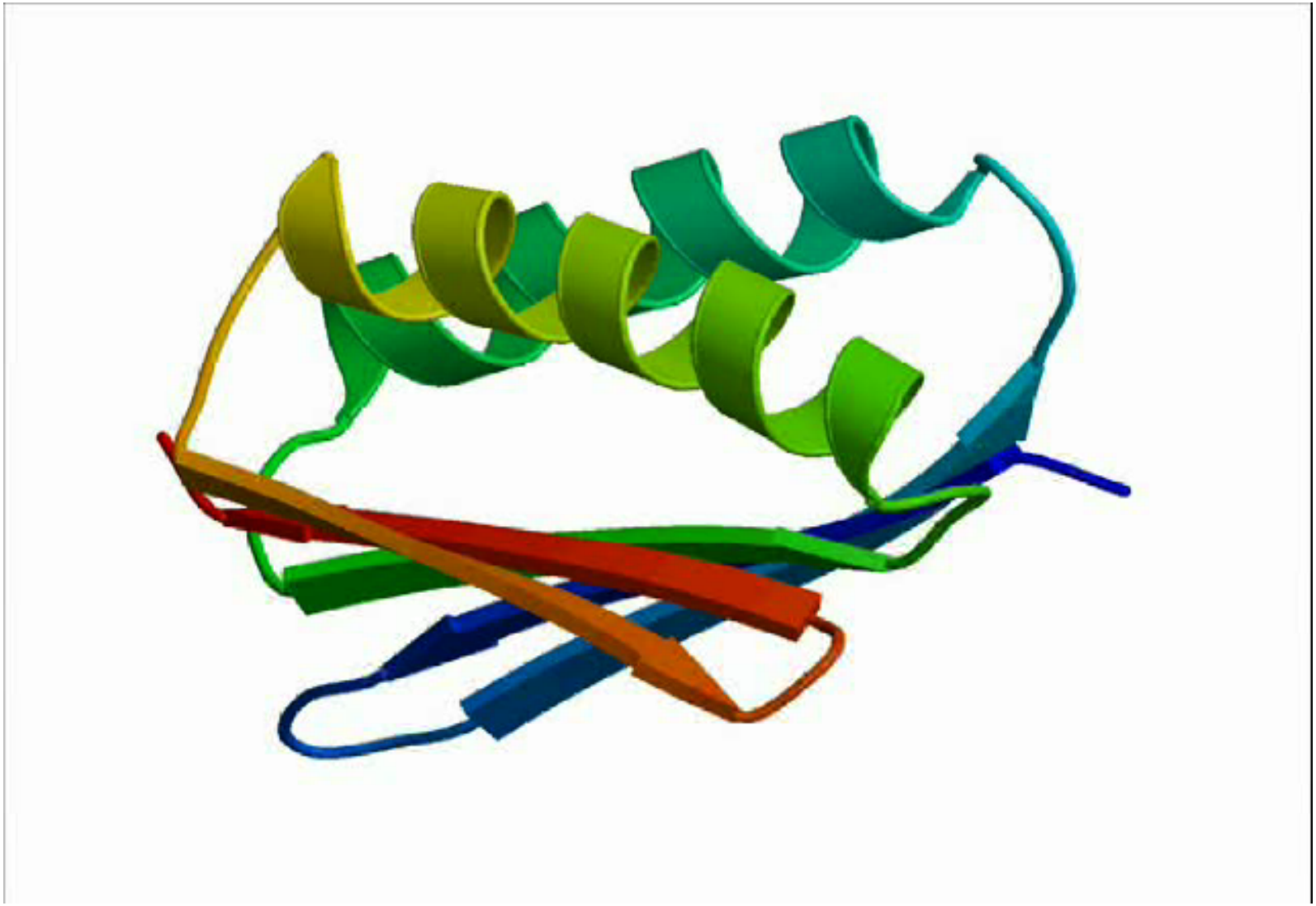
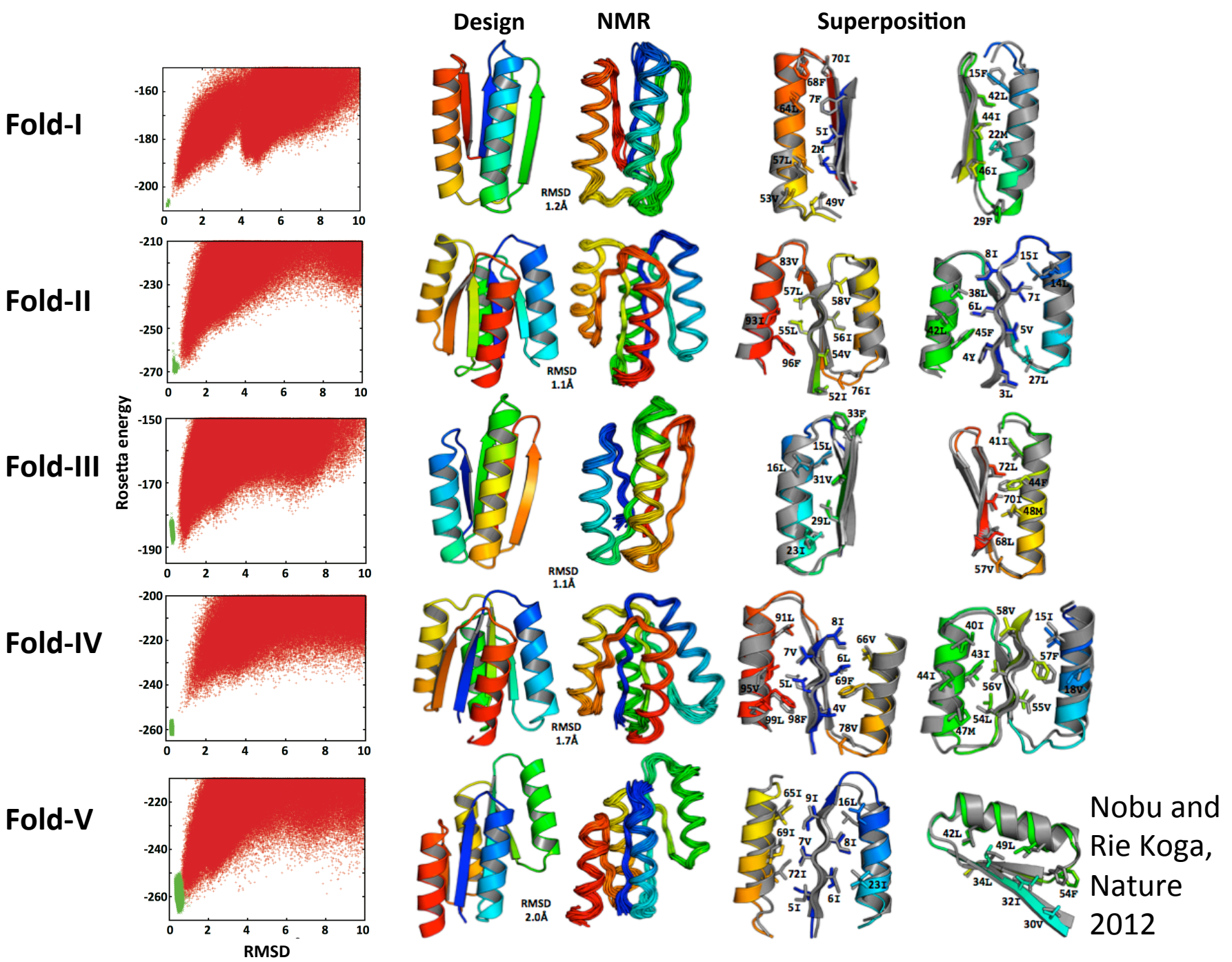
# Computational design of protein structures, functions, and assemblies

# Protein Design Work Flow

- Computer calculation of optimal sequence for desired structure or function
- Read off amino acid sequence of designed protein
- Back translate to DNA sequence, and make gene
- Make protein and assay

**Protein Design: find lowest energy sequence for desired structure and/or function**

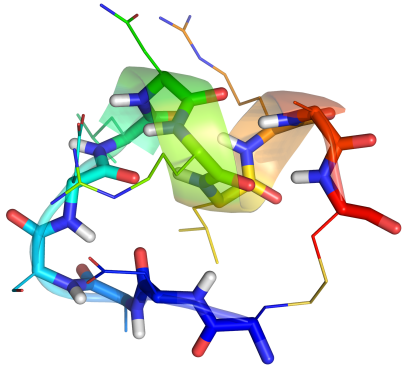


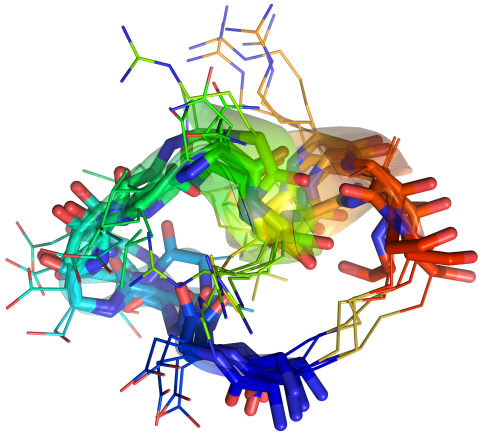



Nobu and Rie Koga, Nature 2012

# Designed cyclic peptides with stable backbone conformations

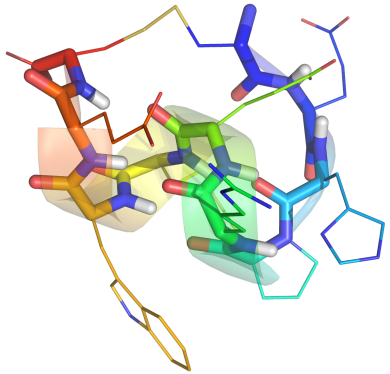
Design 10helix1

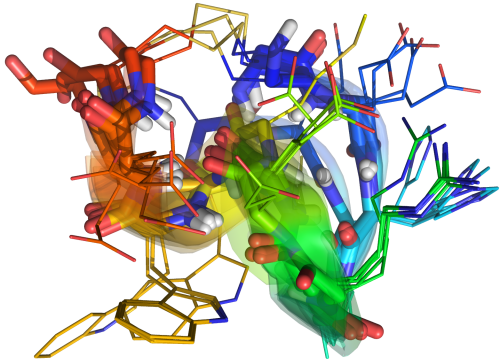




10helix1 NMR ensemble
(undergoing refinement)

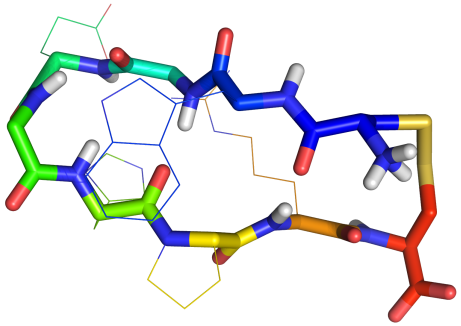
Design 8helix1

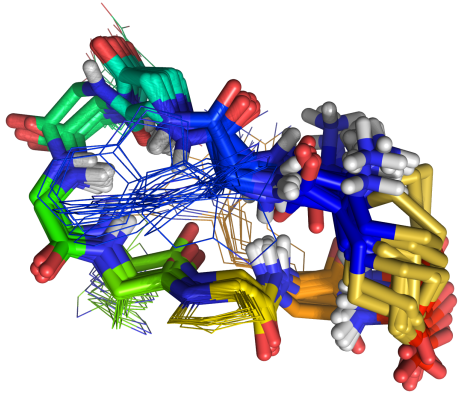




8helix1 NMR ensemble
(undergoing refinement)

Design 5loop1

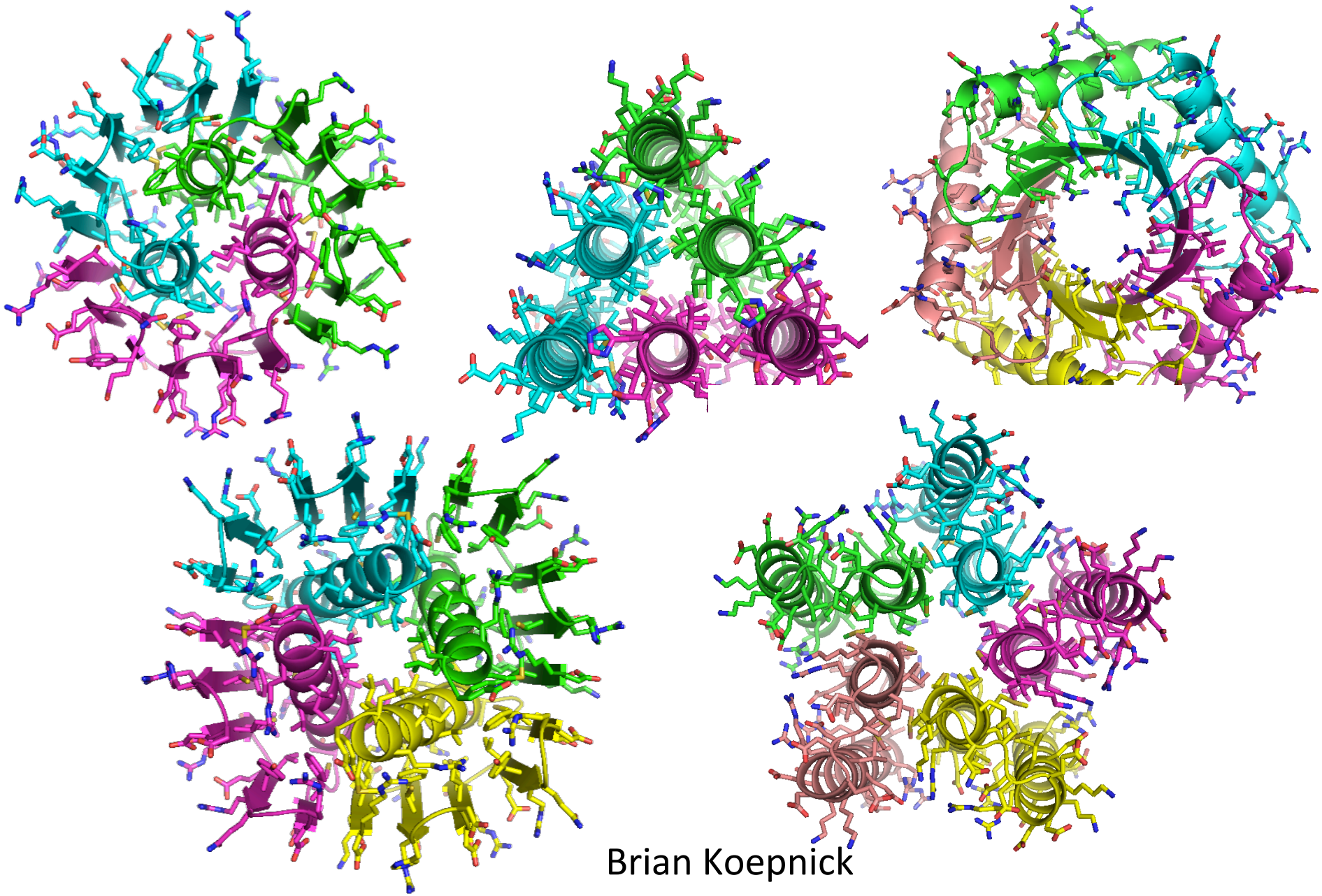




5loop1 NMR ensemble

Vikram Mulligan

# Foldit Symmetric Designs



Brian Koepnick

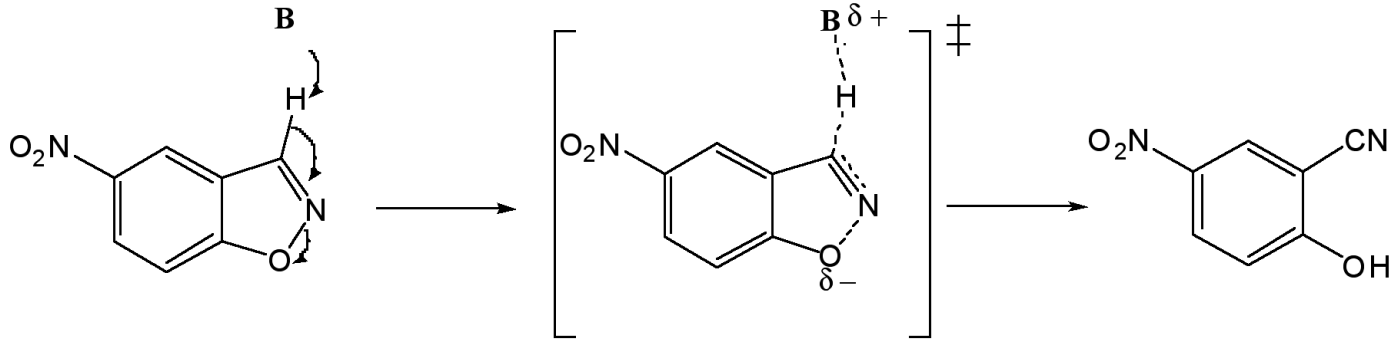
# De Novo active site design

I. Model reaction transition states and intermediates

II. Design disembodied ideal active site around transition states and intermediates

III. Design protein containing ideal active site

# Kemp elimination reaction




Design Process:

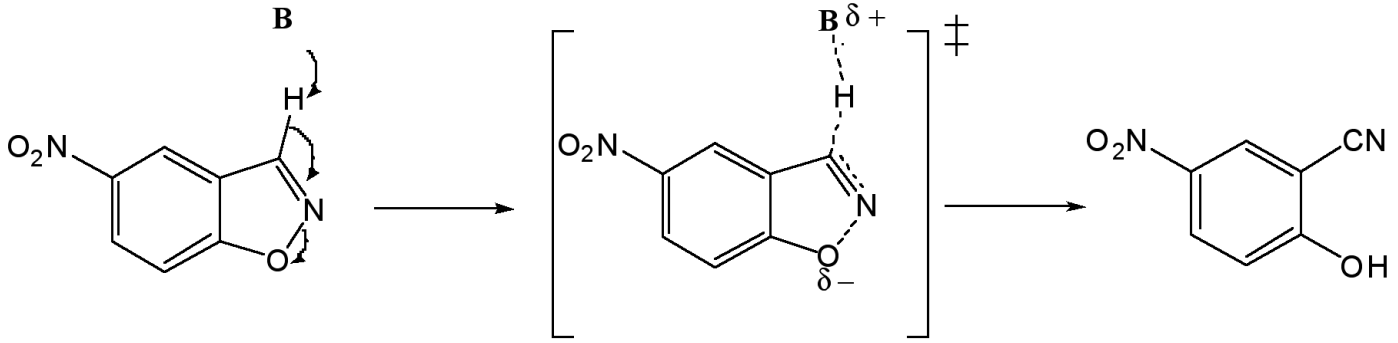
choose catalytic motif

1) match in scaffold

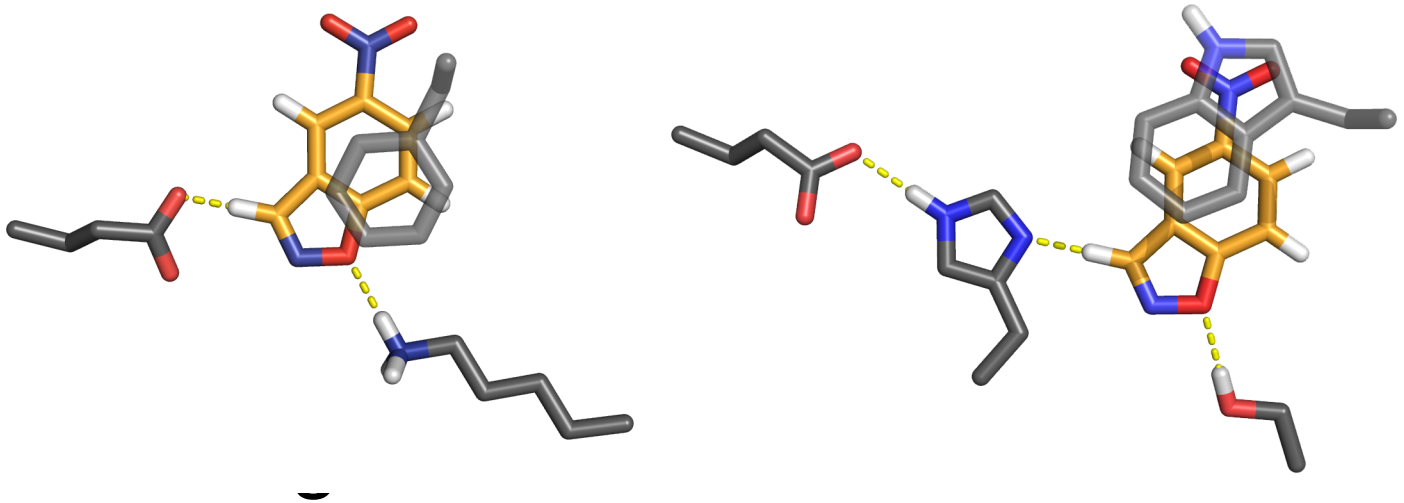
2) design pocket

3) rank designs

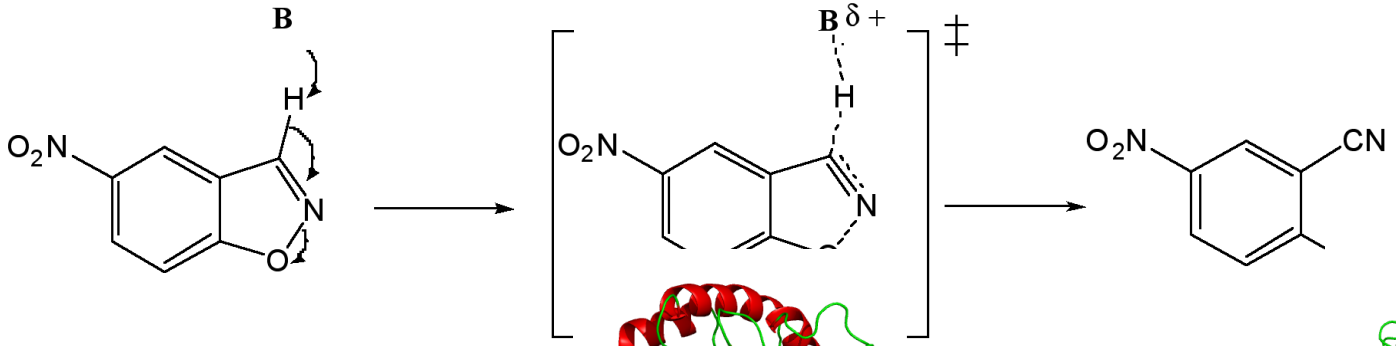
# *de novo* design process




1) choose catalytic motif



# *de novo* design process

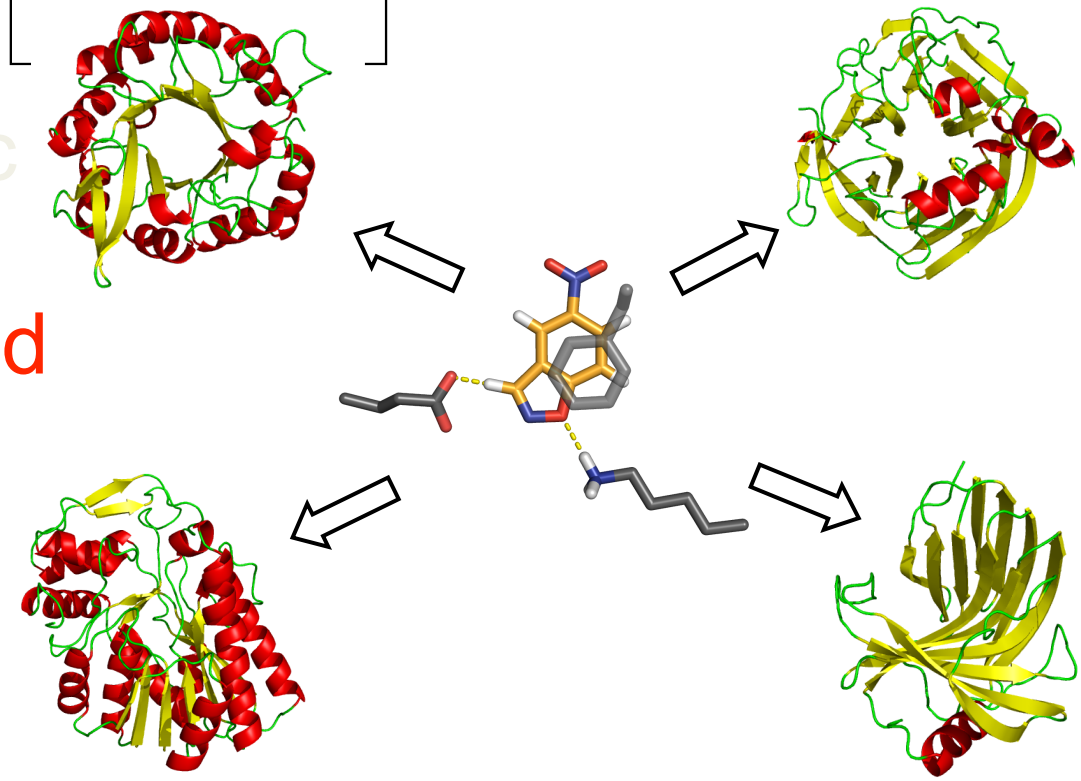



1) choose catalytic

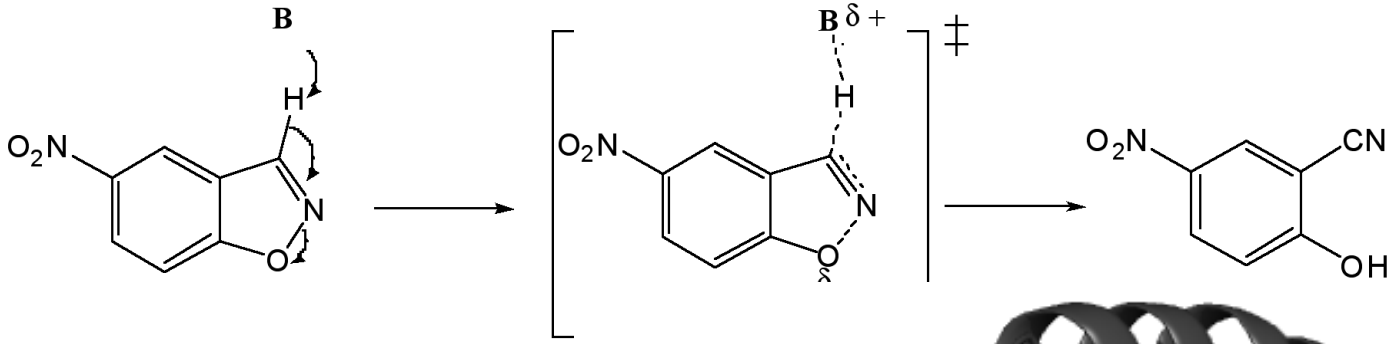
2) match in scaffold

3) design pocket

4) rank designs



# *de novo* design process

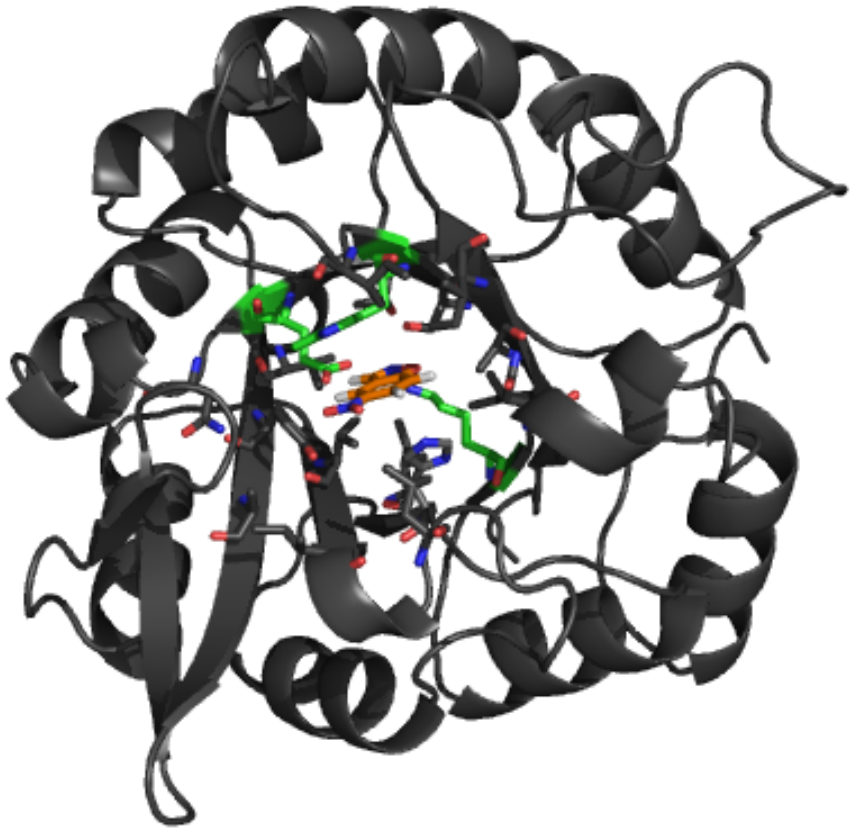



1) choose catalytic m

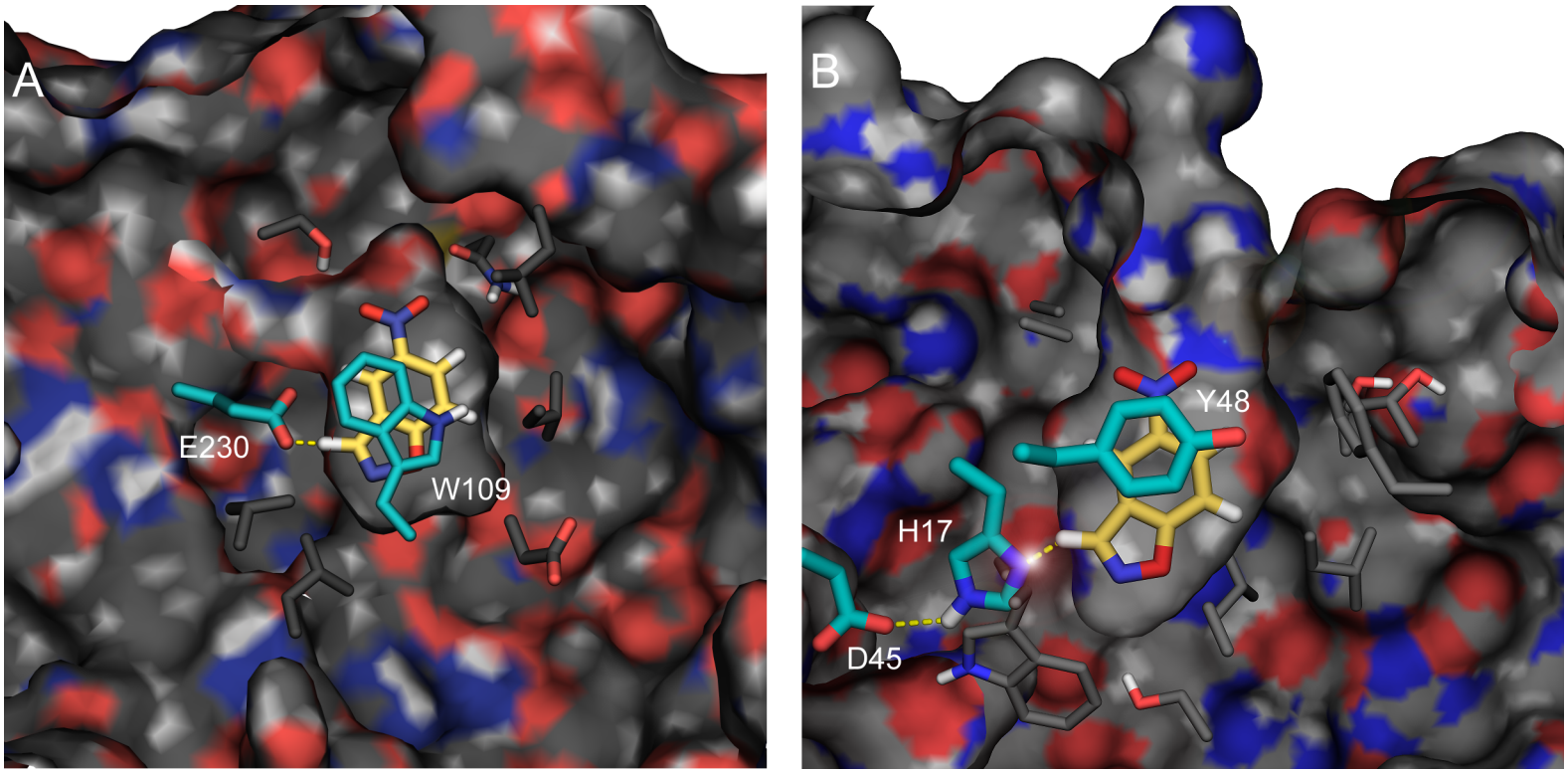
2) match in scaffold

3) design pocket

4) rank designs

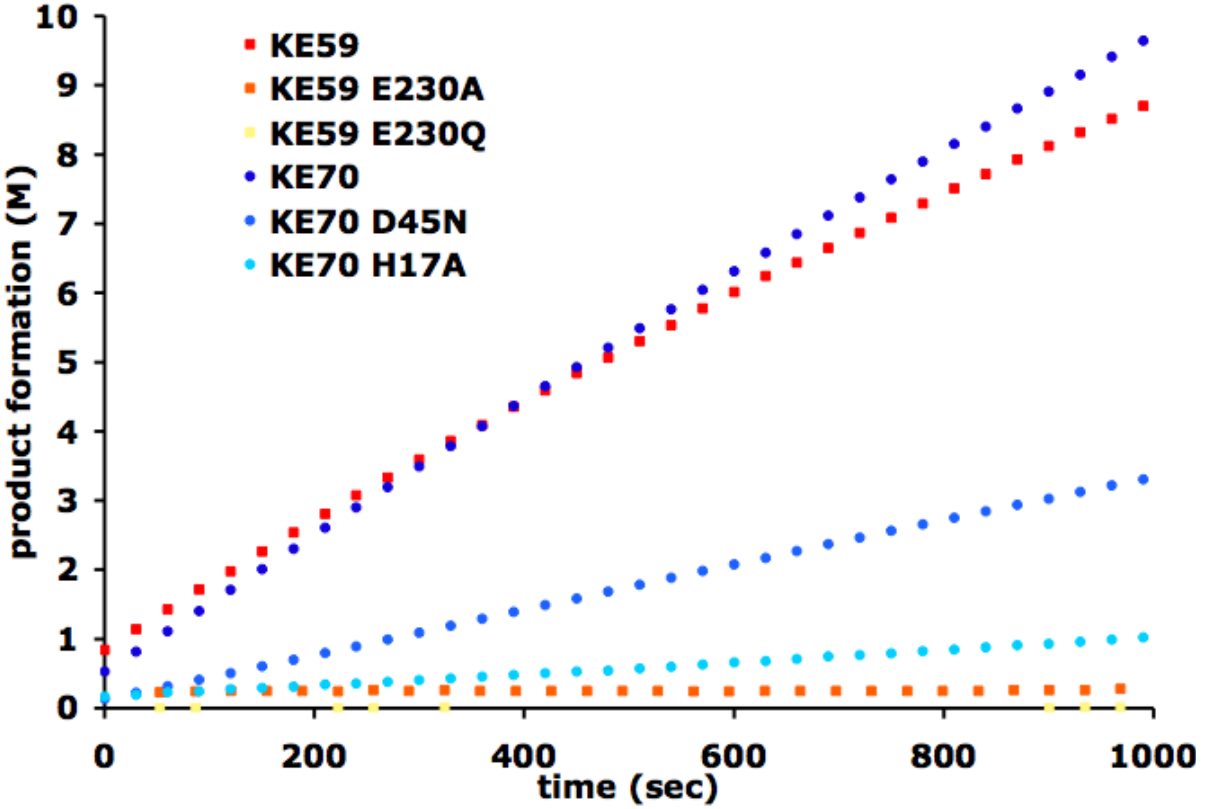


# Examples of design models




Daniela Roethlisberger, Andrew Wollacott

# Catalytic residue dependent activity!



# De novo enzyme design-- Successes thus far

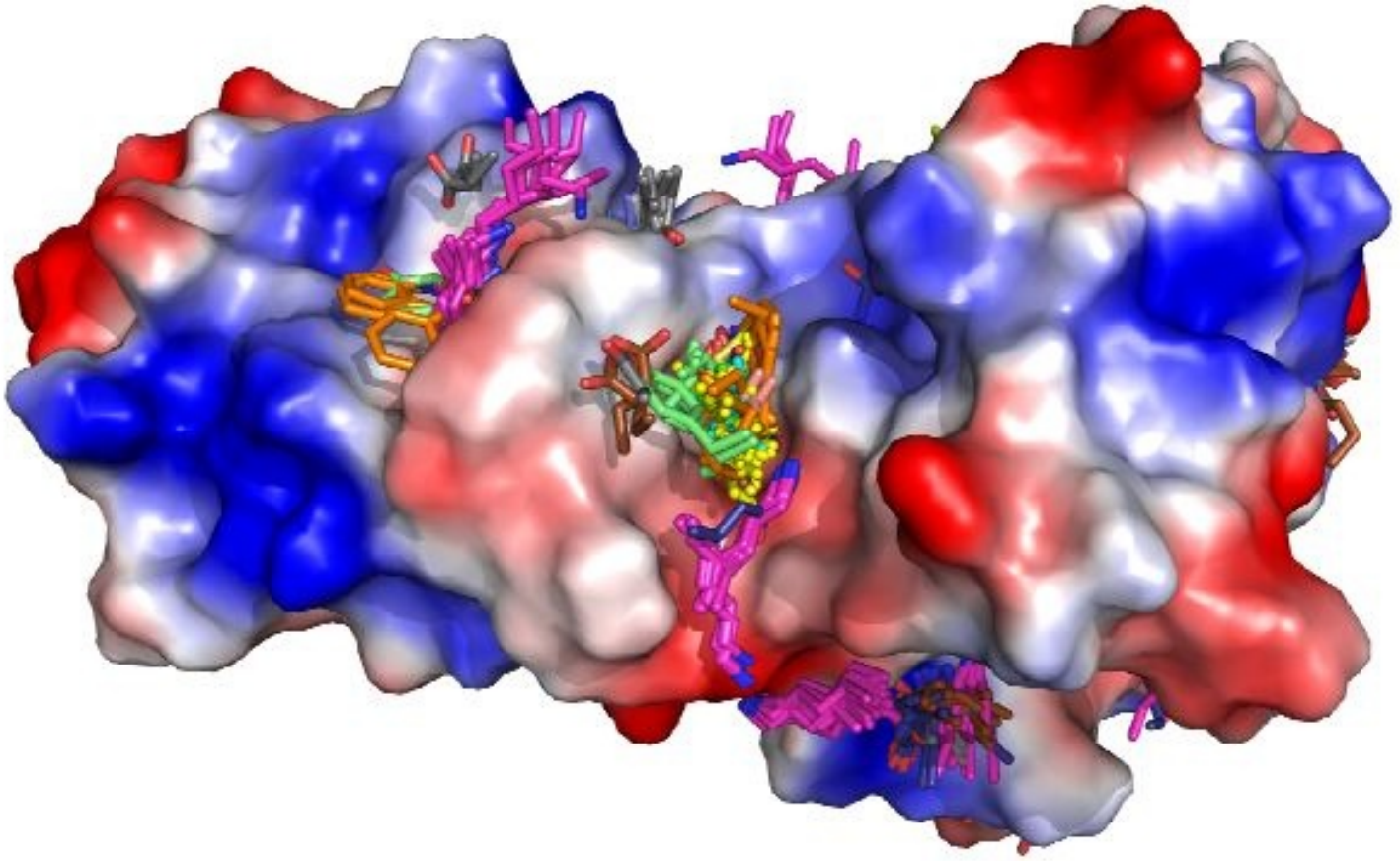
- General acid-base catalysis:  Kemp elimination (Nature 2008)
- Covalent catalysis:  novel aldol catalysts (Science 2008)
- Bimolecular reactions:  Diels Alder (Science 2010), Baylis Hillman
- Polar transition state stabilization: ester hydrolysis

# De novo enzyme design

- Can design active enzymes from scratch!
- Starting activities low, but can be increased readily by directed evolution (evolved Kemp kcat/Km ~ $5x10^5$)
- Need more precise positioning of catalytic groups, elimination of competing reactions, dynamics (?), etc.
- Enzymes are masters of art of compromise--have to do everything well!

| | Activity | Development Time |
|---|---|---|
| Designed Enzymes +directed evolution | +++ | < 5 years |
| Catalytic Antibodies | +++ | ~25 years |
| Natural Enzymes | +++++++ | ~$10^8$ years |

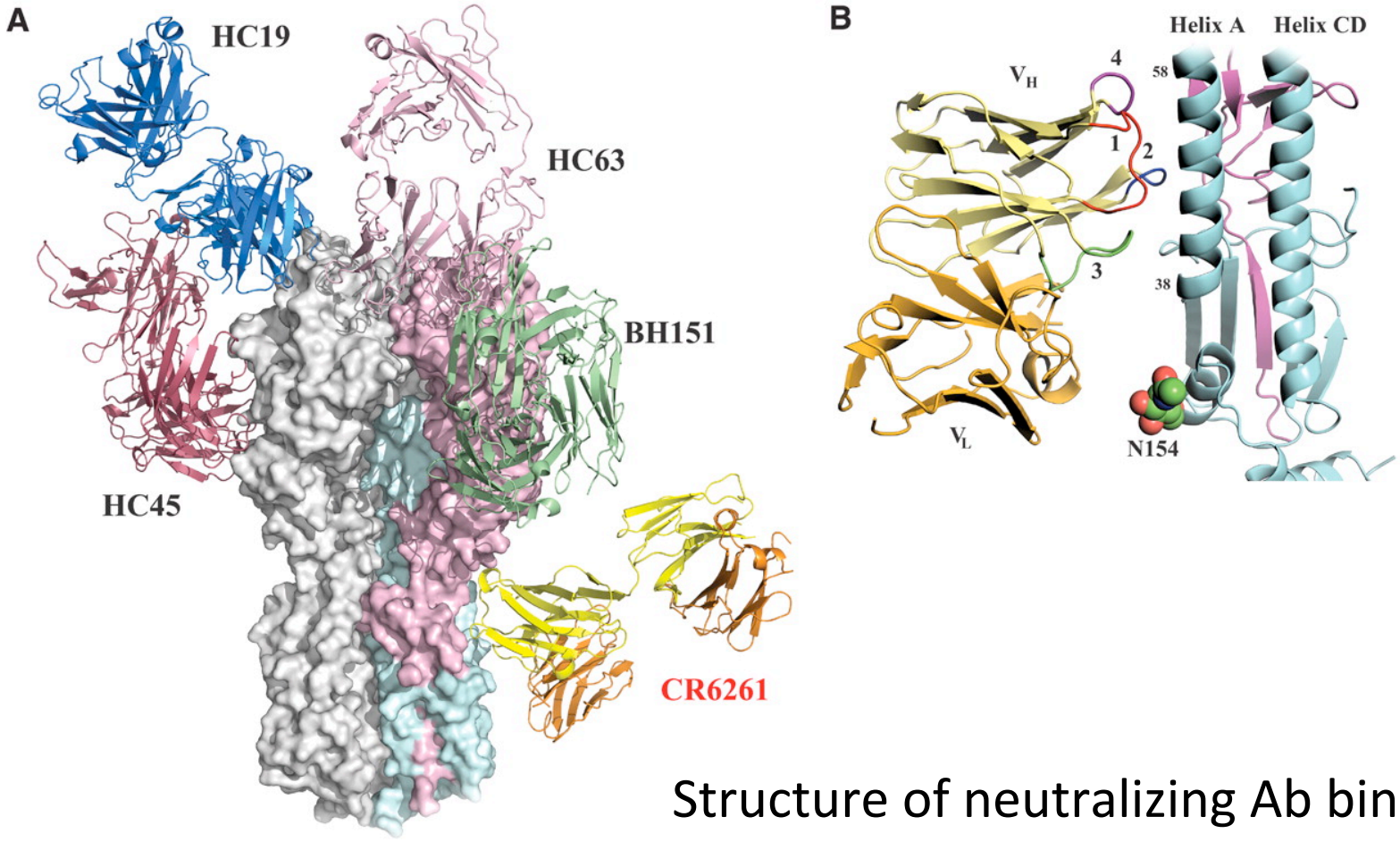
# Design of Binding



# “Hot-spot” centered approach to de novo protein-protein interface design

- Generate a disembodied hot-spot residue map of the target surface patch
- Dock large set of scaffolds against the surface patch, favoring configurations that support multiple hotspots.
- Build hot spots onto docked scaffolds:
  - Superimpose scaffold on hot-spot interaction requiring highest precision
  - Build on additional hot-spots by minimizing the scaffold rigid body, sidechain, and backbone degrees of freedom
- Optimize interface for binding affinity
- Filter designs on computed binding energy and shape complementarity

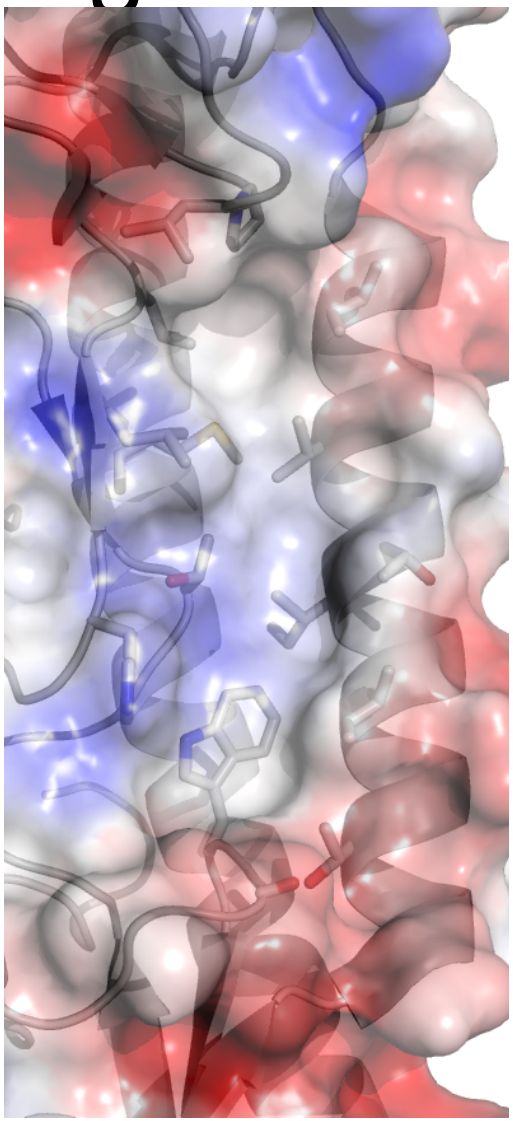
# Design of binders to conserved epitope on Spanish Flu hemagglutinin

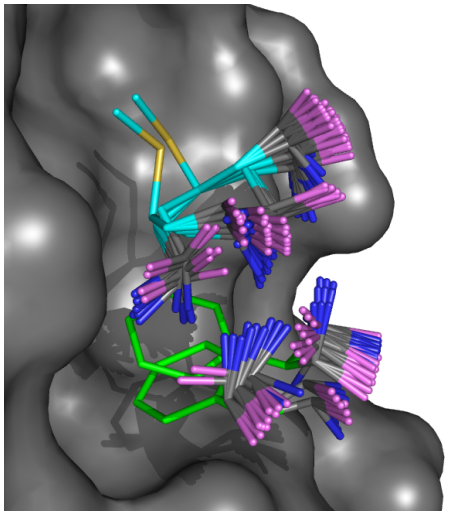


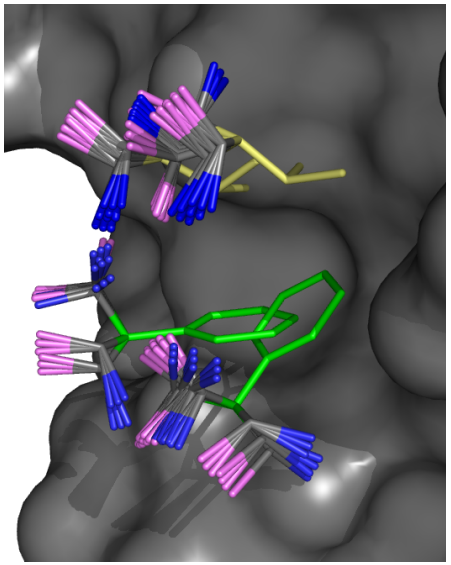

Structure of neutralizing Ab binding to stalk region of influenza HA

Ekiert, Wilson *et al*. Science 324:246

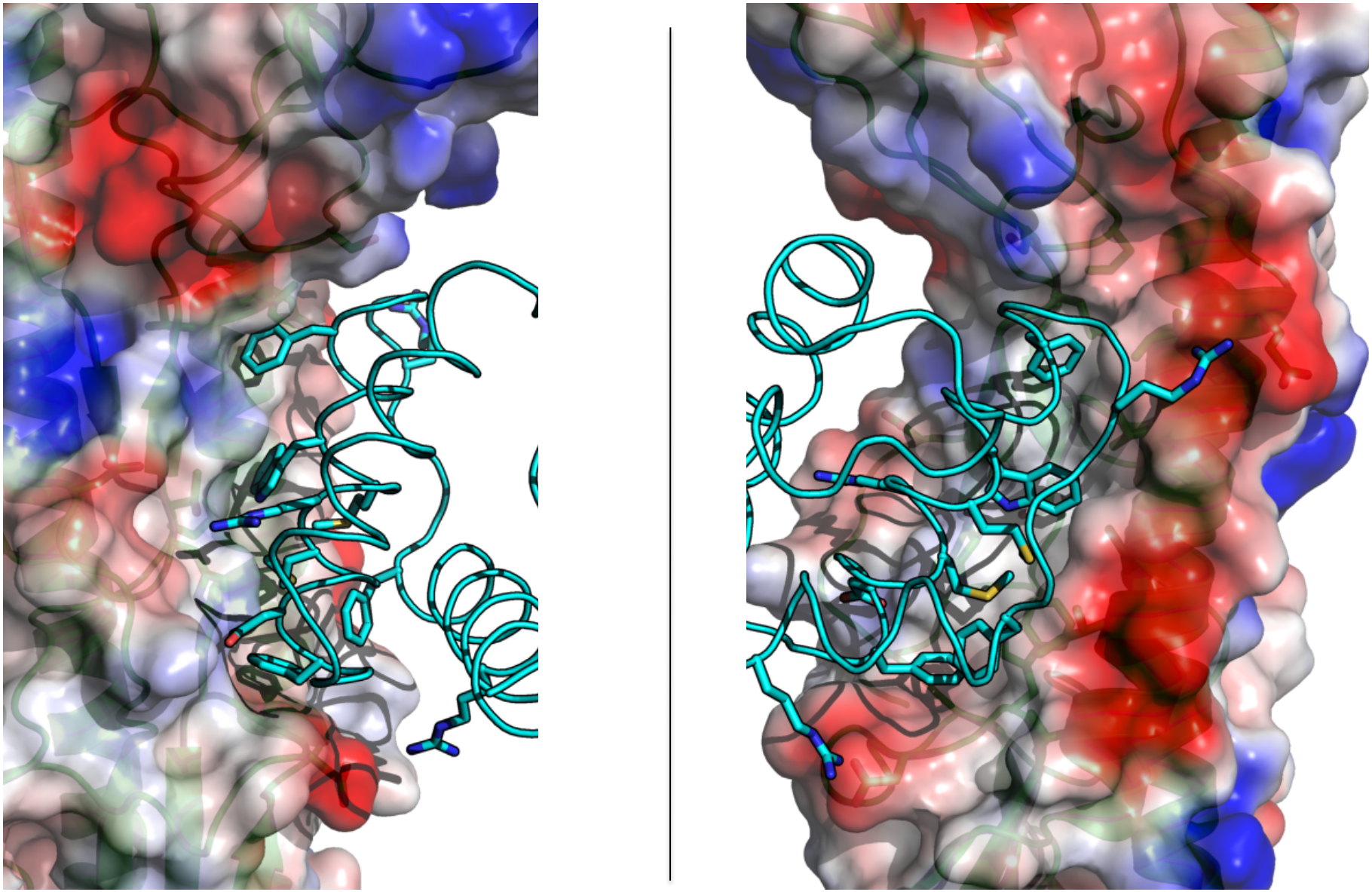
# First, dock disembodied residues against HA surface





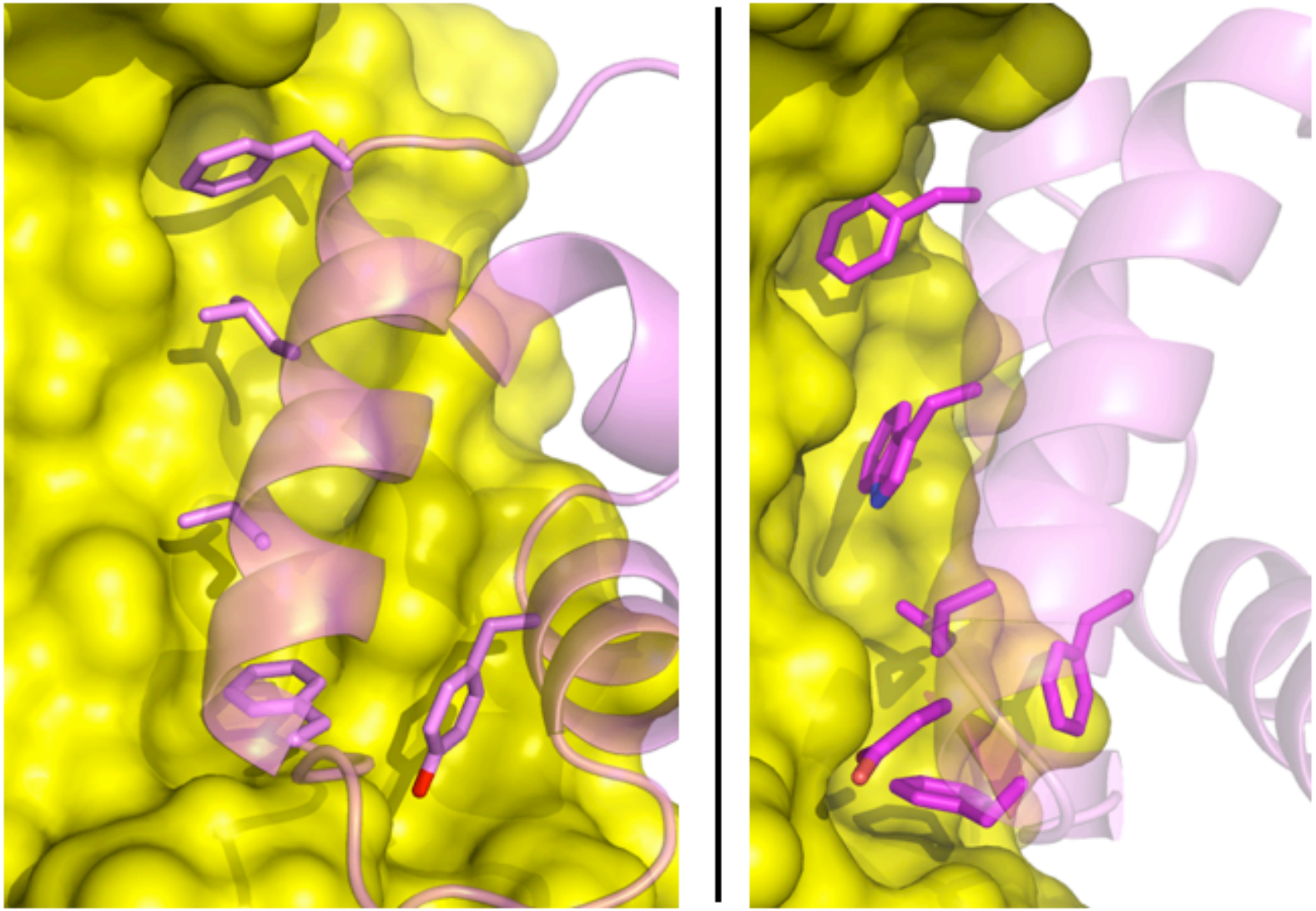


# Second, find/build scaffold supporting the interacting residues

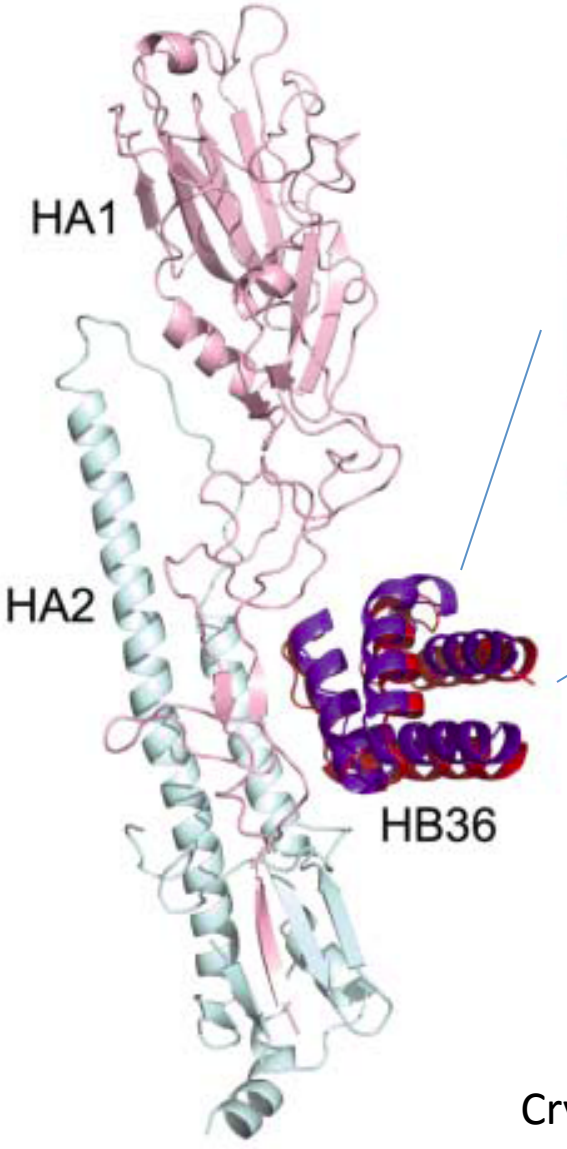


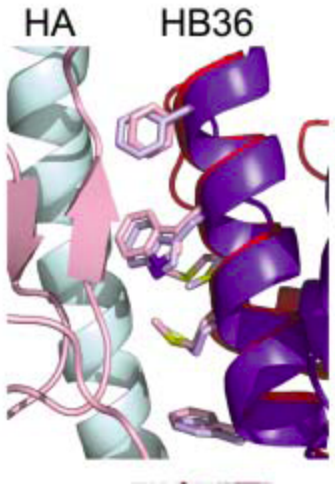
HB80

HB36

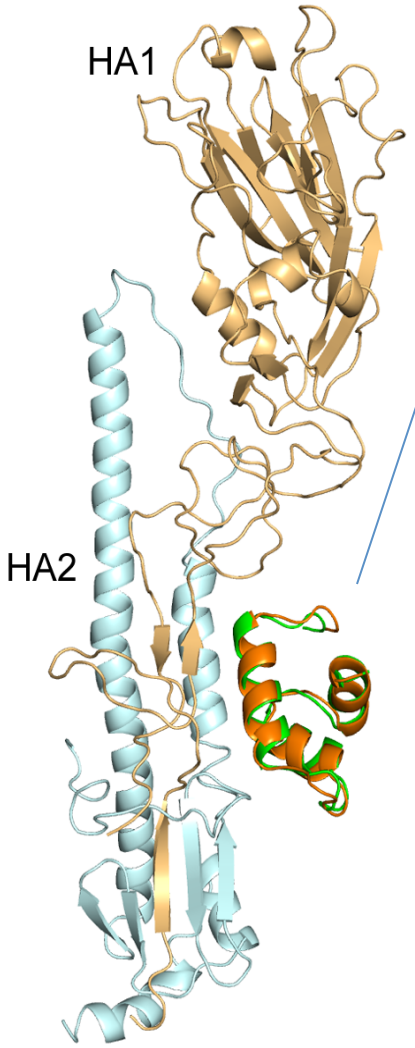


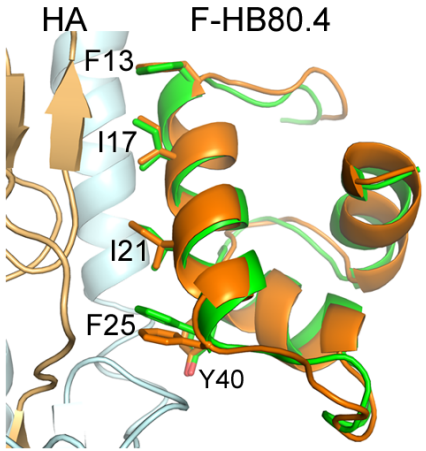
# Crystal structures of designed binders bound to HA closely match design models




Red: xtal
Purple: design




Orange: xtal
Green: design

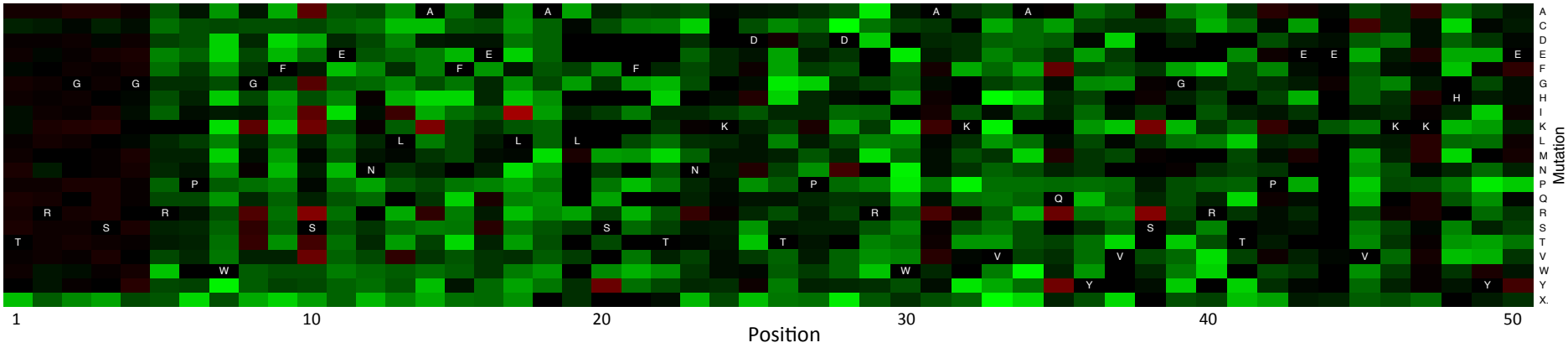
Crystal structures: Damian Ekiert, Cyrille Dreyfus, Ian Wilson

The method is far from perfect-only 2/80 designs bind the virus, and even these bind weakly

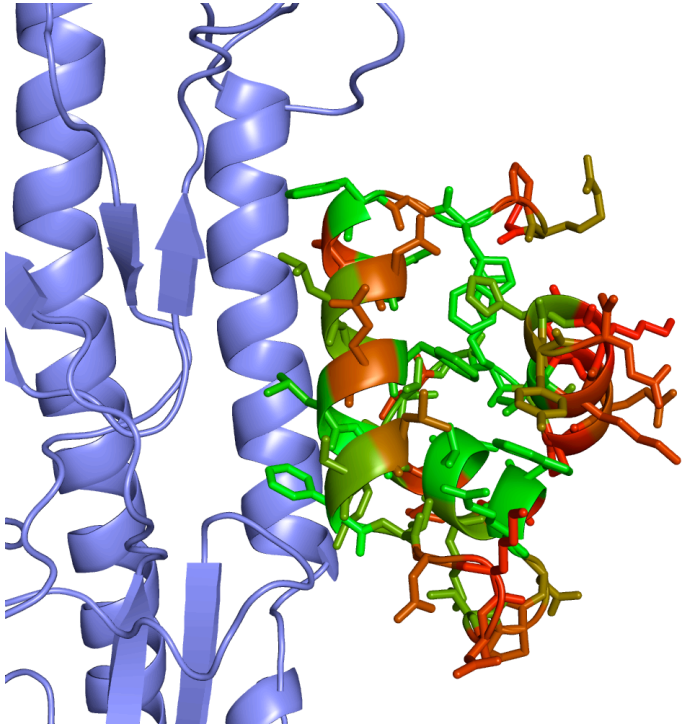
How do we improve the energy function used in the design calculations?
How do we make tighter binders?

Library selection plus next gen sequencing (Doug Fowler, Stan Fields)

# Use next-gen sequencing to comprehensively map optimality of designed sequence (HB80)




1. Create library of all point mutants
2. Select for binding
3. Deep sequence (Illumina PE-76)
4. Compute ratio for each mutant of population in selected and unselected pool
5. Hotspot residues are largely invariant; opportunities for improving designs are revealed



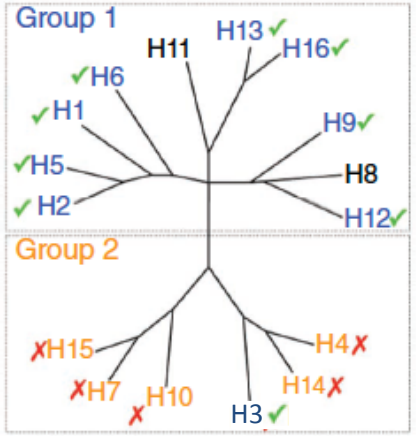
Positions colored by Shannon entropy

Aaron Chevalier
Tim Whitehead

# Binders to multiple sites on the Influenza HA enable unique readout of HA identity

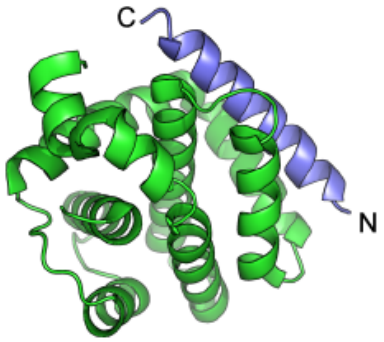
| Group I | HB80.4 | HB80.3_H5 | HB36.5 | HB36.5_H2 | S24* | S52* |
|---|---|---|---|---|---|---|
| **H1** | yes | yes | yes | yes | yes | yes |
| **H2** | yes | no | no | yes | yes | yes |
| **H5** | yes | no | yes | yes | no | Yes |
| **H6** | yes | n.d. | yes | yes | n.d. | n.d. |
| **H9** | yes | n.d. | yes | yes | n.d. | n.d. |
| **H12** | yes | n.d. | no | no | n.d. | n.d. |
| **H13** | yes | n.d. | yes | yes | n.d. | n.d. |
| **Group II** | | | | | | |
| **H3** | no | no | no | no | yes | yes |

HB80.4 binding profile

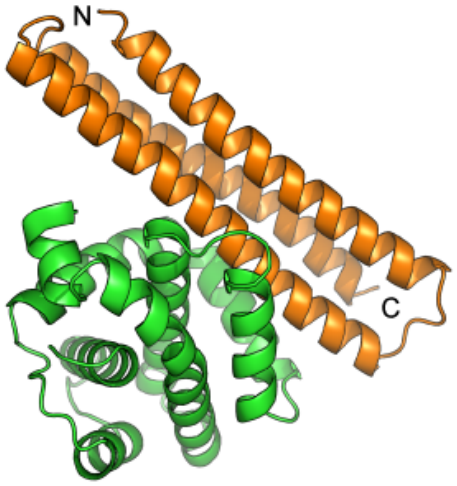



Eva Strauch
David La
Aaron Chevalier

# Engineered inhibitor of a viral Bcl-2 protein (BHRF1) associated with lymphoma




**Bim-BH3**
**$K_D$ 12 ± 4 nM**
**(Non-specific)**

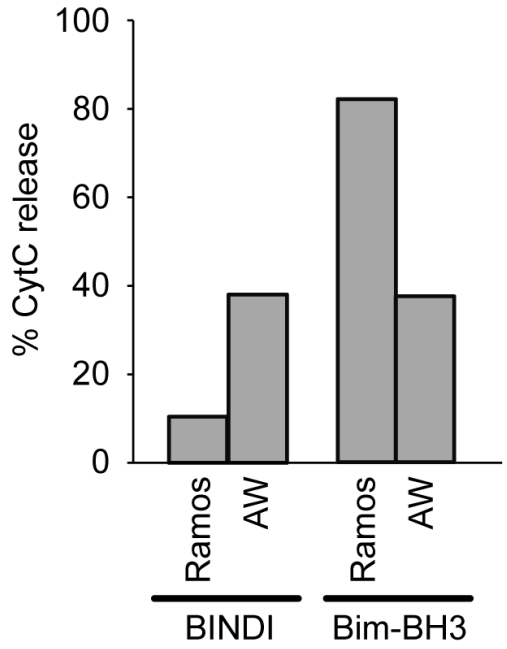



**BINDI**
**$K_D$ 0.22 ± 0.05 nM**
**(Specific)**

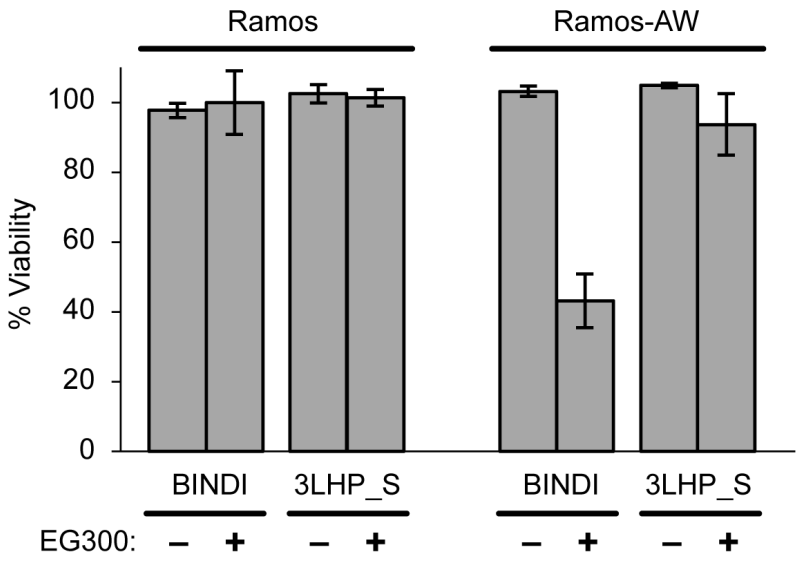
**Dissociation constants (nM) for prosurvival Bcl-2 family members**

| Protein | BHRF1 | Bcl-2 | Bcl-W | Mcl-1 | Bfl-1 | Bcl-XL | Bcl-B |
|---|---|---|---|---|---|---|---|
| Bim-BH3 | 12 ± 4 | 2.02 ± 0.08 | 2.1 ± 0.1 | 0.6 ± 0.2 | 2.1 ± 0.3 | 3 ± 1 | 12.2 ± 0.1 |
| BINDI | 0.22 ± 0.05 | 2,100 ± 100 | 870 ± 40 | 40 ± 10 | 2,600 ± 800 | 810 ± 80 | > 10,000 |

# Engineered inhibitor of a viral Bcl-2 protein (BHRF1) associated with lymphoma



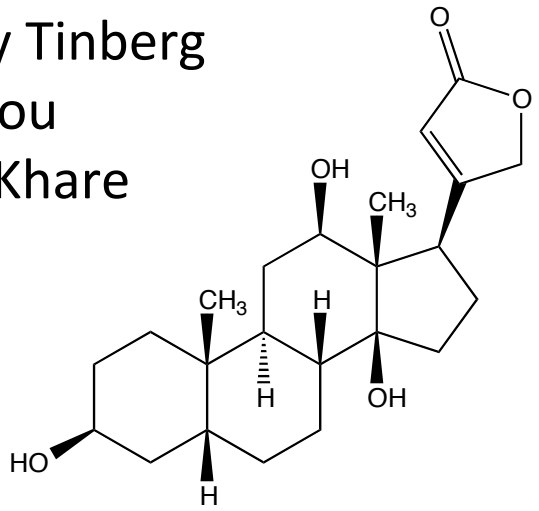

Mitochondrial CytC Release

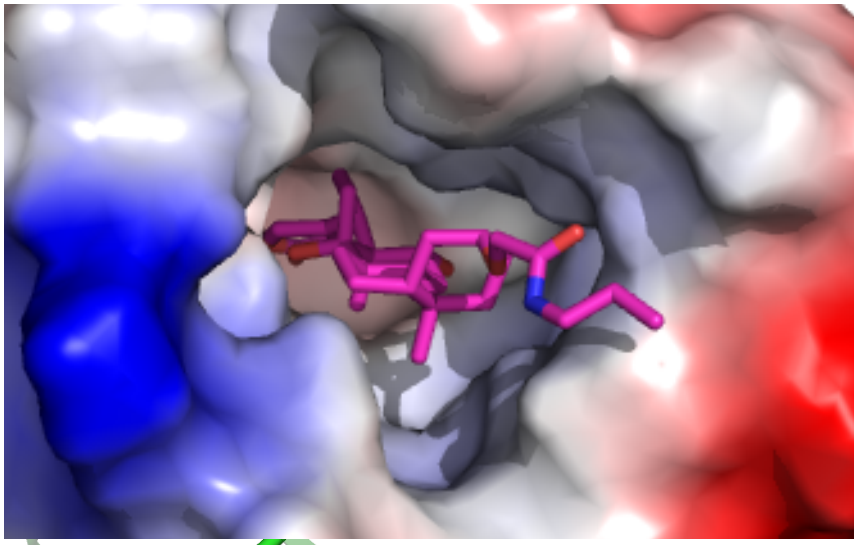


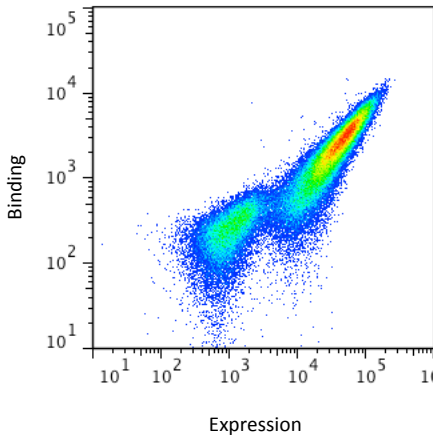

Intracellular Delivery and Cell Death

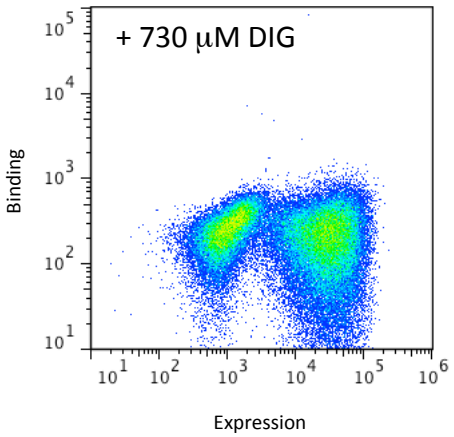
# De novo design of small molecule binding proteins

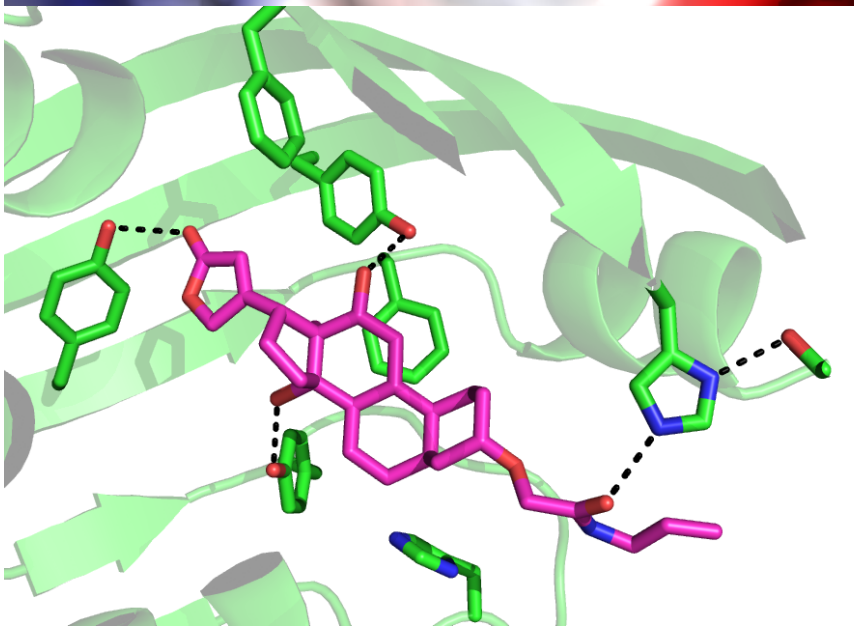
Christy Tinberg
Jiayi Dou
Sagar Khare



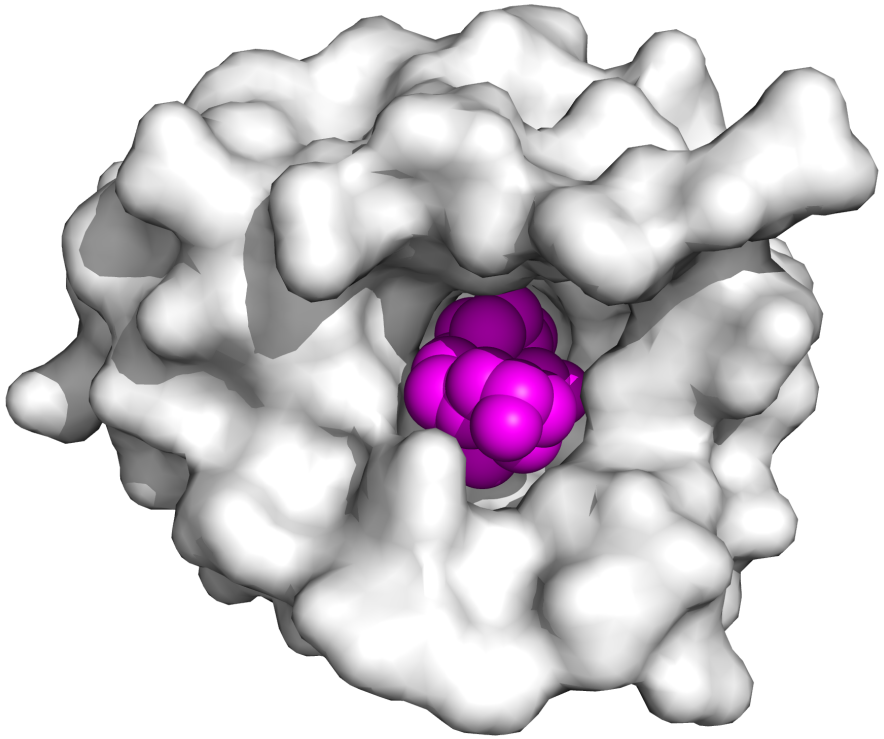



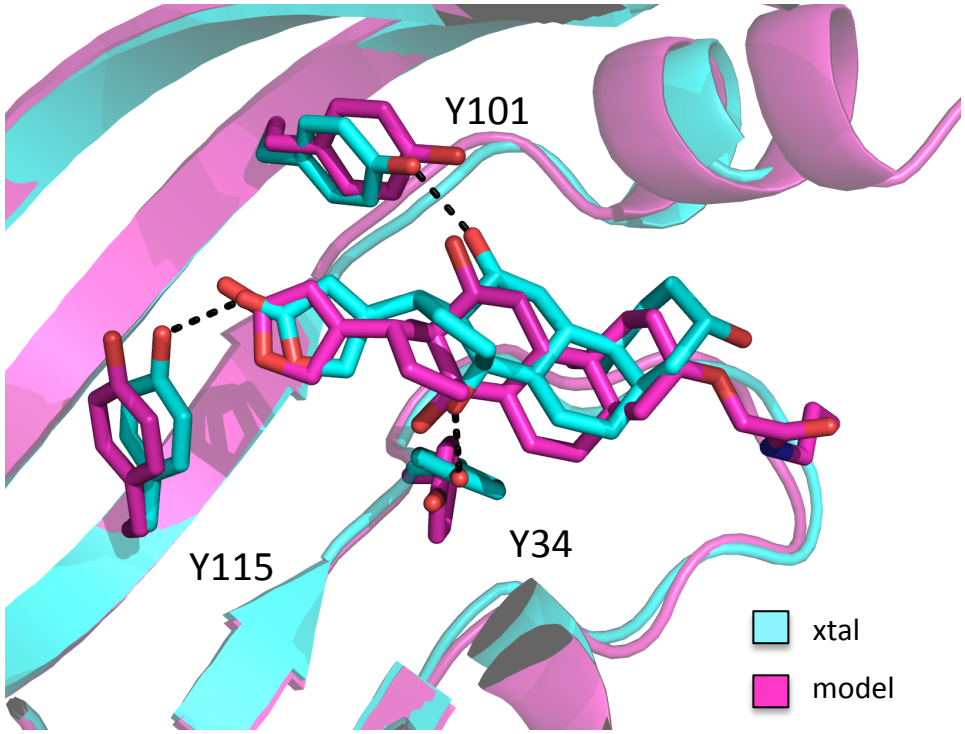


# DIG10.2 Crystal Structure Confirms Model





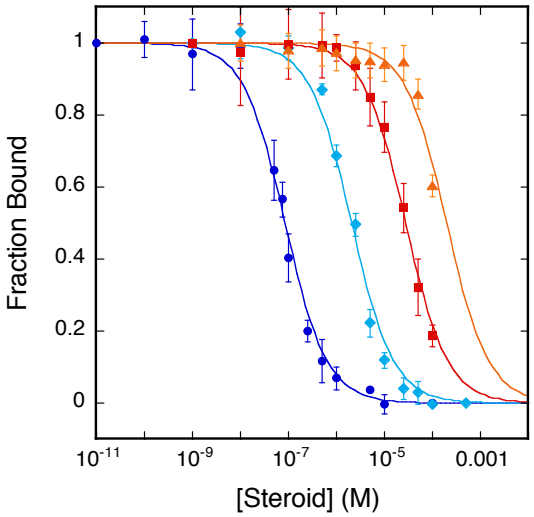

$S_c$ = 0.67

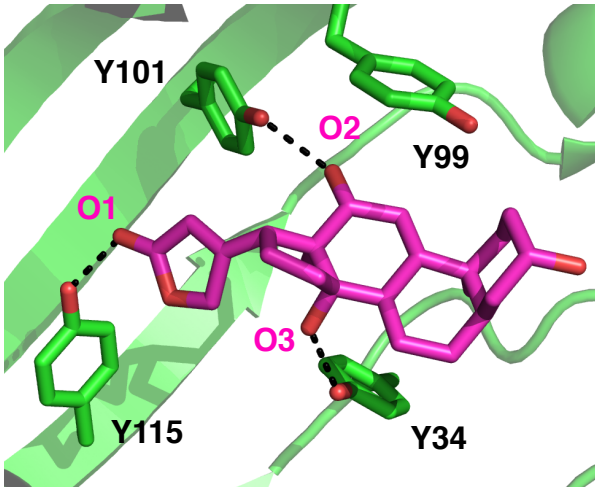
Backbone RMS = 0.460 Å

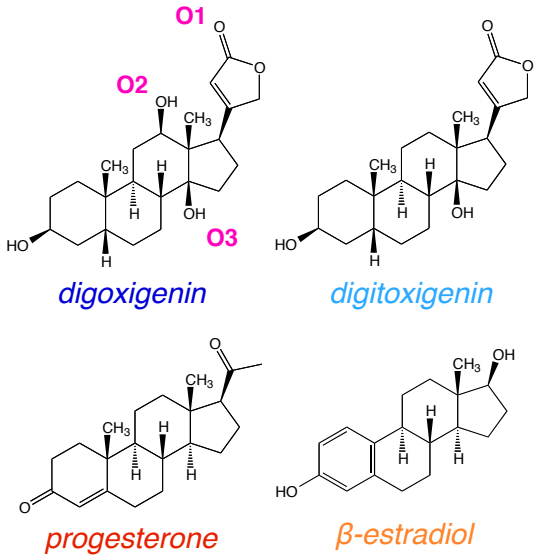
All-atom RMS = 0.53 Å

Ligand RMS = 1.00 Å

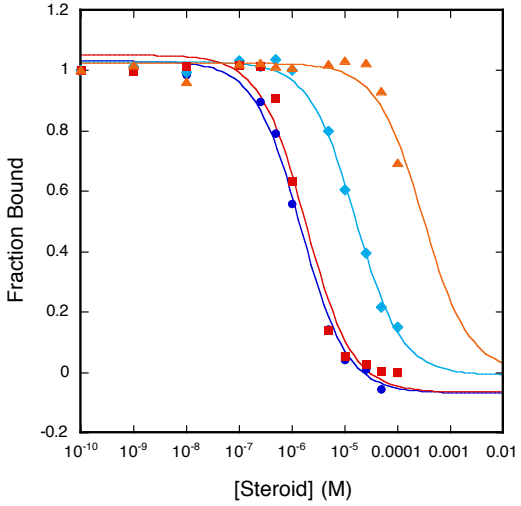
# Precise control of ligand binding selectivity



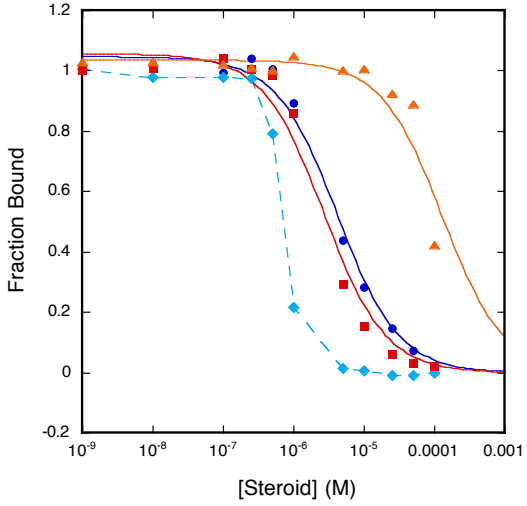



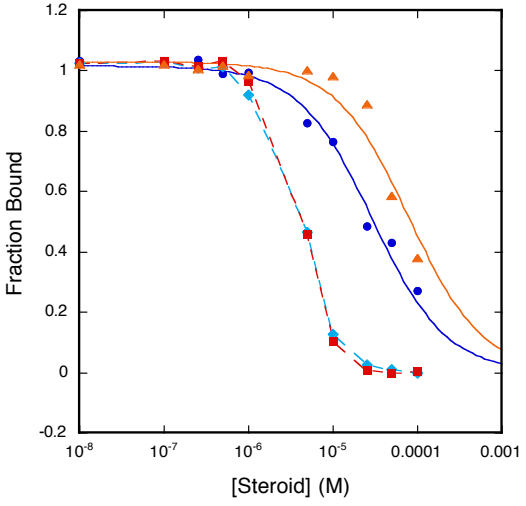

Y34F



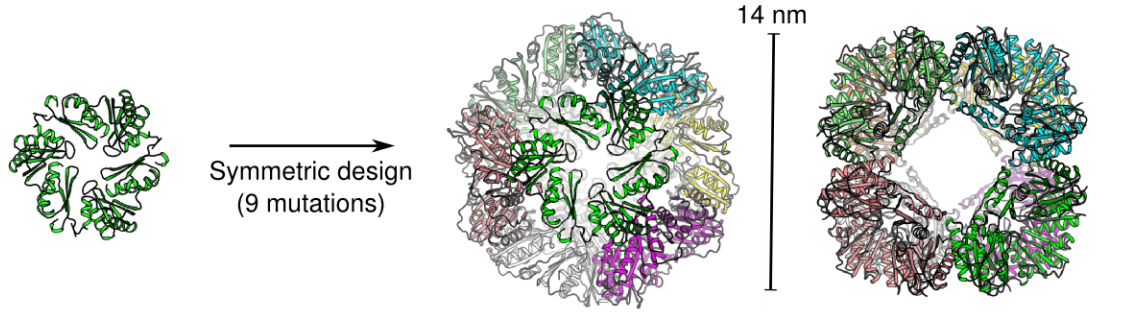

Y101F




Y34F/Y99F/Y101F



# Rosetta MatDes: A general method for designing protein-based materials

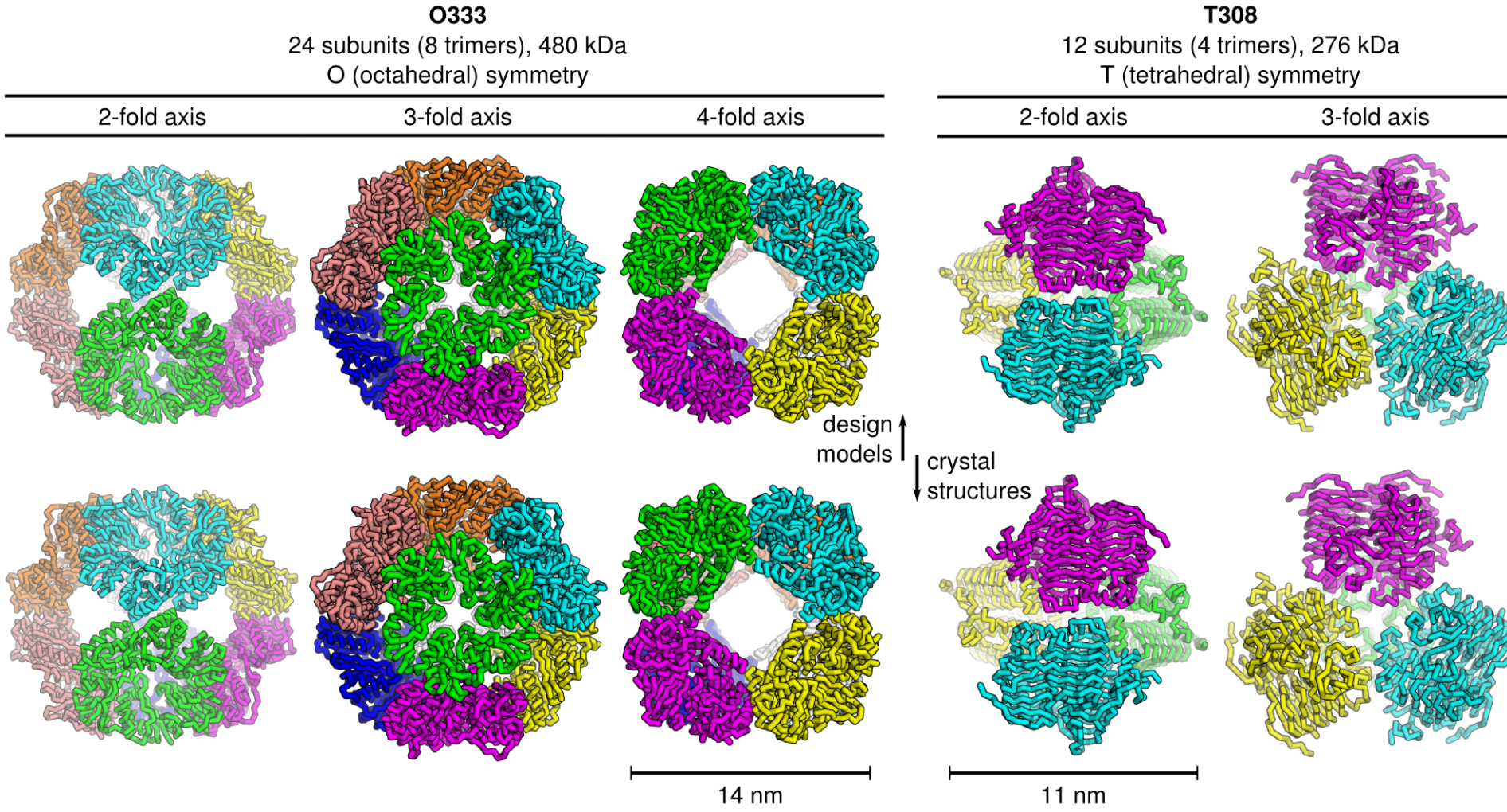



**Wild-type protein**: *S. enterica* PduT
3 subunits, 60 kDa
$C_3$ symmetry

**Designed self-assembling protein**
24 subunits (8 trimers), 480 kDa
O (octahedral) symmetry

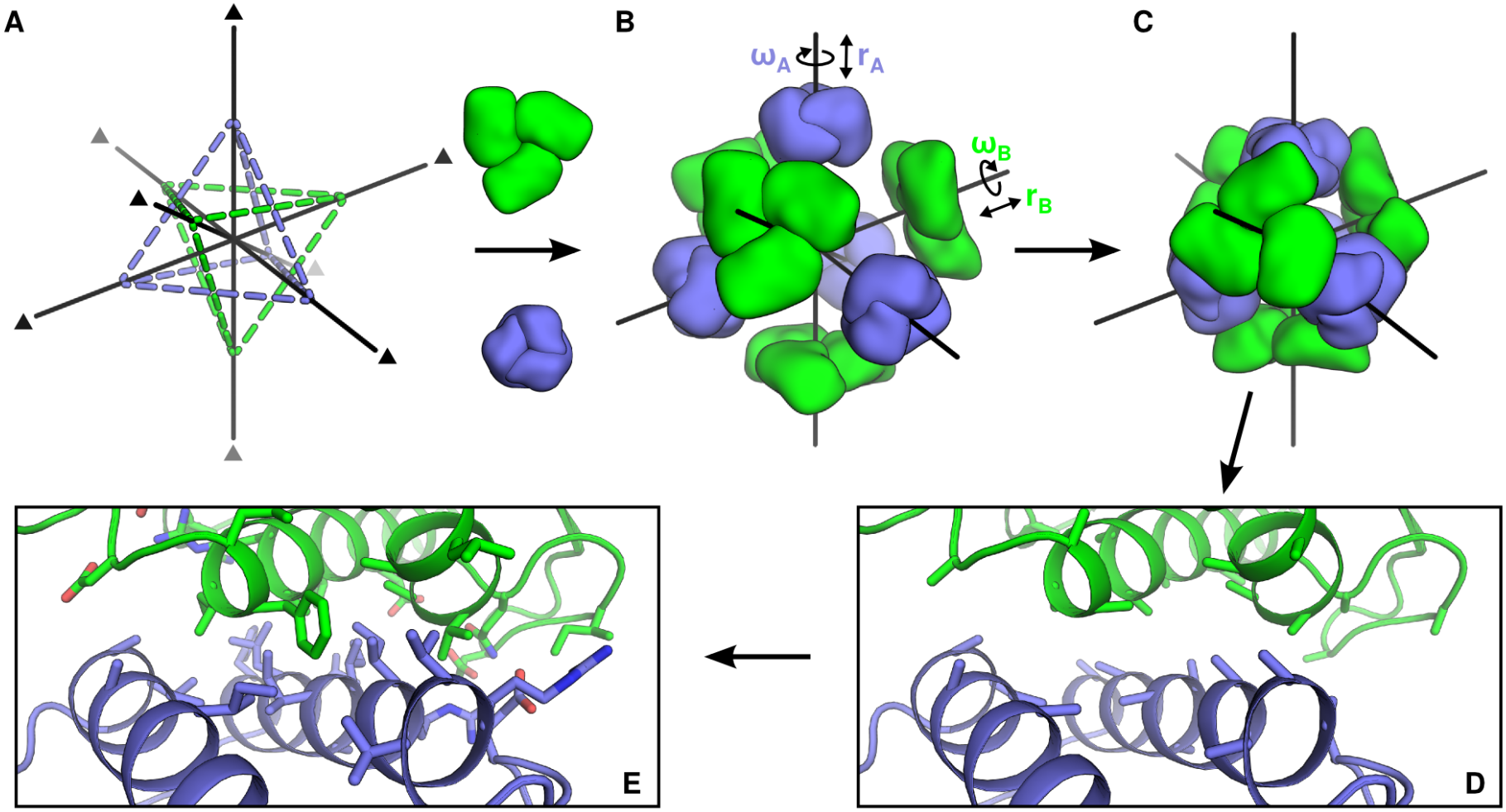
Neil King, Breanna Vollmar, Michael Sawaya

# Crystal structures closely match design models



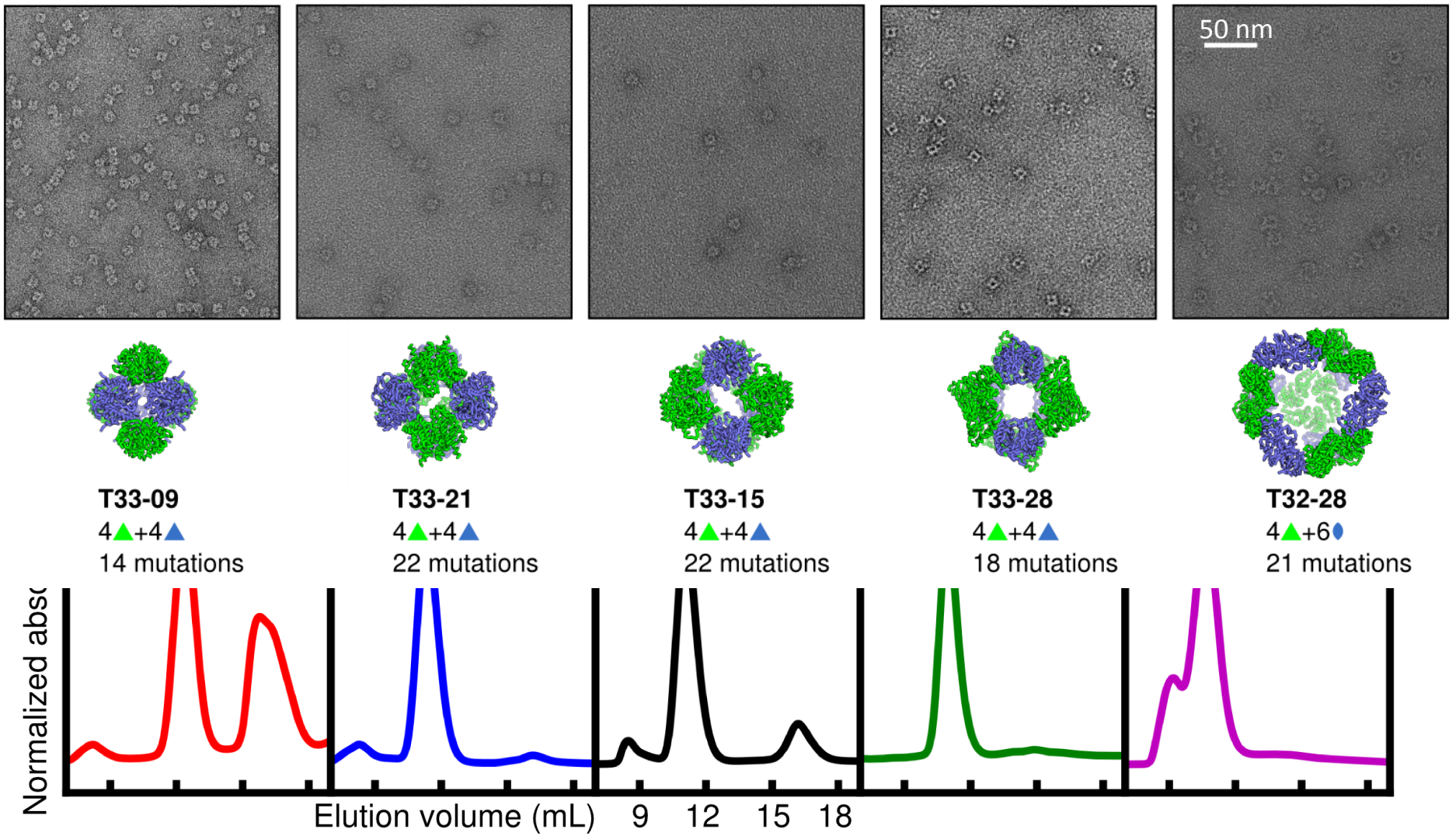

Neil King, Will Sheffler Science 2012

# Design of multi-component materials



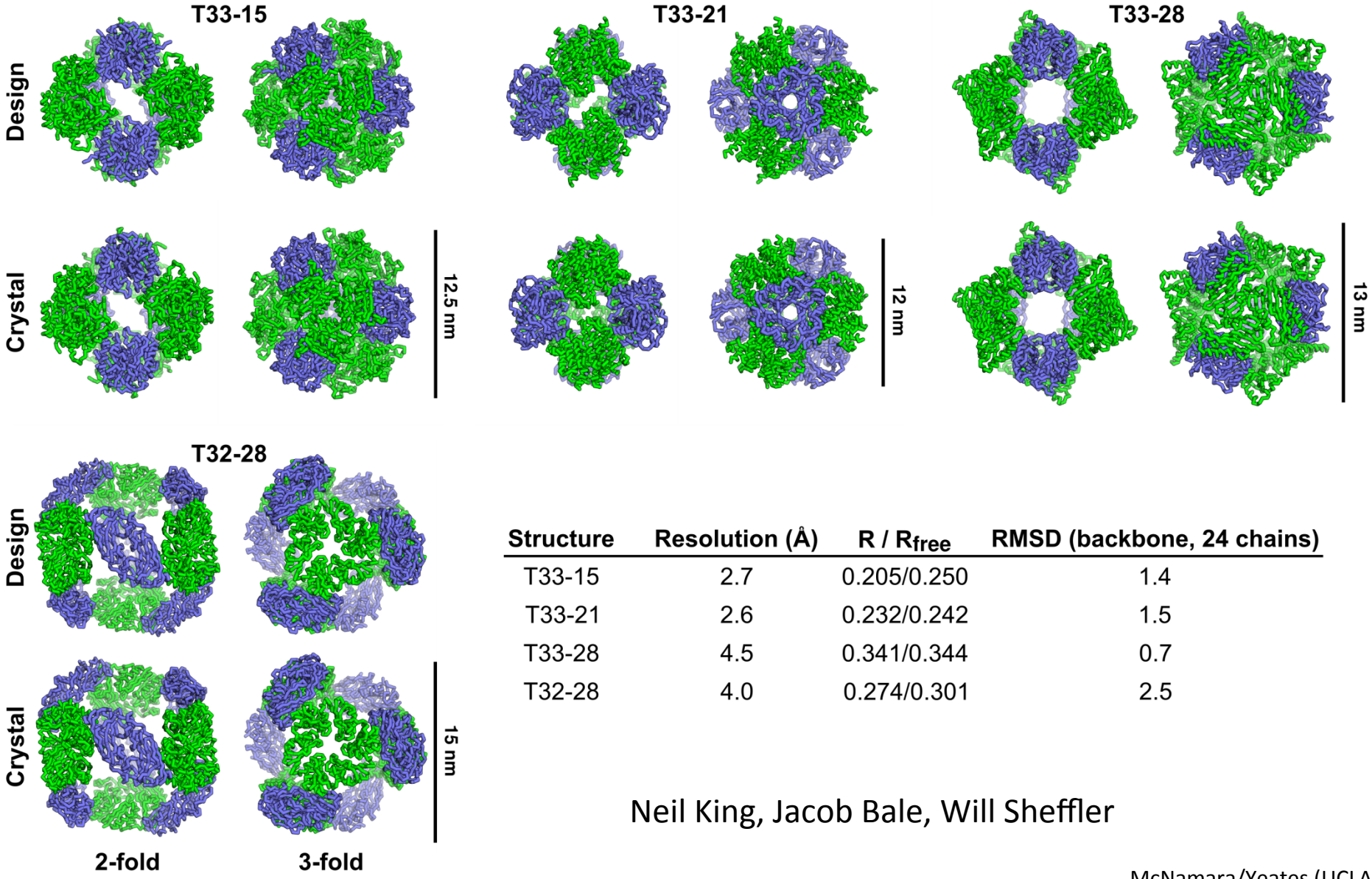

Neil King Jacob Bale, Will Sheffler

# Characterization of designed two component assemblies




Gonen/Gonen (Janelia)

# Crystal structures very close to design models

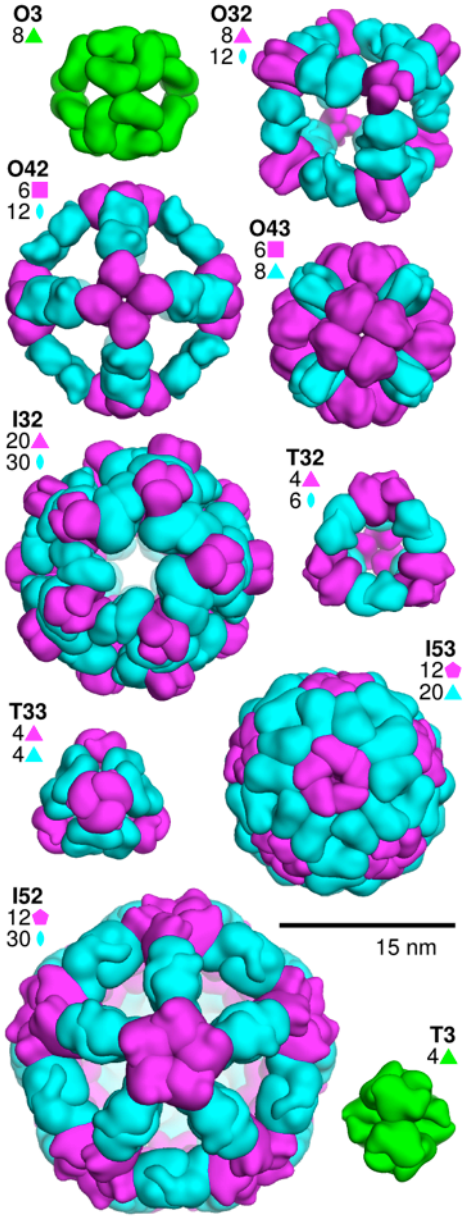



| Structure | Resolution (Å) | R / $R_{free}$ | RMSD (backbone, 24 chains) |
|---|---|---|---|
| T33-15 | 2.7 | 0.205/0.250 | 1.4 |
| T33-21 | 2.6 | 0.232/0.242 | 1.5 |
| T33-28 | 4.5 | 0.341/0.344 | 0.7 |
| T32-28 | 4.0 | 0.274/0.301 | 2.5 |

Neil King, Jacob Bale, Will Sheffler

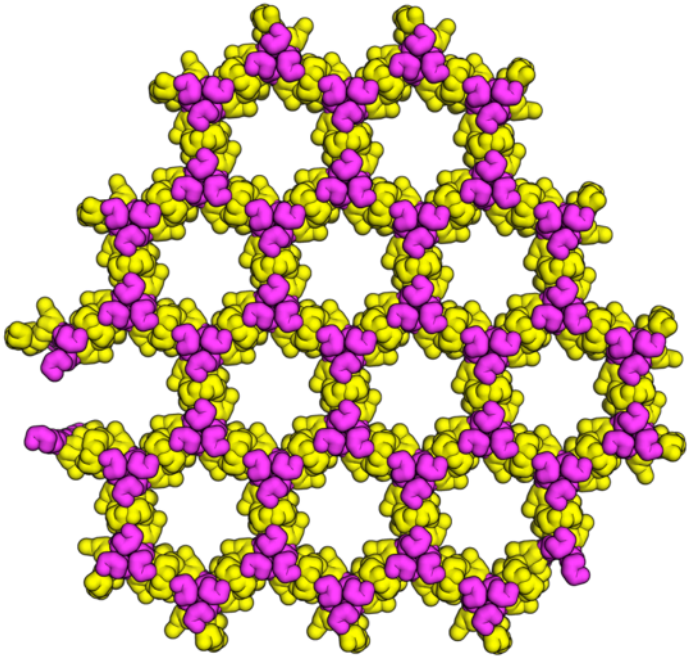
McNamara/Yeates (UCLA)

# Route to improved vaccines and targeted delivery?



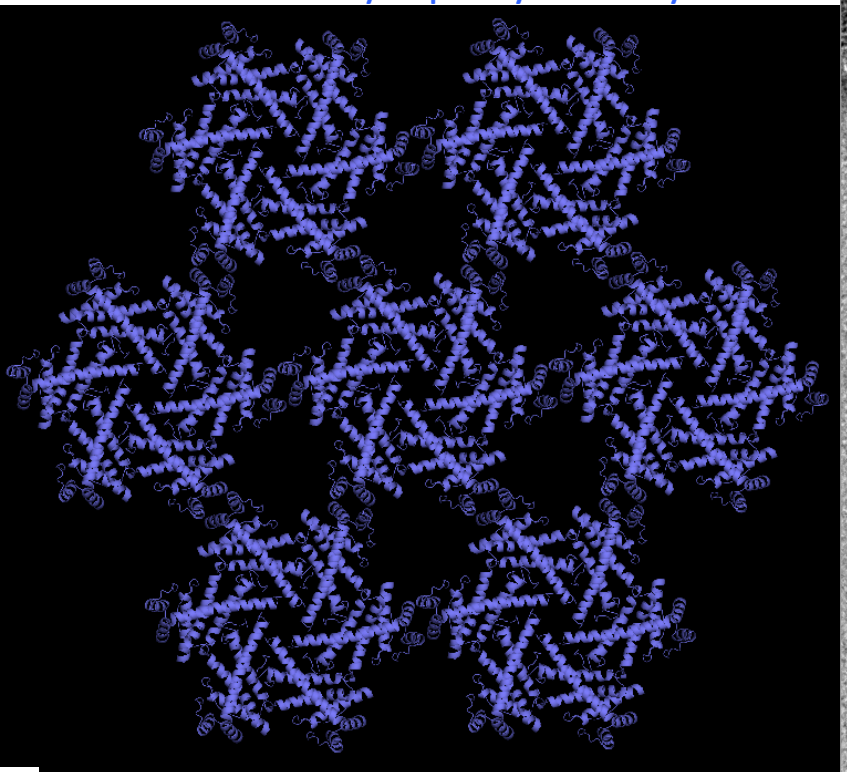

**Nanocages** for targeted delivery, vaccine design, synthetic biology

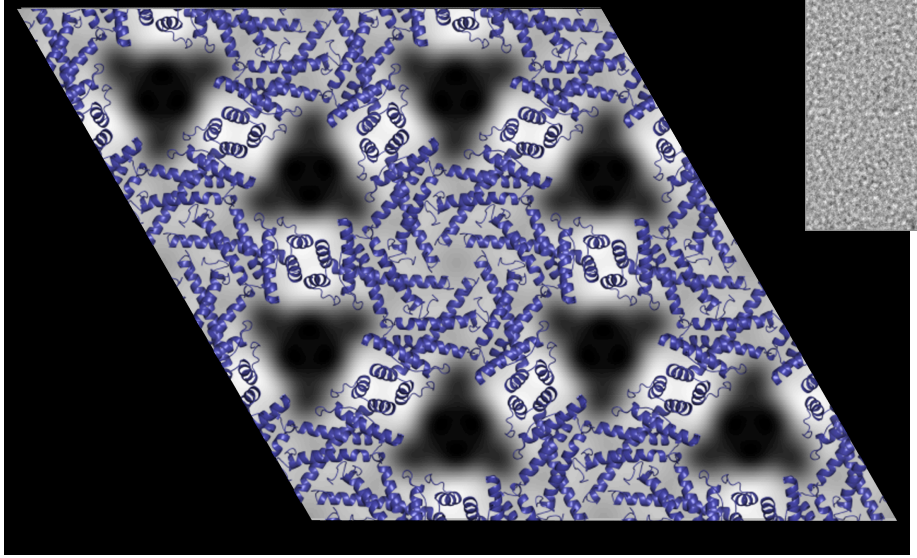
**Nanolayers** for bioactive materials and diagnostics

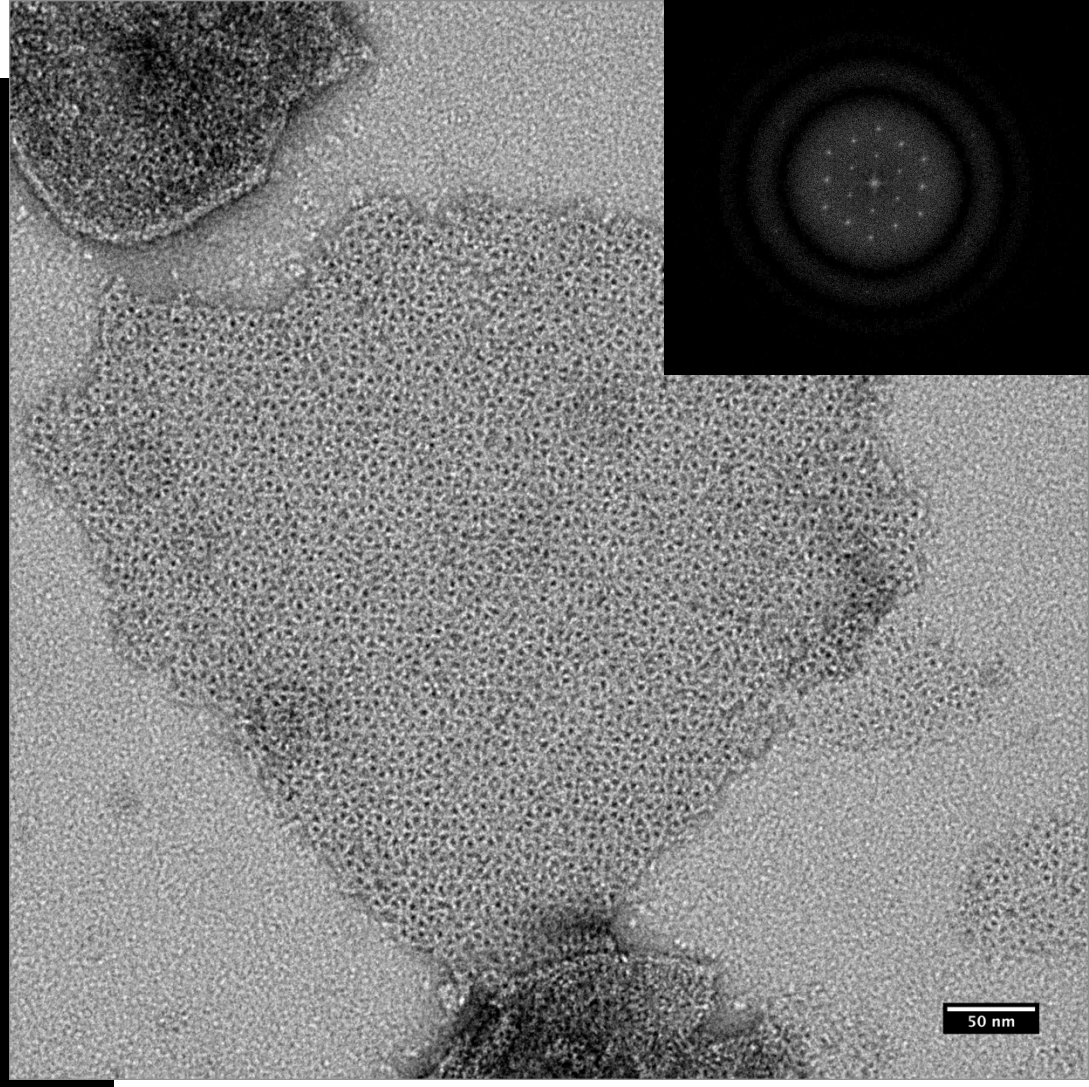


**Nanowires** for molecular or electronic transport

# 2-D Array – p6 Symmetry







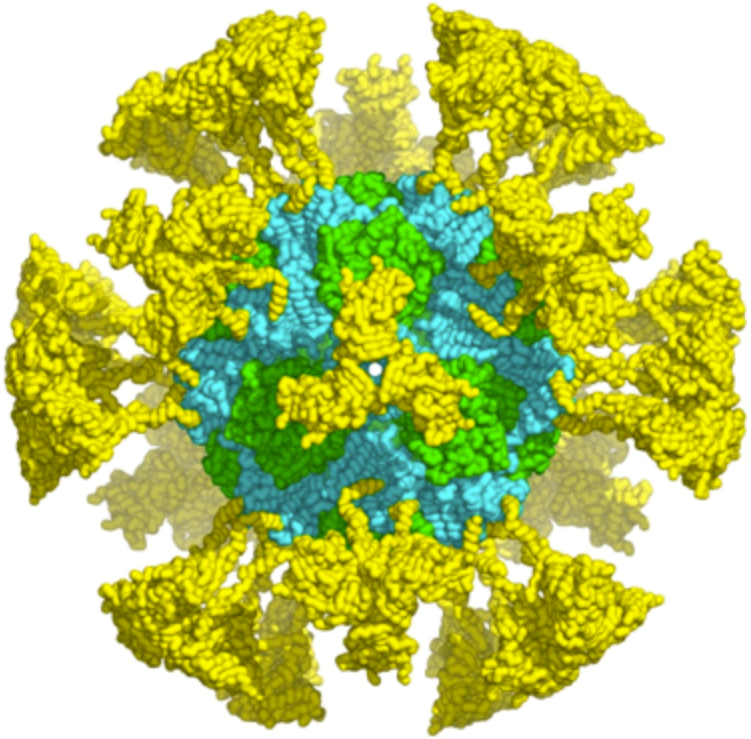

Final Resolution = 14Å
Unit Cell Dimentions, a=b=120A γ=120°
p6 symmetry

# Towards Next Generation Vaccines

- Computationally designed stabilized epitopes that elicit broadly neutralizing antibodies
- Engineered into self assembling two-component virus-like nanoparticles.

SOSIP HIV epitope Trimer

2-Component self-assembling nano-particle.



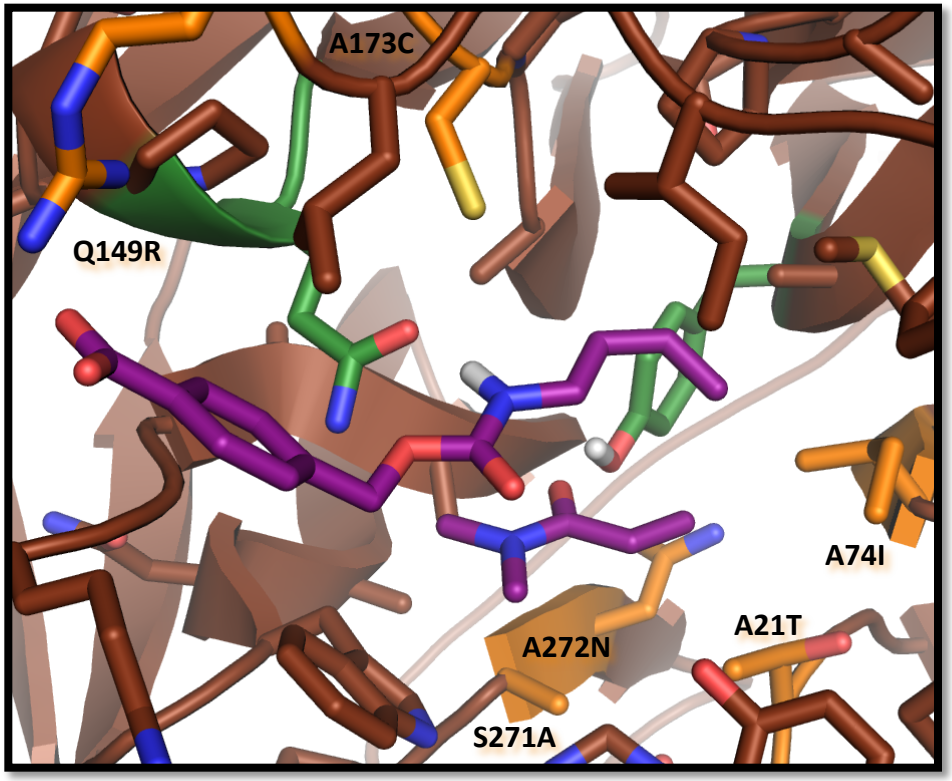
**Can enlist the general public to solve design problems! (FoldIt)**

- Protein structure determination
- Algorithm discovery
- Radical enzyme backbone redesign

# *de novo* designed Diels-Alderase

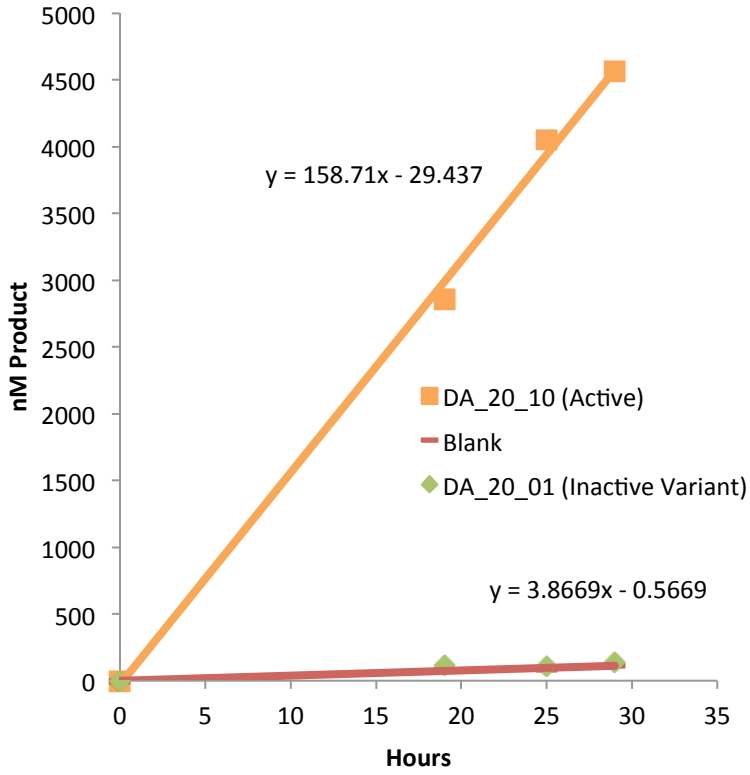
DA_20_10 Active Site View

, catalytic residues



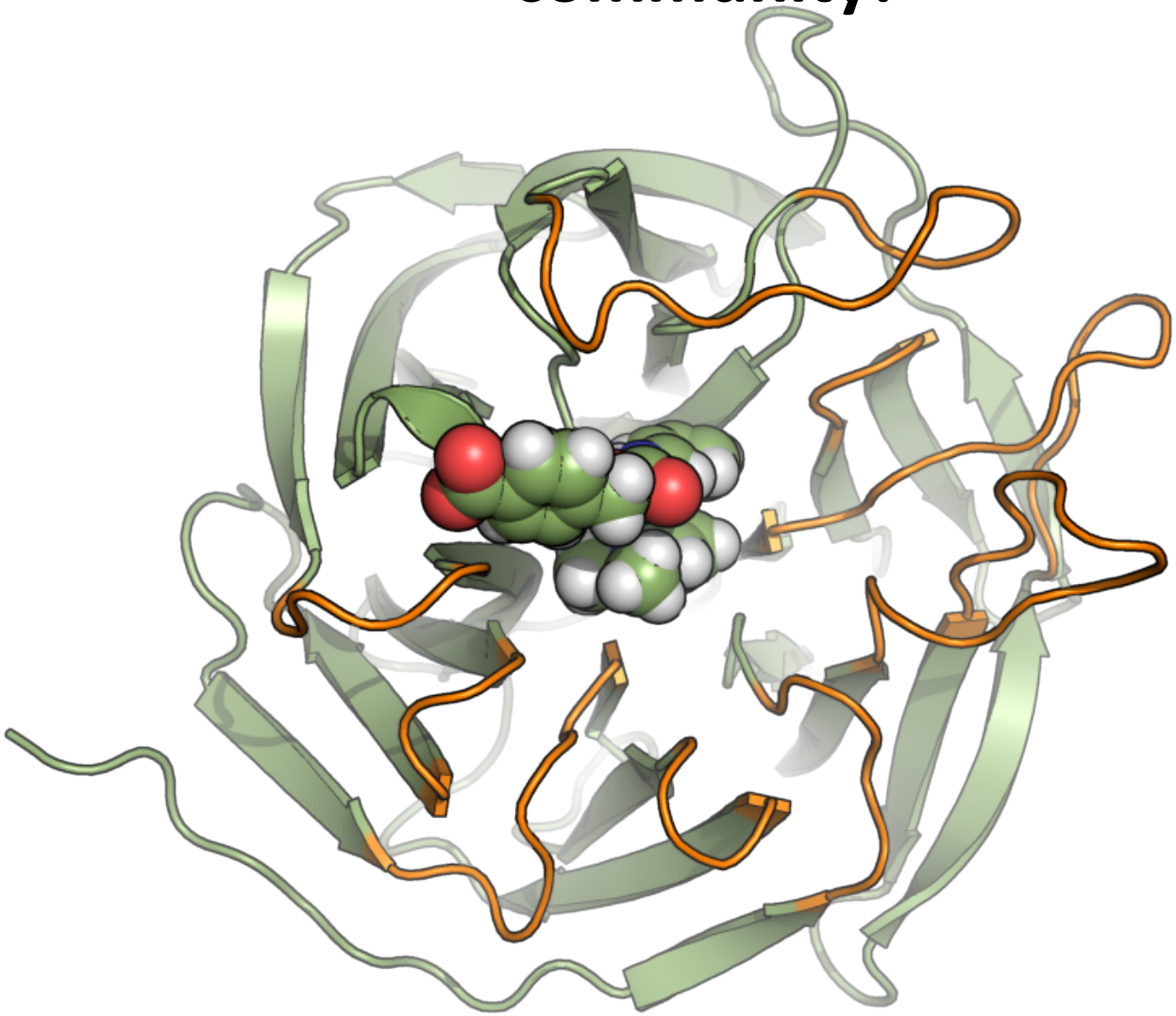

Diels-Alder Reaction Progress Curve

(1x PBS, 298K, 0.1mM Diene, 3mM Dienophile, 20uM Protein)

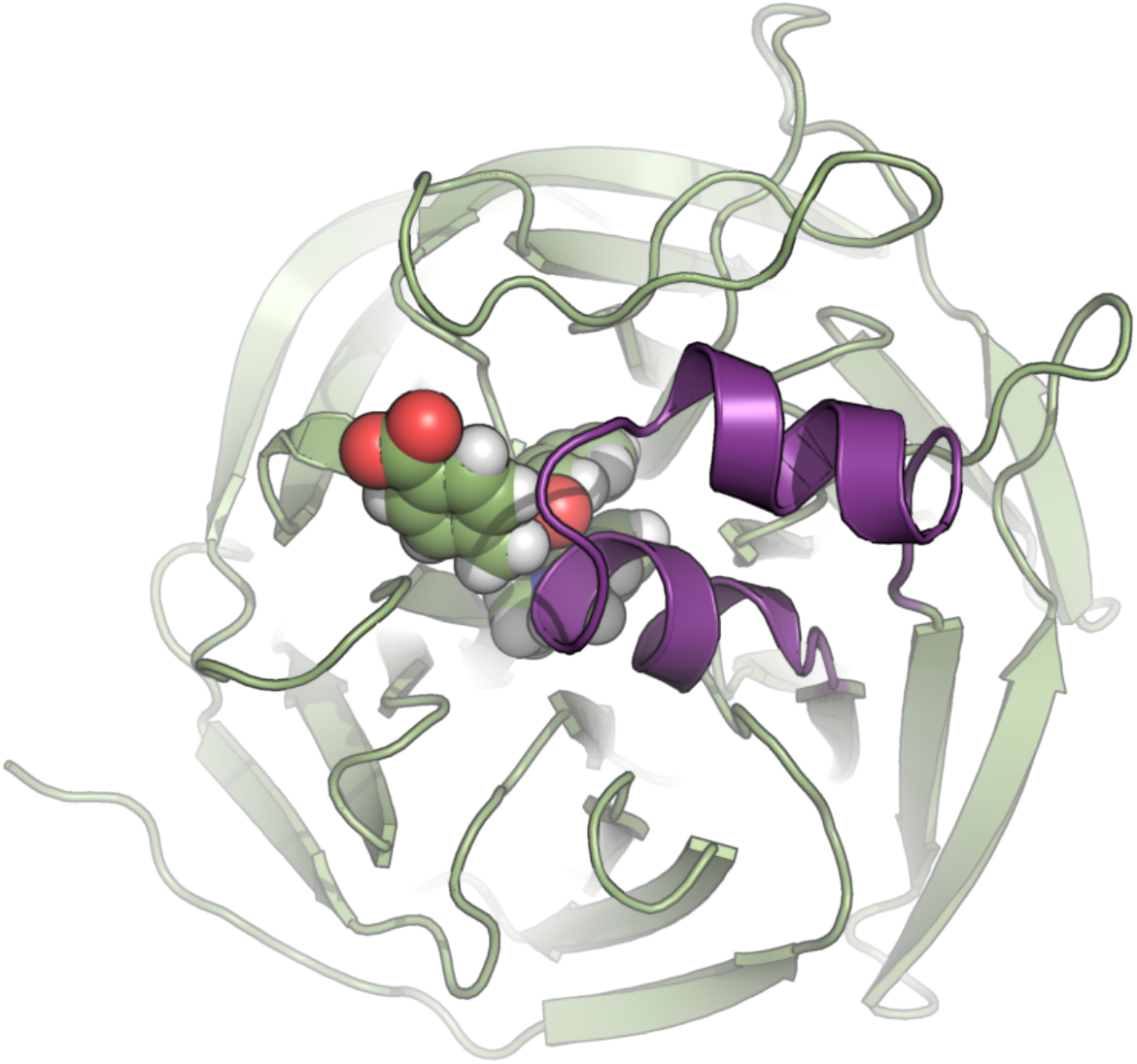



Justin Siegel, Alex Zanghellini, Science 2010

# Can we improve activity of designed Diels Alderase by remodeling active site loops? Lets ask the Fold.it community!



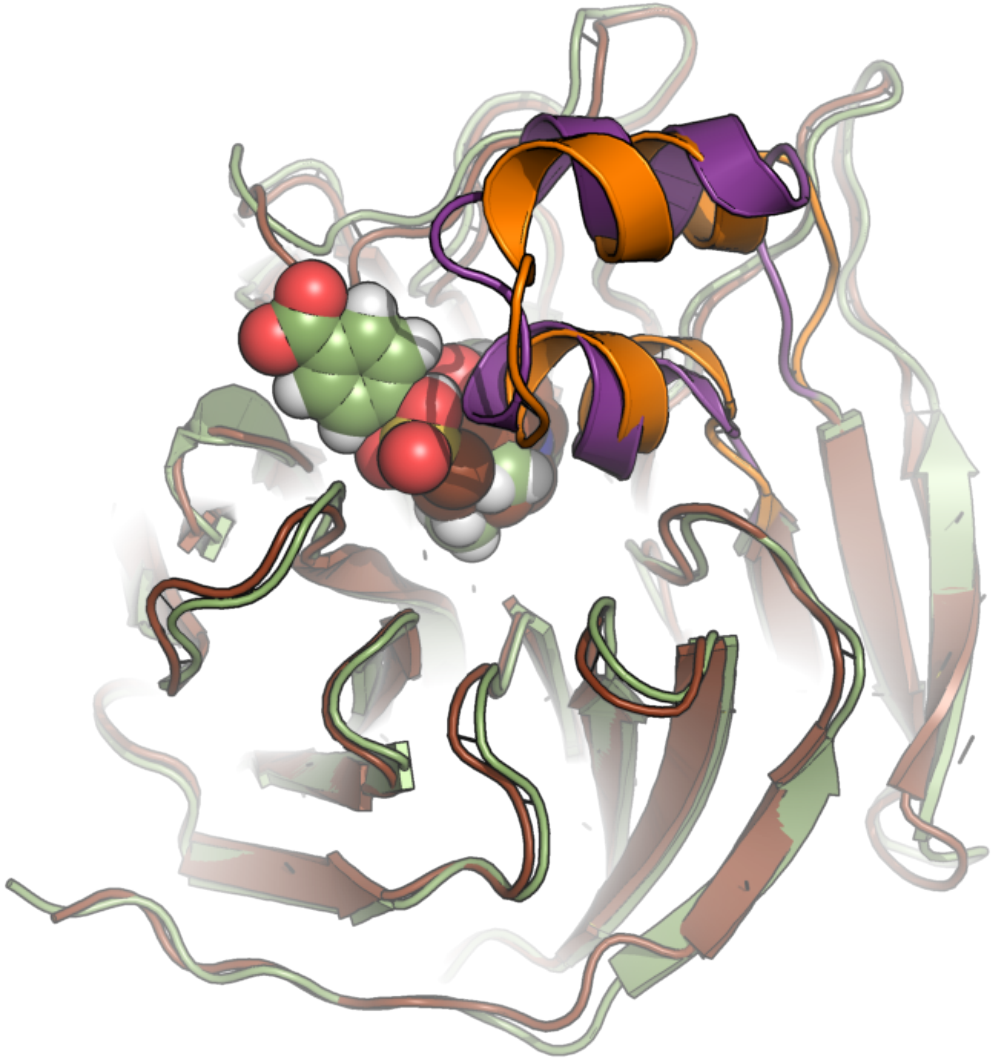
# Helical hairpin insertion leads to 18-fold greater catalytic activity greater than DA_20_10



Chris Eiben
Justin Siegal
Foldit Players!

# Crystal structure shows both helices in 24-amino acid designed loop are placed correctly

DESIGN (GREEN/PURPLE) vs. CRYSTAL STRUCTURE (BROWN/GOLD)



Jacob Bale
Barry Stoddard