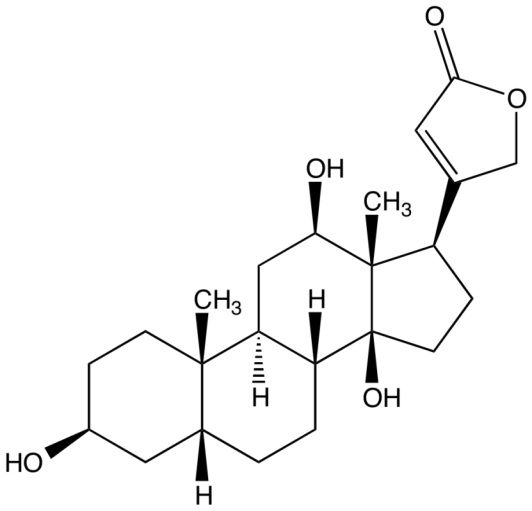
De Novo Design of Ligand Binding

Christy Tinberg

with Sagar Khare, Jiayi Dou, Jorgen Nelson, and David Baker

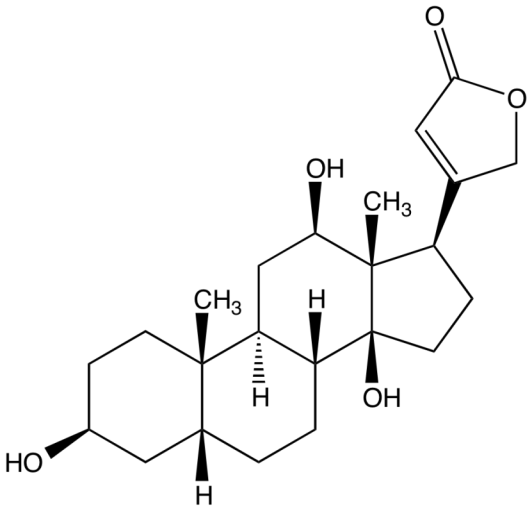
RosettaCon 2012

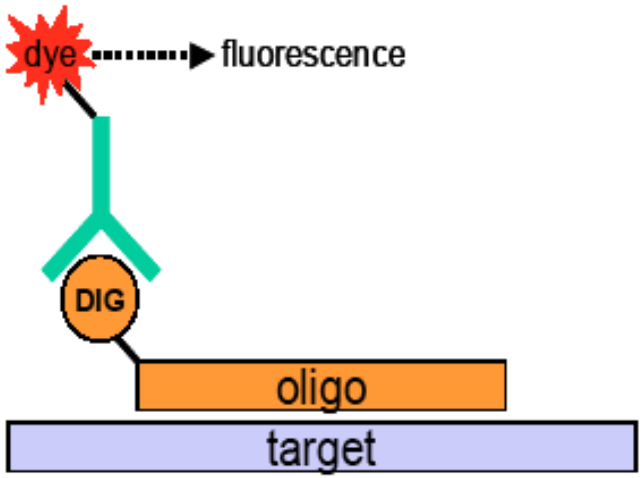
Target: Digoxigenin



cardioactive glycoside (GSK)

Target: Digoxigenin

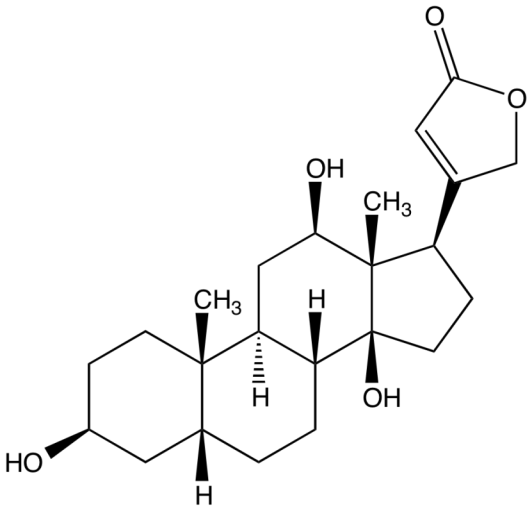


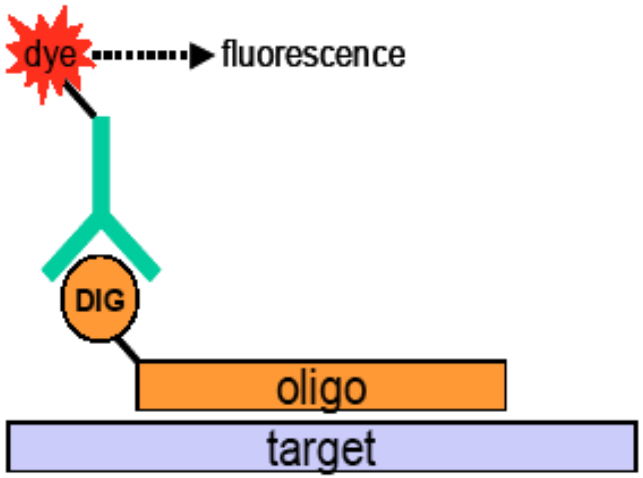


cardioactive glycoside (GSK)

biochemical non-radioactive labeling reagent

Target: Digoxigenin





cardioactive glycoside (GSK)

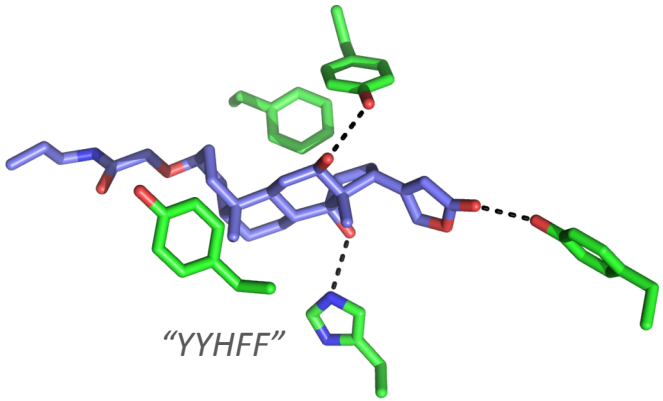
biochemical non-radioactive labeling reagent

- Large and rigid
- Relatively hydrophobic, but has 4 hydrogen bonding groups
- 3 PDB structures available (2 antibodies, 1 engineered lipocalin)
- Experimentally feasible (good positive controls; reagents are commercially available)

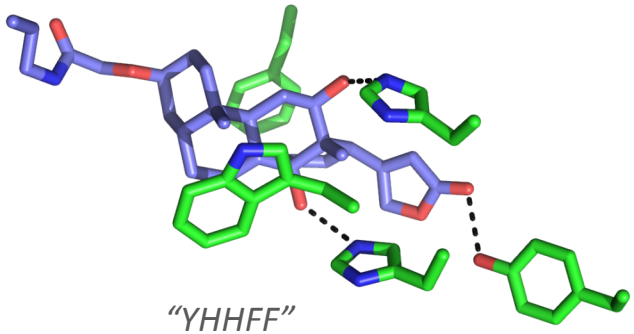
Digoxigenin Design Considerations

- Protein-ligand interaction energy
- Shape complementarity of binding site
- Pre-organization of binding site in a “binding-competent” conformation
- Protein stability in the absence of ligand

Binding Site Geometry



"YYHFF"

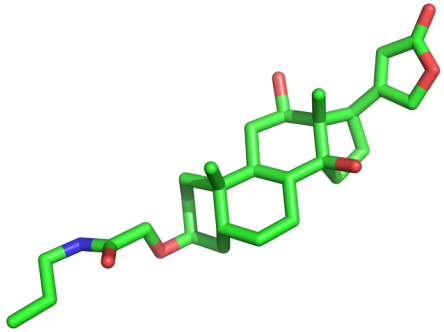


"YHHFF"

Scaffold Library (401)

Enzdes Scaffold set
DALI homologs of previously
well-expressed scaffolds
from Baker lab

Ligand



Matches

Fixed-bb Design

Raw Designs

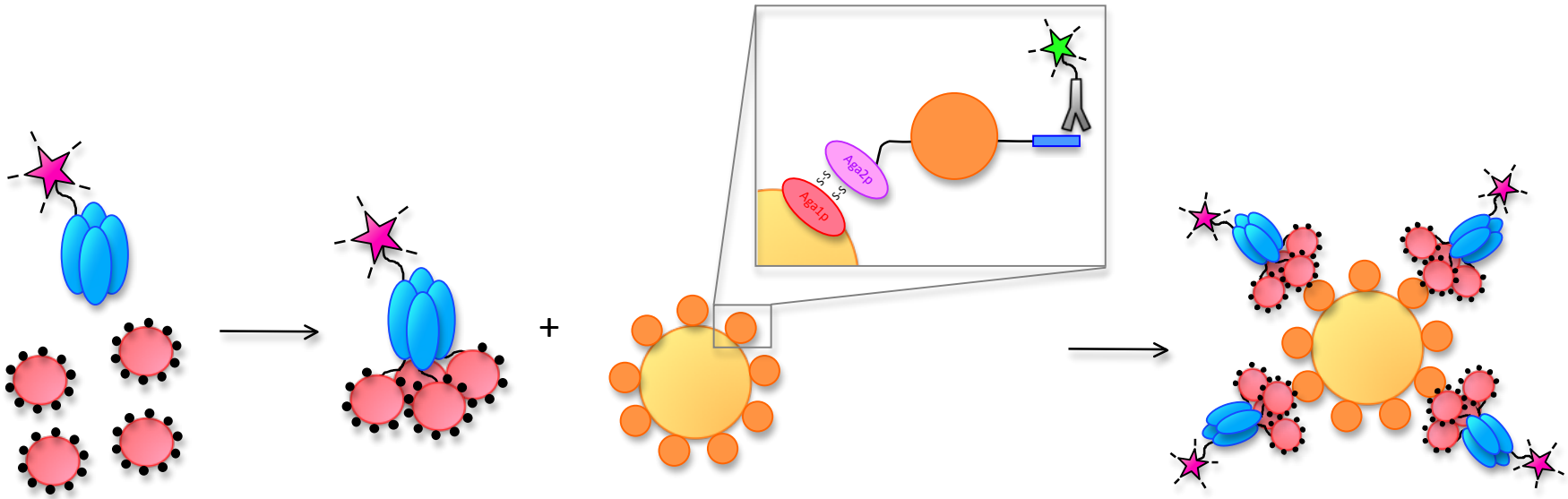
Foldit manipulation

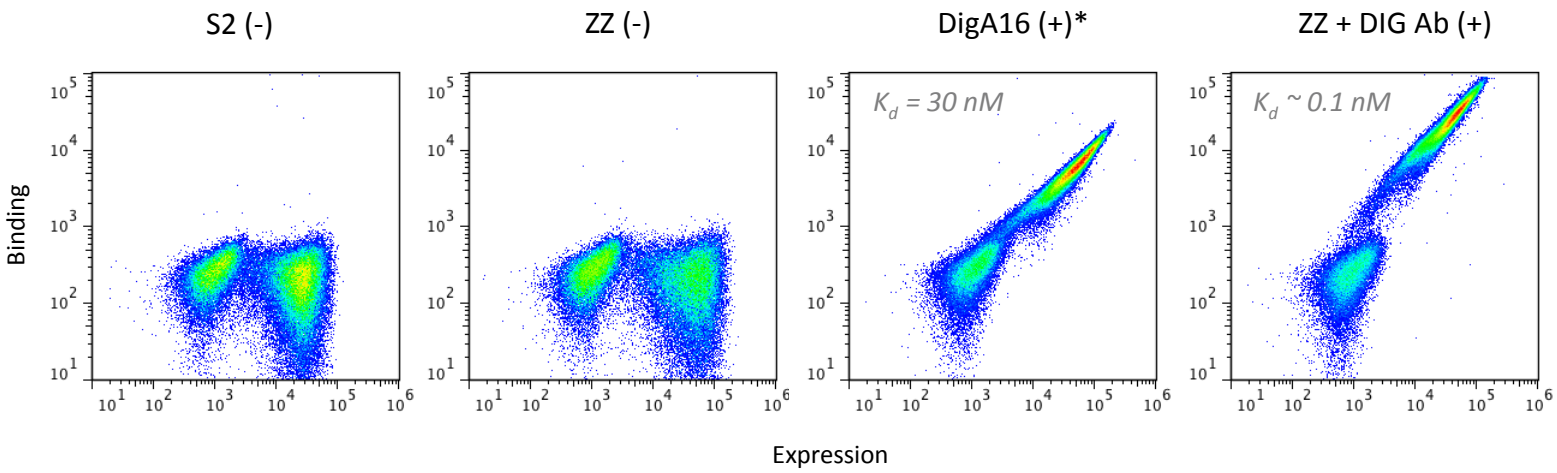
Refined Designs

17 designs in 14 scaffolds ordered

w/ Sagar Khare, Jiayi Dou

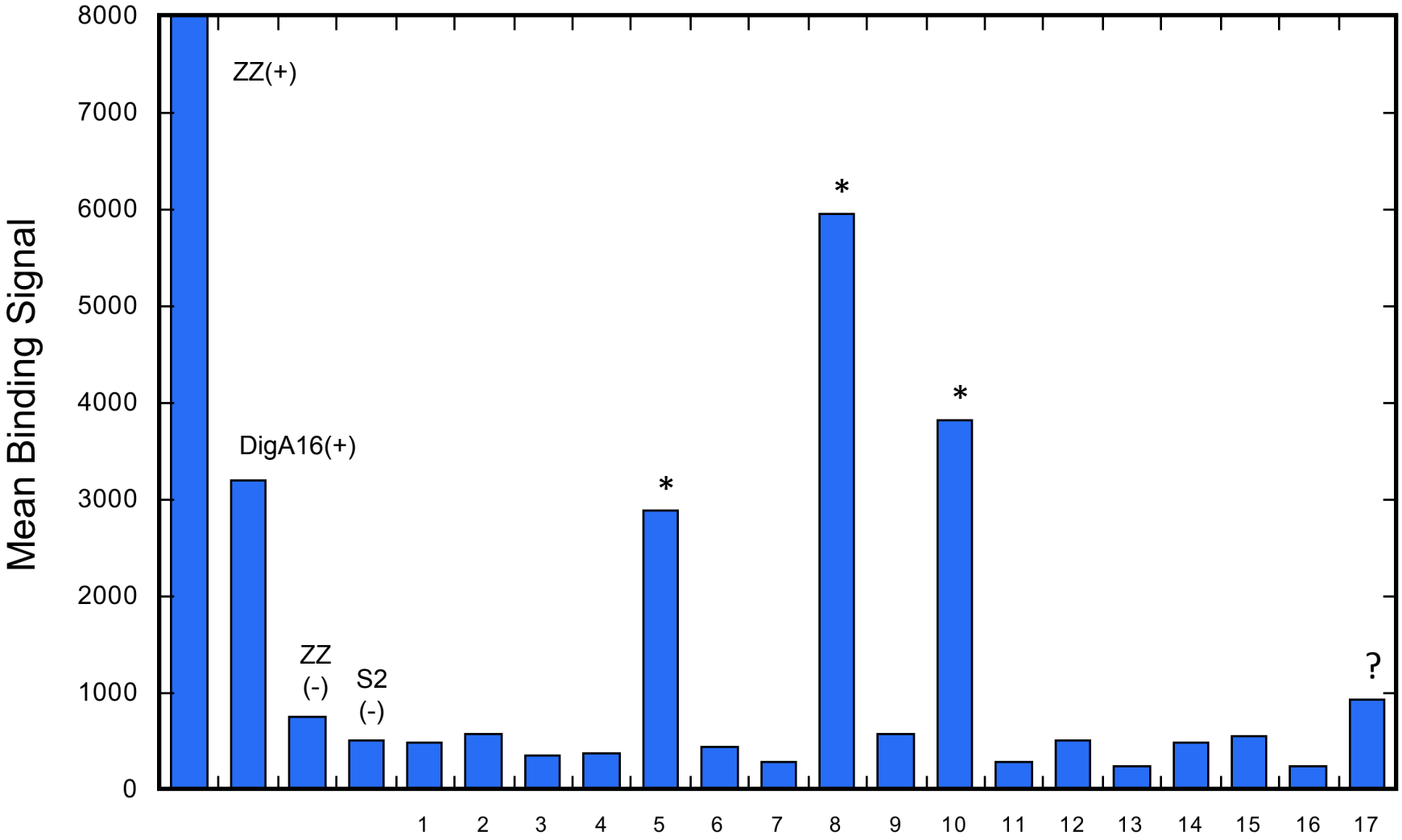
Avidity-based Yeast Surface Display Design Screening



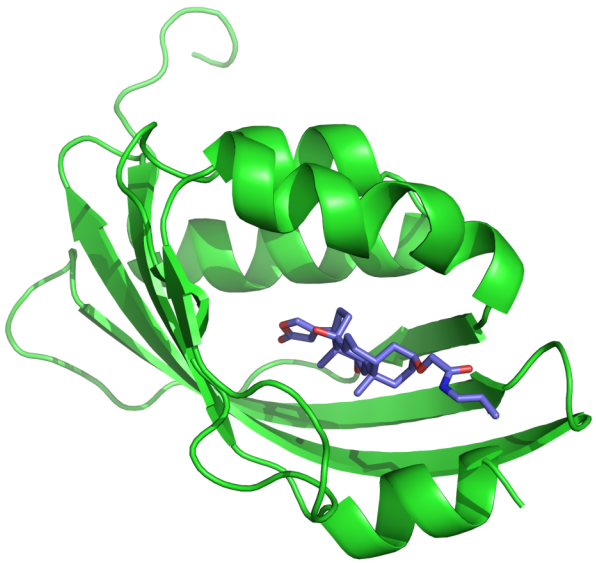


*S. Schlehuber, G. Beste, and A. Skerra. *JMB*, 2000

Three Hits Identified by Yeast Surface Display

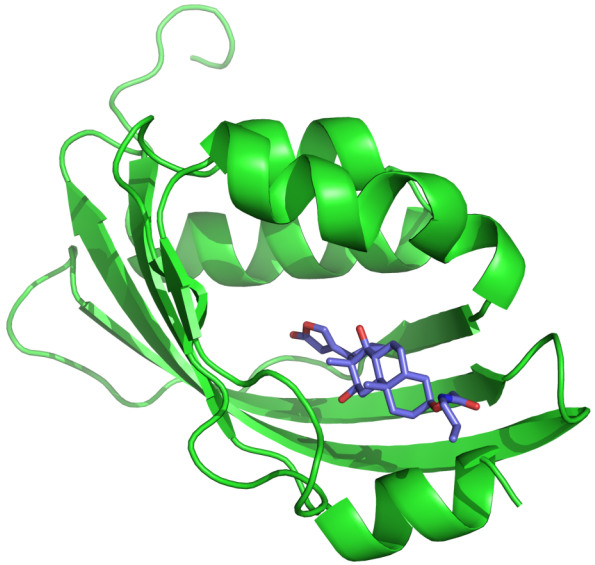


Three DIG binders



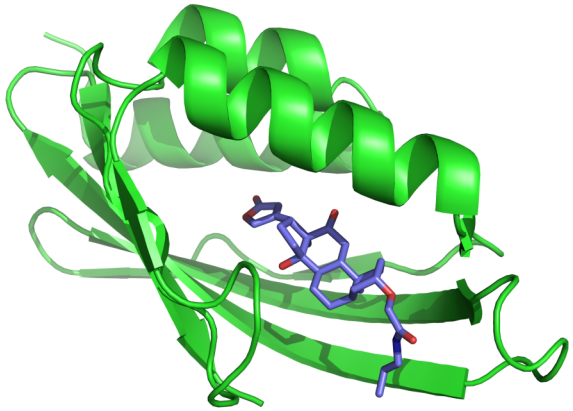
DIG_10
PDB 1z1s

10 mutations



DIG_5
PDB 1z1s

11 mutations



DIG_8
PDB 3hk4

19 mutations

Two scaffolds are homologs in the nuclear transport factor 2 (NTF2) fold-class

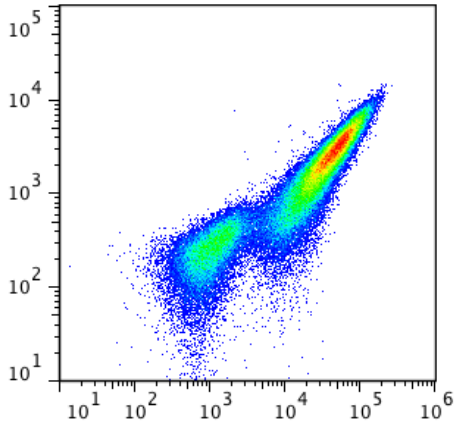
Rosetta Energy Function Predicts the Top Binders

Design	Scaffold	# Mutations	Sc	DIG_total_score
DIG_8	3hk4	19	0.67	-8.46
DIG_10	1z1s	10	0.59	-8.38
DIG_1	1gy7	16	0.73	-8.28
DIG_5	1z1s	11	0.68	-7.84
DIG_11	1zo2	13	0.64	-7.74
DIG_7	3gwr	13	0.67	-6.82
DIG_15	3fmz	13	0.67	-6.76
DIG_3	1pvx	12	0.69	-6.45
DIG_2	1mve	15	0.60	-6.30
DIG_4	3b4o	15	0.46	-6.02
DIG_16	3gwr	10	0.70	-5.74
DIG_14	3e5z	16	0.60	-5.73
DIG_13	2ox1	18	0.63	-5.57
DIG_6	3cu3	14	0.58	-5.42
DIG_17	3cu3	7	0.56	-4.80
DIG_9	1i60	19	0.51	-4.44
DIG_12	2owp	12	0.67	-4.24

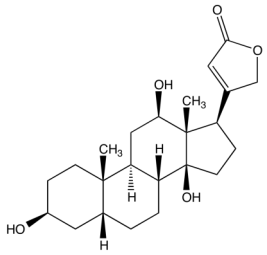
Control Experiments Confirm Authenticity of Hits

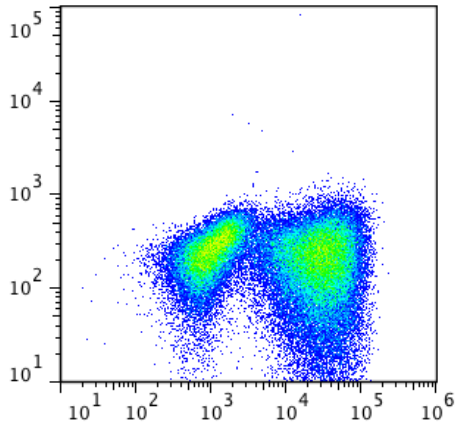
DIG 10

- free DIG

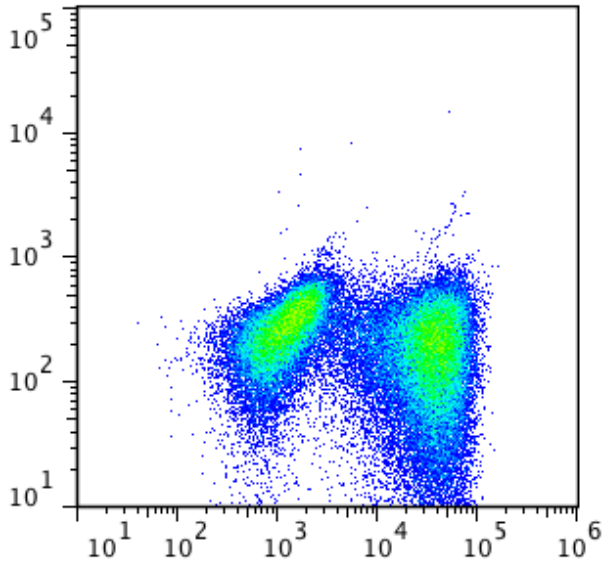


+ 760 µM free DIG

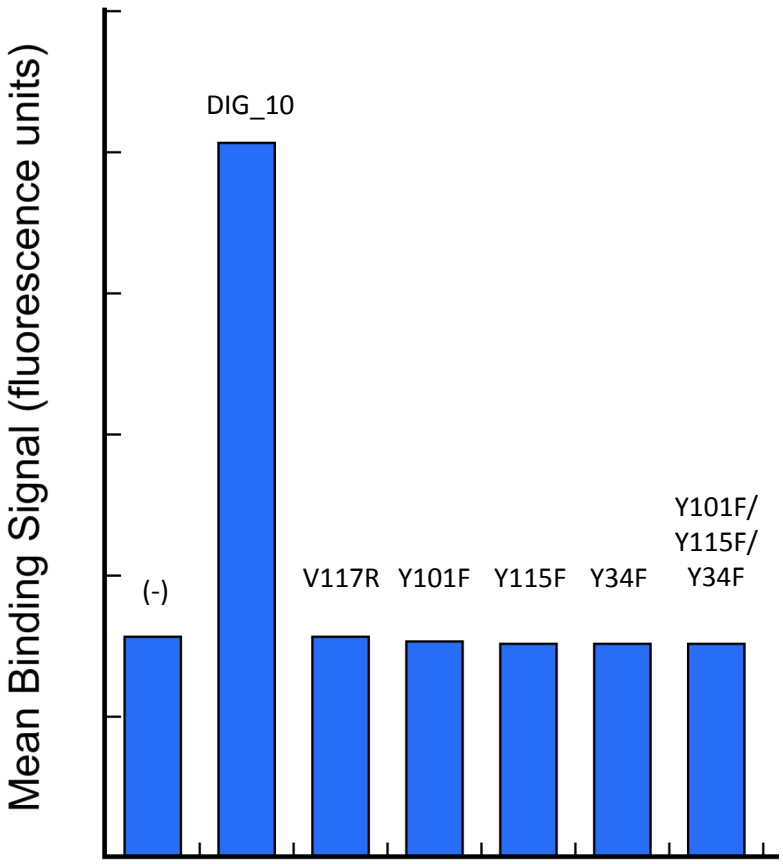


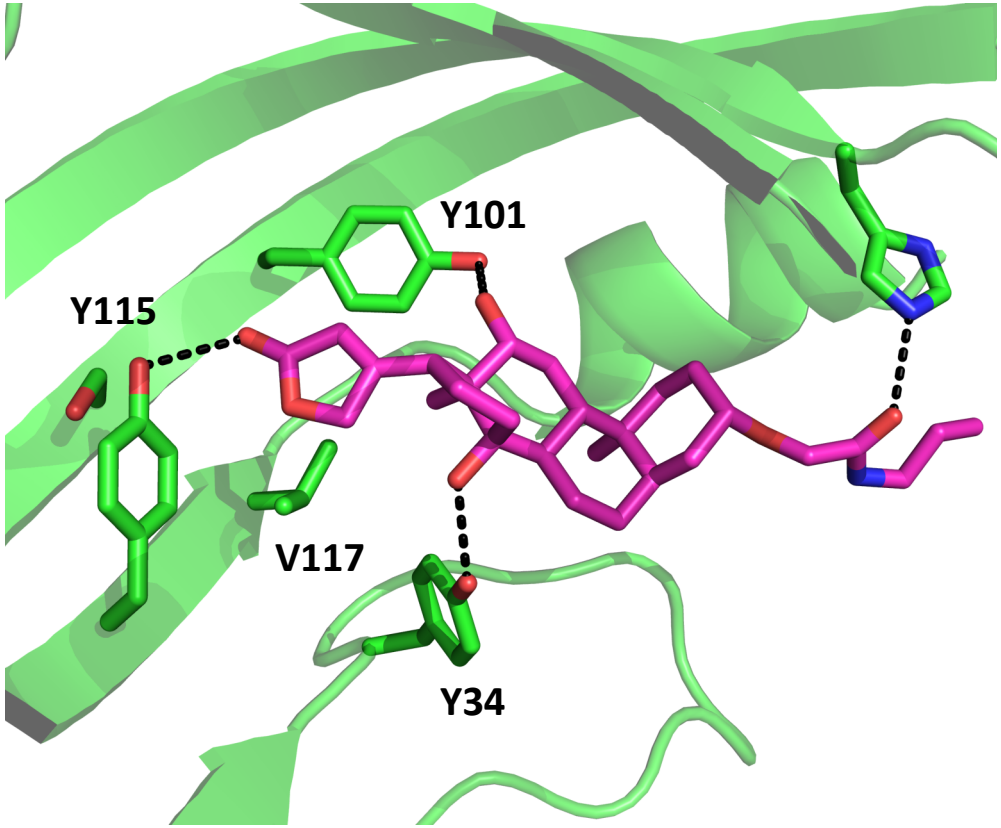


Scaffold (1z1s) Control



Knockout Mutations Support the DIG_10 Binding Model

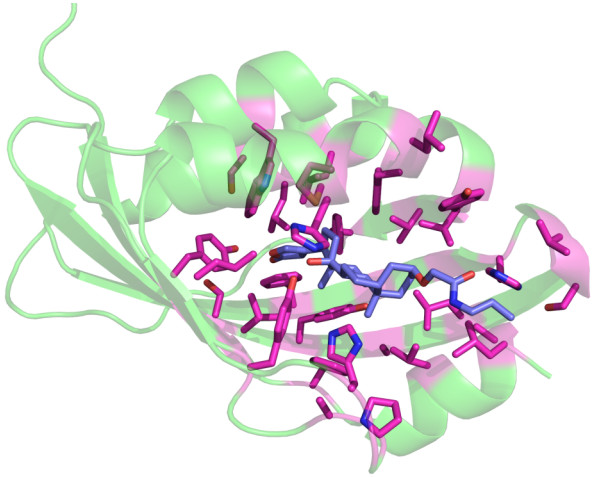




DIG binds in the intended pocket

All three H-bonding tyrosines are important for binding

Affinity Maturation Strategy



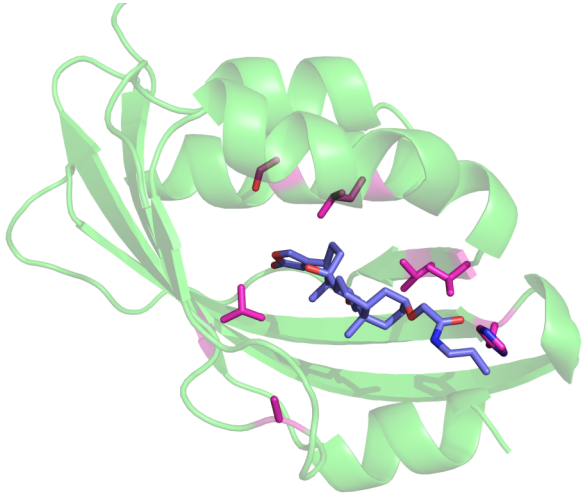
Site Saturation Mutagenesis

Amino acids within 7 Å of DIG mutated to every amino acid

3-4 permissive rounds of FACS

Combinatorial Mutagenesis

Important positions identified in SSM library combined

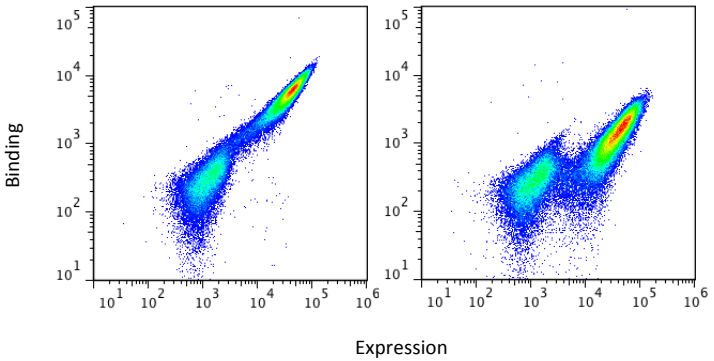


3-4 rounds of FACS
Selection stringency increased in each round

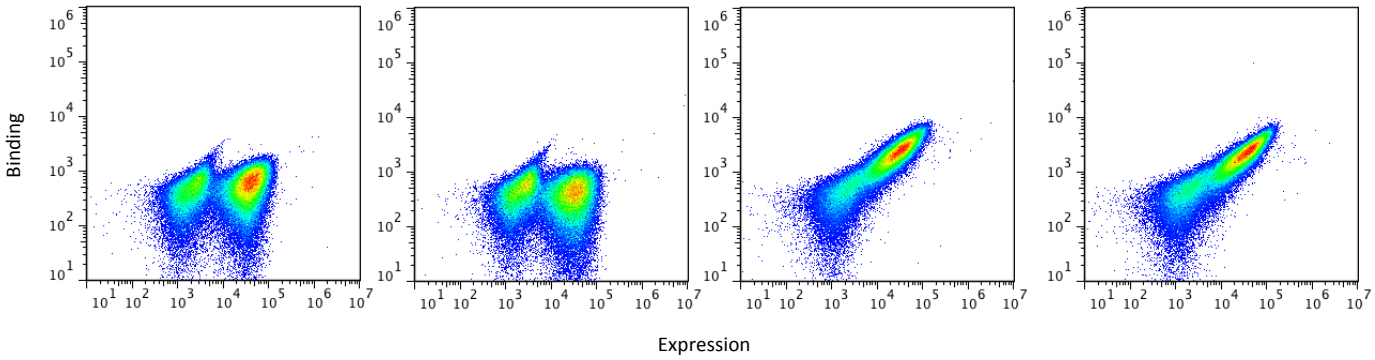
1 (or a few) winner(s)

Two rounds of DIG_10 directed improves binding affinity on yeast

Before directed evolution (3 µM DIG-BSA label)



After directed evolution (5 pM DIG-RNase label)



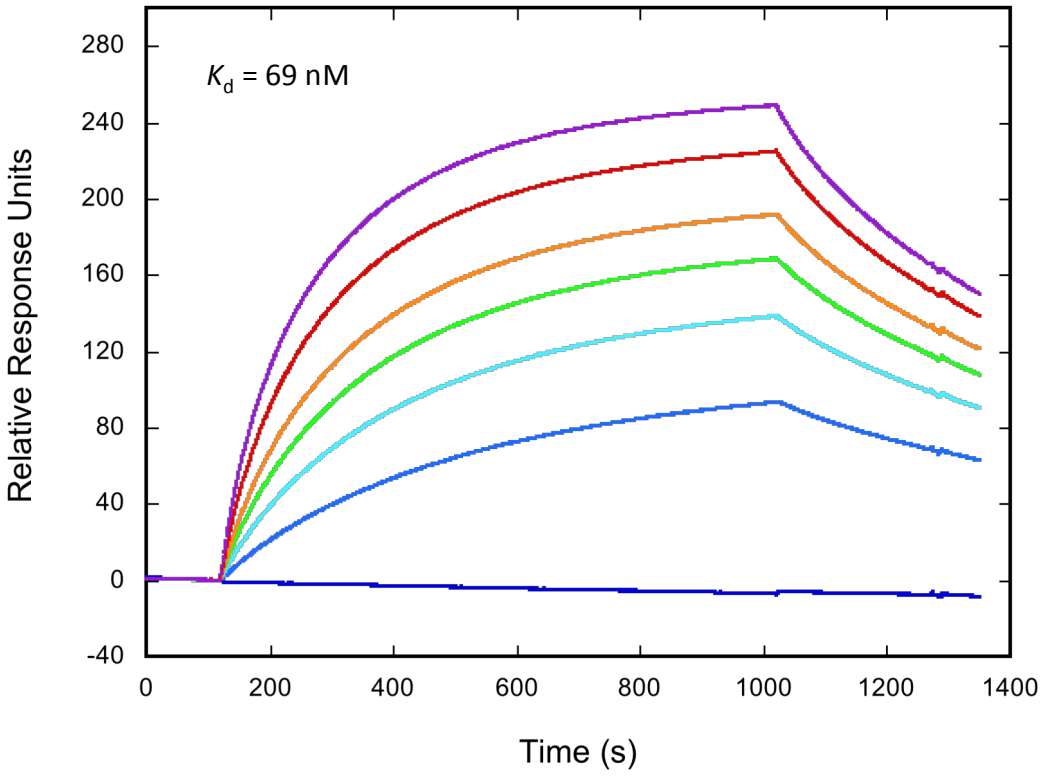
DigA16
+ control (K_d = 30 nM)

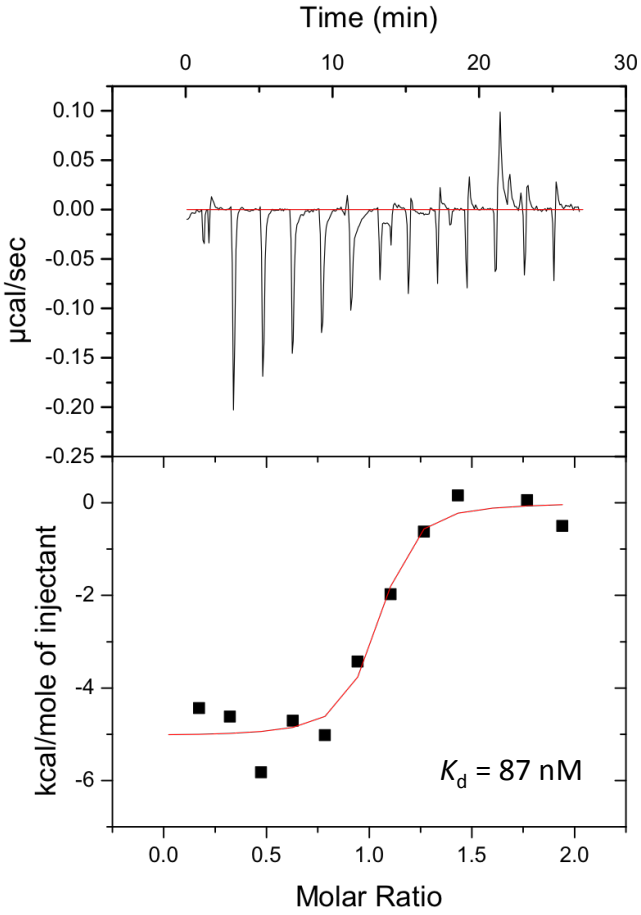
DIG10

DIG10.1a
evolved variant (S10A; V62M; V64I; V117L; A127P)

DIG10.1b
evolved variant (S10M; V62M; V64I; V117L; A127P)

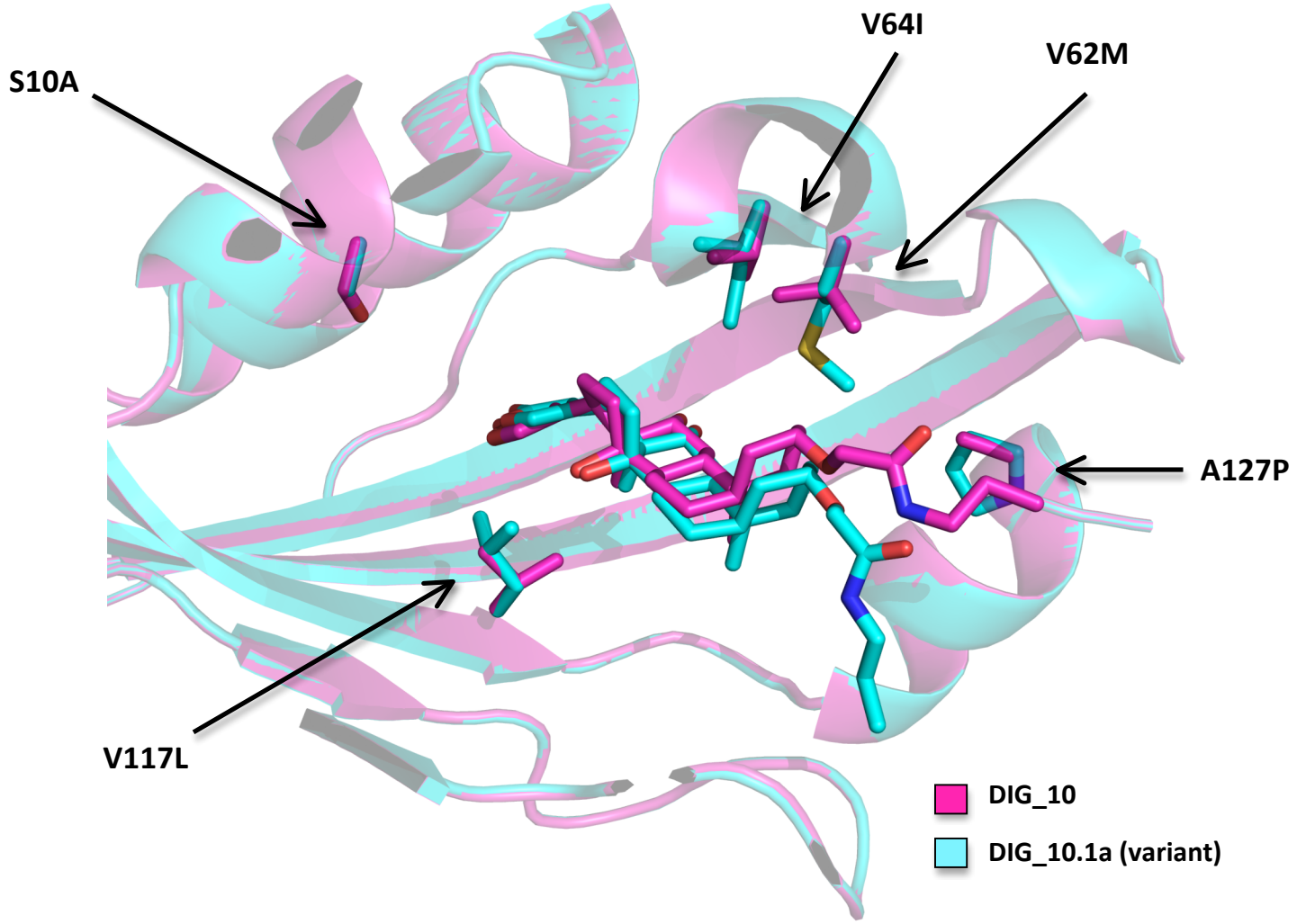
Evolved Variants Bind DIG with Nanomolar Affinity





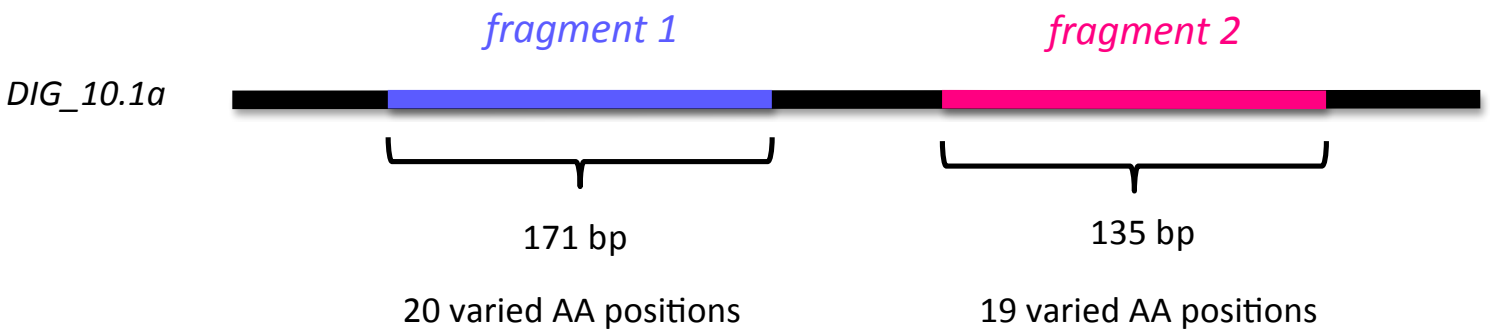
Preliminary solution experiments suggest mid-nanomolar affinities for evolved variants

Evolved Variant Mutations Improve Hydrophobic Packing



Design Strategy for Next-generation Sequencing Libraries

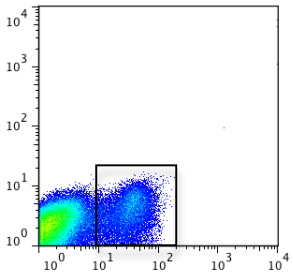
Next-generation sequencing can provide us with a “binding fitness ” landscape

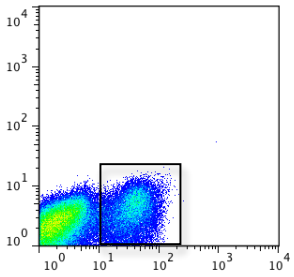


We hope to achieve all possible single point mutations plus a random subset of intra-fragment doubles

Selection Strategy for Next-gen Sequencing Libraries

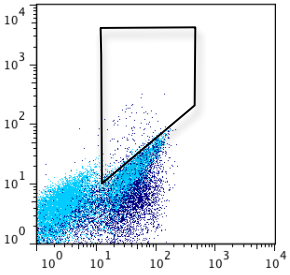
fragment 1 library *fragment 2 library*

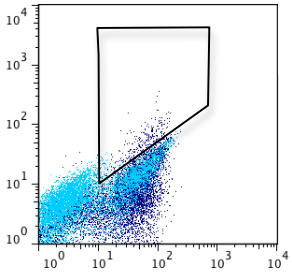




Selections for expressing populations

“naïve library”

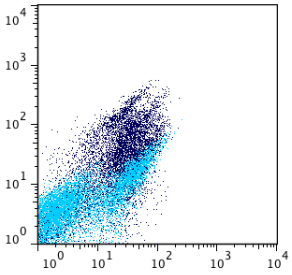


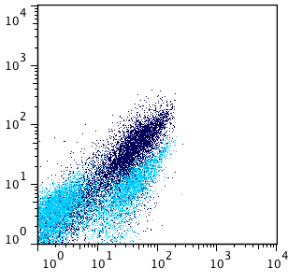


Library

DIG_10.1a (input clone)

Selections for variants with binding signals greater than that of the input clone (two rounds)



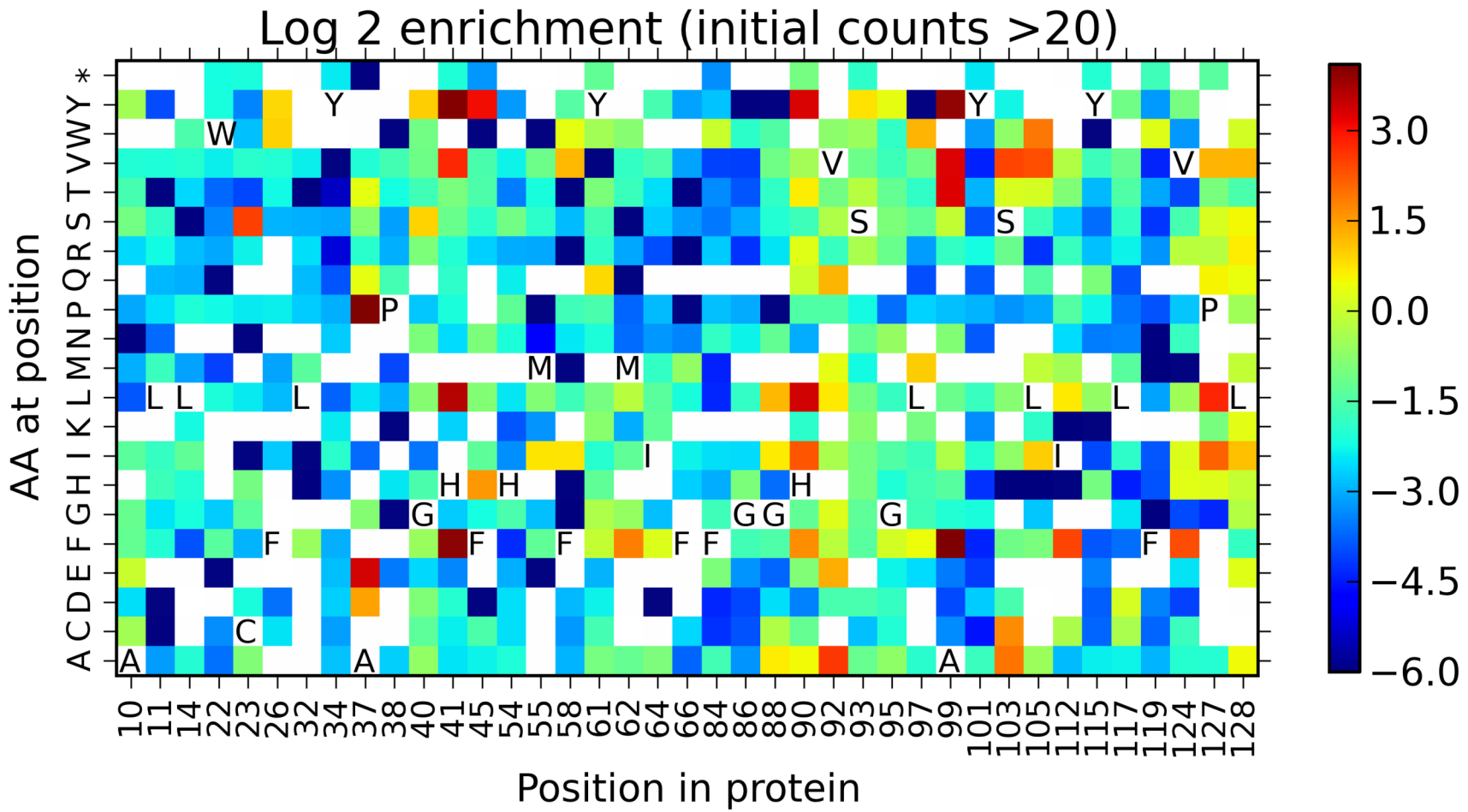


Library

DIG_10.1a (input clone)

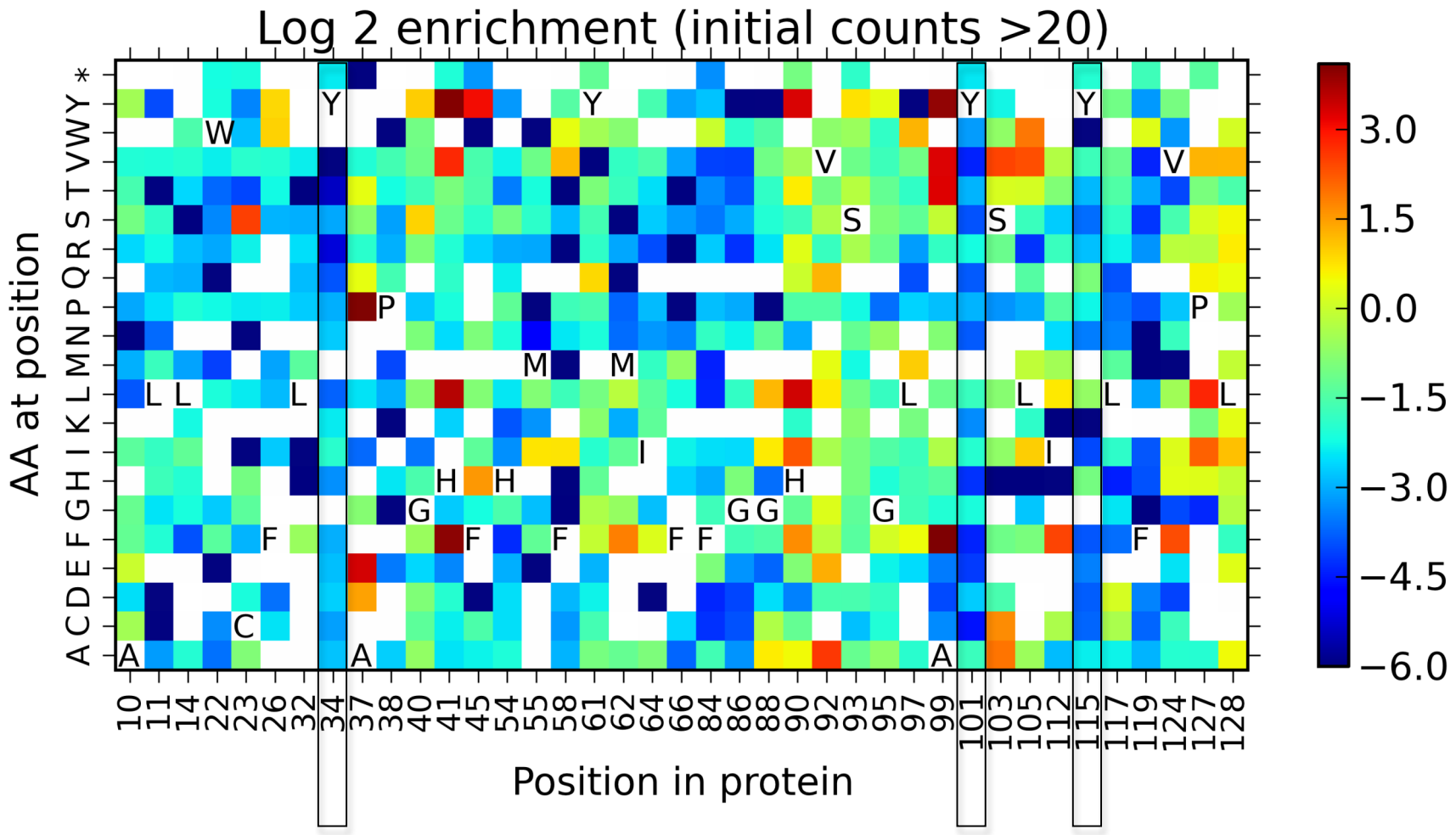
“selected population”

Next-gen Data are Consistent with Model



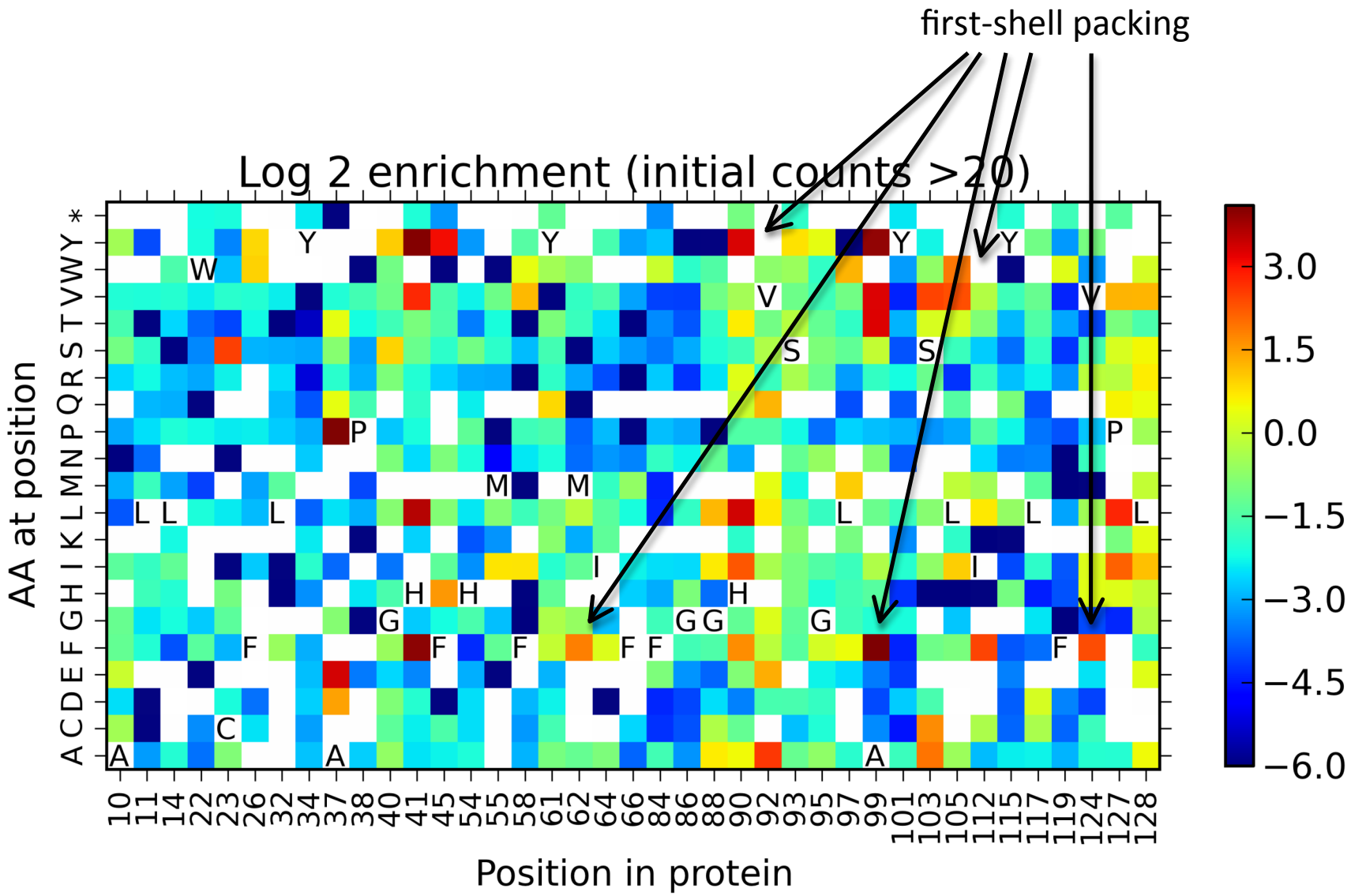
w/ Jorgen Nelson

Next-gen Data are Consistent with Model

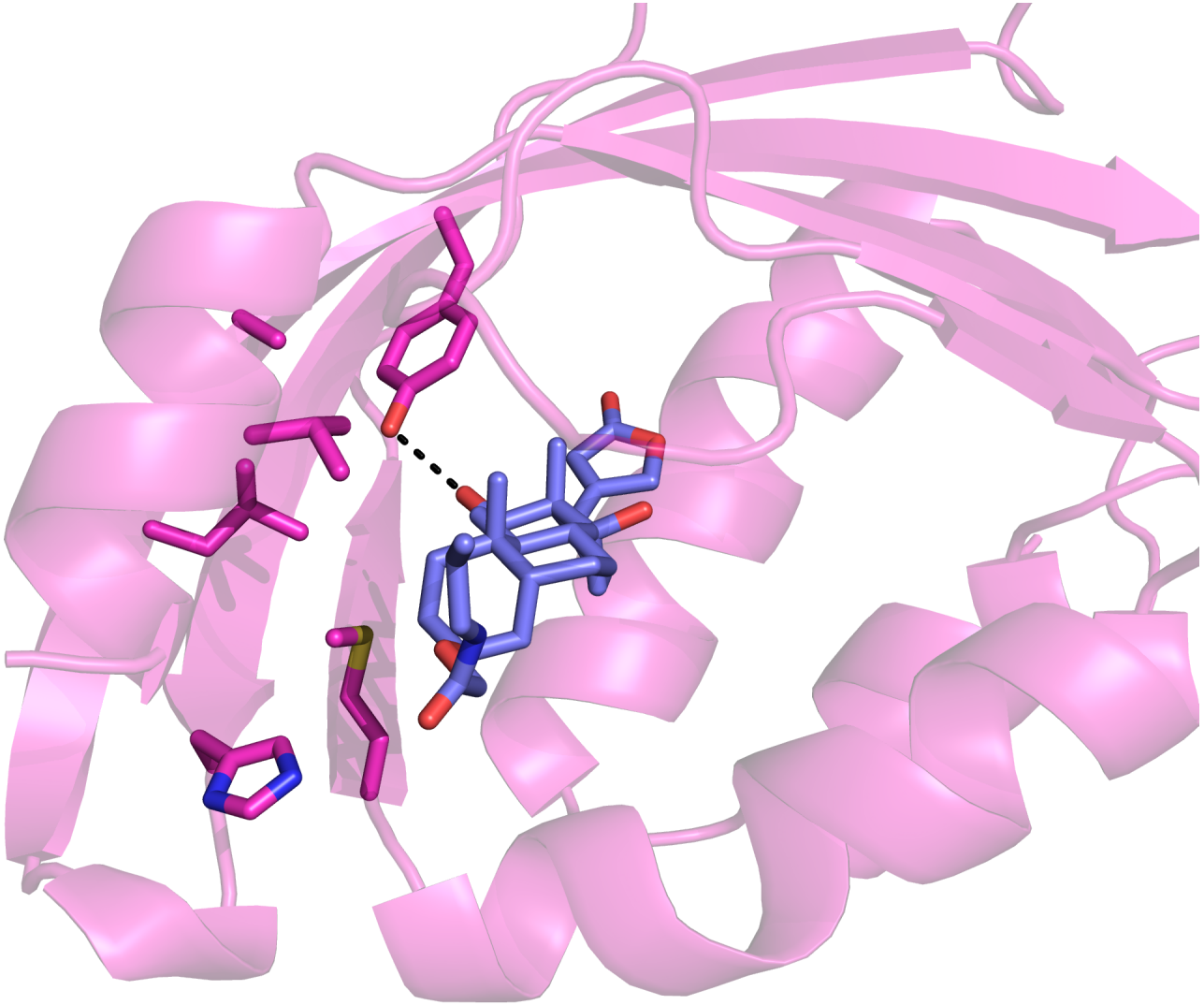


All three predicted H-bonding tyrosines are optimal

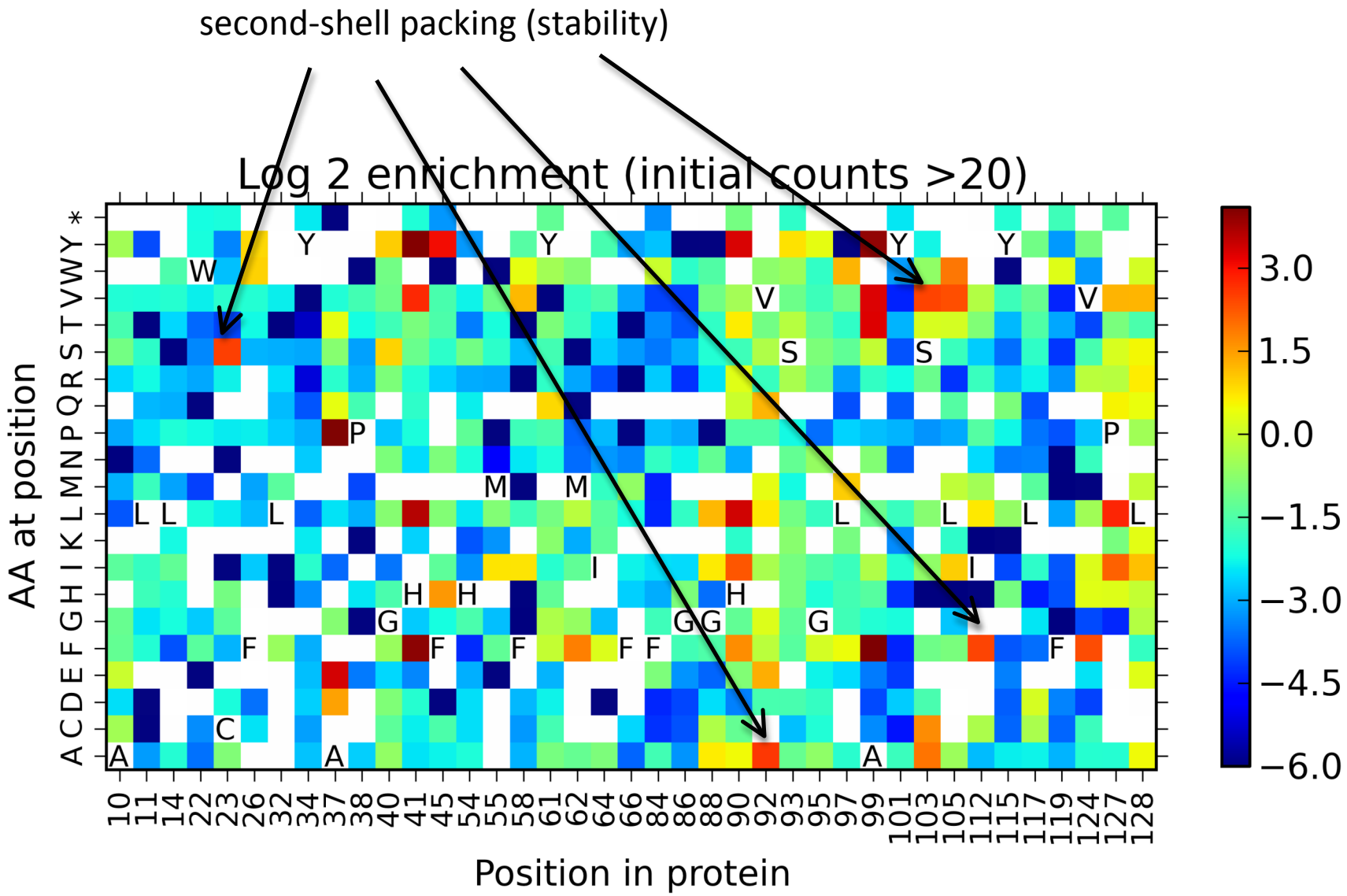
Next-gen Data Suggest Avenues for Improvement



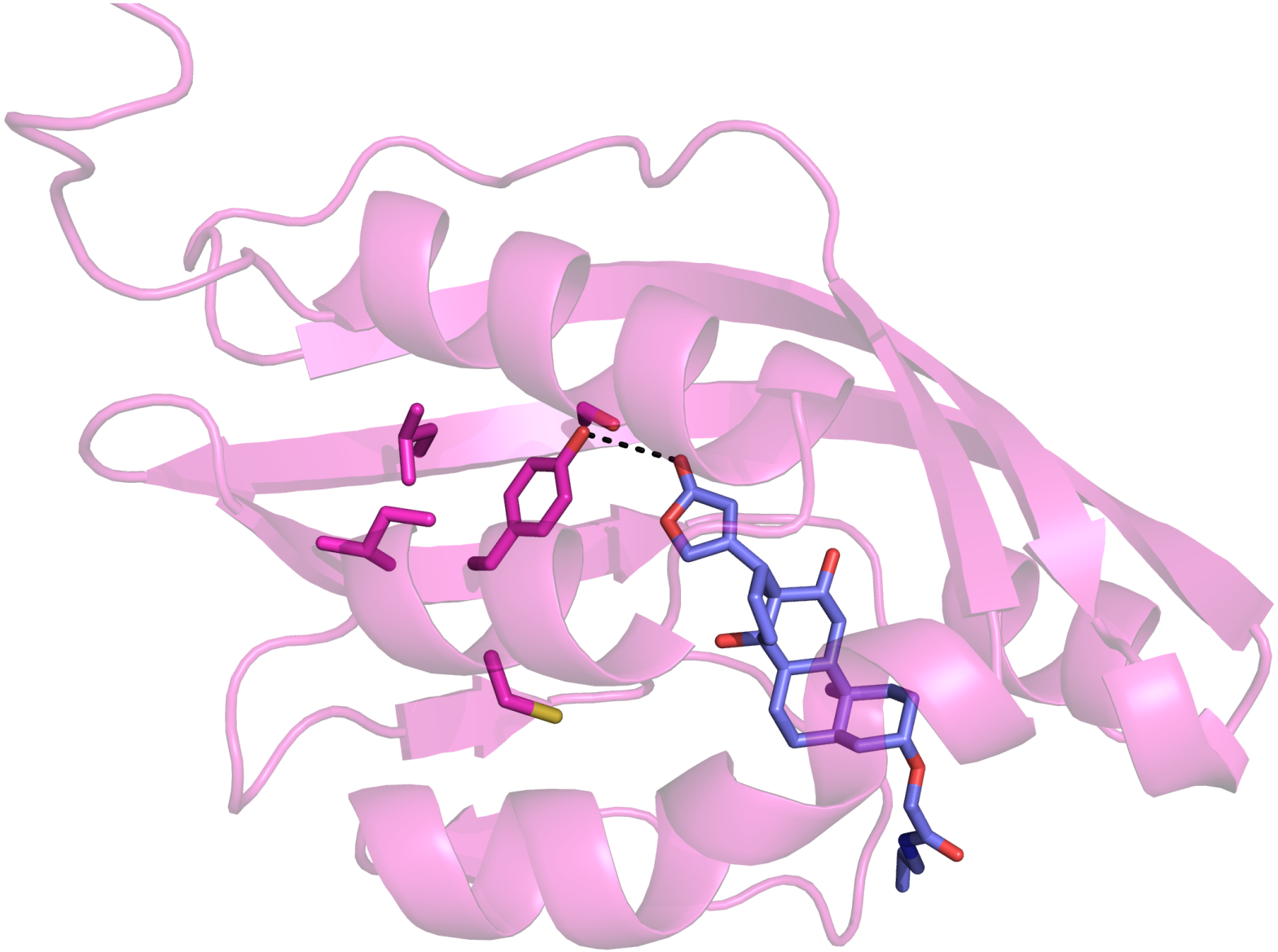
Next-gen Data Suggest Avenues for Improvement



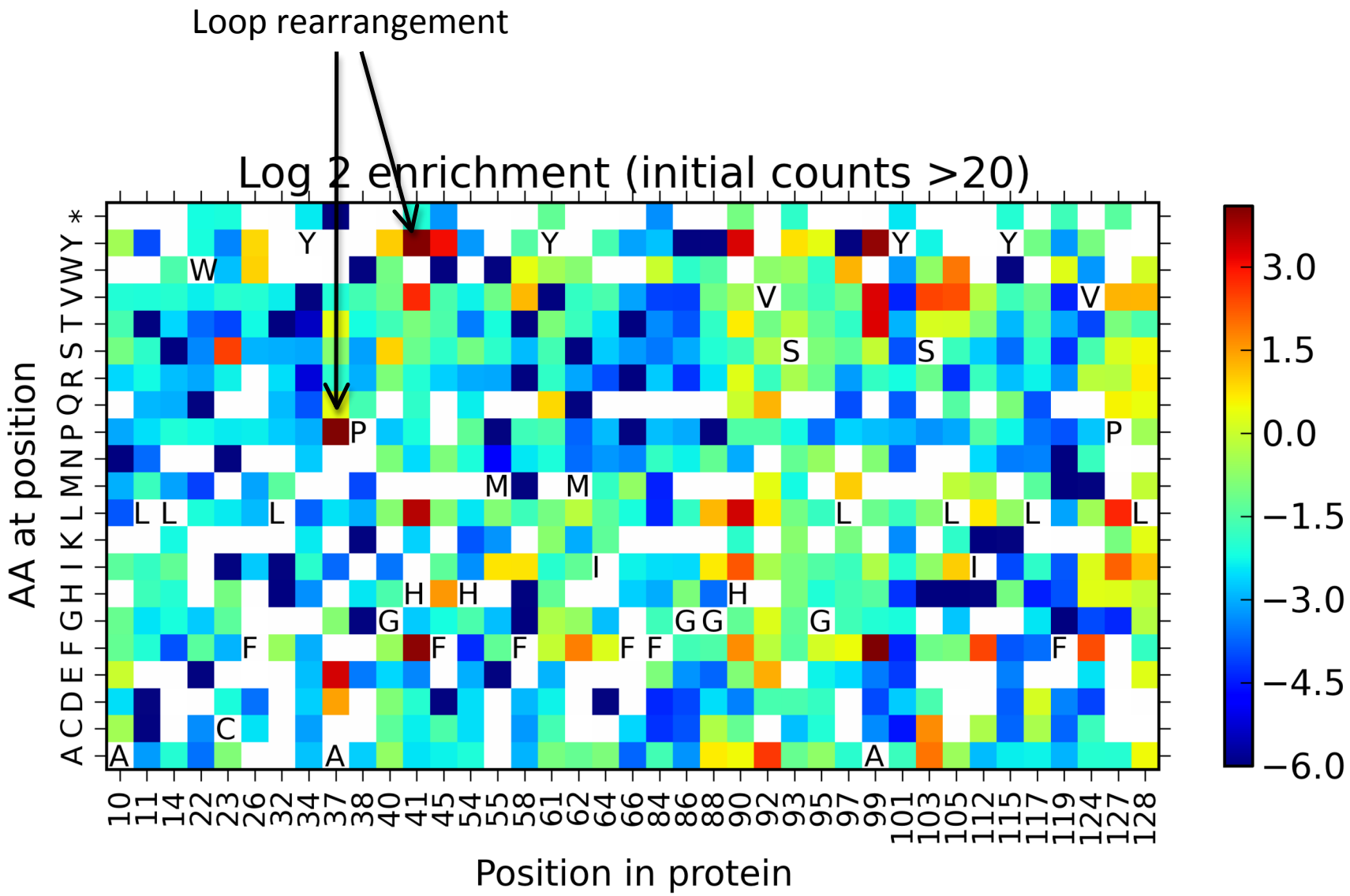
Next-gen Data Suggest Avenues for Improvement



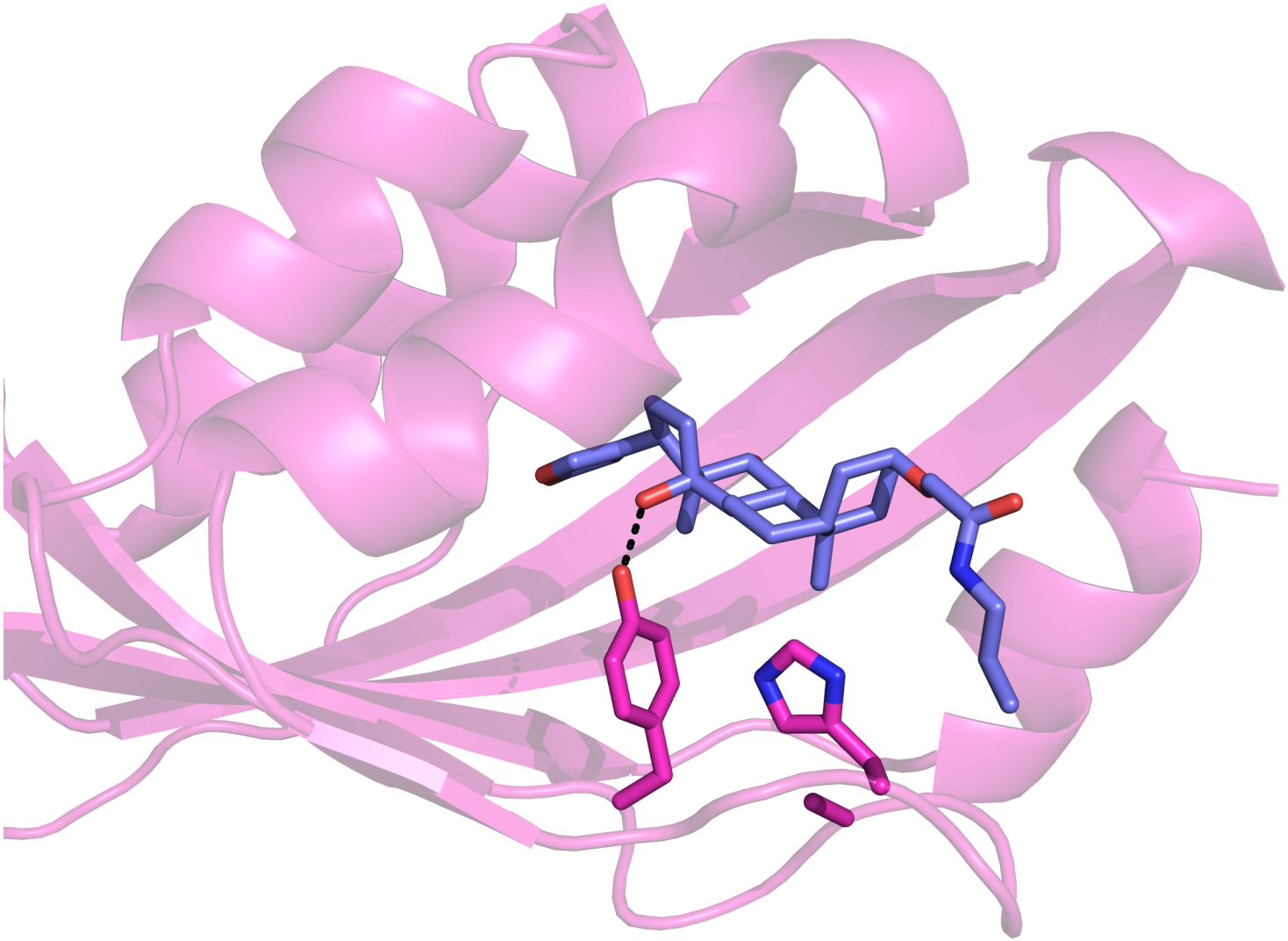
Next-gen Data Suggest Avenues for Improvement



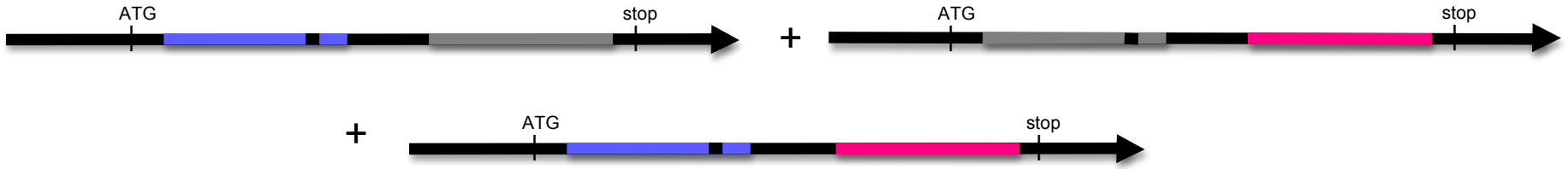
Next-gen Data Suggest Avenues for Improvement



Next-gen Data Suggest Avenues for Improvement



Finding the library “winner”



Transform into yeast

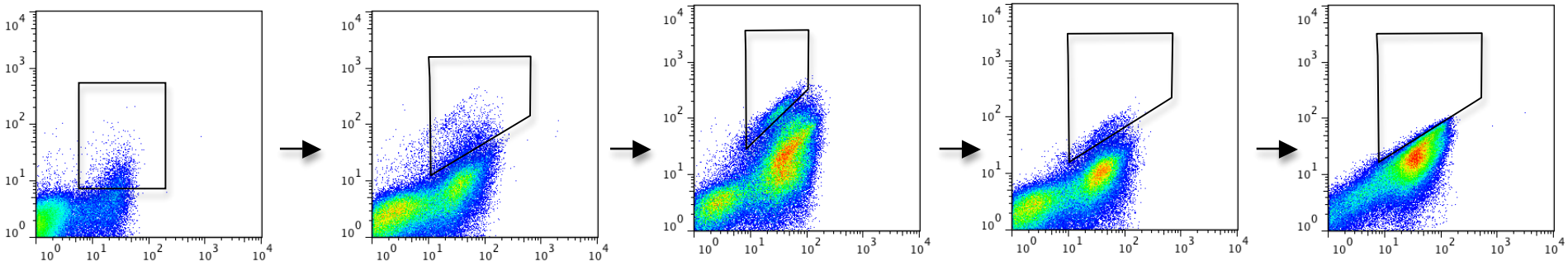
1 *80 nM label*

2 *80 nM label*

3 *50 nM label*

4 *1 nM label*

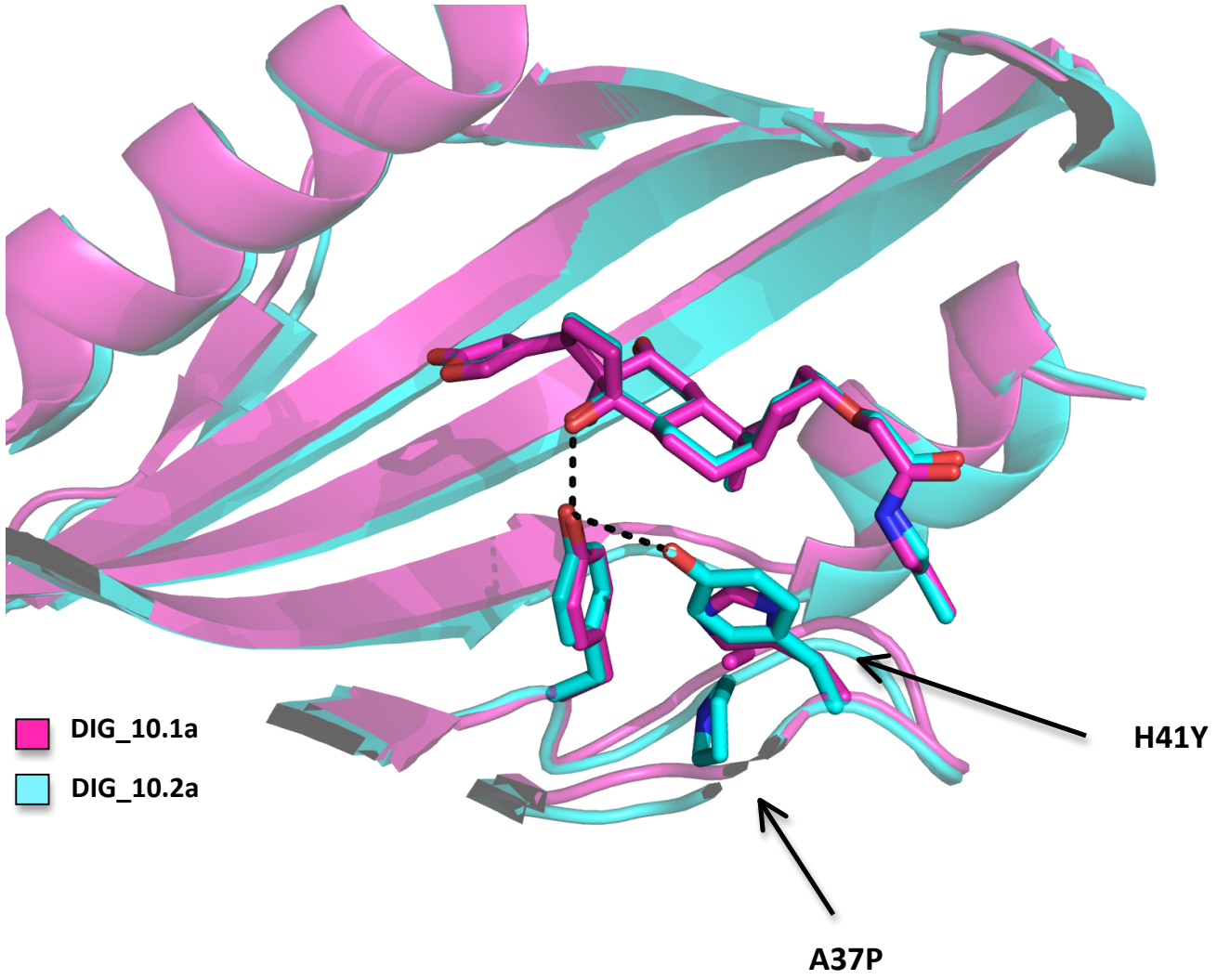
5 *1 nM label*



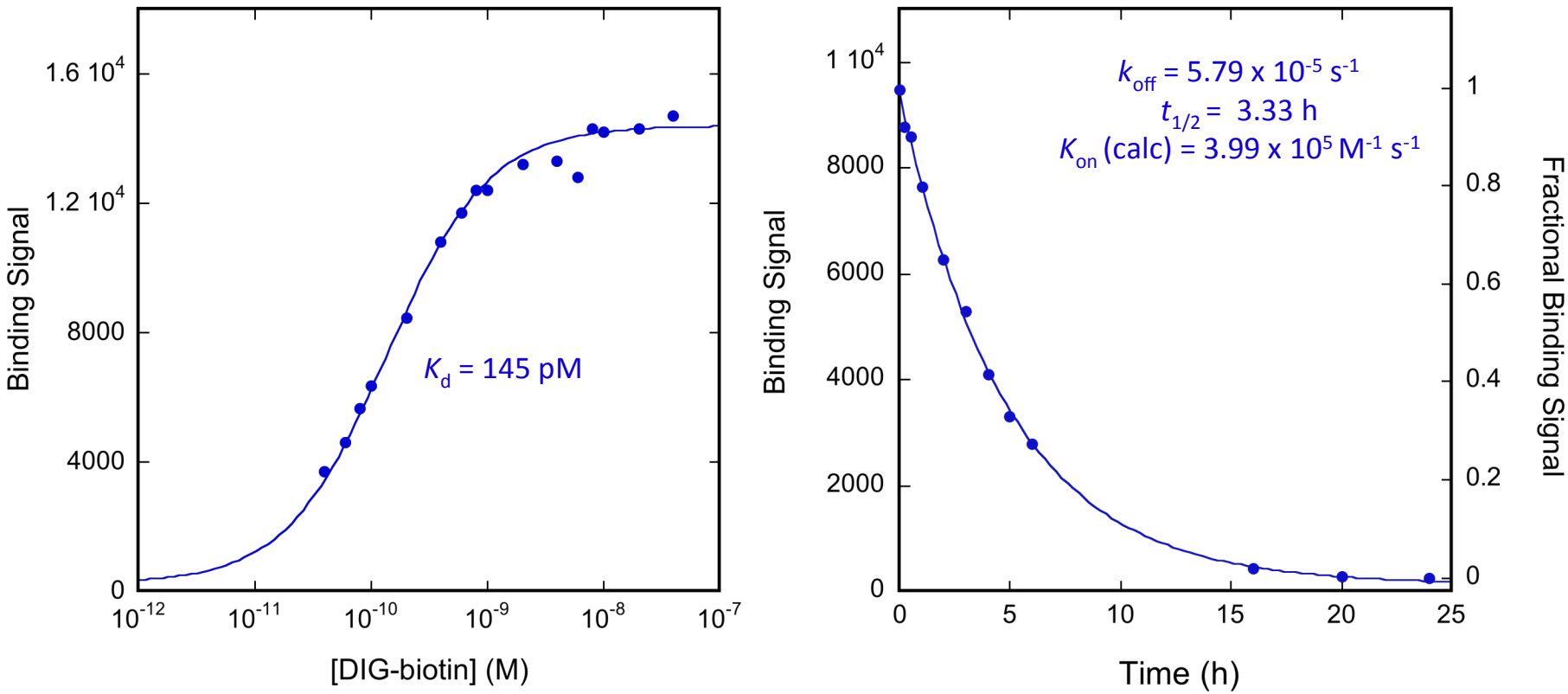
Sanger sequencing

“winner”

The “Winner” is a Double Mutant with Two Loop Mutations

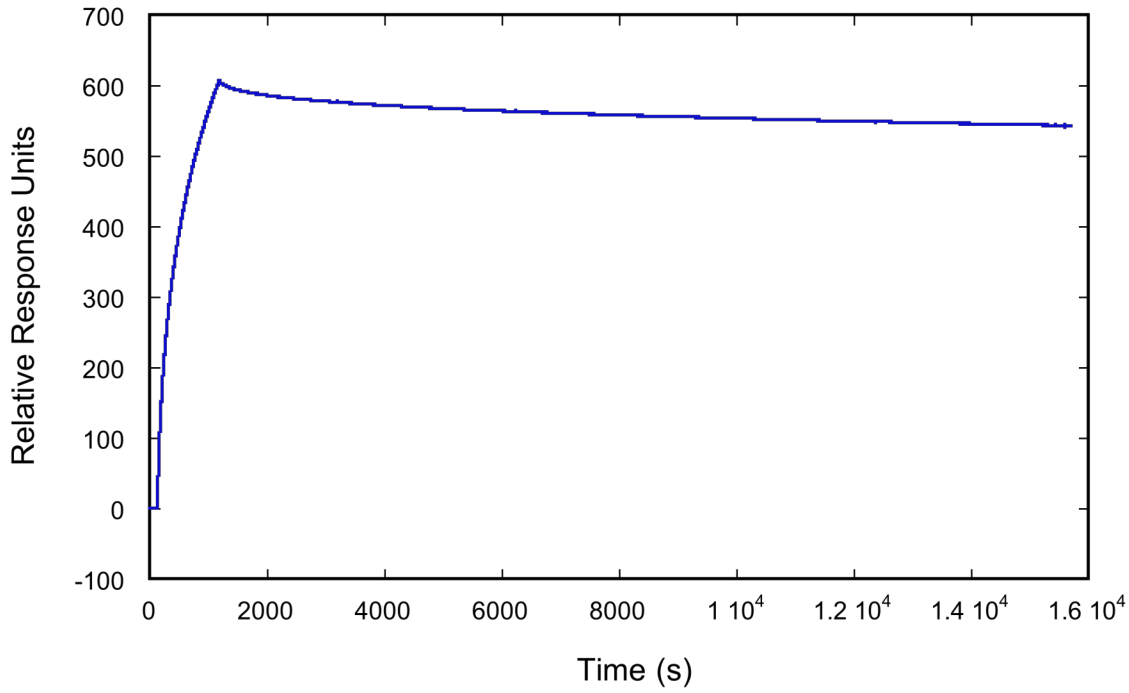


Evolved Variant Binds DIG with Sub-Nanomolar Affinity



Preliminary on-yeast experiments suggest a sub-nanomolar affinity for the evolved variant

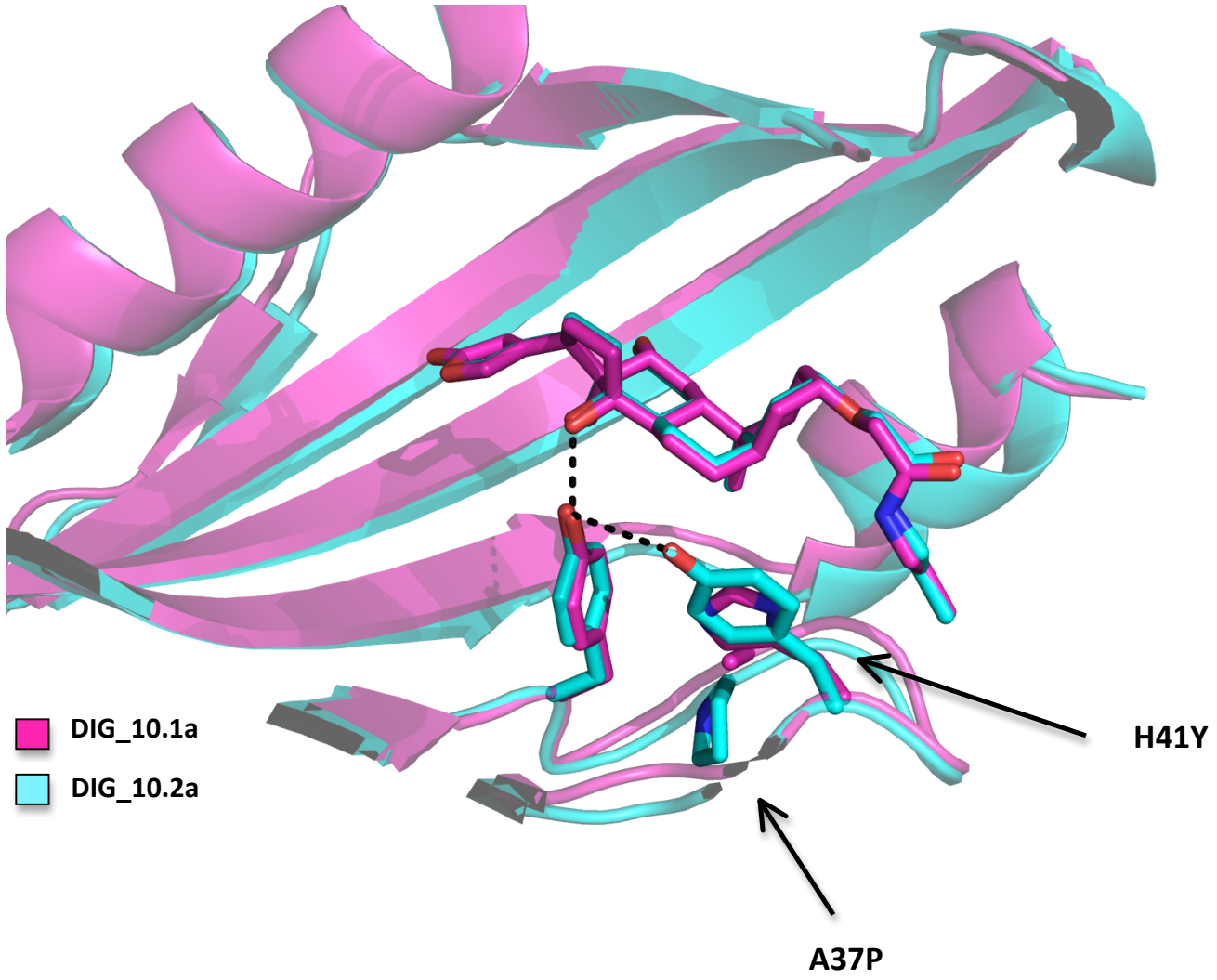
Evolved Variant Binds DIG with Sub-Nanomolar Affinity



SPR experiments also suggest high affinity in solution

k_{on} and k_{off} both appear to be low

The “Winner” is a Double Mutant with Two Loop Mutations



Conclusions and Outlook

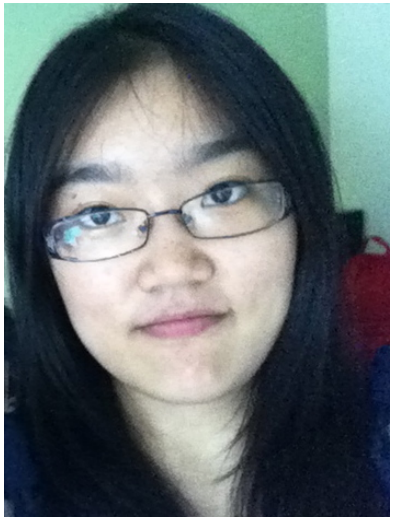
- Sub-nanomolar steroid binding was achieved using a combination of computation and directed evolution

- Multiple interactions are required for binding; precision is key

- A major shortcoming of the working designs was under-packing of the binding sites

- For DIG_10, higher binding affinity was obtained by altering the cavity entrance loop; might act as a lid that closes over DIG

Thanks!



David Baker

Sagar Khare
Jiayi Dou
Jorgen Nelson

Alberto Schena
Kai Johnsson
} Synthesis of “DIG-PEG-biotin”

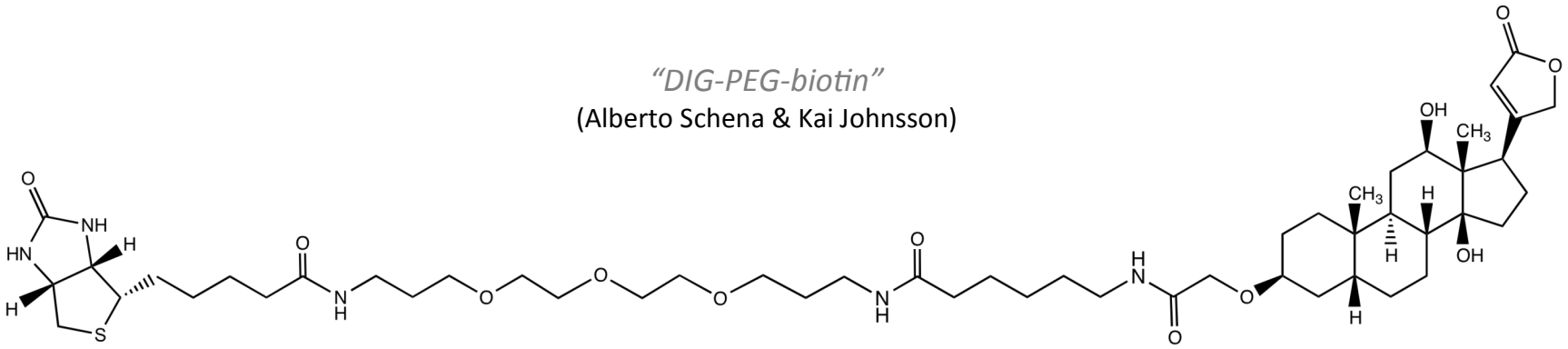
Lindsey Doyle
Barry Stoddard
} Crystallography

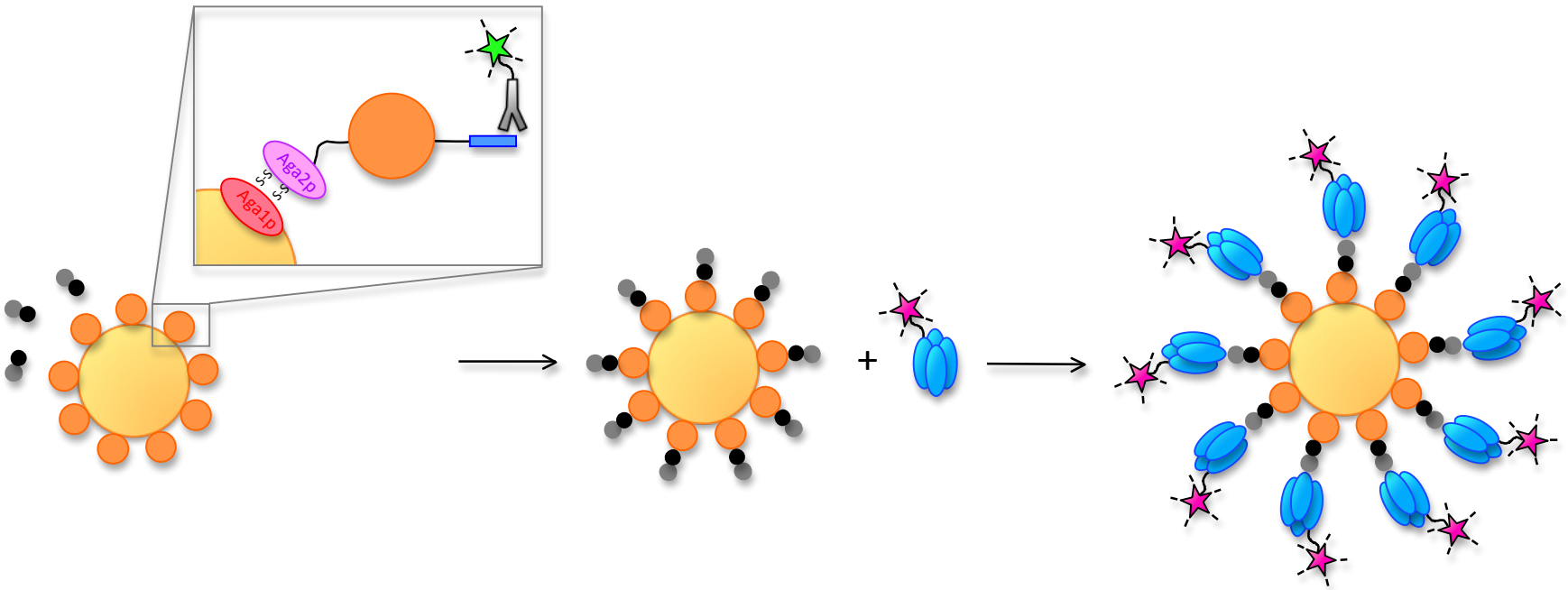
Lewis Kay - NMR

Selection Strategy for Next-gen Sequencing Libraries

"DIG-PEG-biotin"

(Alberto Schena & Kai Johnsson)



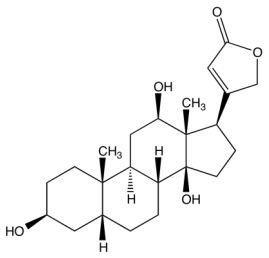


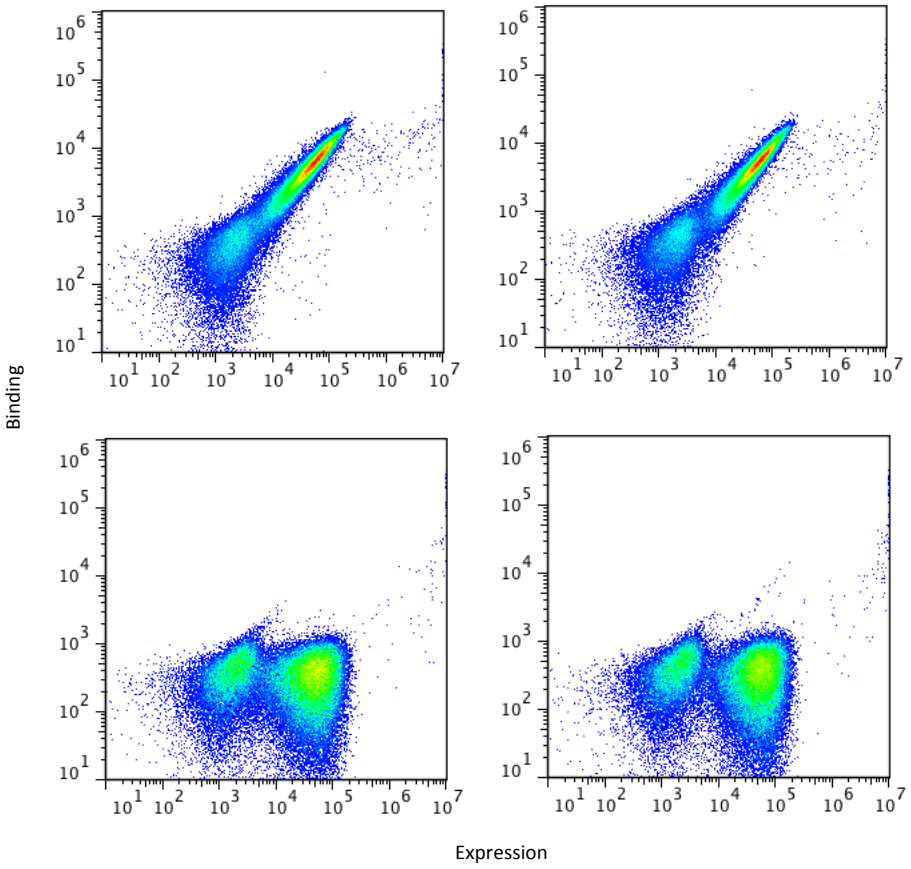
Evolved Variant Signals are Competed Off by free DIG

(1 nM DIG-RNase label)

- free DIG

+ 980 µM free DIG

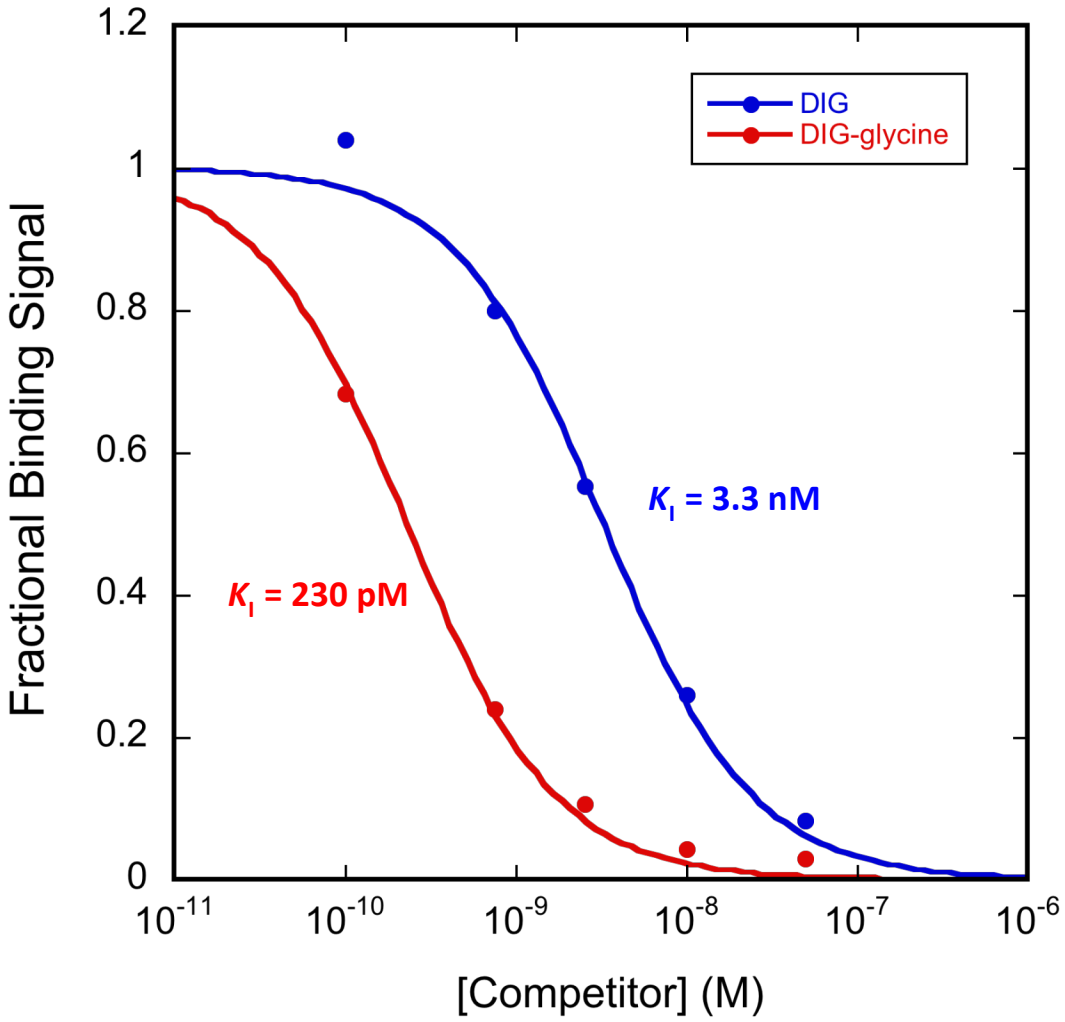


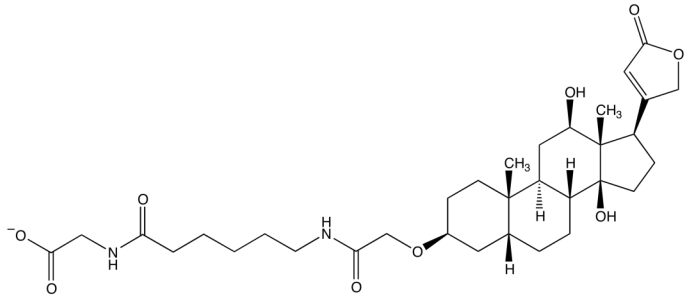


DIG10.1a
evolved variant (S10A; V62M; V64I; V117L; A127P)

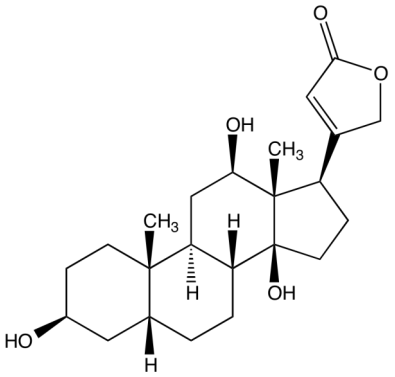
DIG10.1b
evolved variant (S10M; V62M; V64I; V117L; A127P)

DIG10_2a Linker Effects





DIG-glycine



DIG

variable fragment 1:

… TTC aaa ACG GGT TTT atg gcc CGG cgg tgc GTA CGG ata TTC cag …

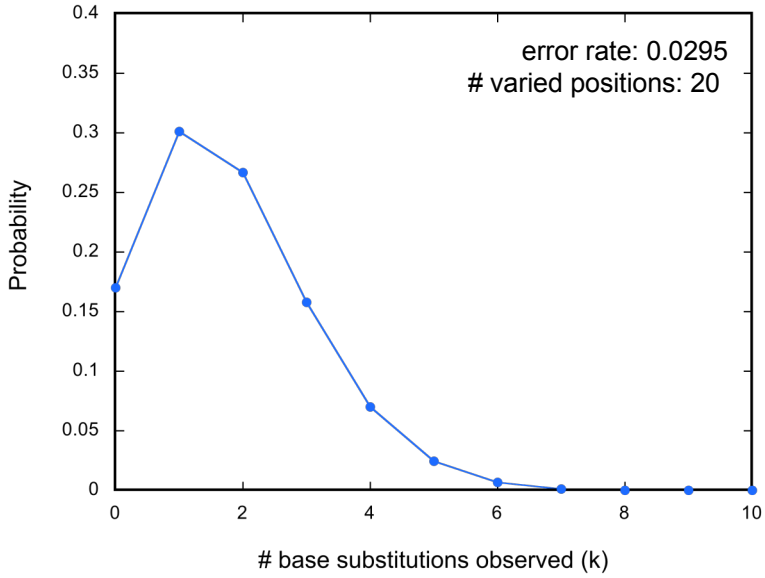
96.952% A
1.016% C
1.016% T
1.016% G

Varied positions are synthesized by doping in a small amount of each non-wild type base with the wild-type base

 # base substitutions observed

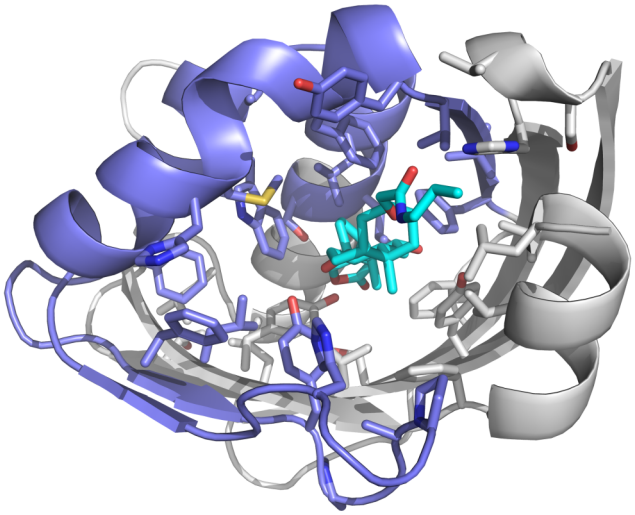
$$P(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

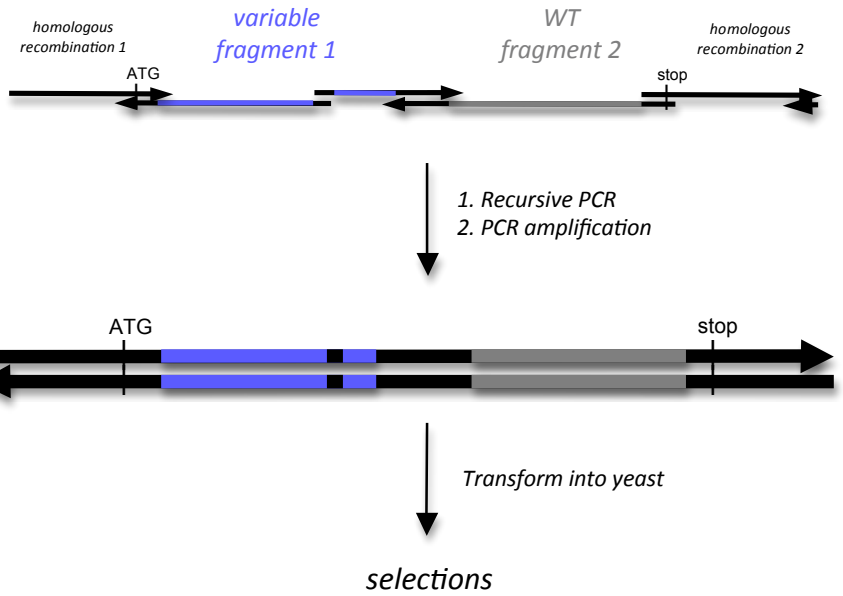
 # base substitutions expected
(error rate times the number of varied positions)



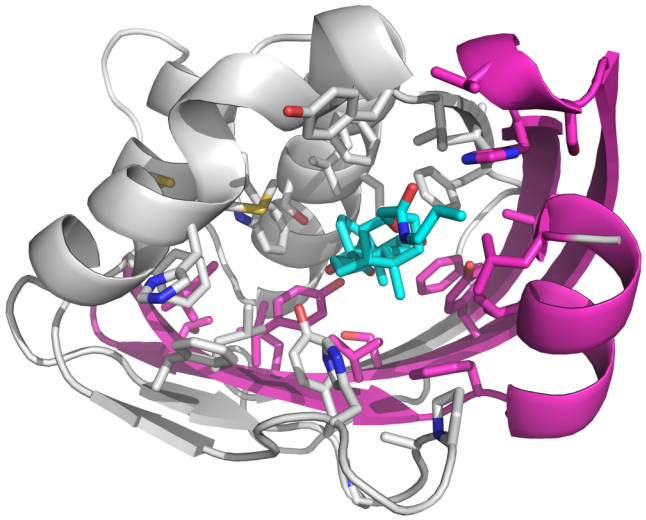
Calculated error rates are corrected for synonomous mutations by an empirical factor

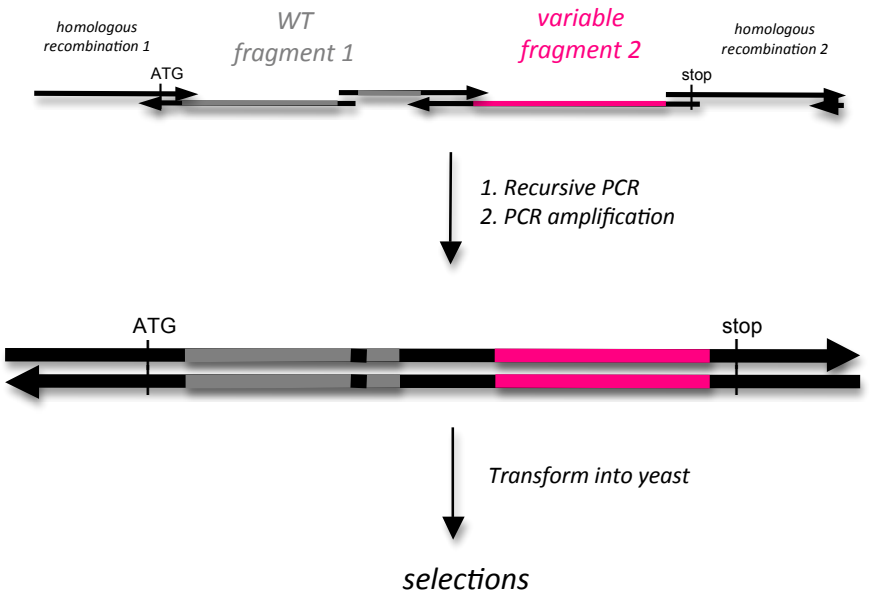
fragment 1 library

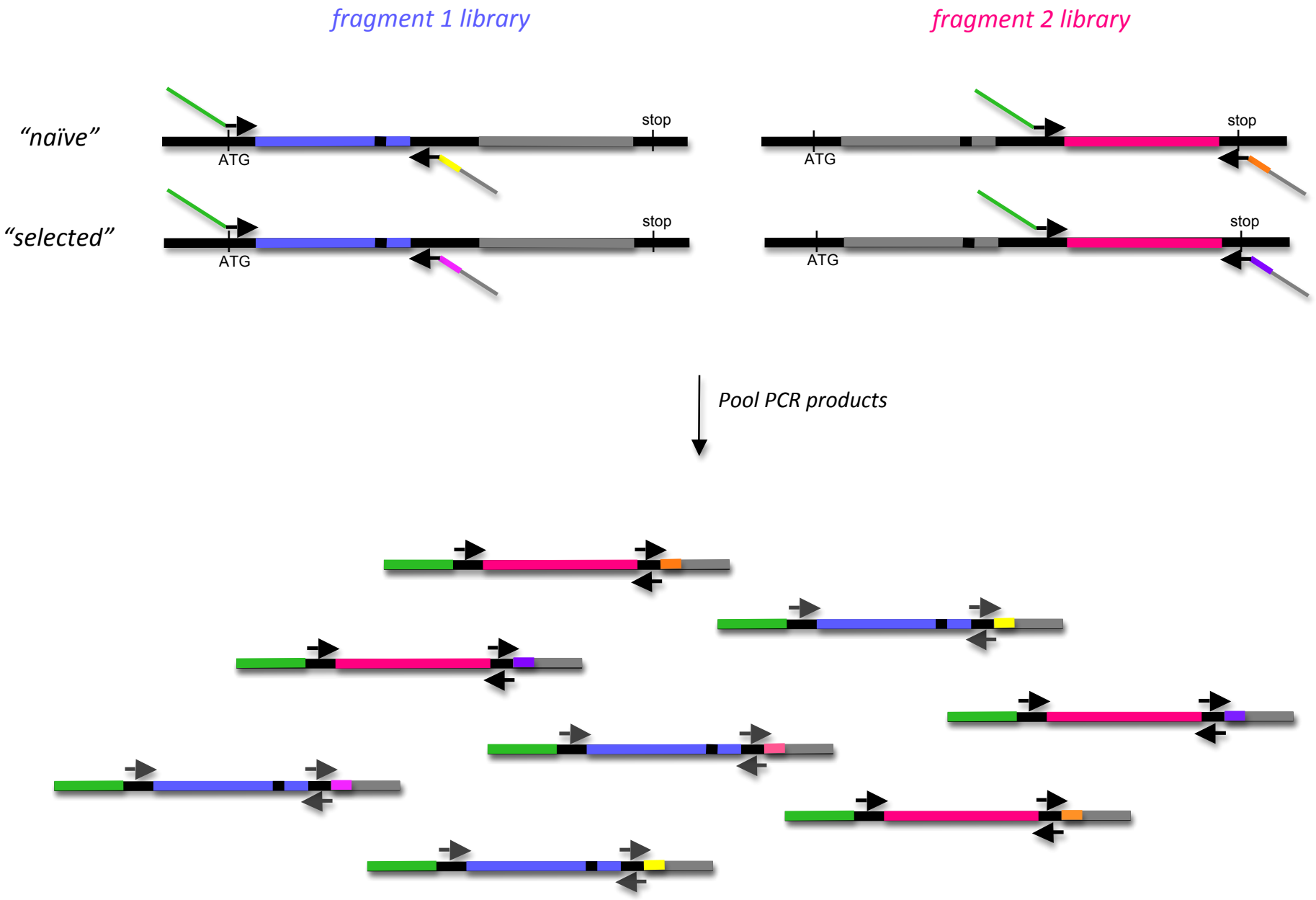


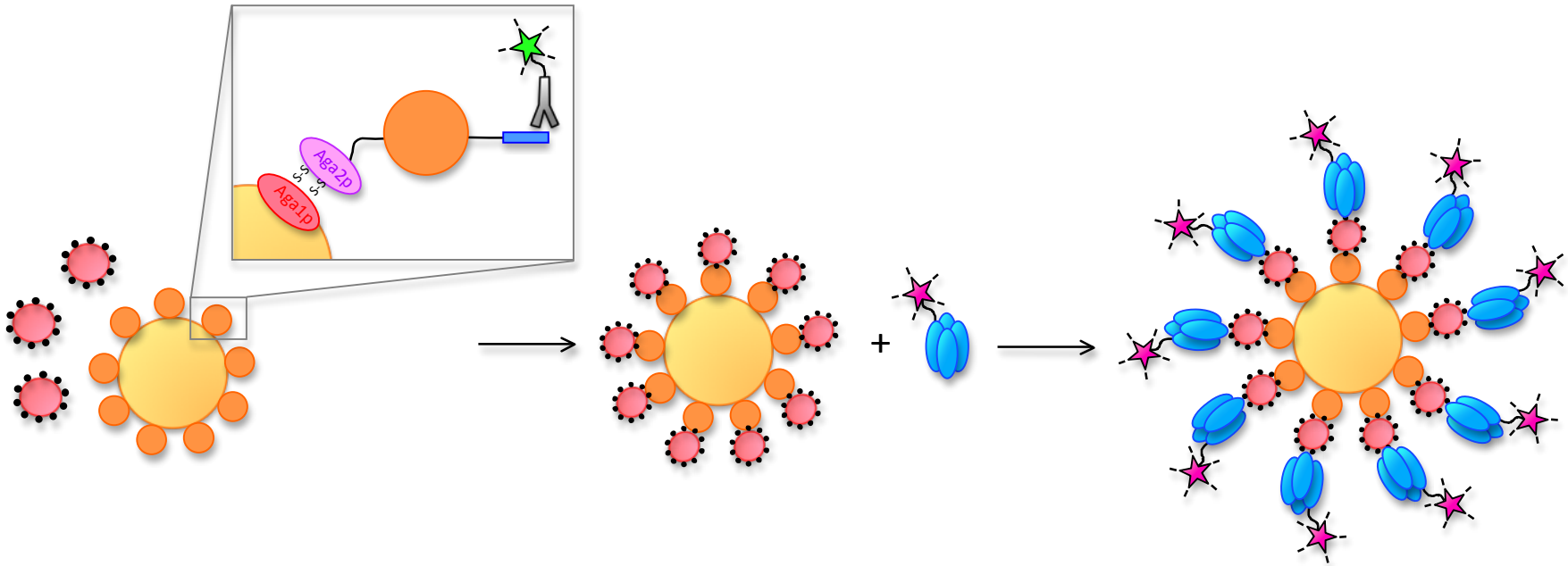


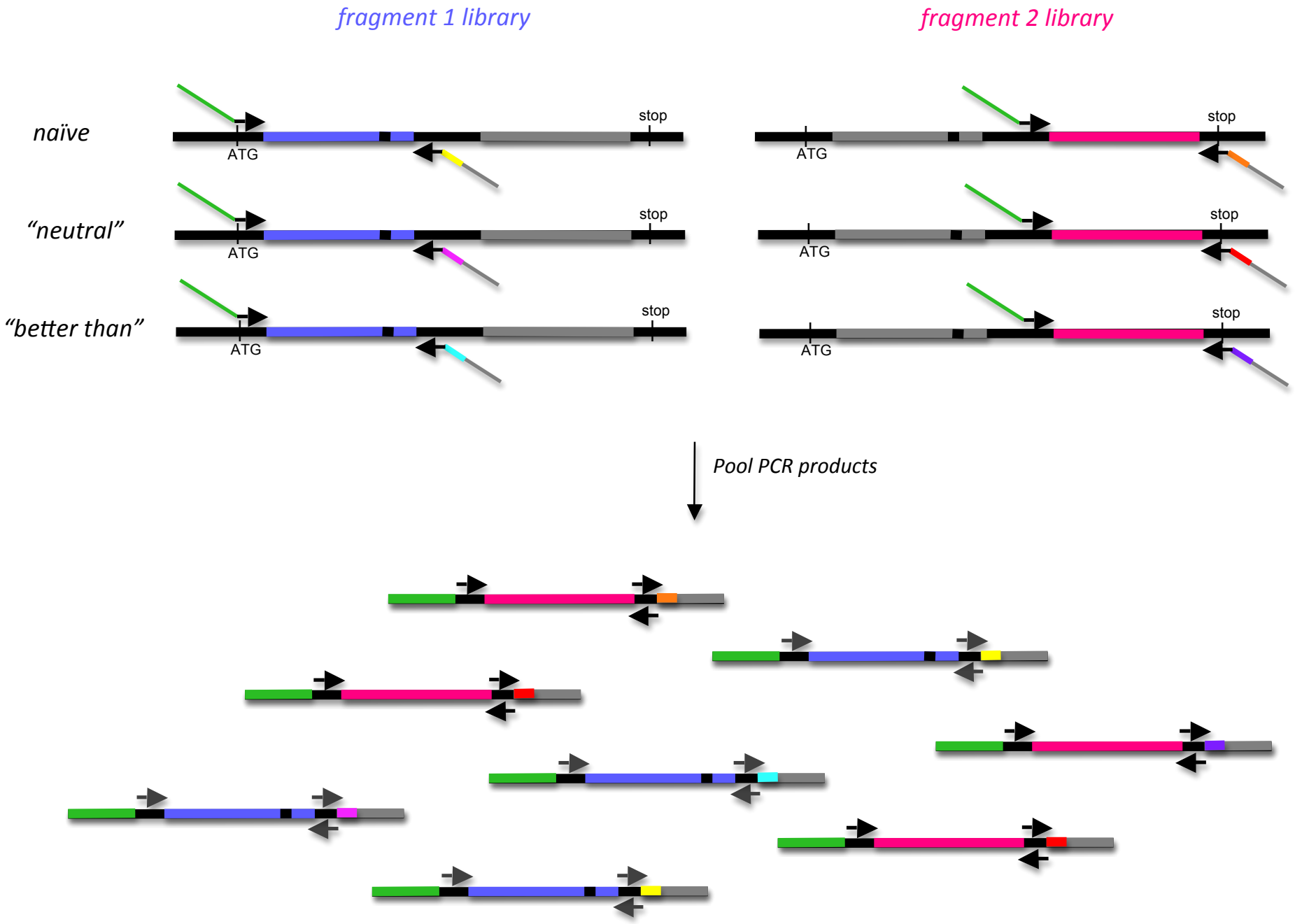
fragment 2 library

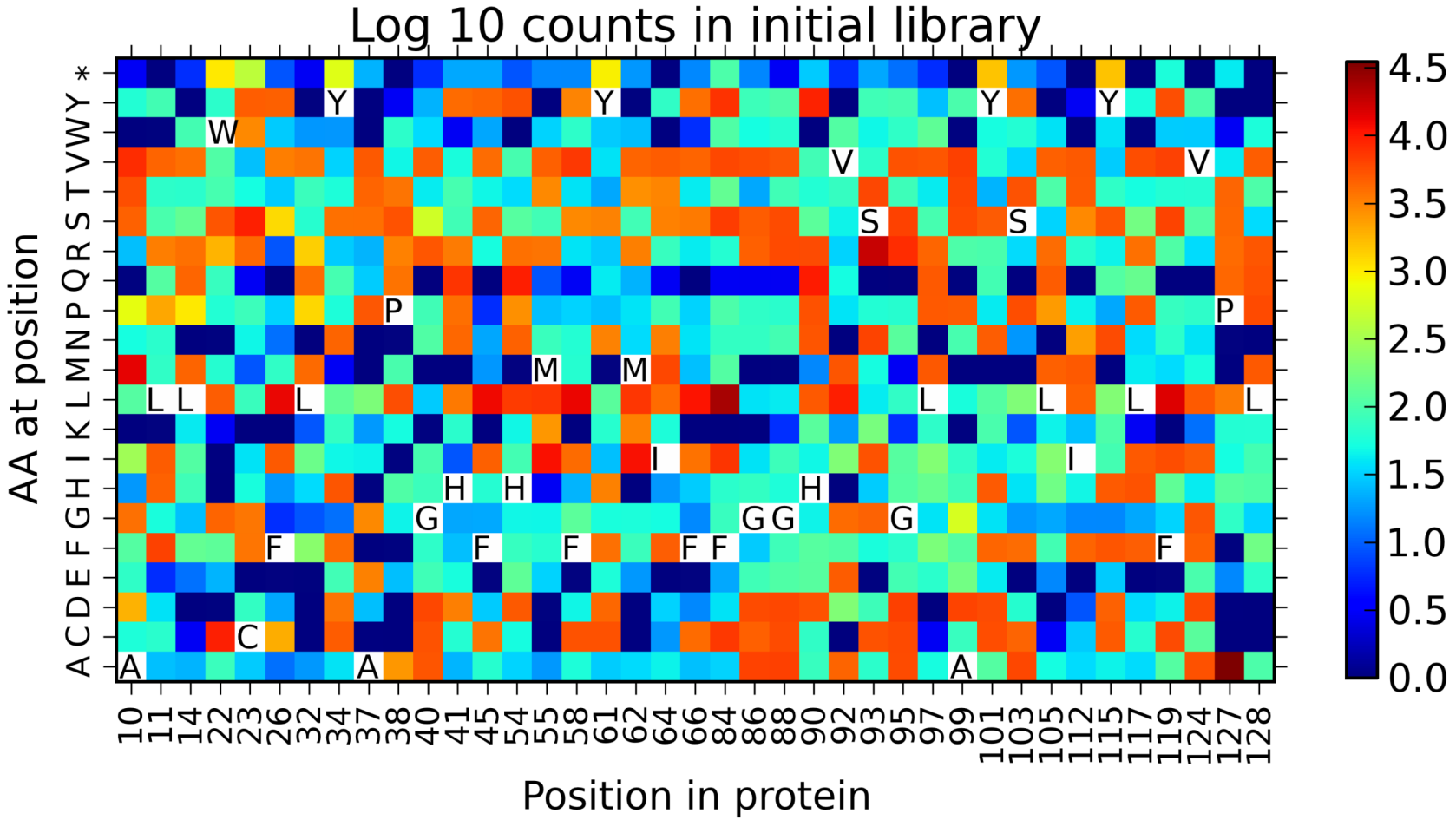





fragment 1 library
fragment 2 library
"naïve"
"selected"
ATG
stop
ATG
stop
ATG
stop
ATG
stop
Pool PCR products


Aga2p
Aga1p
S
S
S
S
+


fragment 1 library
fragment 2 library
naïve
"neutral"
"better than"
ATG
stop
Pool PCR products


Log 10 counts in initial library
AA at position
A C D E F G H I K L M N P Q R S T V W Y *
10 11 14 22 23 26 32 34 37 38 40 41 45 54 55 58 61 62 64 66 84 86 88 90 92 93 95 97 99 101 103 105 112 115 117 119 124 127 128
Position in protein
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0