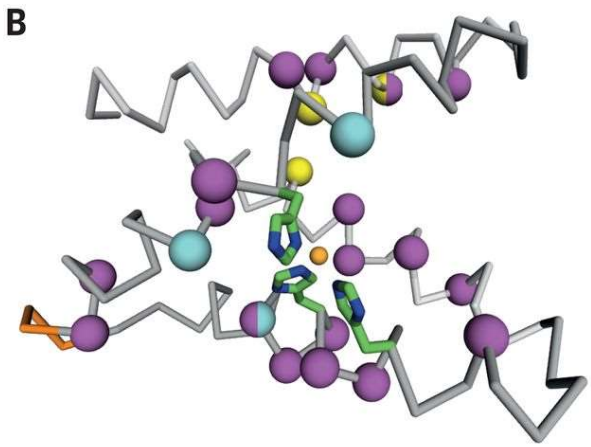


Zinc-mediated assembly of helix-turn-helix fragments, followed by fusion and asymmetric diversification, afforded MID1sc10, an efficient metalloesterase.

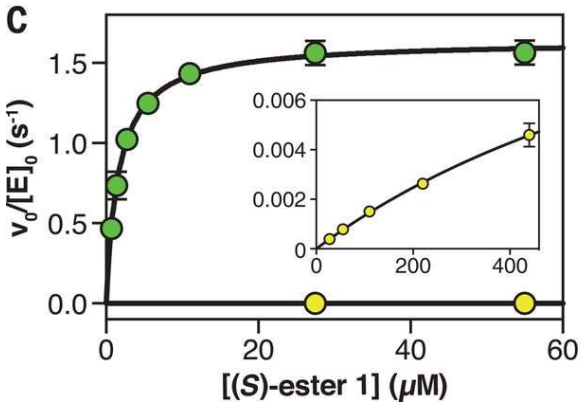
S. Studer, D.A. Hansen, Z. Pianowski, P.R.E. Mittl, A. Debon, S.L. Guffy, B.S. Der, B. Kuhlman, D. Hilvert *Science,* **2018**, *362,* 1285-1288

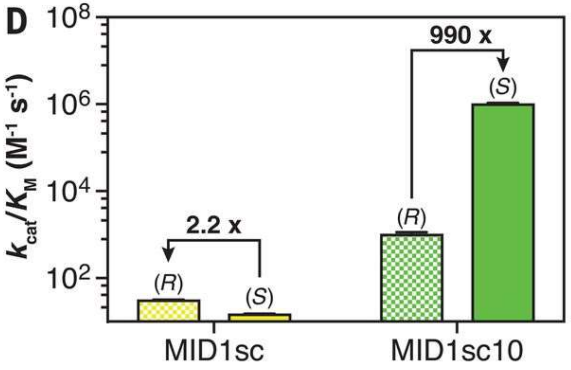
Evolution of a metalloenzyme from short peptides





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).



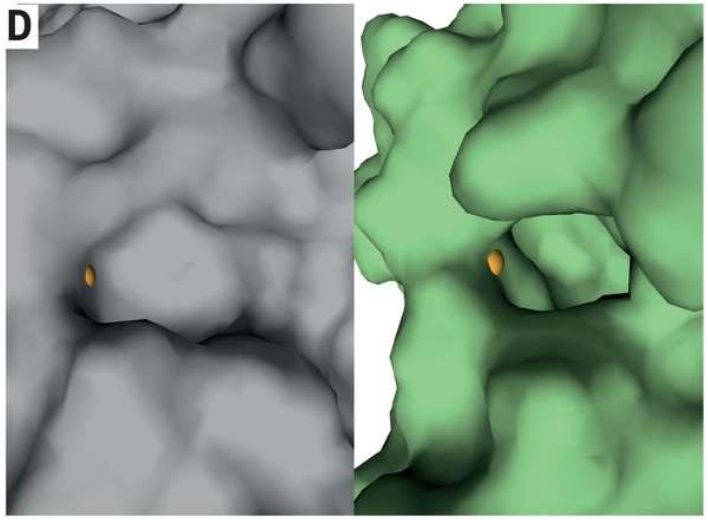


The evolved variant MID1sc10 is highly enantioselective as a consequence of a 2200-fold specificity switch from the modestly (R)-selective starting catalyst MID1sc

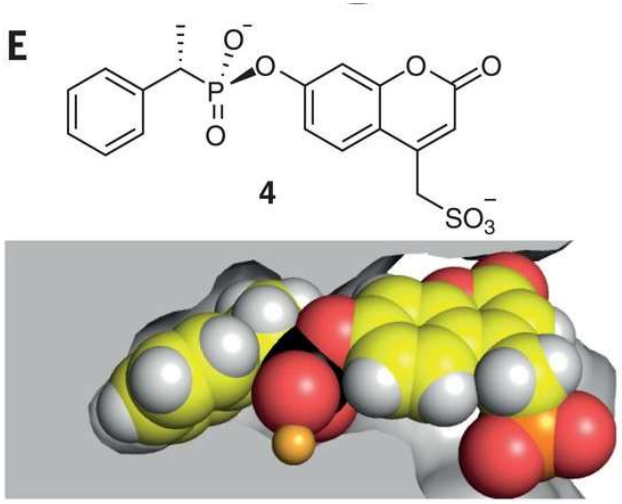
Michaelis-Menten plots for **MID1sc** (yellow and inset) and **MID1sc10** (green) show a 70,000-fold improvement in hydrolysis efficiency for (S)-configured **1** after optimization.

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

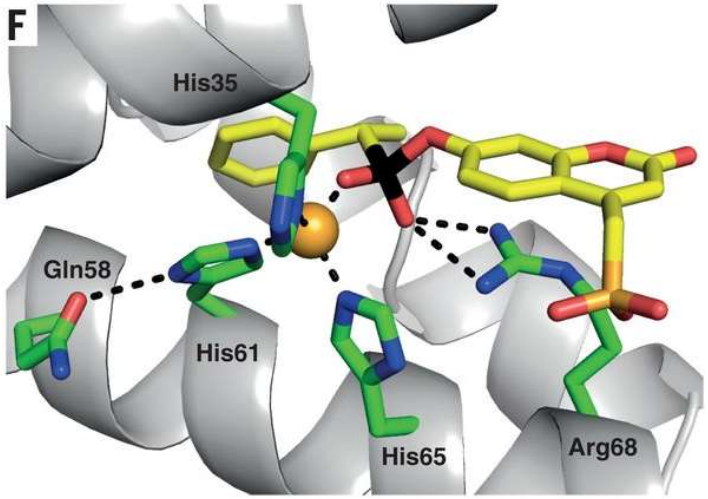
Evolution of a metalloenzyme from short peptides



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

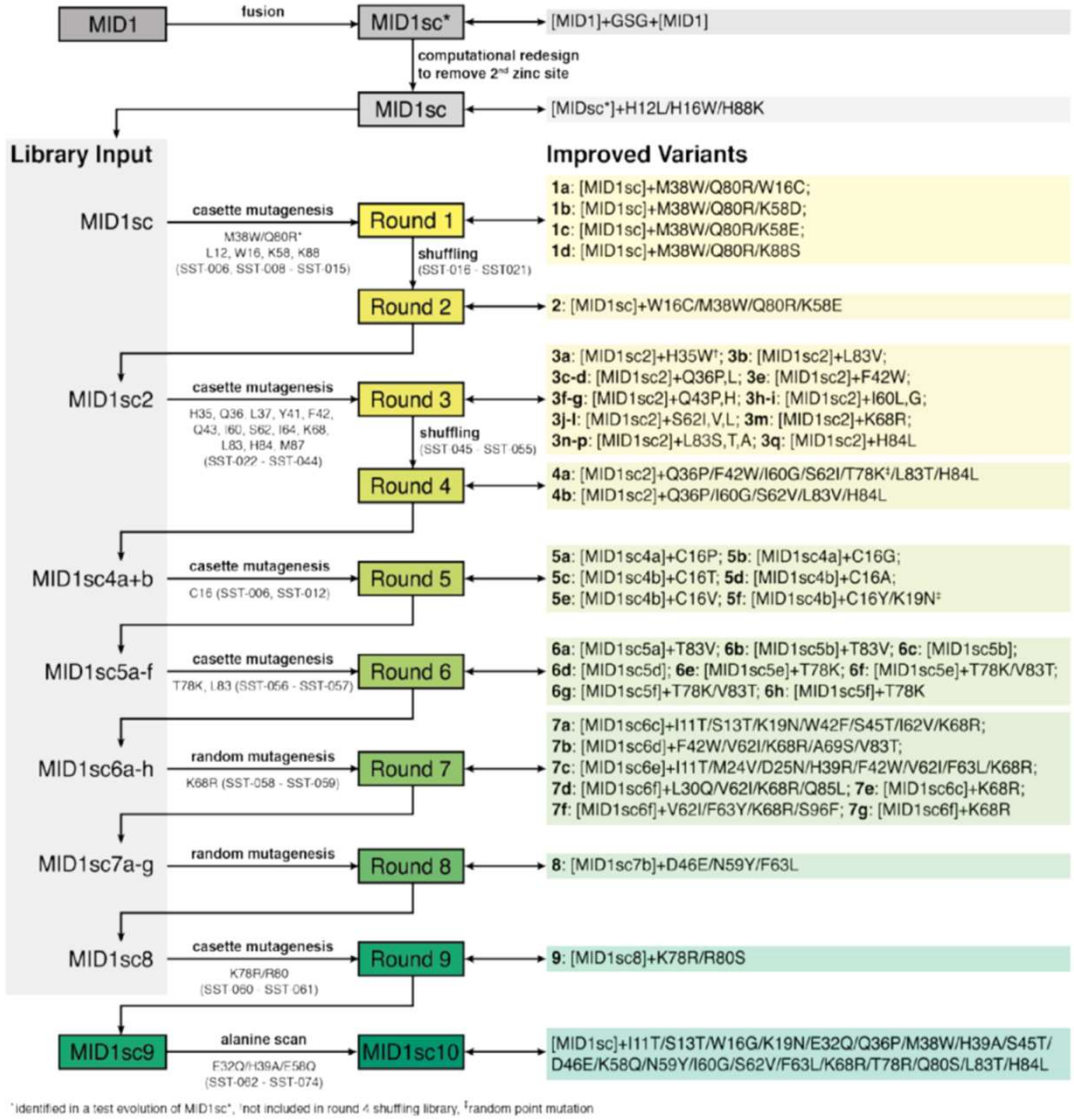


Cut-away view of the active site, showing the snug fit of phosphonate **4** in the binding pocket. (zinc ion – orange)



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

Synthetic life

www.genome.gov/about-genomics/policy-issues/Synthetic-Biology

How new functions can emerge from known biological building blocks?

Overview of the course

artificial ribozymes and aptamers for efficient catalysis and recognition (SELEX, DNAzymes, foldamers);

unnatural base pairing – expansion of the genetic alphabet;

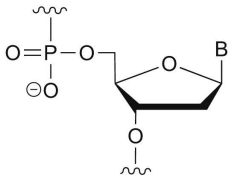
artificial genetic polymers and oligonucleotide analogues (XNA);

biosynthetic incorporation of **unnatural aminoacids (UAAs)** into proteins;

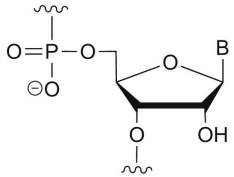
Artificial lipid vesicles as models for protocell multiplication;

enzyme engineering – production of enzymes with unknown or unnatural properties, *ab initio* protein design, directed evolution, theozymes;

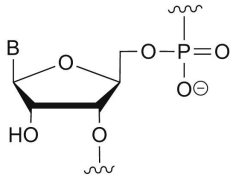
Artificial genetic polymers



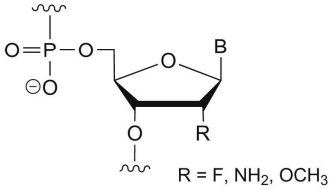
DNA



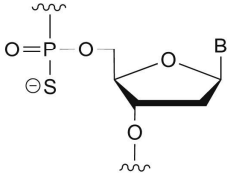
D-RNA
(natural)



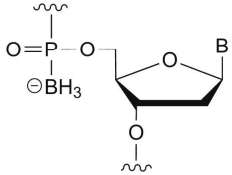
L-RNA
(unnatural)



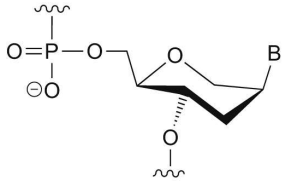
2'-modified RNA



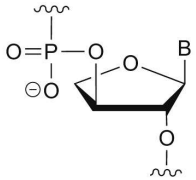
Phosphorothioate



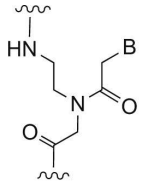
Boranophosphate



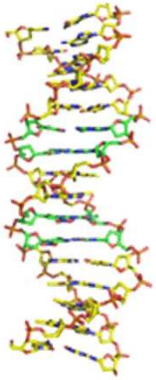
Hexitol Nucleic Acid
(HNA)



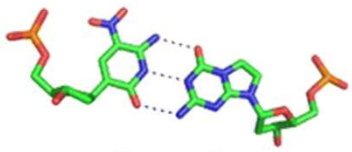
Threose Nucleic Acid
(TNA)

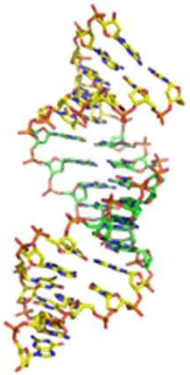


Peptide Nucleic Acid
(PNA)

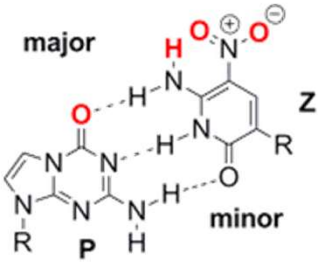


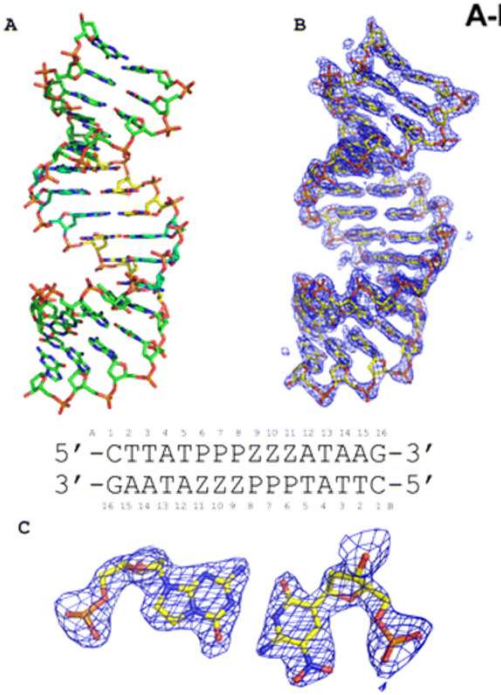
B-DNA

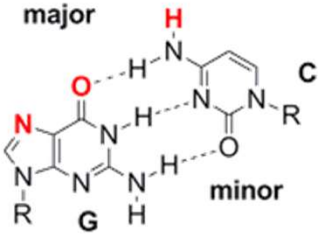


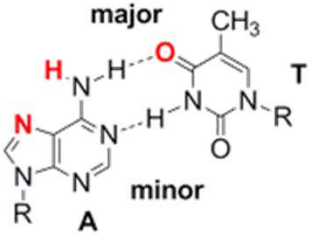


A-DNA

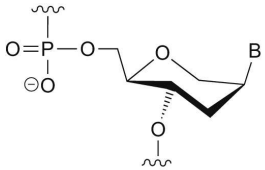




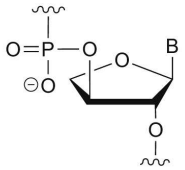




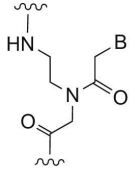
Artificial genetic polymers



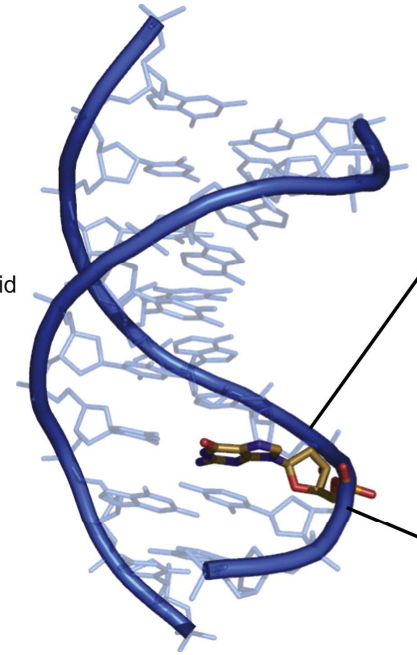
Hexitol Nucleic Acid
(HNA)

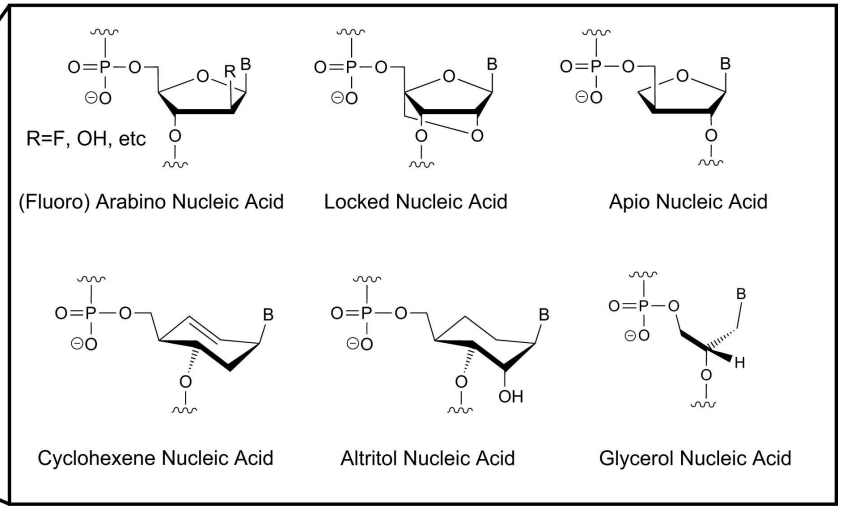


Threose Nucleic Acid
(TNA)



Peptide Nucleic Acid
(PNA)





(Fluoro) Arabino Nucleic Acid

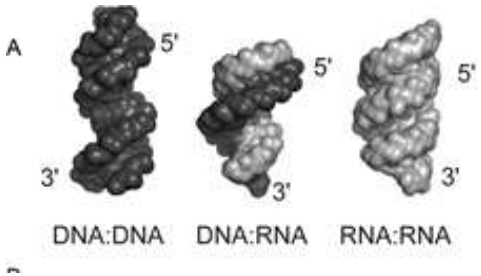
Locked Nucleic Acid

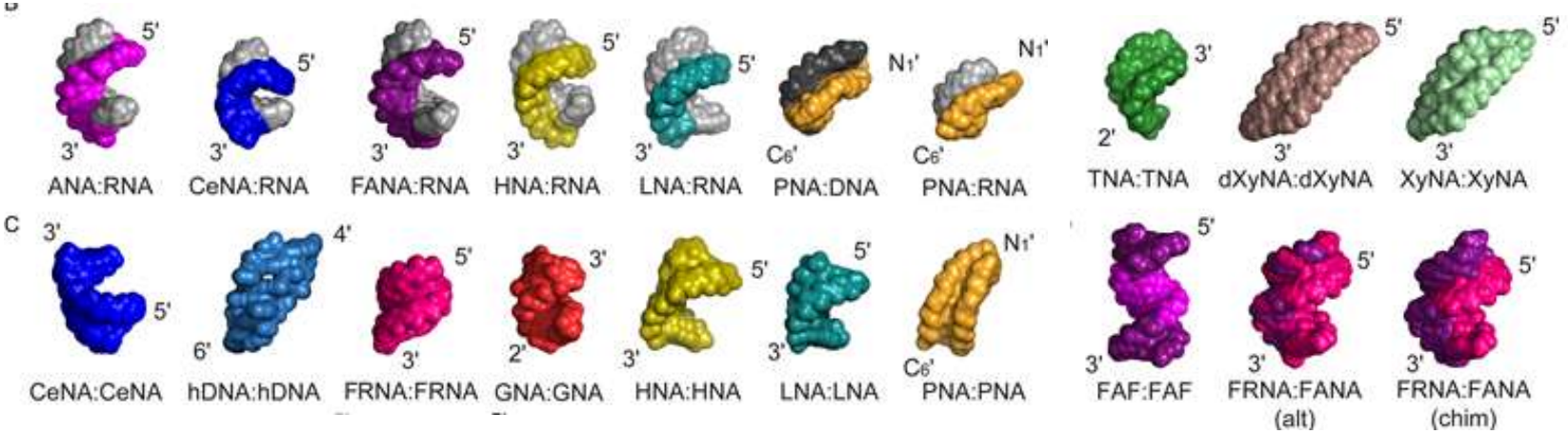
Apio Nucleic Acid

Cyclohexene Nucleic Acid

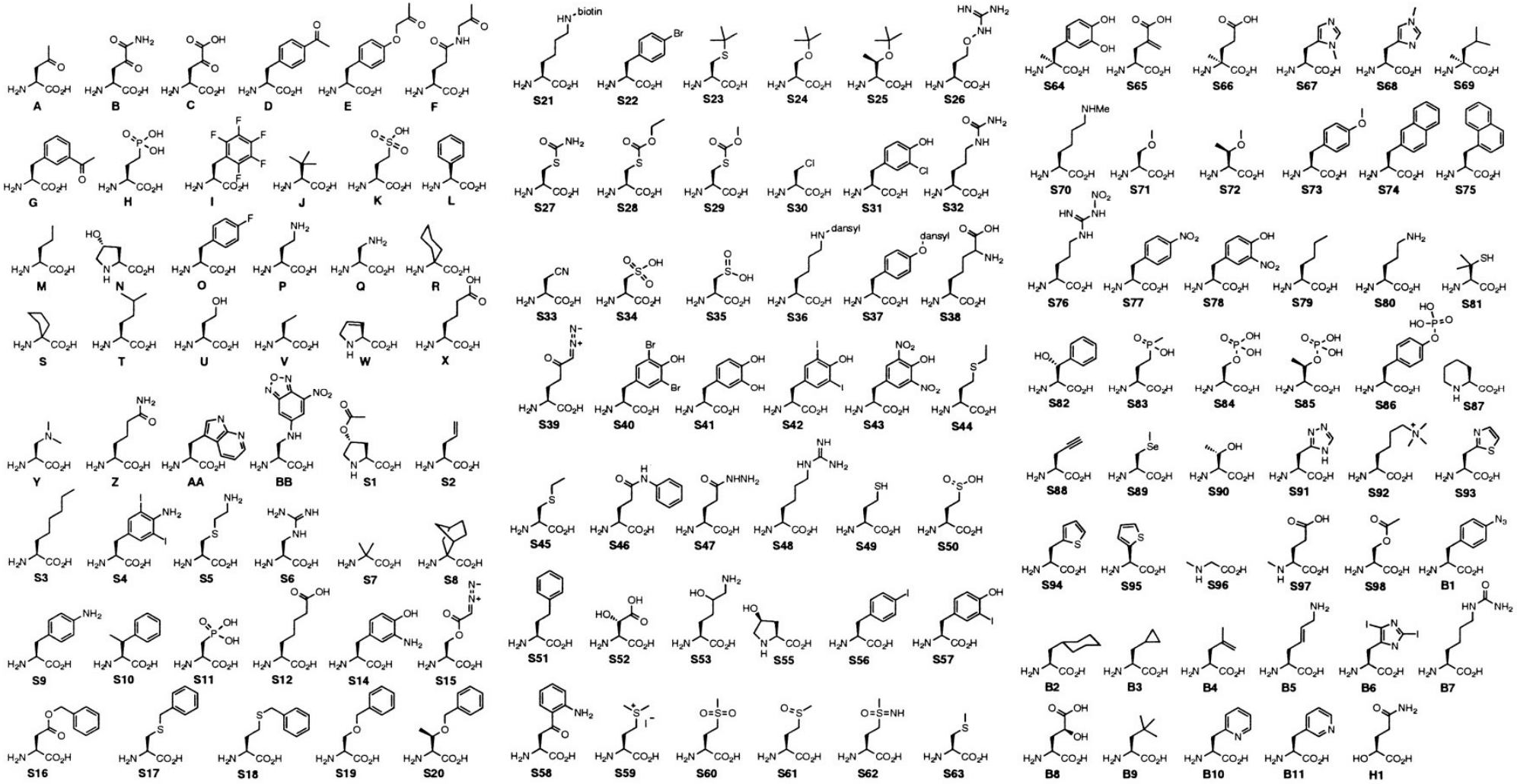
Altritol Nucleic Acid

Glycerol Nucleic Acid

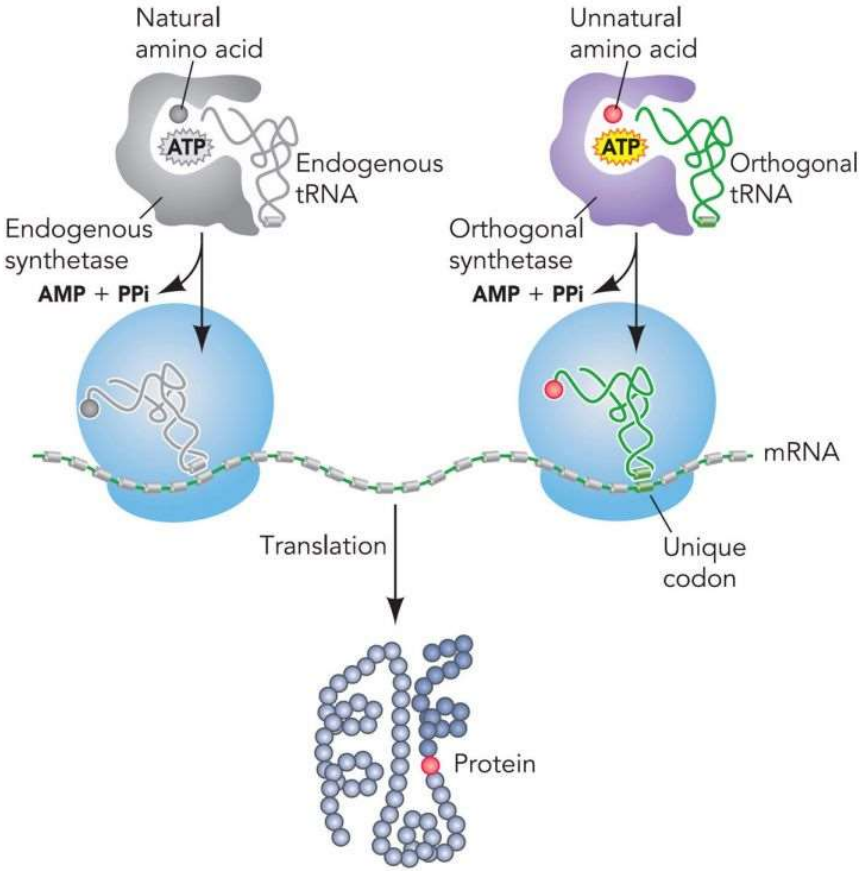


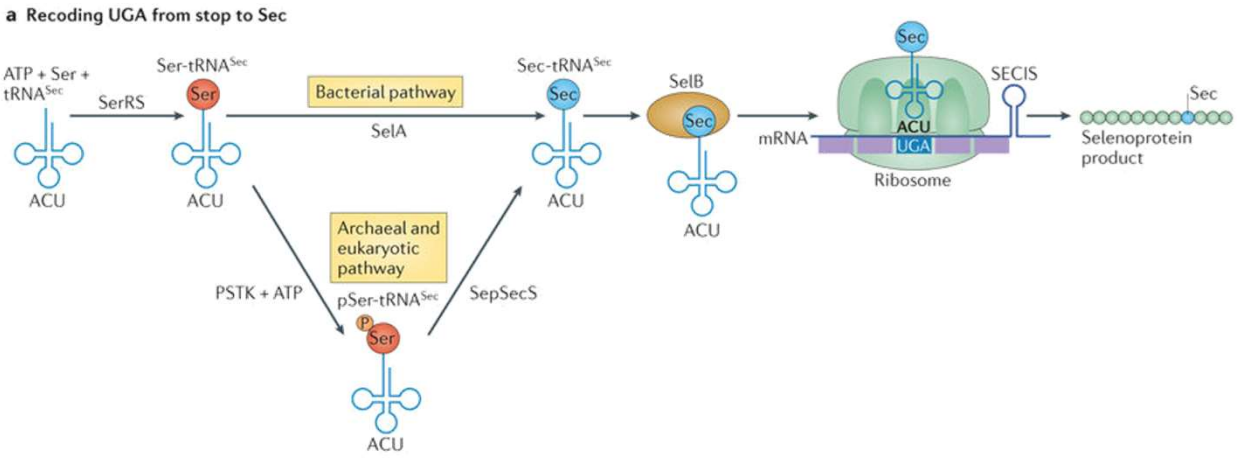


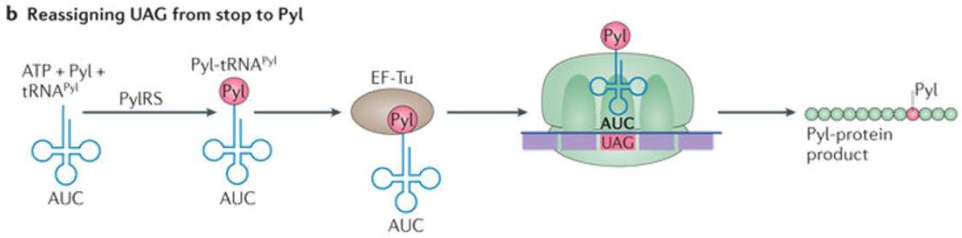
Expanded genetic code



Expanded genetic code

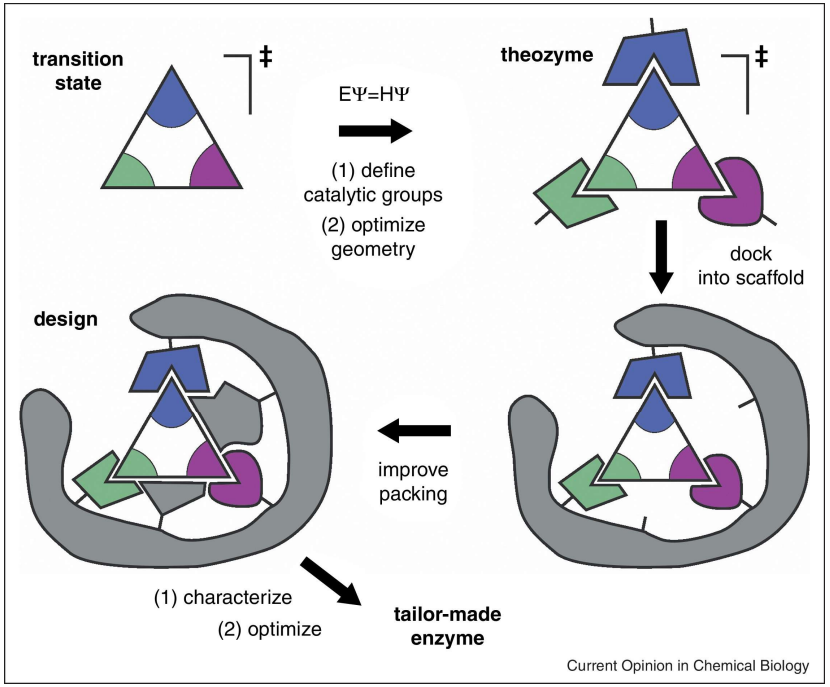


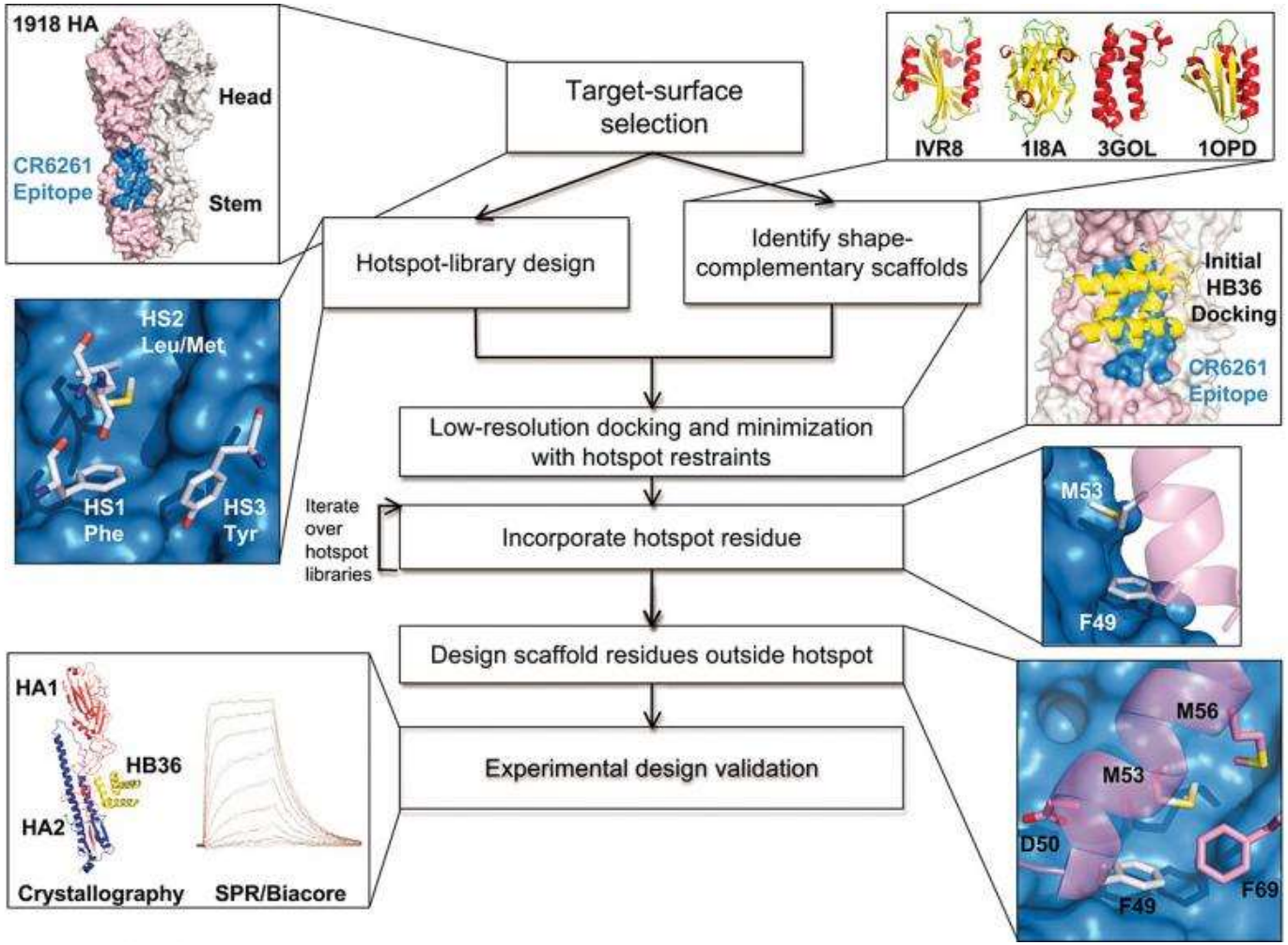




Nature Reviews | Microbiology

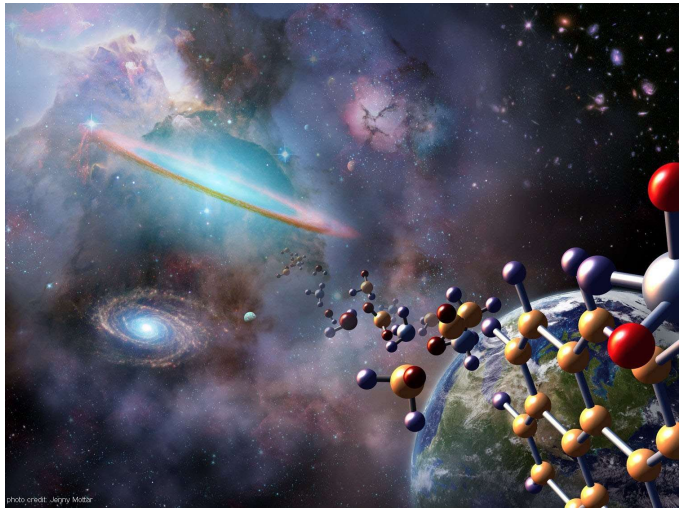
Protein engineering and de novo enzyme design





The molecular origins of life

Life is a self-replicating chemical system capable of evolution (NASA, 2009)



Origin of the Universe – stars, planets, elements

Origin of biorelevant monomers – primordial soup

Complex chemical processes on the way to living systems

Protocells and LUCA