BIOENG-320
Synthetic Biology

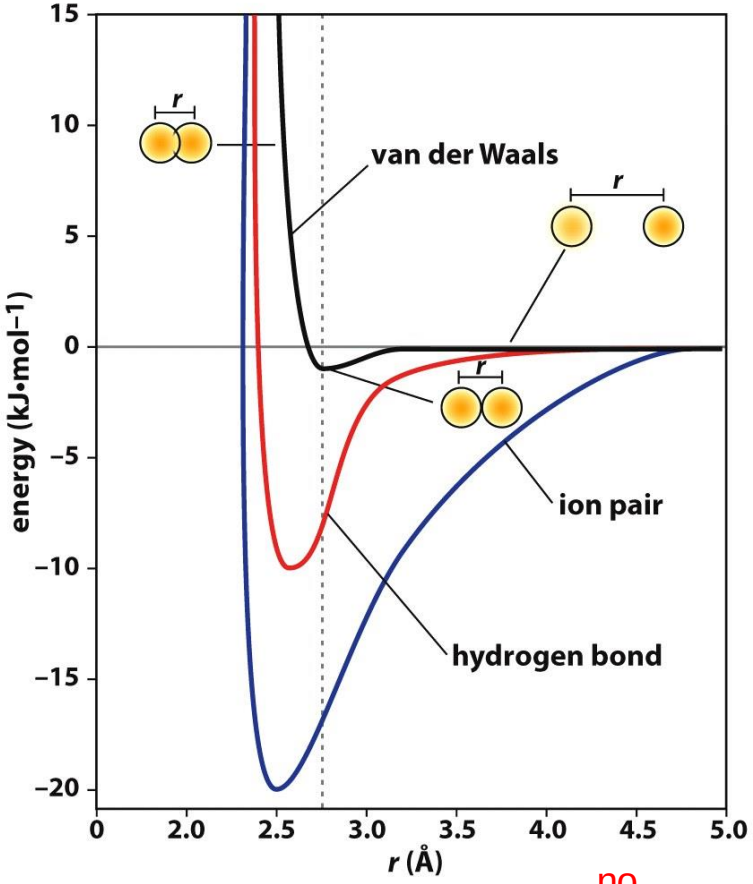
Protein design Lecture 3
March 10, 2025

Design of protein binding

Patrick Barth
EPFL

Protein Design – the binding problem

Atomic interactions



repulsive attractive no interactions

Sum over all pairs of atoms

Protein-ligand interactions

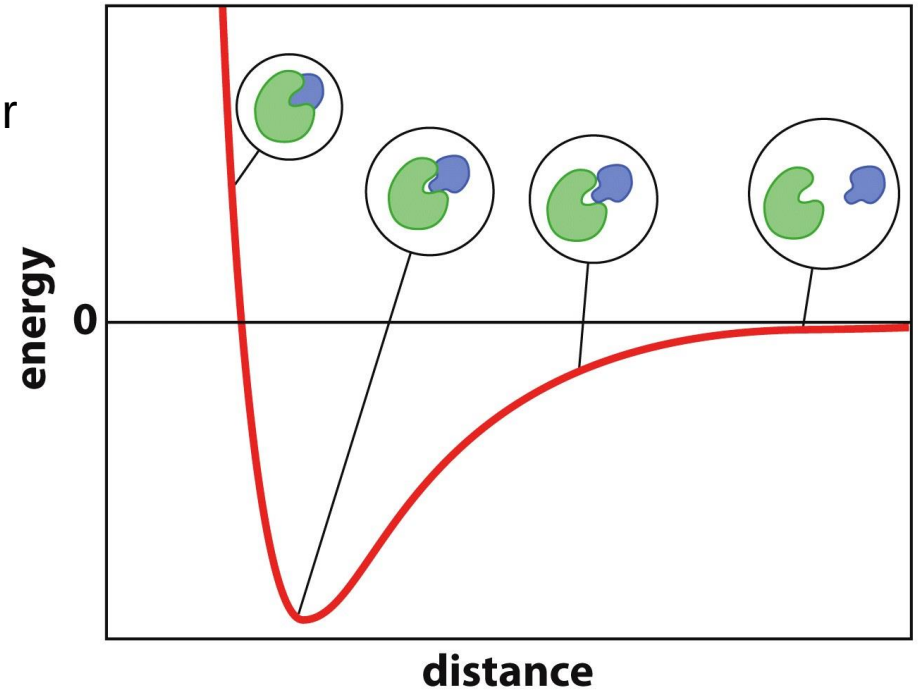


Figure 1.2 The Molecules of Life (© Garland Science 2013)

Protein Design – the binding problem

Protein-ligand interactions

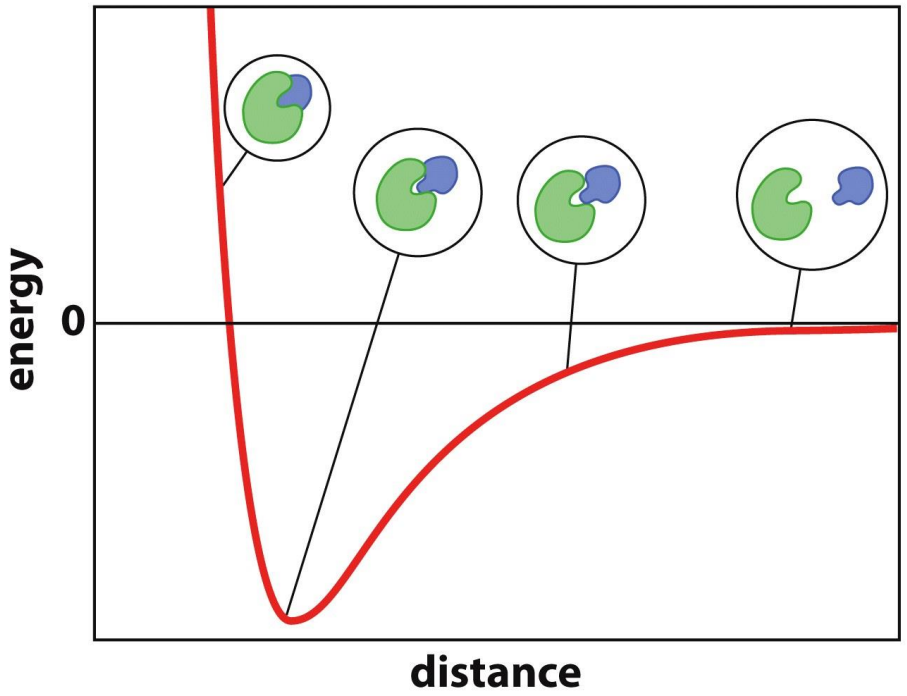


Figure 1.2 The Molecules of Life (© Garland Science 2013)

Similar forces than during protein folding:

1. Desolvation of 2 protein surfaces

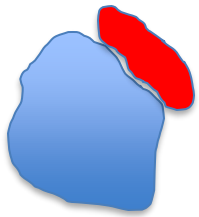
2. Creation of hydrophobic and polar interactions at the binding interface

Balance between 1 and 2 will dictate the strength of binding

→ Protein surface complementarity and surface areas are key!

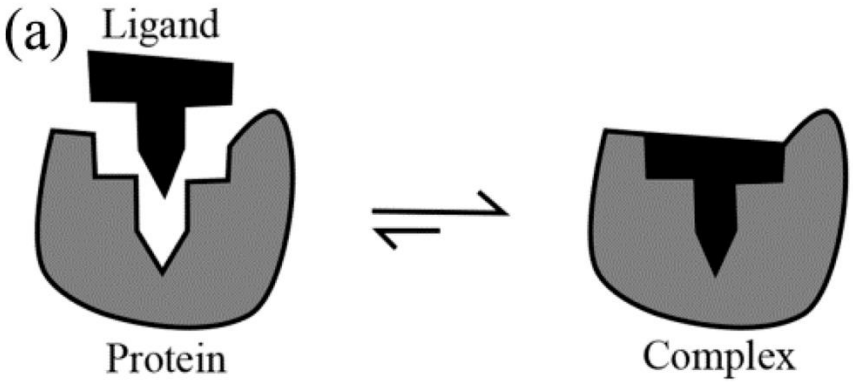
Buried cavity

Flat binding surface



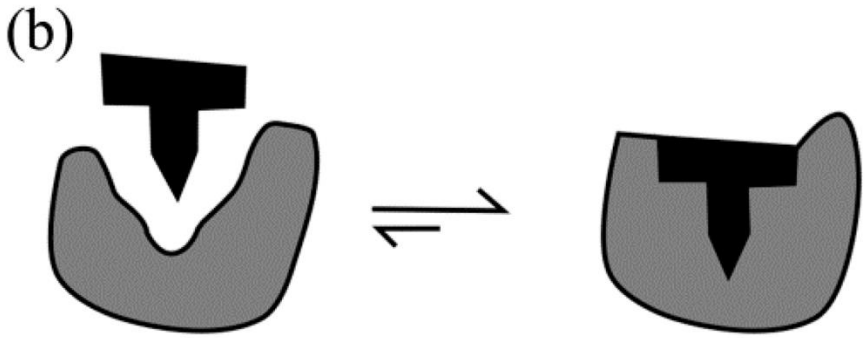
Protein binding – a conformational flexibility problem

Lock and key



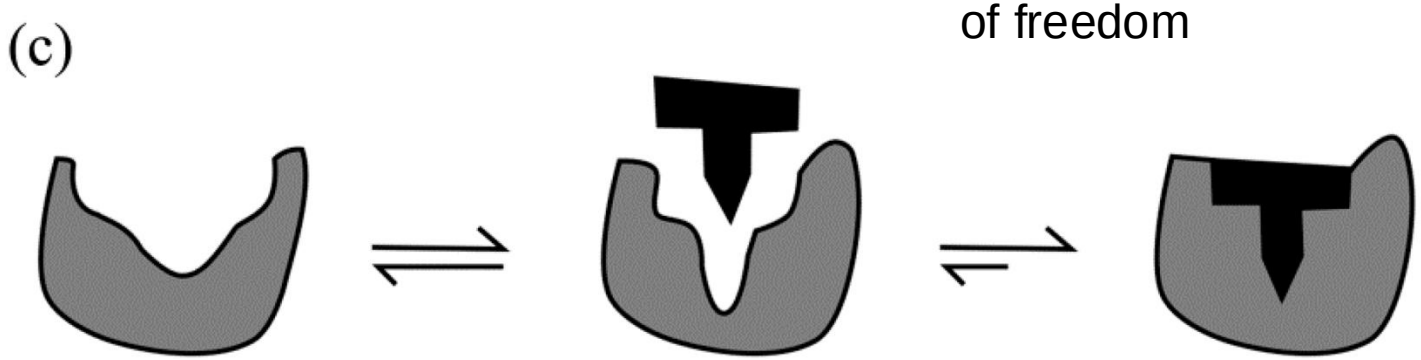
Rigid body conformational degrees of freedom

Induced fit

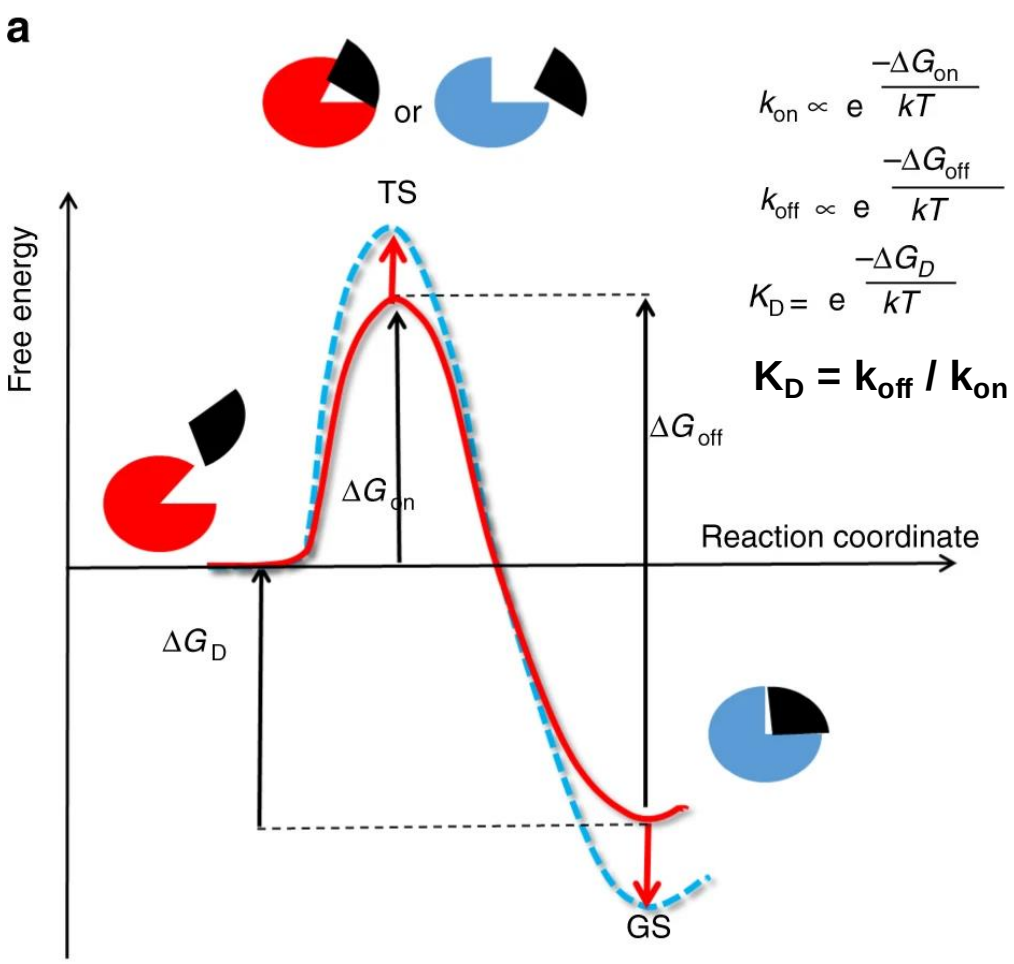


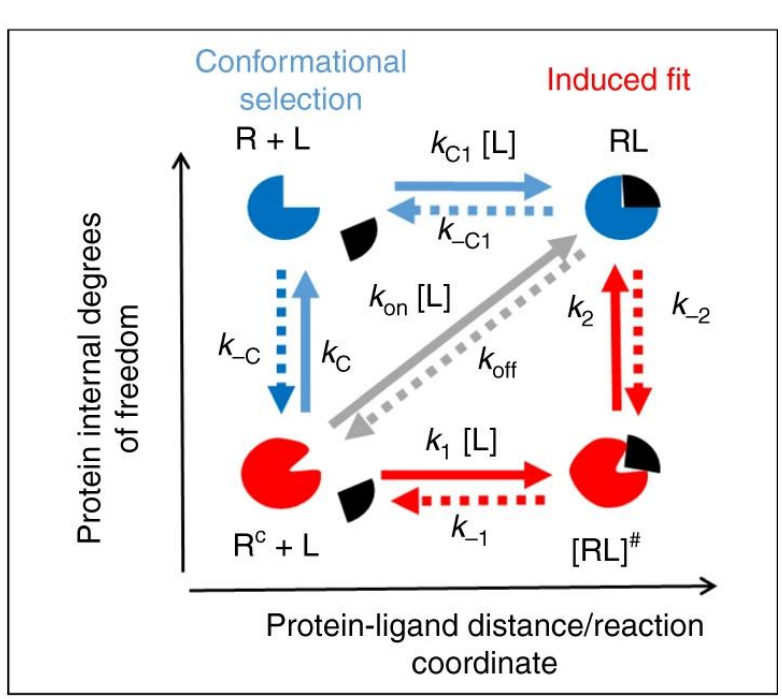
Rigid body + intraprotein conformational degrees of freedom

Conformational selection



Protein binding – a kinetic problem





Conformational selection

$$R^c + L \underset{k_{-C}}{\overset{k_C}{\rightleftarrows}} R + L \underset{k_{-C1}}{\overset{k_{C1}[L]}{\rightleftarrows}} RL$$

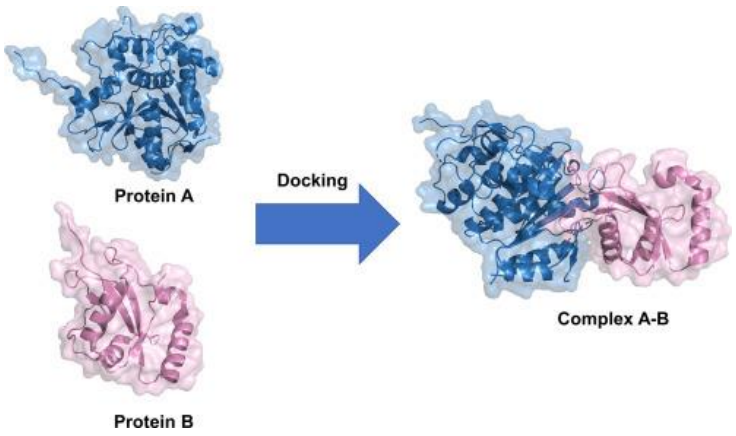
Induced fit

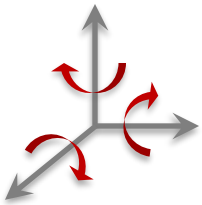
$$R^c + L \underset{k_{-1}}{\overset{k_1[L]}{\rightleftarrows}} [RL]^{\#} \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} RL$$

Experimentally observable

$$R^c + L \underset{k_{off}}{\overset{k_{on}[L]}{\rightleftarrows}} RL$$

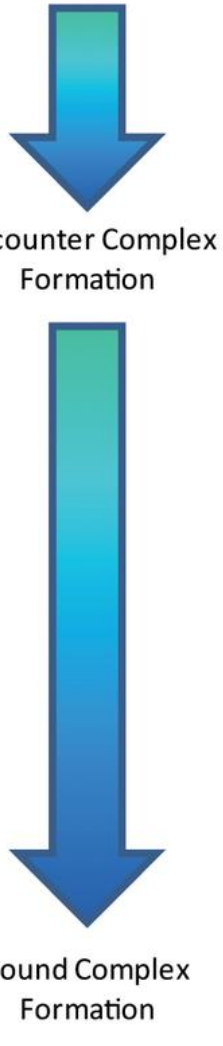
Protein docking: prediction of binding complexes

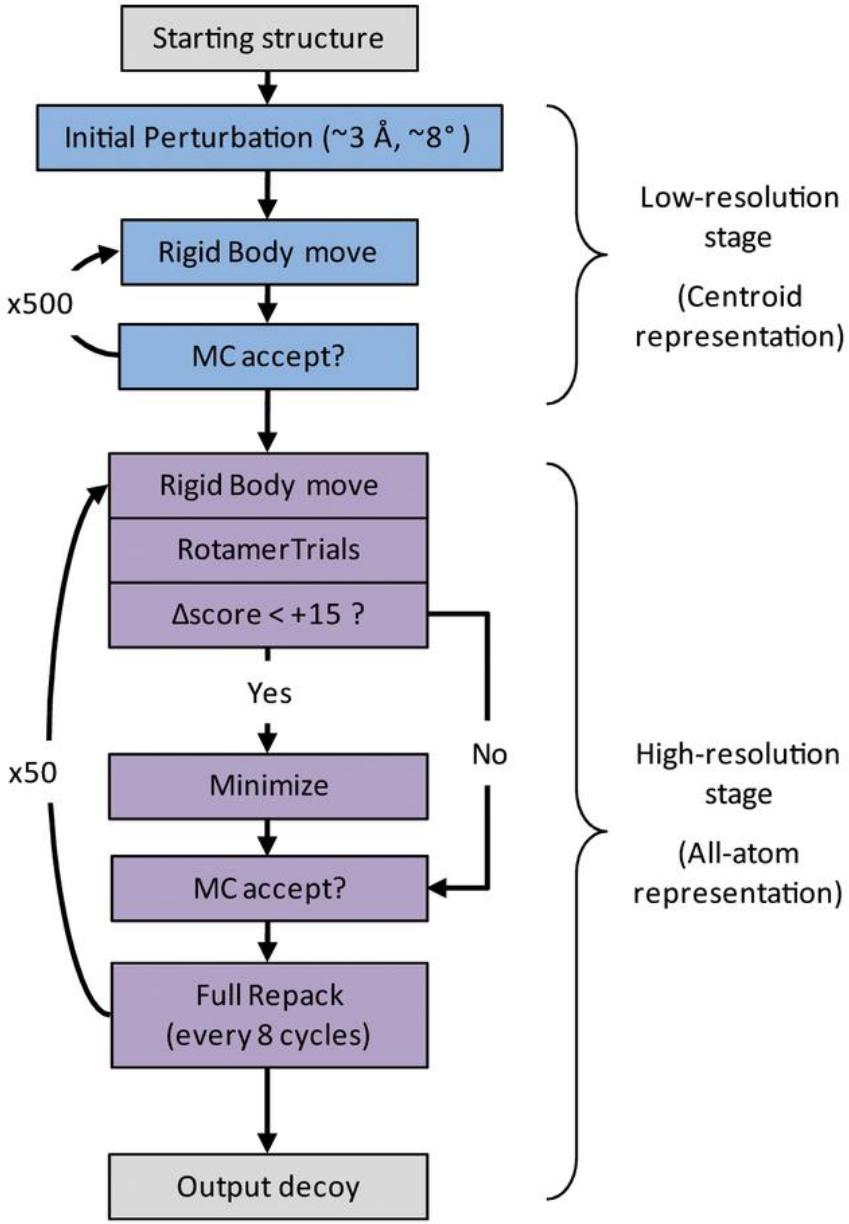




Rigid body degrees of freedom:

3 translations
3 rotations

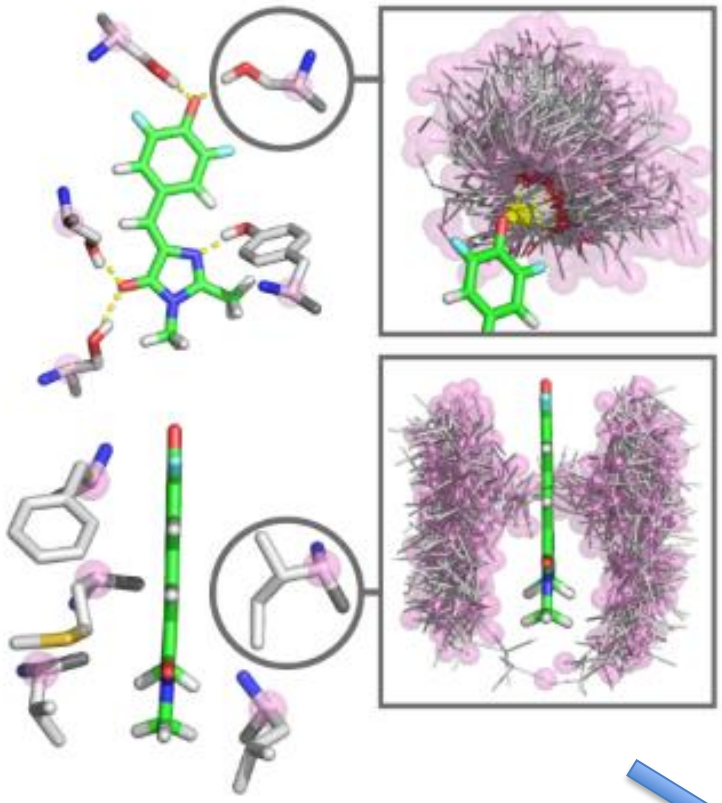




De novo design of ligand binding

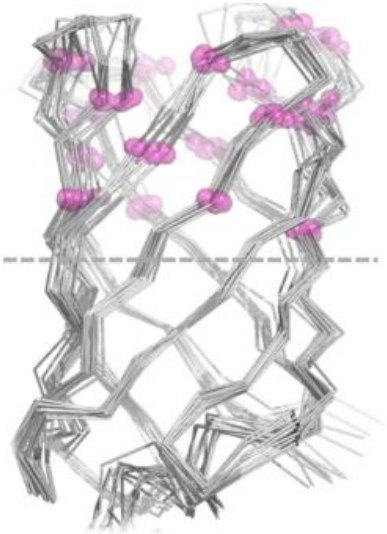
Step 1:
Pre-defined optimal binding interactions between ligand & residue side-chains

Cα atom



Step 2:
Inverse rotamer generation from ligand-side-chain contacts

Step 3:
Matching and grafting of ligand-binding residues onto a protein scaffold



Protein Design – Examples overview

1. De novo protein functional fold

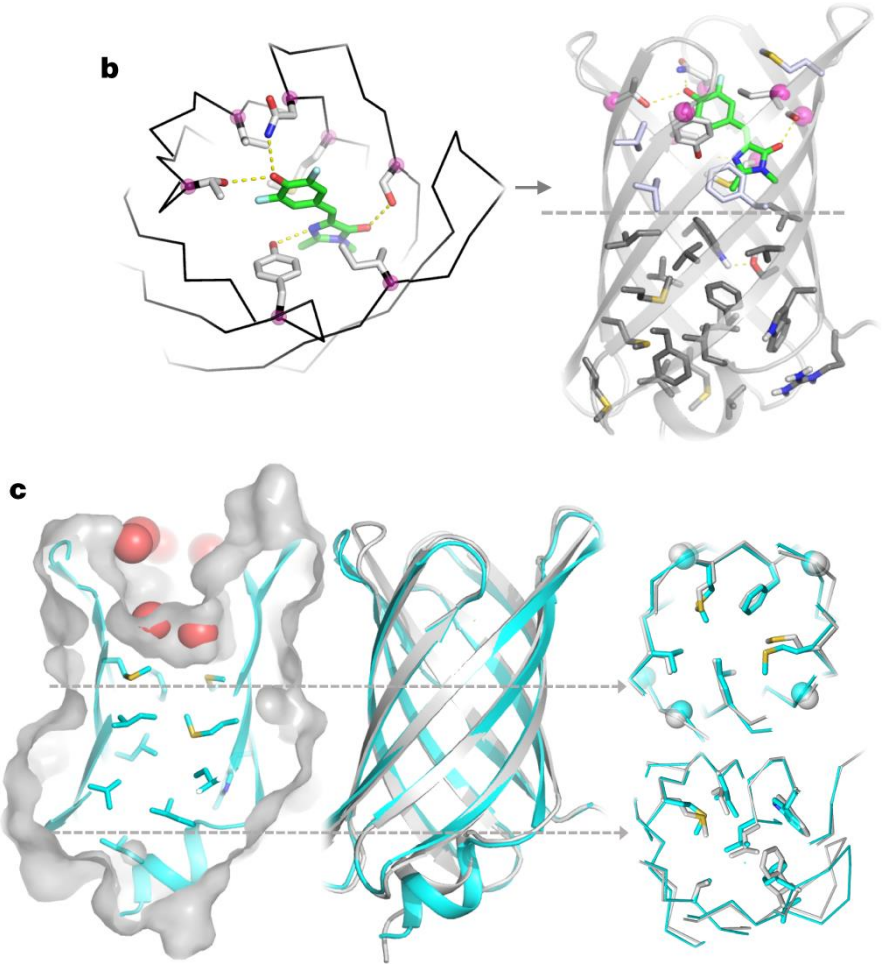
2. Enzyme

3. Ligand biosensor design

Design of new functional folds: fluorescence activating beta barrel

The beta barrel scaffold

Ideal for ligand binding

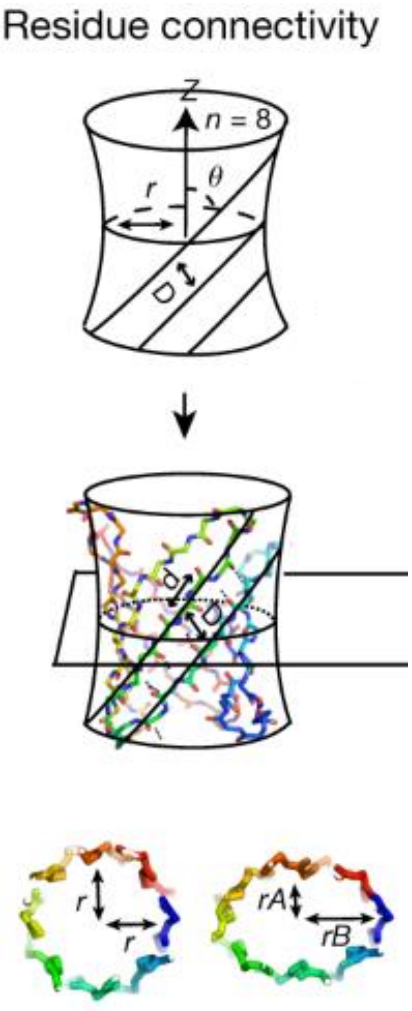


(Dou et al., Nature 2018)

Design of new functional folds: fluorescence activating beta barrel

3D ideal geometry

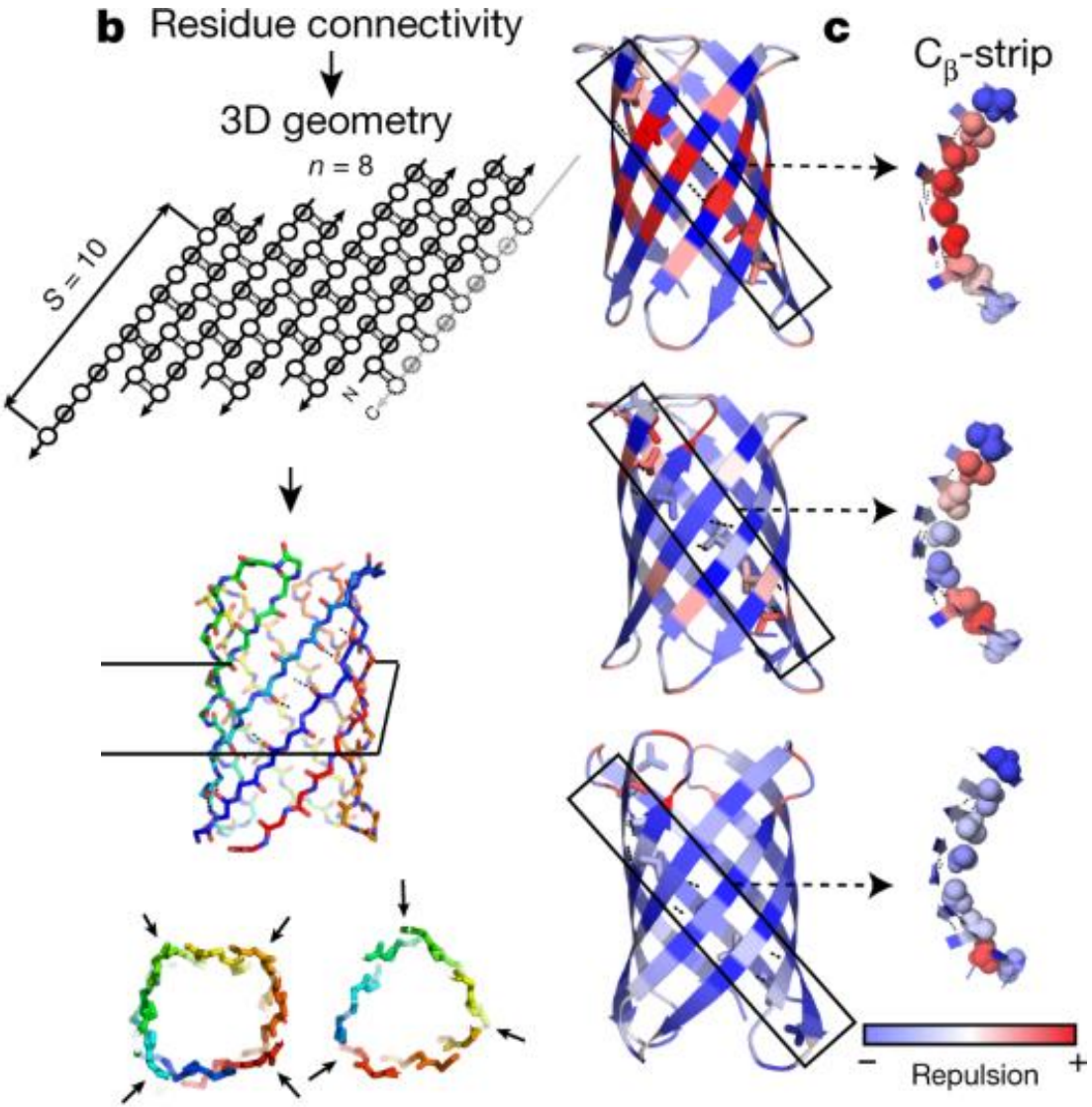
parametrization using an equation for an elliptic hyperboloid of revolution



No folded designs !

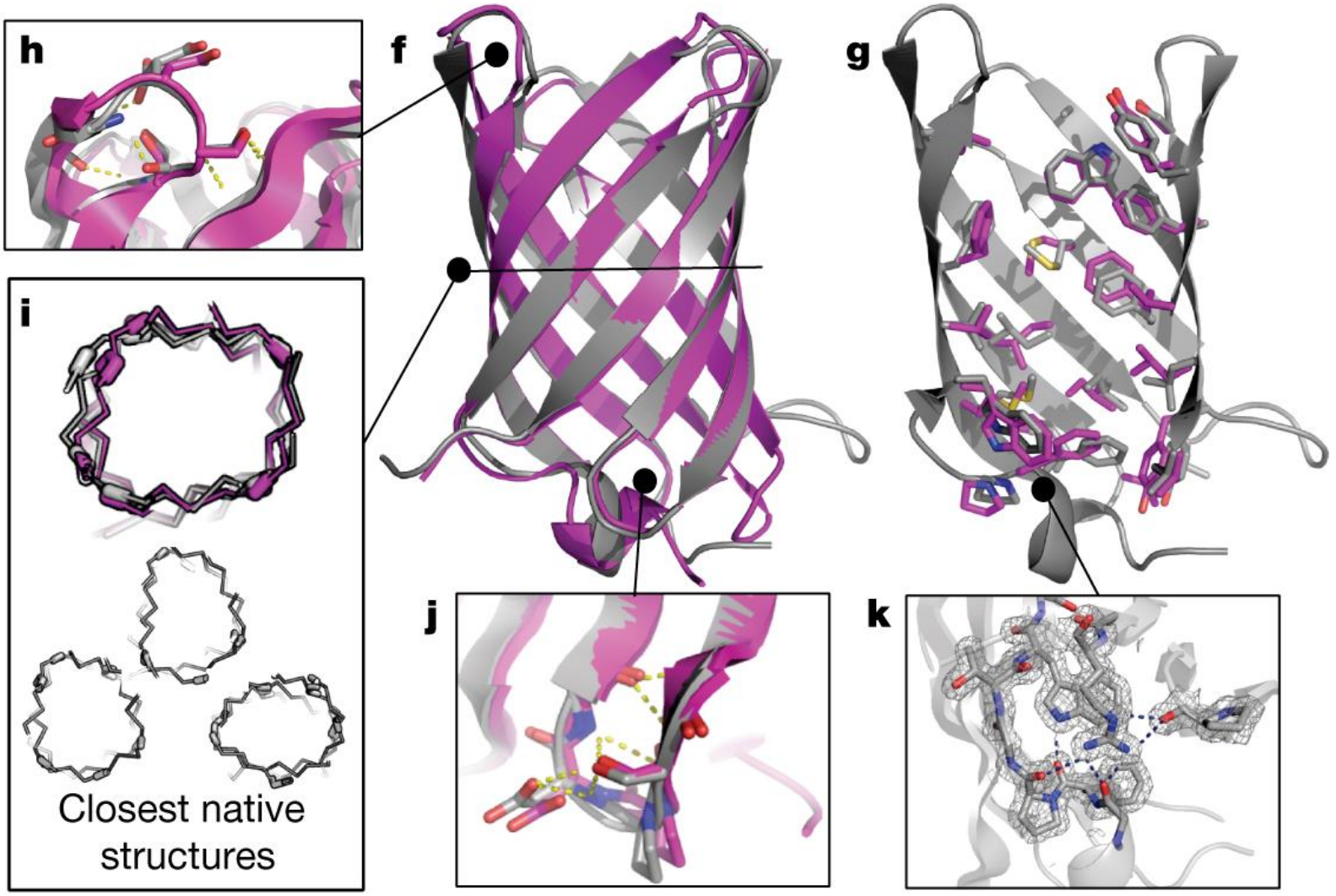
(Dou et al., Nature 2018)

Design of new functional folds: fluorescence activating beta barrel



(Dou et al., Nature 2018)

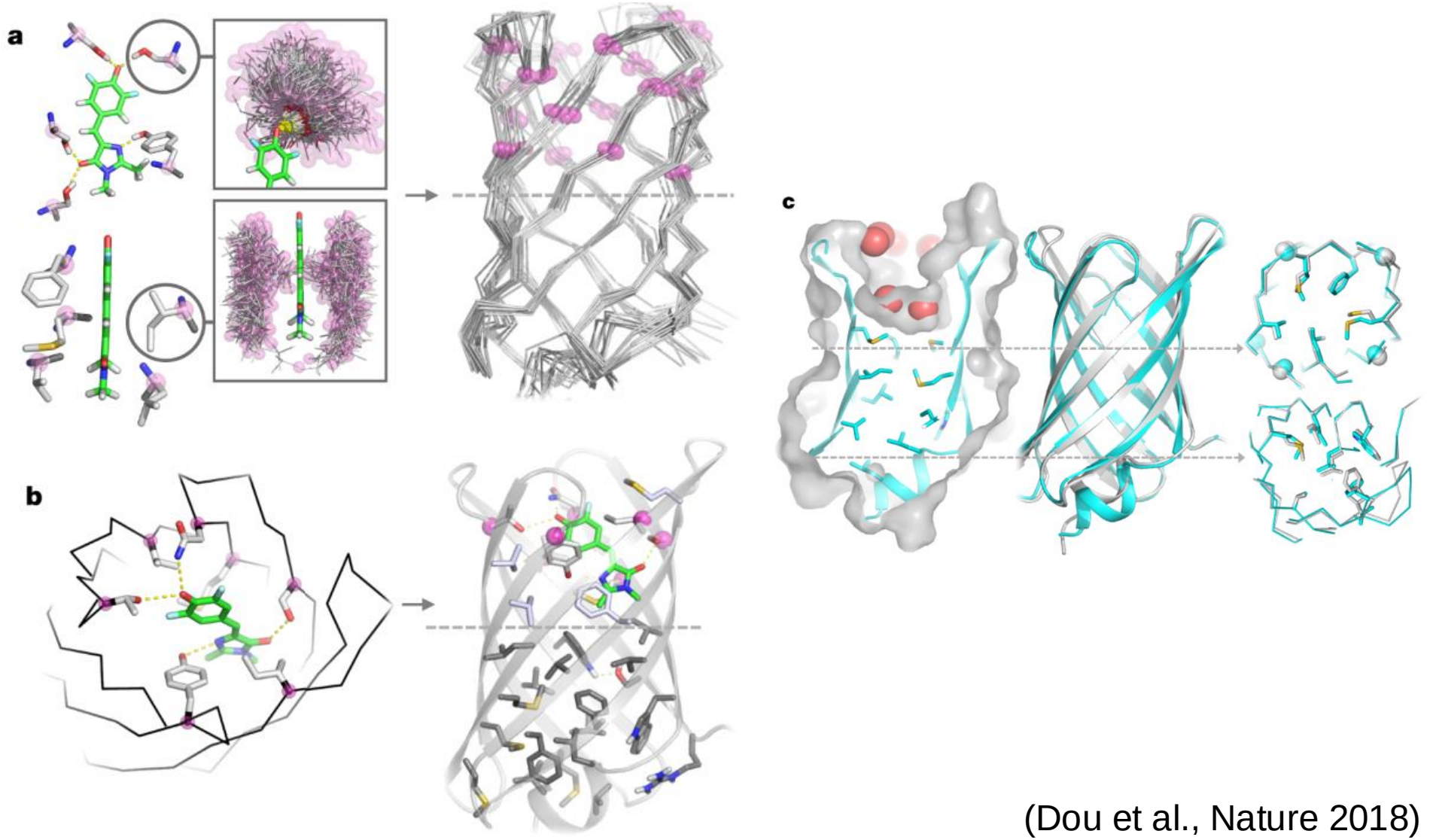
Design of new functional folds: fluorescence activating beta barrel



(Dou et al., Nature 2018)

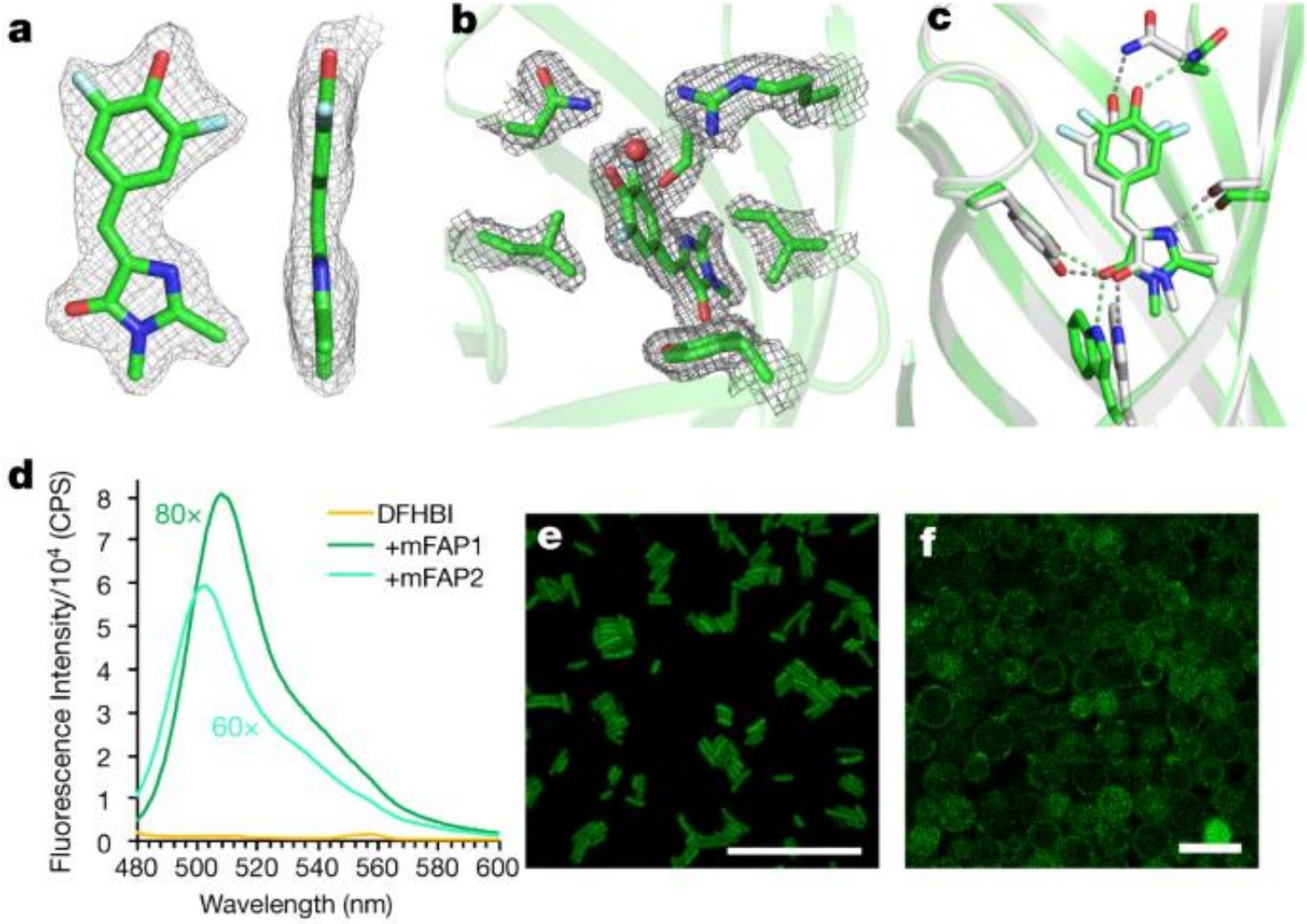
Design of new functional folds

Barrel design with ligand binding cavity



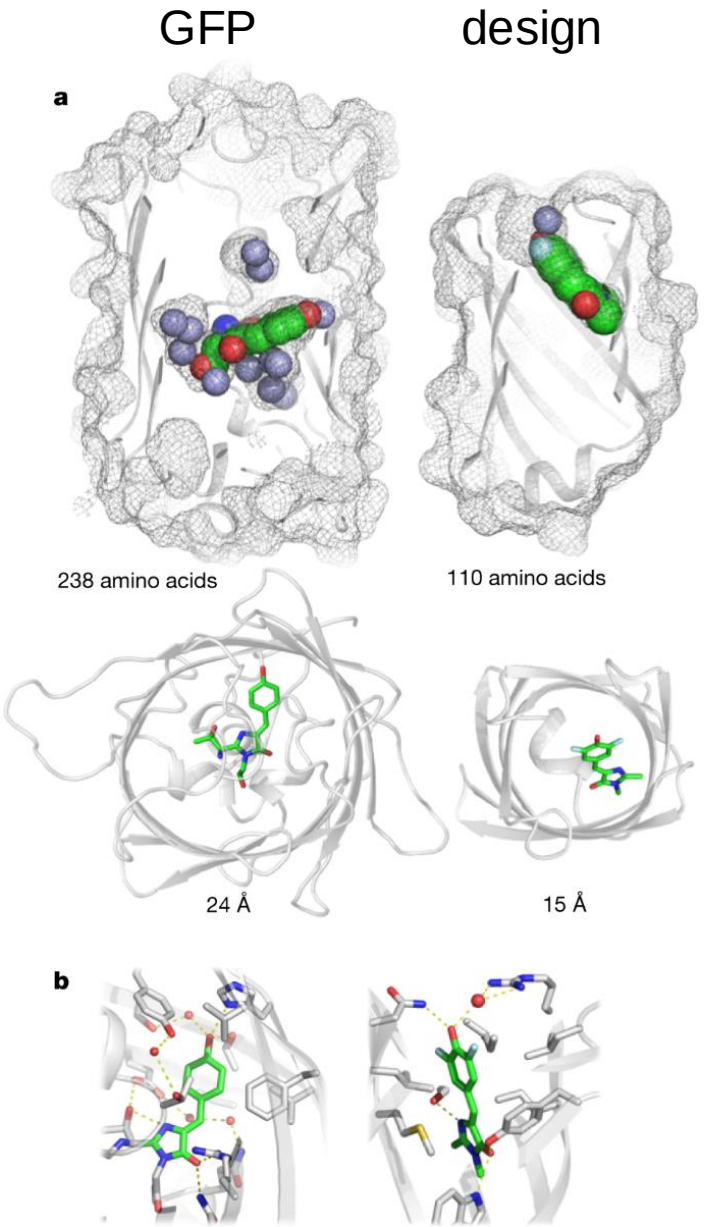
(Dou et al., Nature 2018)

Design of new functional folds



Design of new functional folds

Comparison with naturally evolved GFP



(Dou et al., Nature 2018)

Take home messages

1. De novo design of beta barrels is challenging

2. Are the Gly kink bulge rules universal and applicable for all barrels?

3. Precise ligand binding is possible when starting from a hyperstable and rigid scaffold: destabilizing ligand binding cavity carved into a large hydrophobic stabilizing core

(Dou et al., Nature 2018)

Protein Design – Examples overview

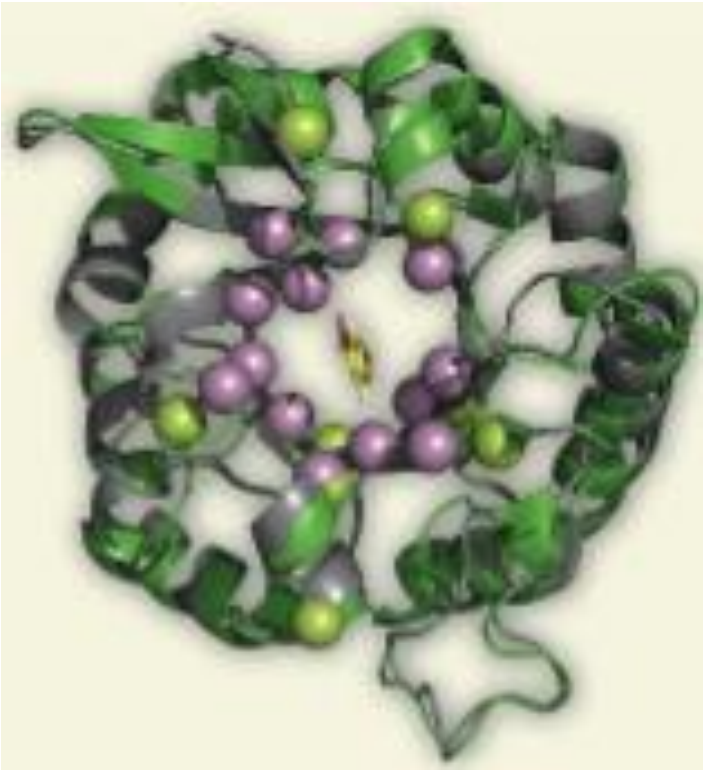
1. De novo protein functional fold

2. Enzyme

3. Ligand biosensor design

Design of a novel enzyme

- Goal: design artificial enzymes that catalyze unnatural reactions
- Enzymes:
  - lower the activation barrier, by
  - stabilizing transition state
  - shielding reactants

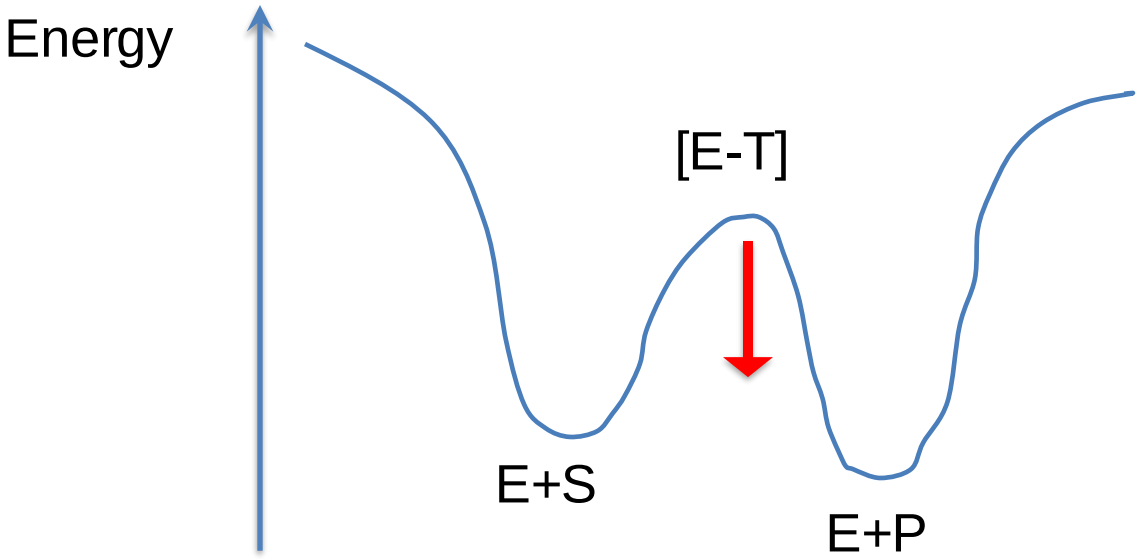


Roethlisberger et al. 2008; Liang et al., 2008

Design of a novel enzyme

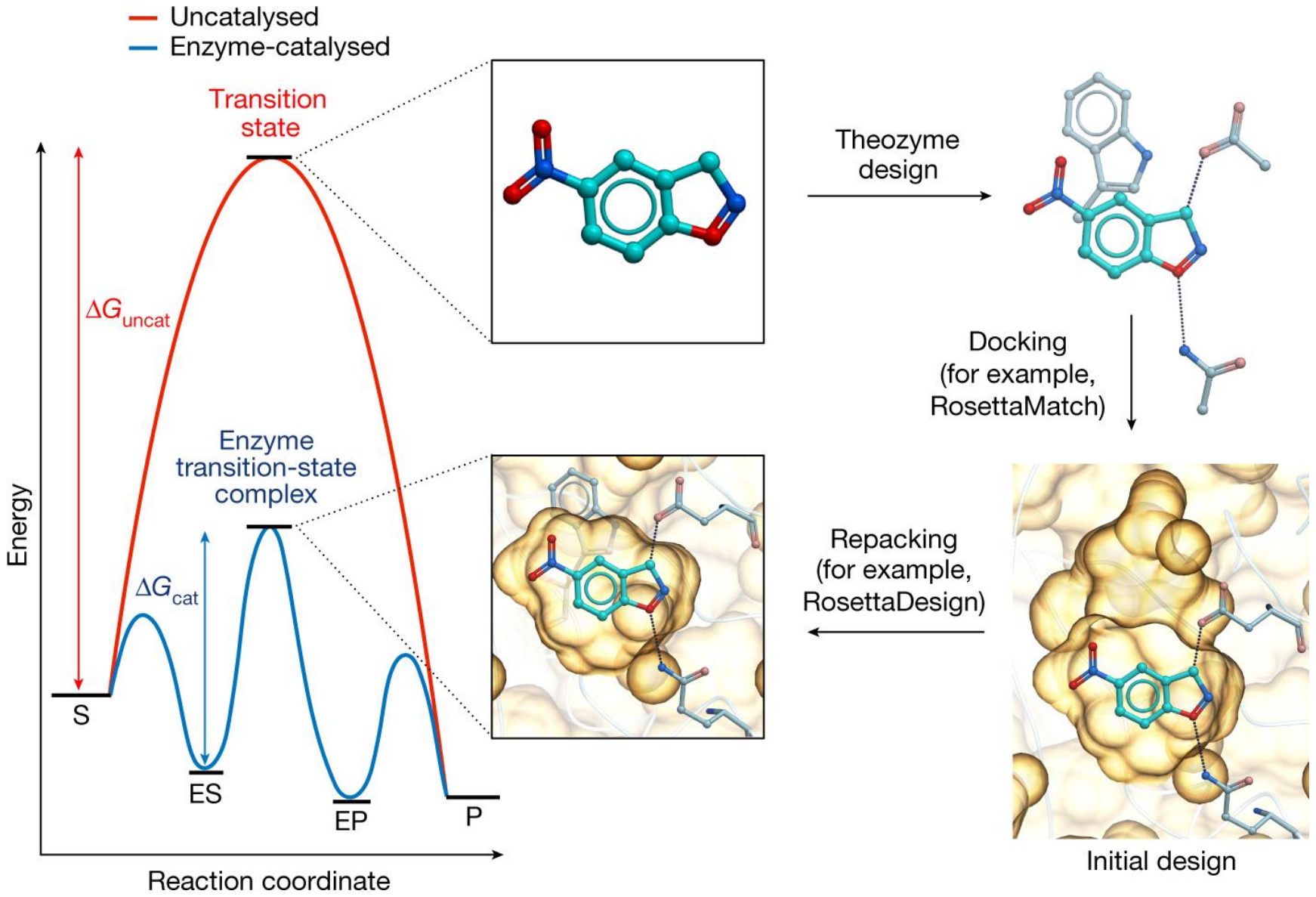
Roethlisberger et al. 2008; Liang et al., 2008

Enzyme + Substrate --> [E-T] --> Enzyme + Product



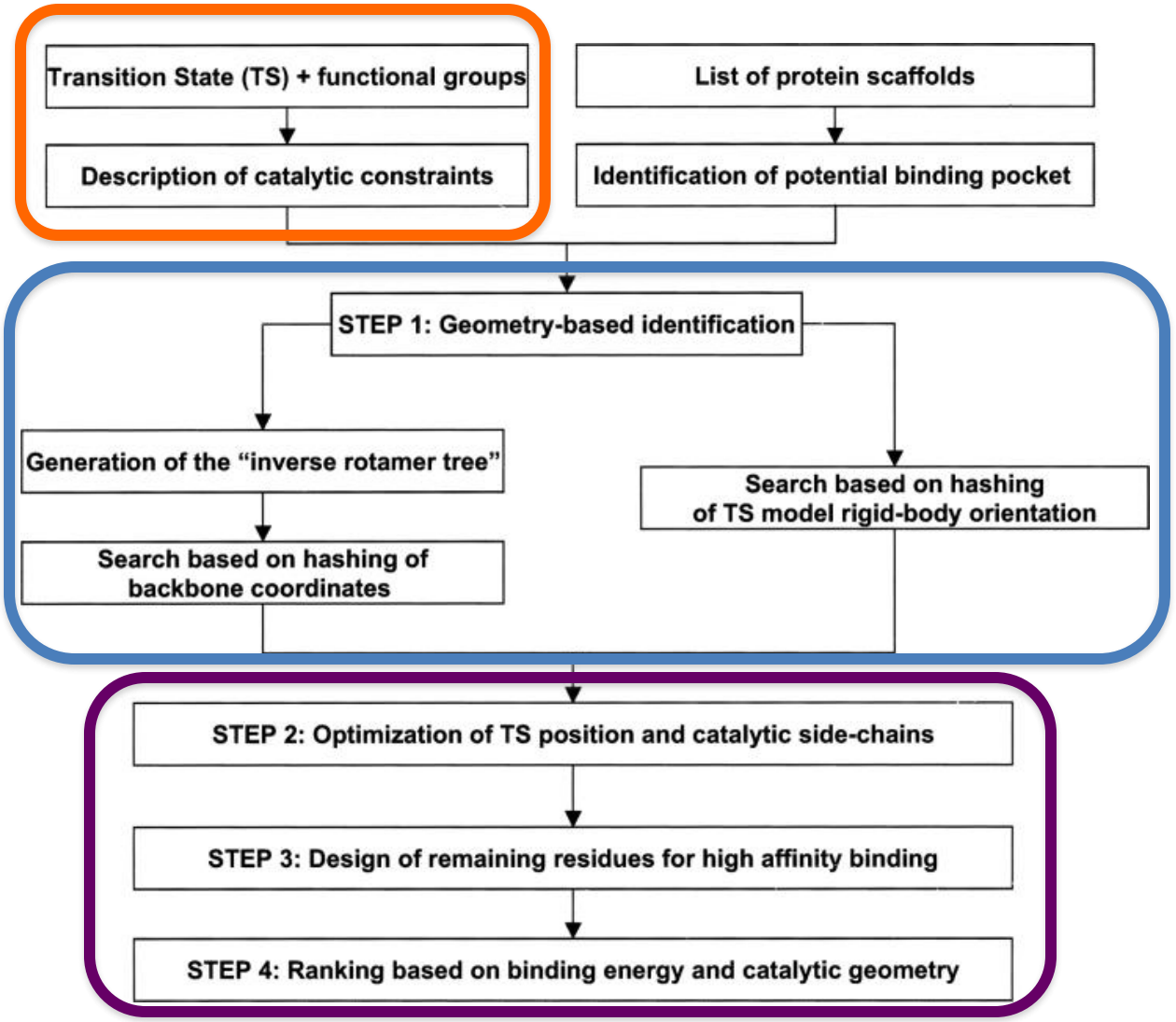
Design of a novel enzyme

Roethlisberger et al. 2008; Liang et al., 2008



Design of a novel enzyme

Roethlisberger et al. 2008; Liang et al., 2008



Approach:

1. Model transition state of reaction (QM)
2. Stabilize with carefully placed chemical groups around it

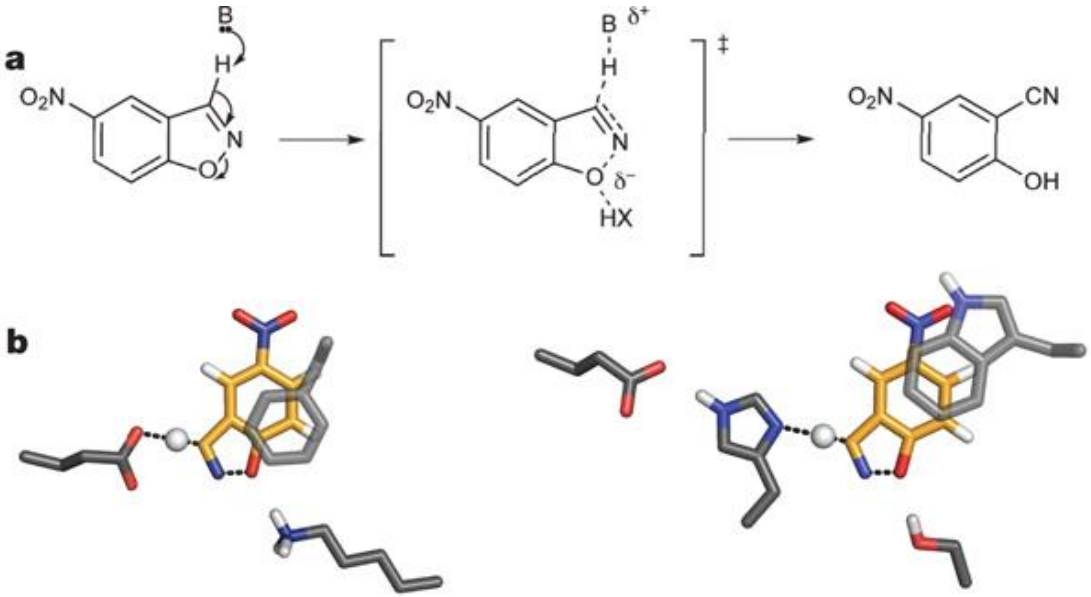
3. Graft resulting active site into an existing protein

4. Alter the sequence of the protein to accommodate the active site

Model transition state

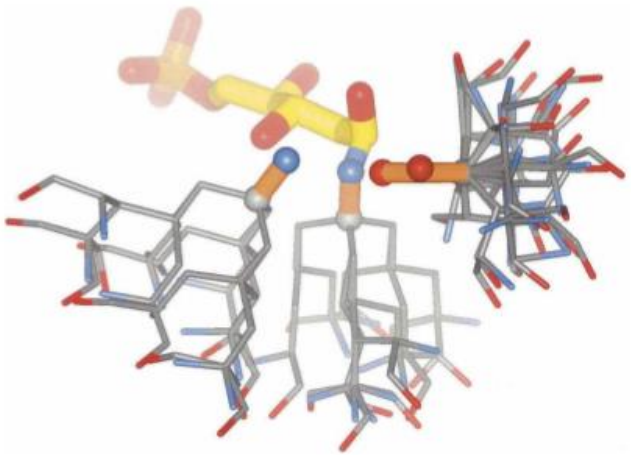
- Kemp Elimination
- Water mediated

Model transition state



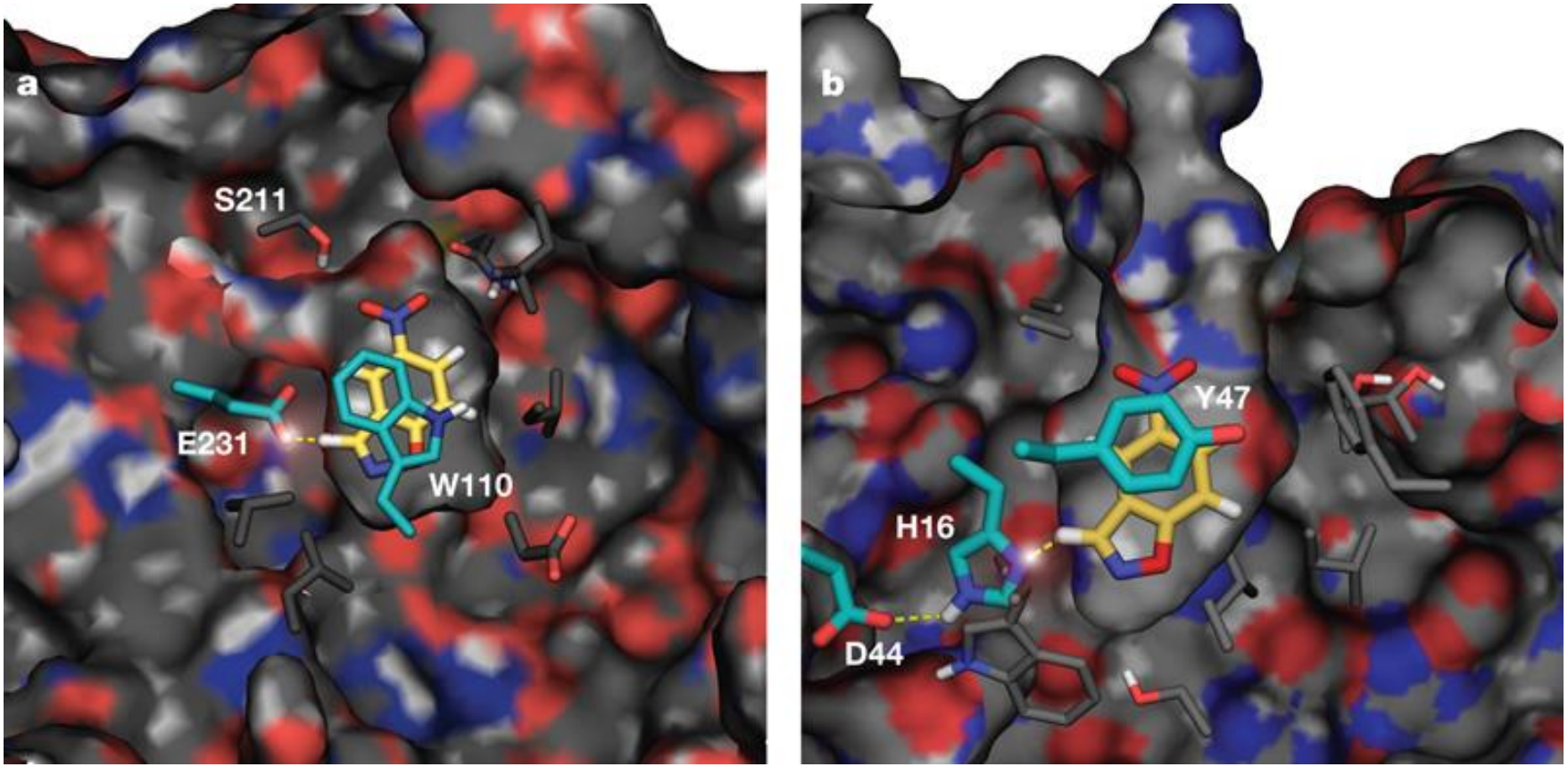
Search for Template

- Inside-out:
  - Build inverse rotamer tree starting from catalytic site
  - Search for fitting backbone templates (geometric hashing)

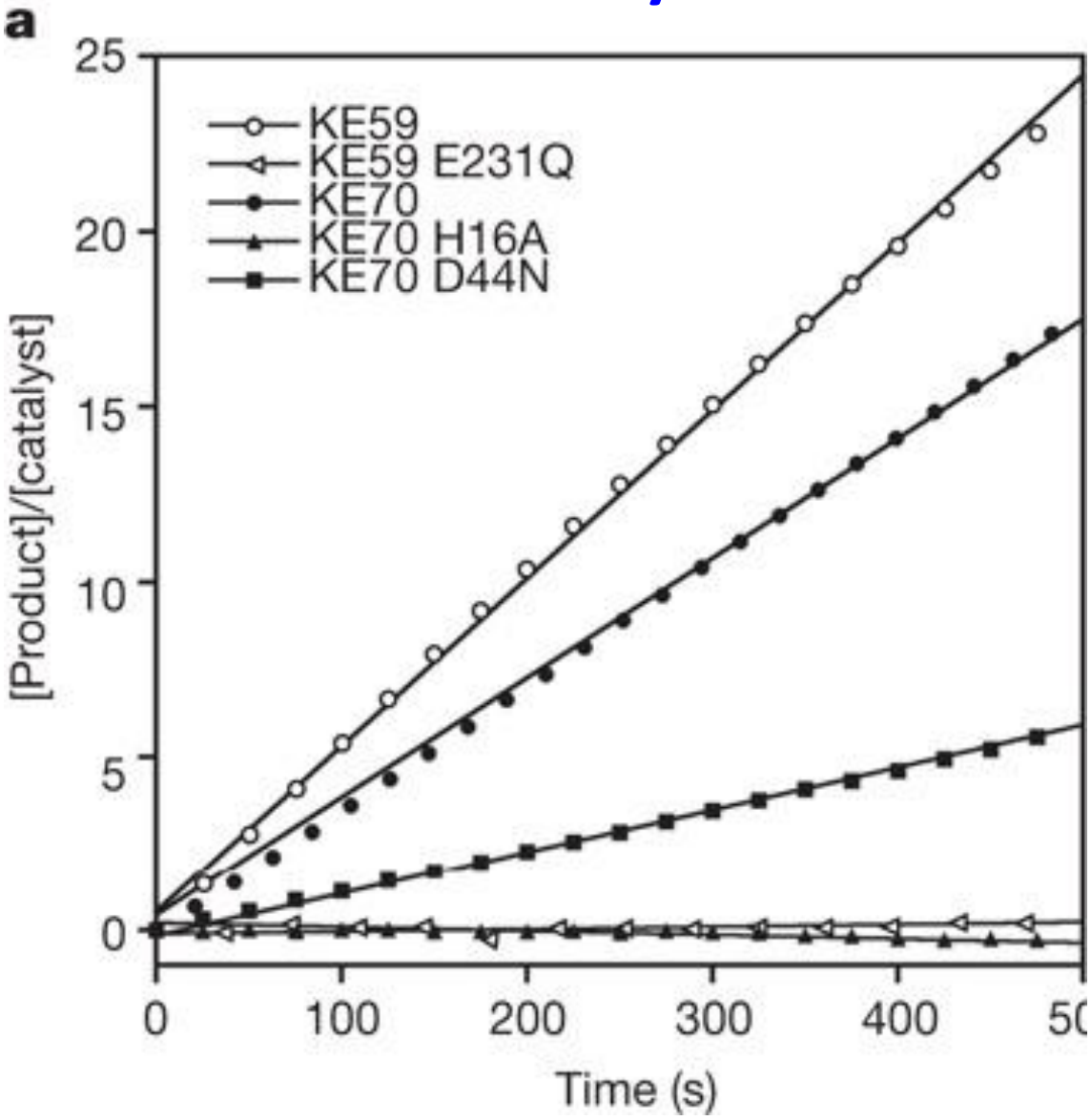


- RosettaMatch: Outside-in:
  - Place side chains and transition state model at each position
  - Search for transition state model orientations that fit several positions

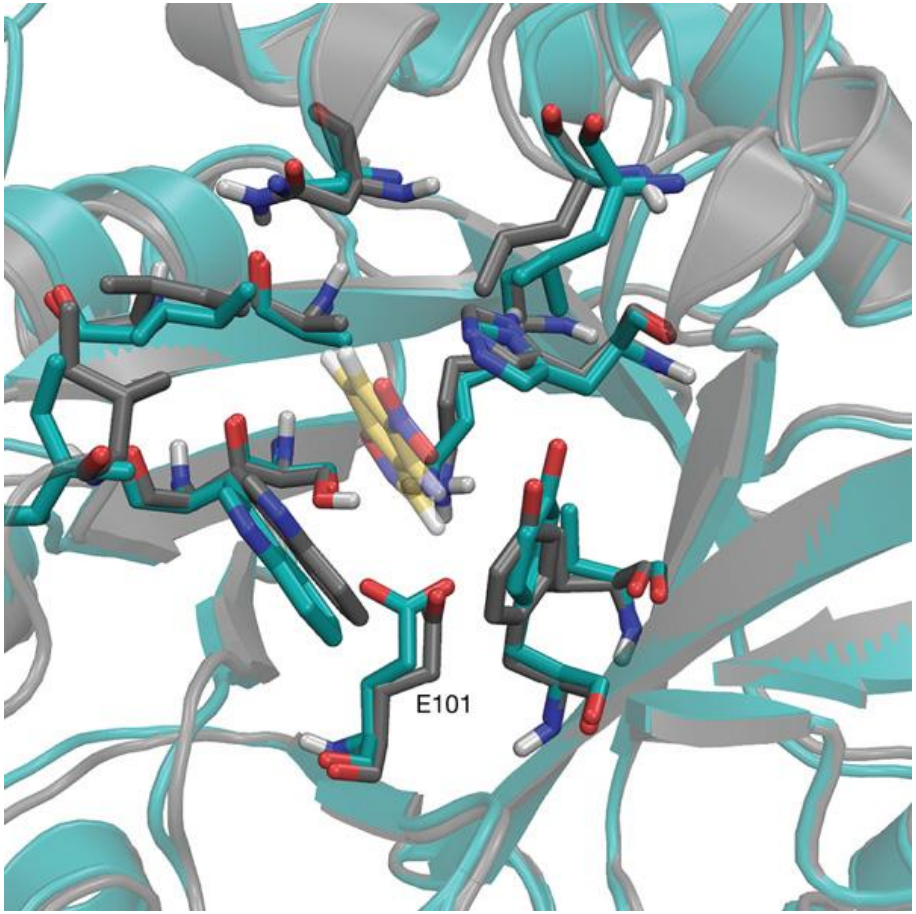
Find match



Validation 1: enzyme is active



Validation 2: accurate structure prediction



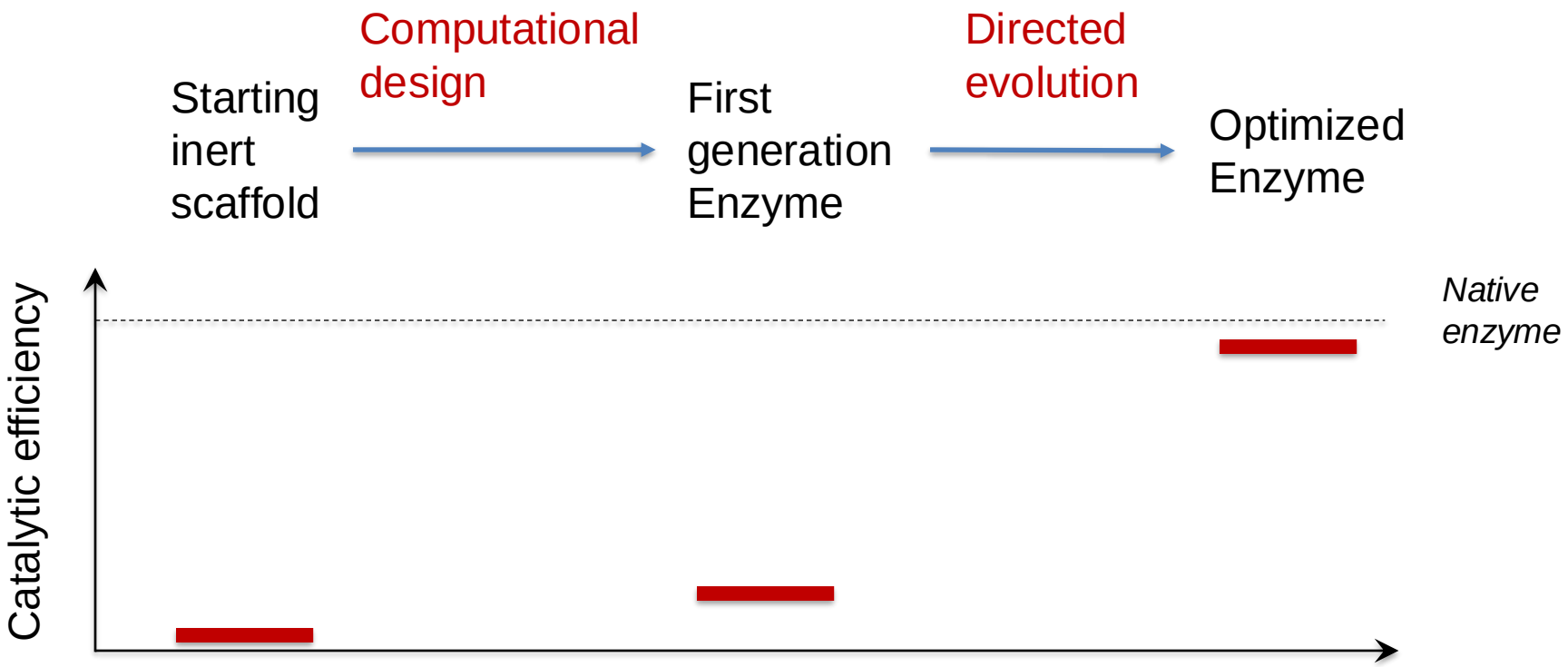
Is this all?

Is the enzyme design problem solved?

Designed HG3: k_{cat} ~ 1 s^{-1}

Native kemp eliminases: k_{cat} ~ 400-500 s^{-1}

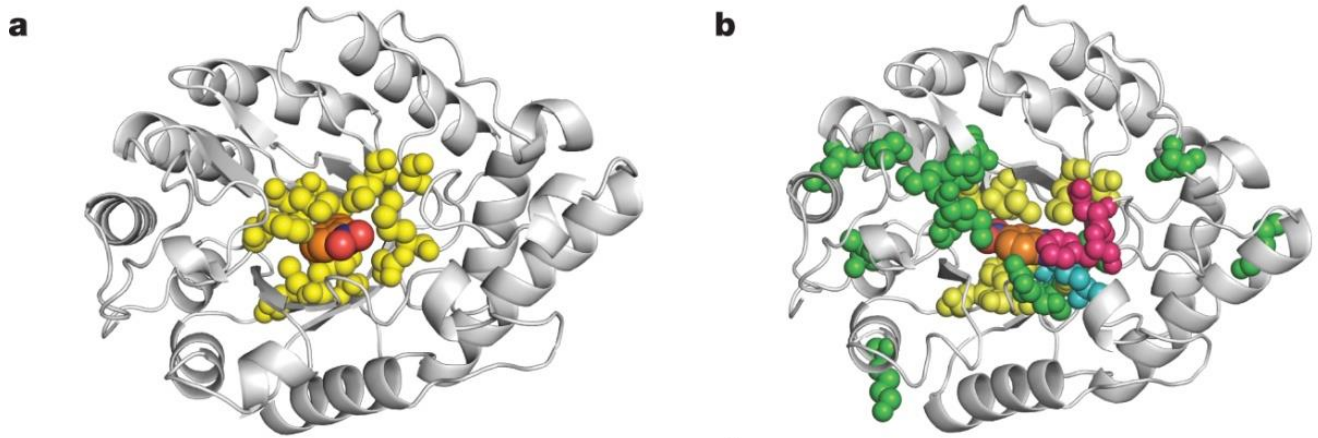
Precision is essential for efficient catalysis in an evolved Kemp eliminase (Blomberg, Nature 2014)



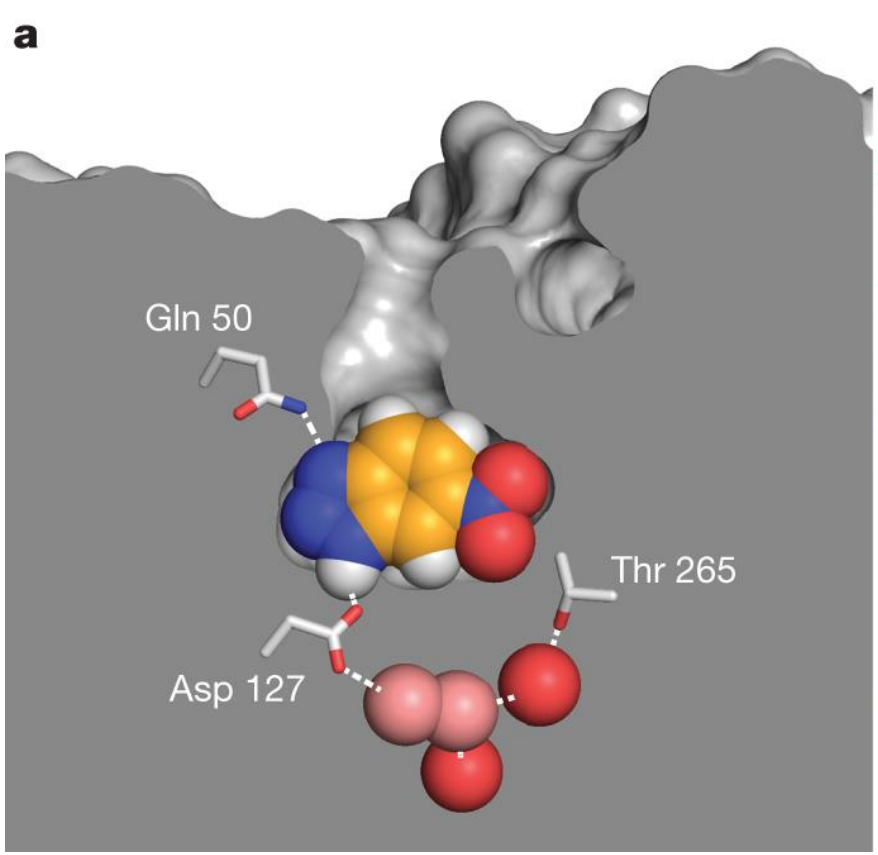
Directed evolution of Kemp eliminase HG3 (Blomberg, Nature 2014)

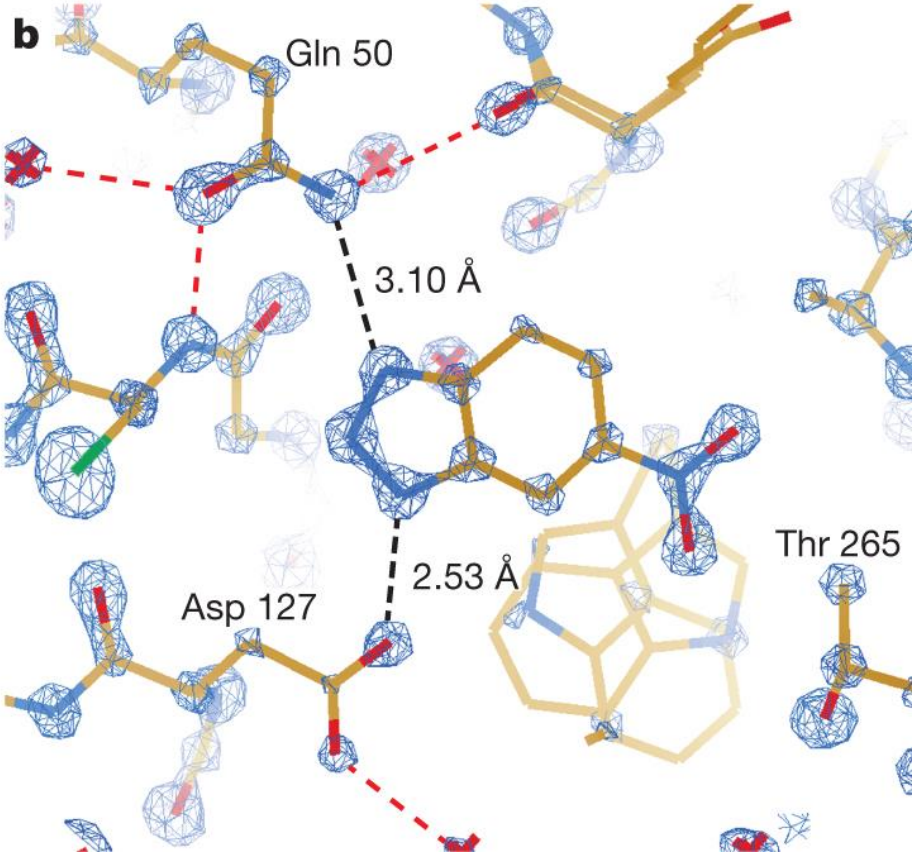
Computational design: HG3

Directed evolution: HG3.17



Crystal structure of HG3.17 (Blomberg, Nature 2014)





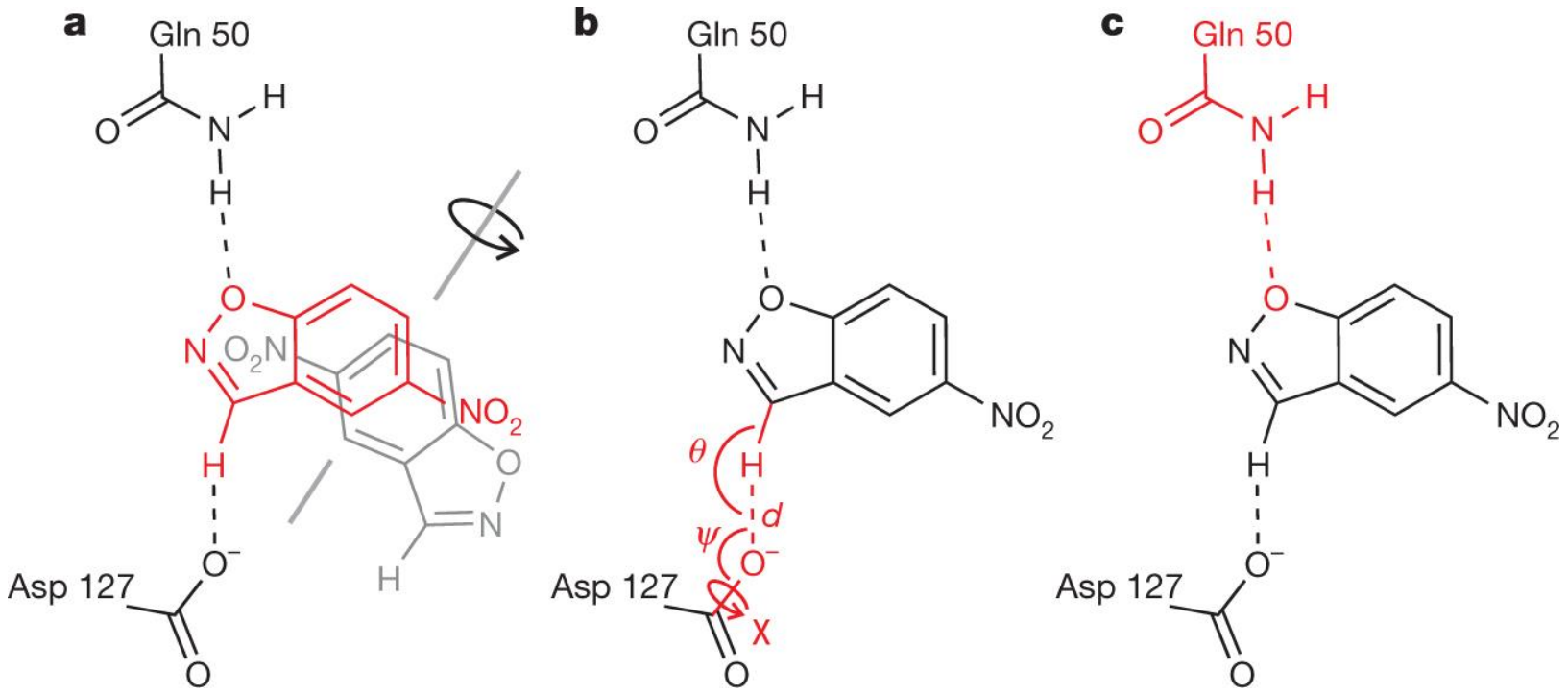
Catalytic improvement of HG3

(Blomberg, Nature 2014)

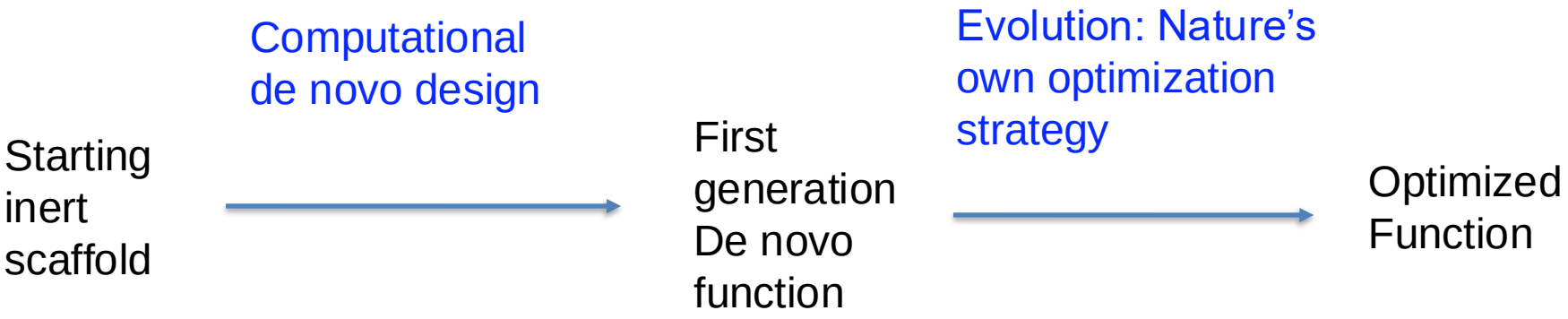
Elimination of a on unproductive binding mode

Efficient proton transfer

Optimal TS stabilization



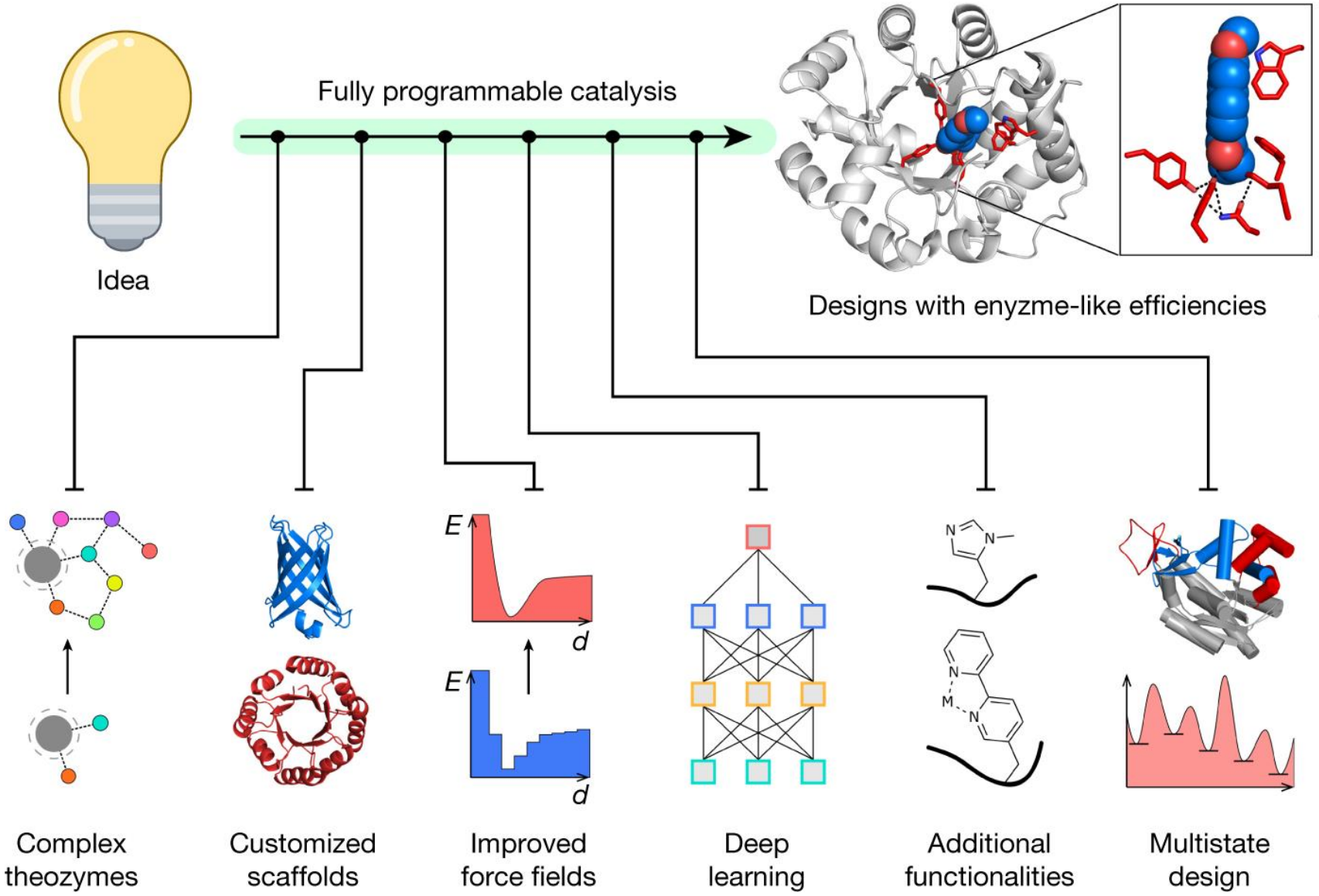
Take home messages



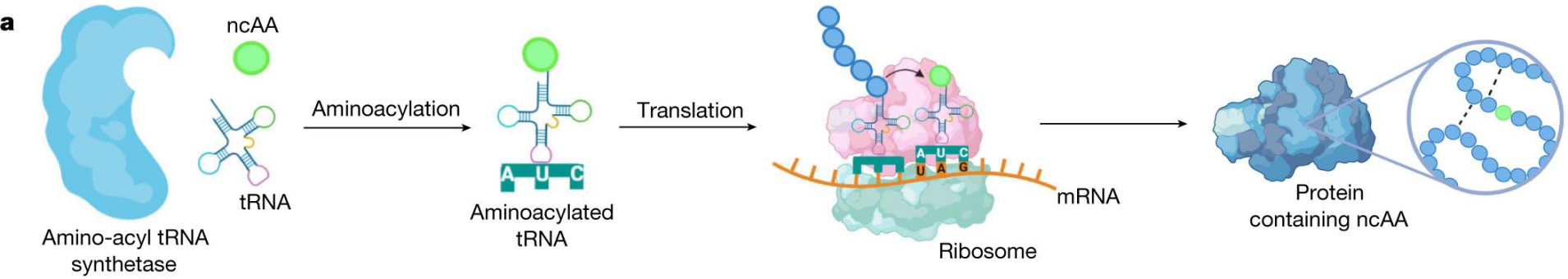
Limited space of de novo functions!

Time consuming and costly !

A road map to better enzymes



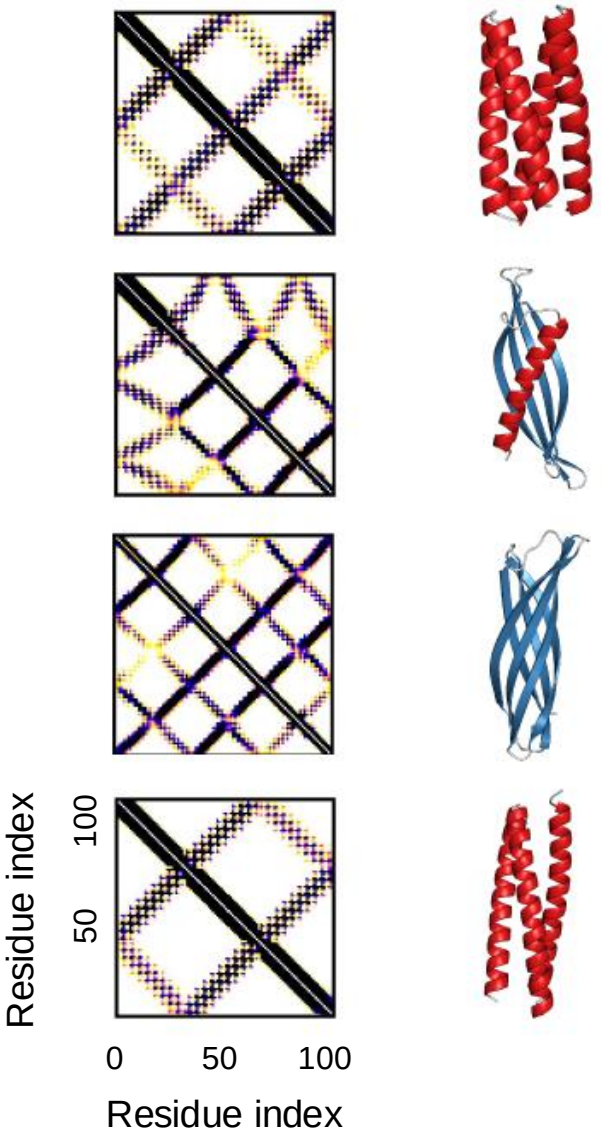
One future direction: non canonical AA for expanded chemical space



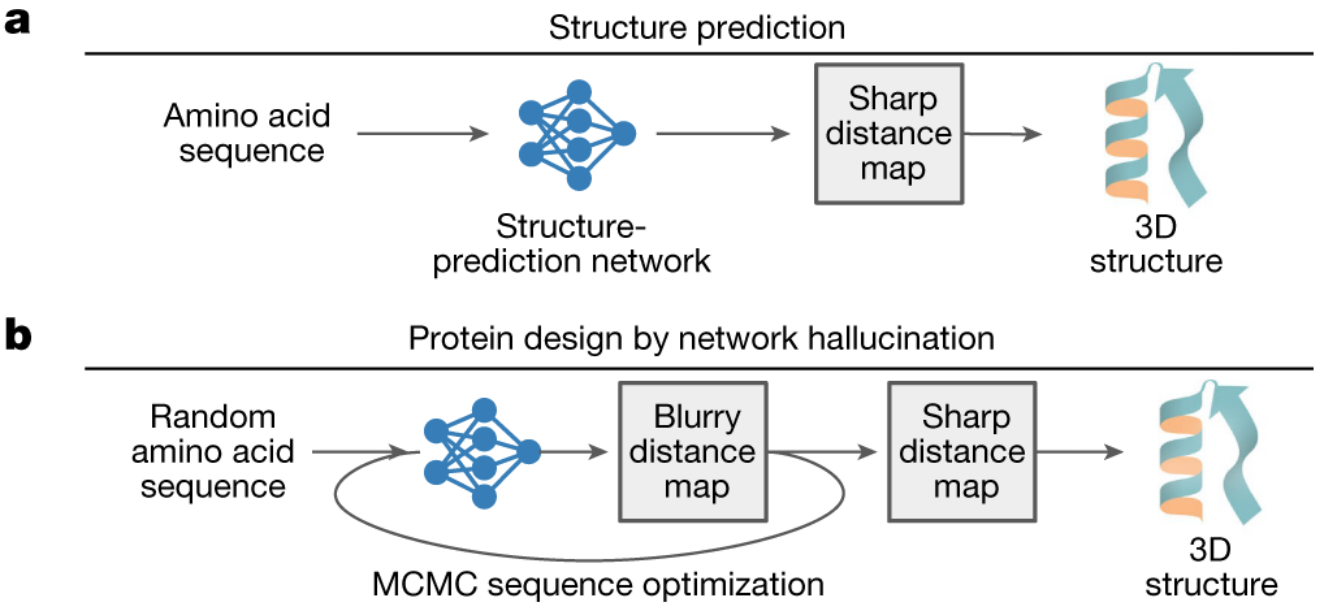
De novo design of luciferases using deep learning

Each protein structure
is characterized
by a matrix of
residue-residue contacts

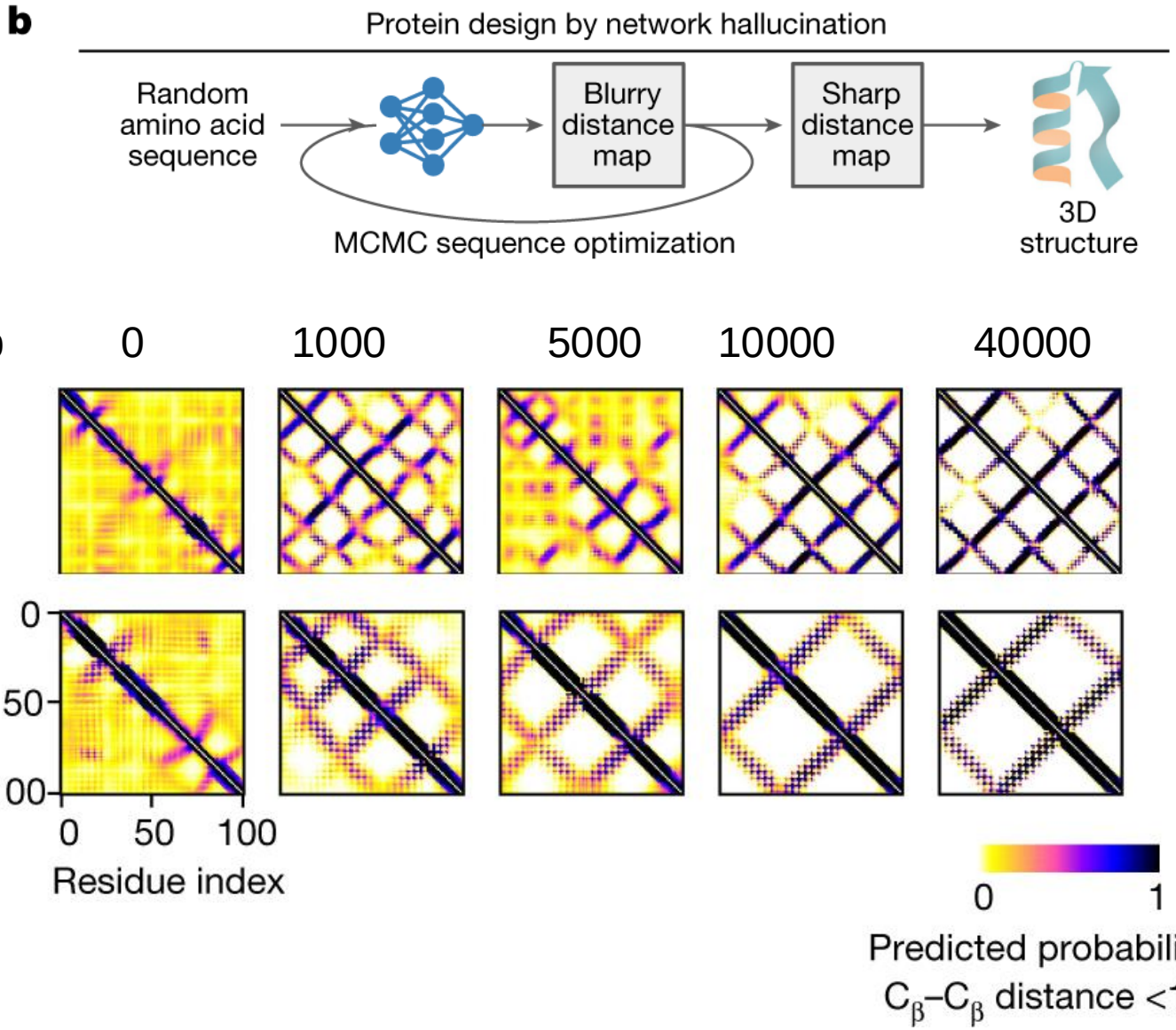
Neural networks learn
these contact patterns



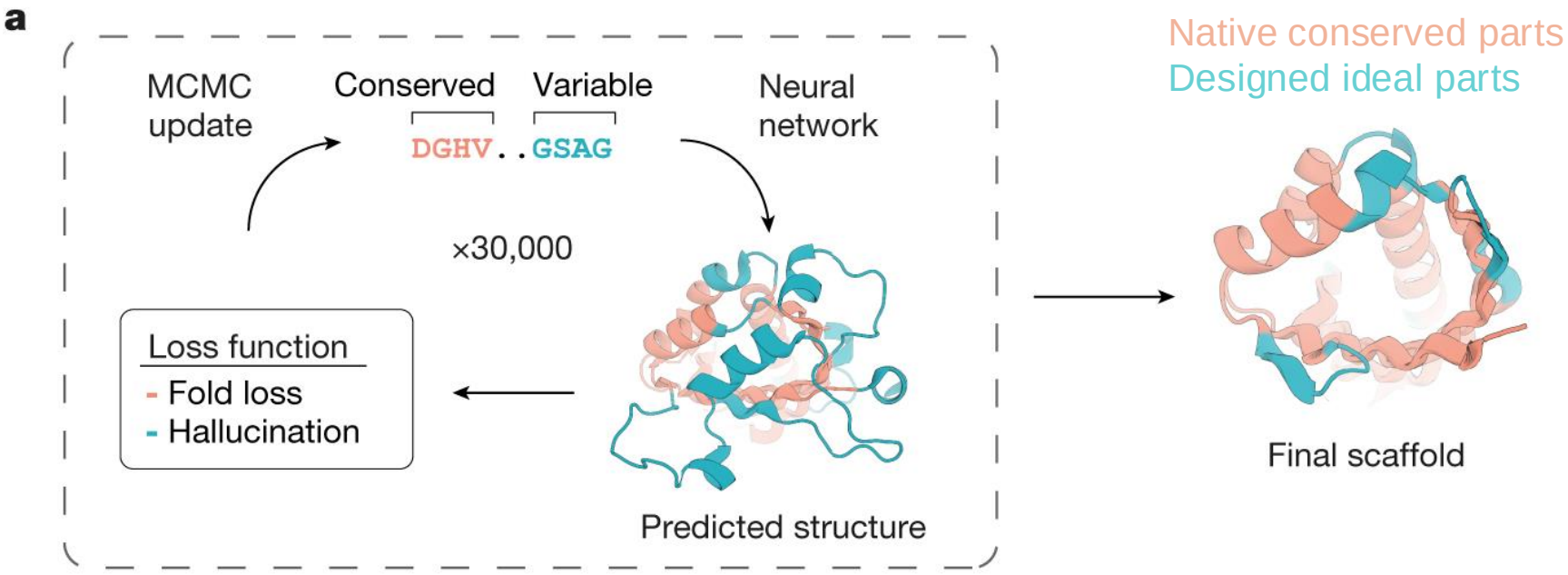
De novo design of luciferases using deep learning



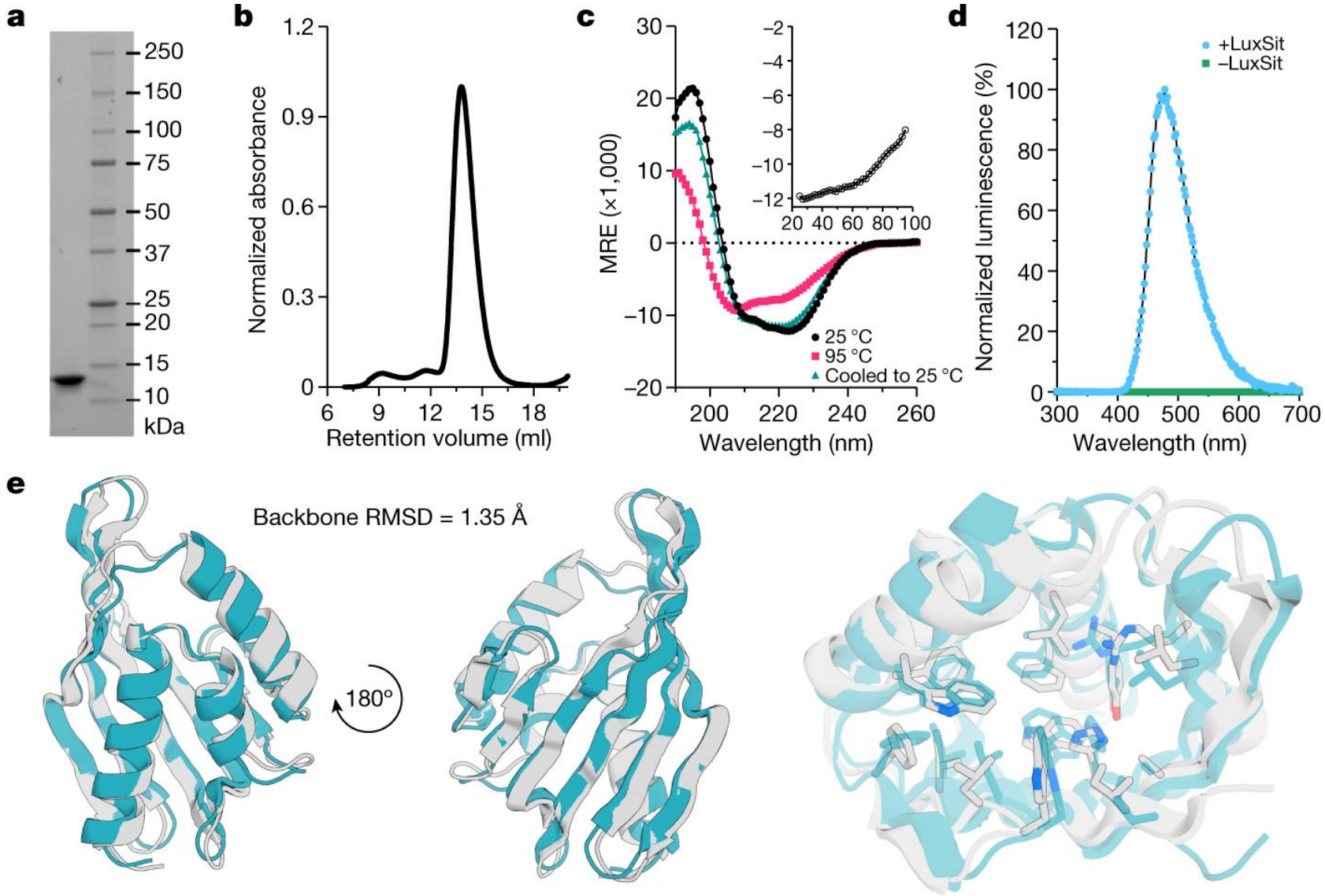
De novo design of luciferases using deep learning



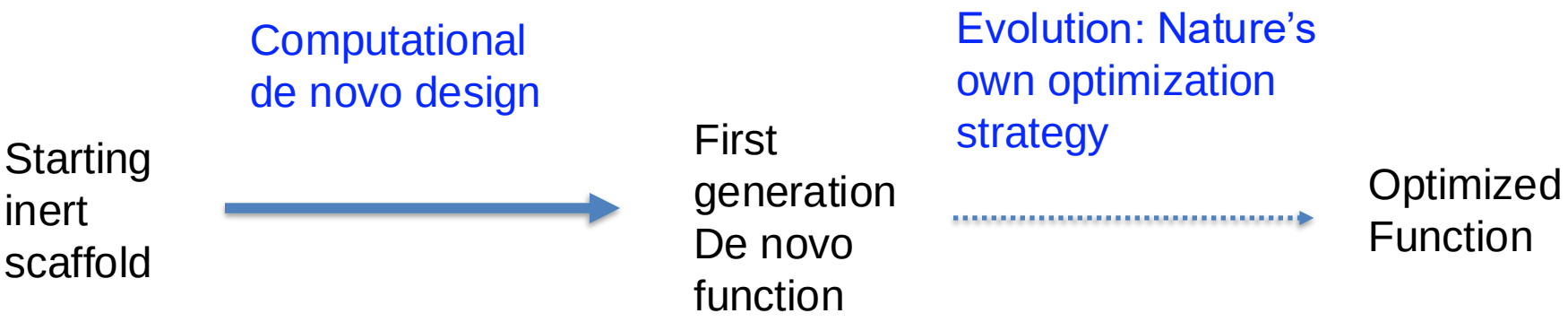
De novo design of luciferases using deep learning



De novo design of luciferases using deep learning



Take home messages



Deep learning

Artificial chemistry

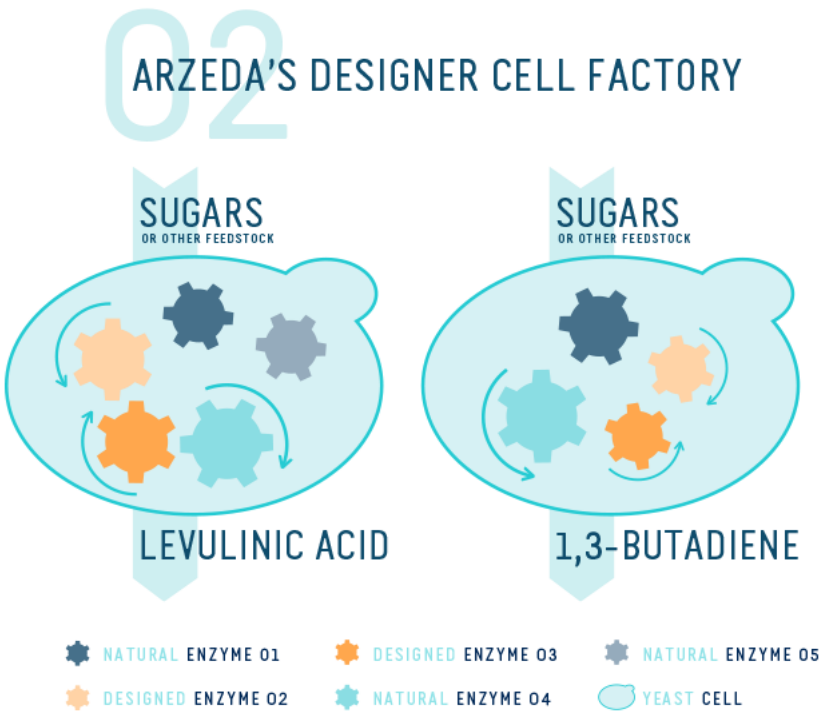
Many others

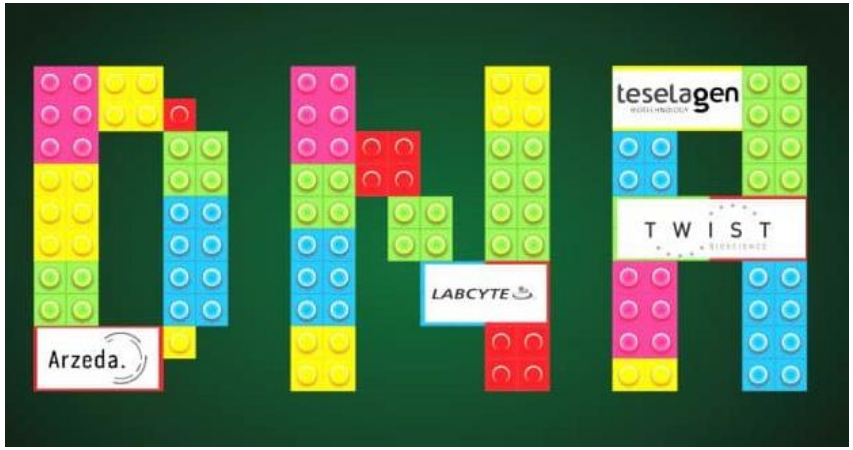
Enzyme Design for a more sustainable world

Arzeda.
THE PROTEIN DESIGN COMPANY™

Arzeda is harnessing the power of computational protein design to build novel enzymes and discover new pathways that enable cost effective, sustainable production of value-added specialty chemicals and ingredients

bp





Protein Design – Examples overview

1. De novo protein functional fold

2. Enzyme

3. Ligand biosensor design

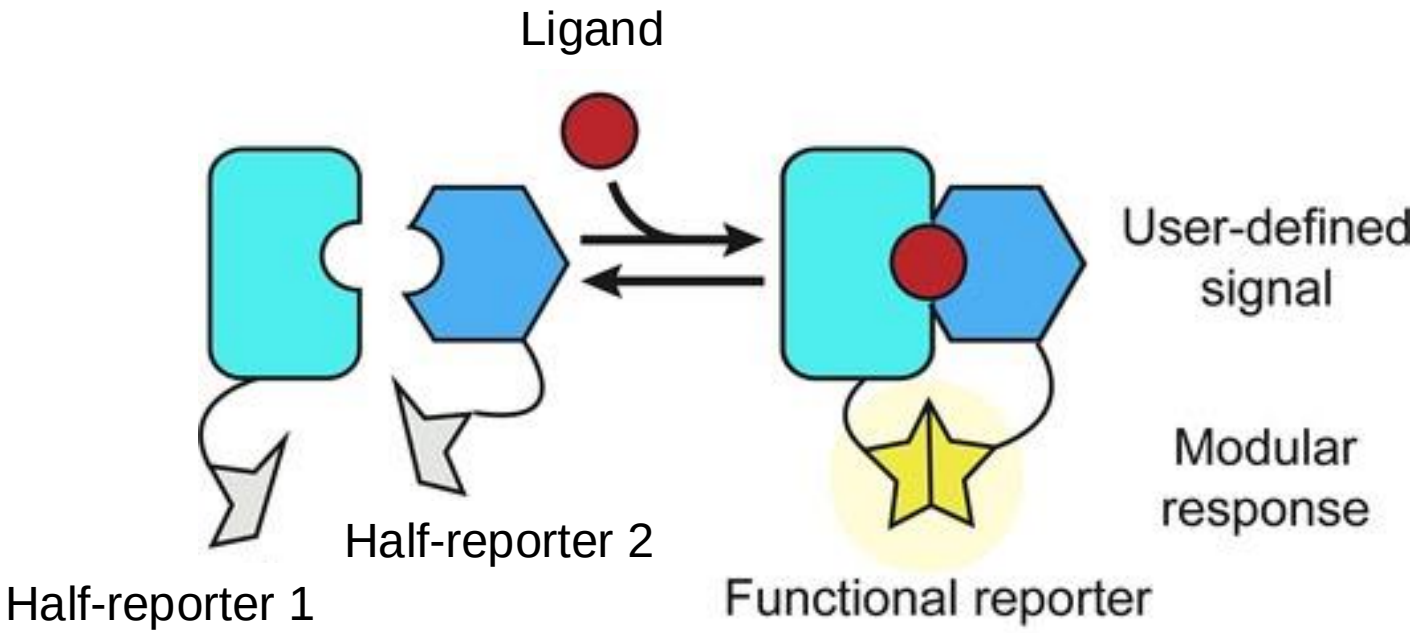
Ligand biosensor design

What is a biosensor?

What can it be useful for?

Why do we care?

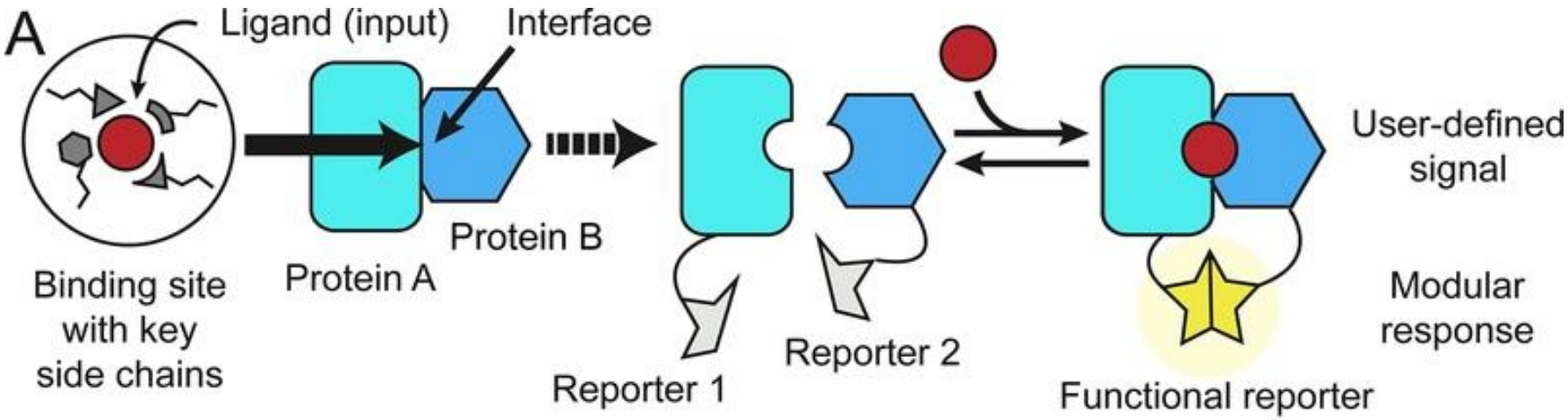
Ligand biosensor design



A wide range of applications

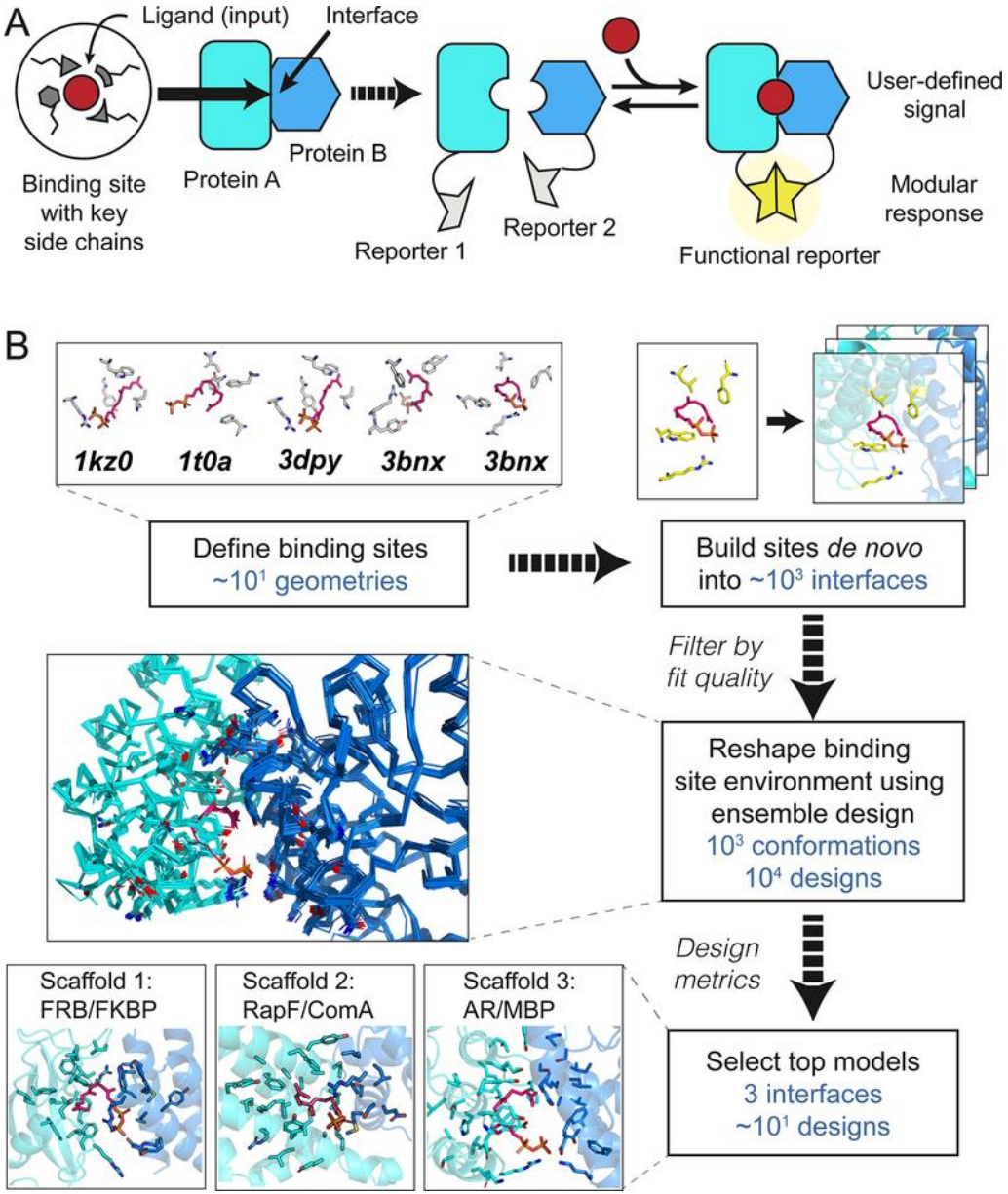
Diagnostics
Detection of contaminants
Data storage

Ligand biosensor design



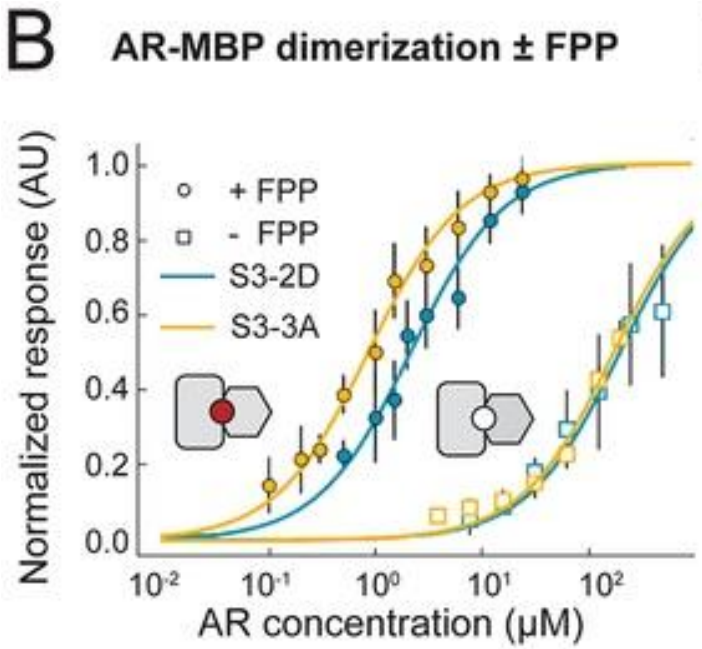
(Glasgow 2020)

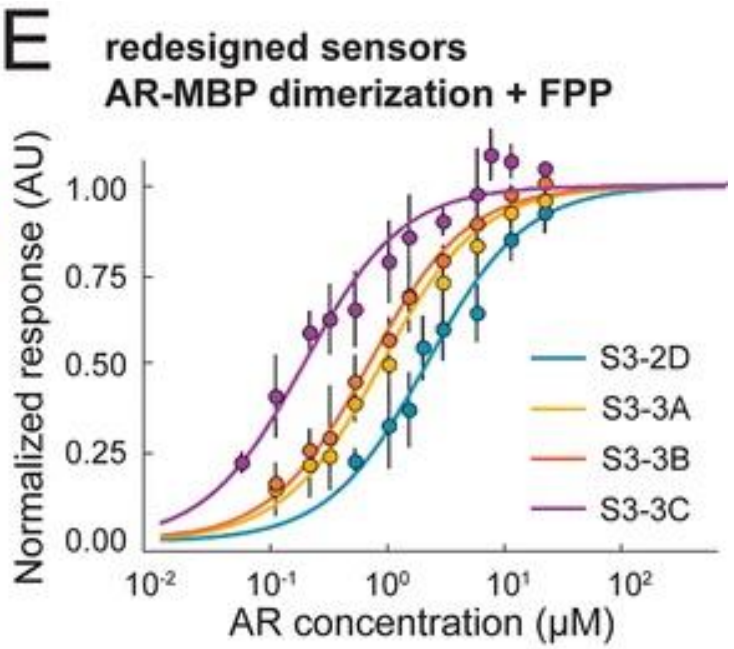
Ligand biosensor design



(Glasgow 2020)

Ligand biosensor design





C

K_D^{app}, AR-MBP dimerization ± FPP (μM)

	S3-2D	S3-3A	S3-3B	S3-3C
+FPP	2.1 ± 0.18	0.87 ± 0.06	0.67 ± 0.03	0.17 ± 0.02
-FPP	>200 #	>200 #	14.0 ± 1.09	6.16 ± 0.31
Fold change	>100	>200	20.8	36.2

 # low estimate

(Glasgow 2020)

Ligand biosensor design

D

K_D, FPP binding to AR or MBP (μM)

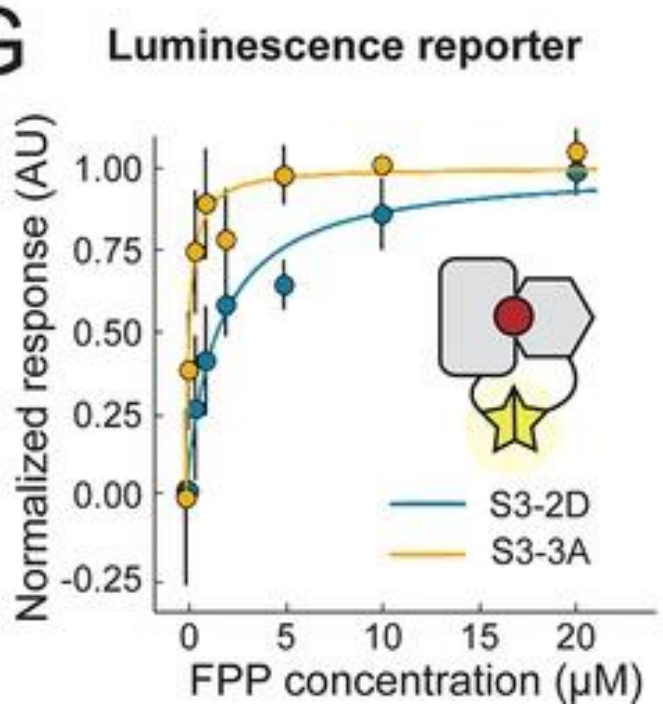
S3-2D AR	6.1 ± 1.4
S3-2D MBP	>1000 #

F

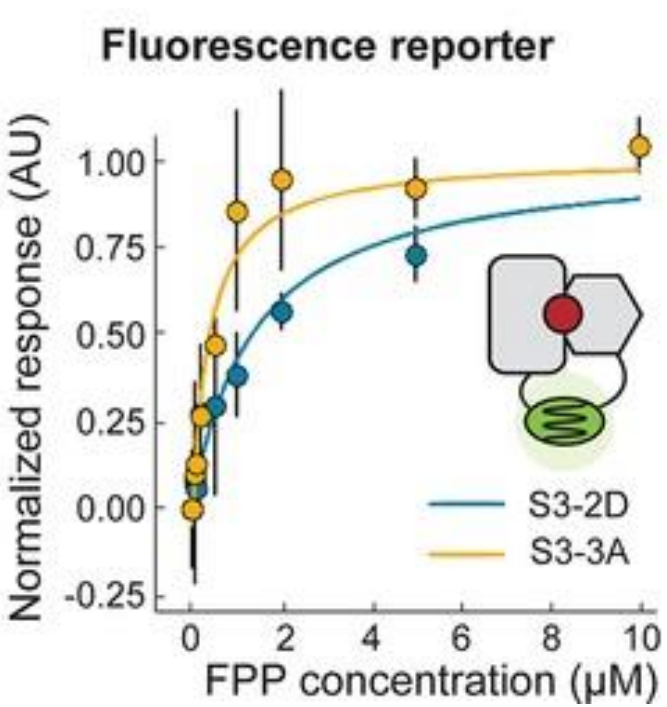
K_D^{app}, FPP sensitivity of sensors (μM)

	S3-2D	S3-3A
Luminescence	1.6 ± 0.47	0.18 ± 0.05
Fluorescence	1.4 ± 0.50	0.33 ± 0.13

G

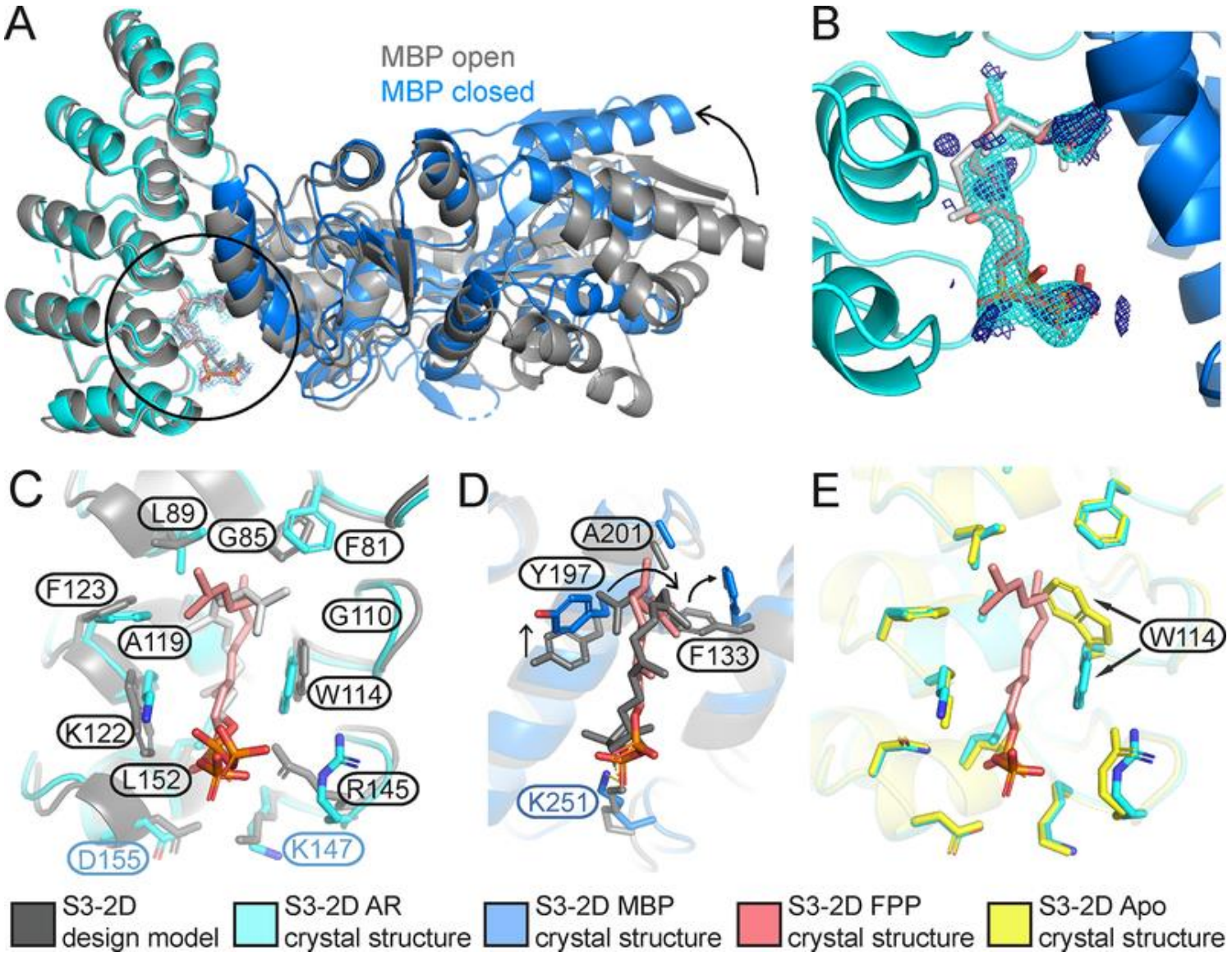


H



(Glasgow 2020)

Ligand biosensor design



(Glasgow 2020)

Ligand biosensor design

Have they solved the biosensor design problem?

Where do you see key limitations?

What could be the next steps?

(Glasgow 2020)

Take home messages

1. Protein binding vs folding differences (interactions, conformational space)

2. Solutions for binding site design (starting from precise ligand conformation)

3. Computational design solutions are often suboptimal. Need time-consuming optimization by directed evolution

4. AI-based methods are emerging that can create protein scaffolds stabilizing precise ligand structures