

# Implementation of Genetic Algorithm on DNA-Encoded Library



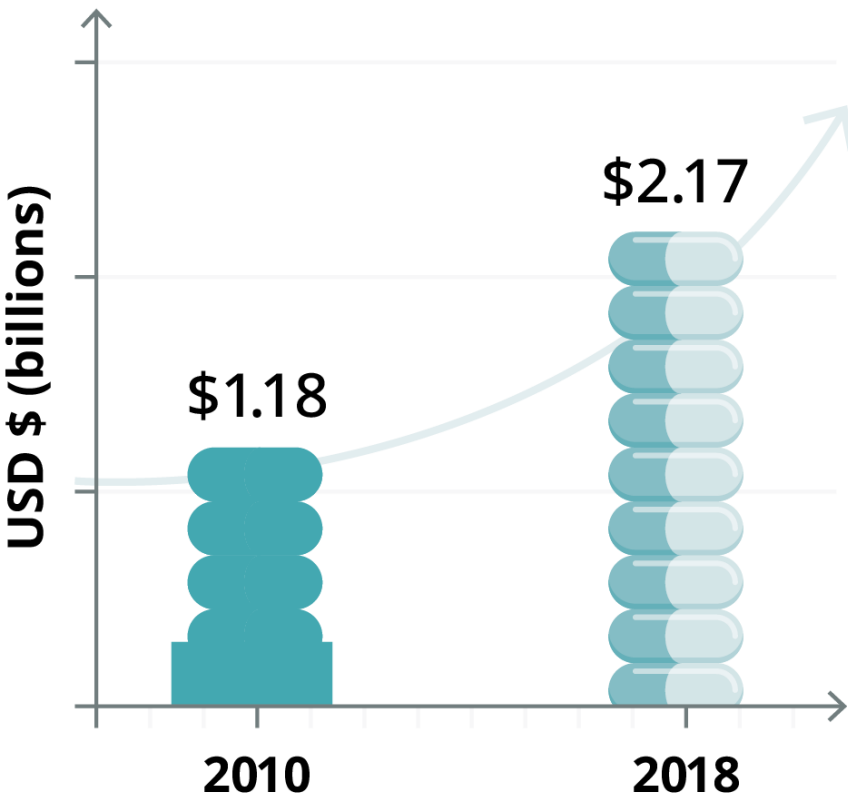
2023. 5. 24. Literature Session  
D3 Takeuchi Aoi

- **Introduction**
  - DNA-encoded library
  - Genetic algorithm
  - PNA
  
- **A mating mechanism to generate diversity for the Darwinian selection of DNA-encoded synthetic molecules**

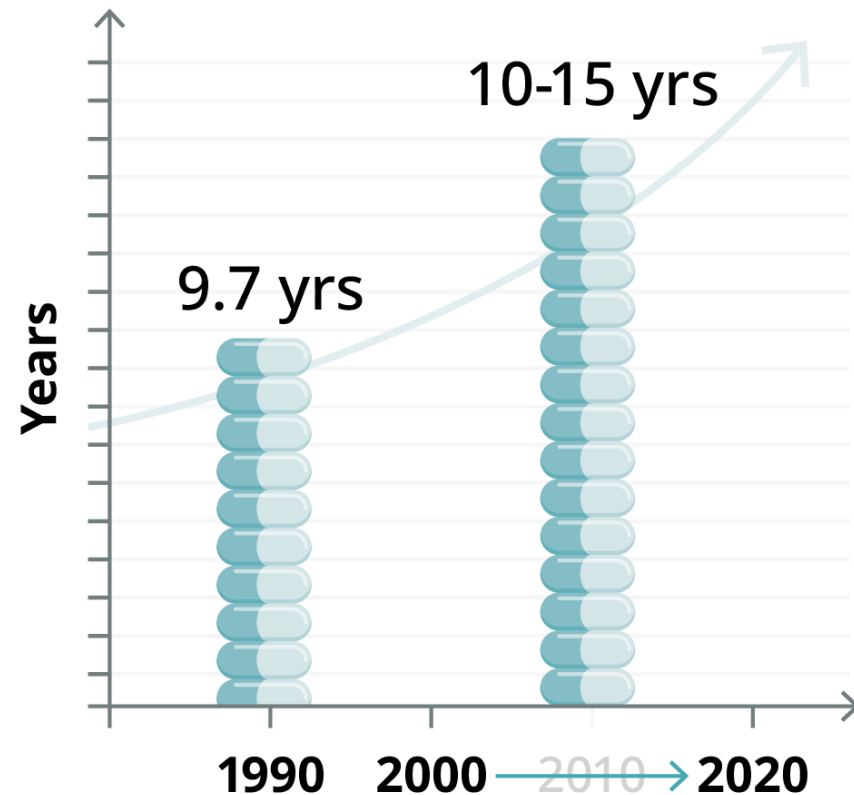
Vummidi, B. R.; Farrera-Soler, L.; Daguer, J.-P.; Dockerill, M.; Barluenga, S.; Winssinger, N. *Nat. Chem.* **2022**, *14*, 141.

# Hardship in Productive Drug Development

Cost of bringing new drug to the market



Average length of drug development

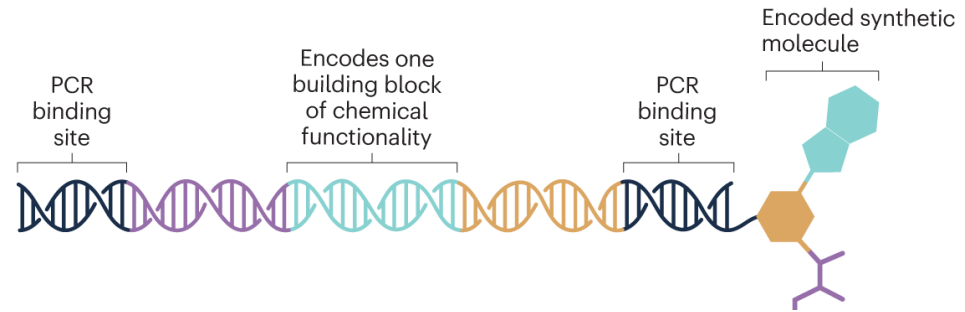


Urgency of expeditious supply of lead compounds has been more serious than ever.

1. Terry, C., Lesser, N. Ten years on, Measuring the return from pharmaceutical innovation 2019. Deloitte.
2. DiMasi, J. A.; Grabowski, H. G.; Hansen, R. W. *J. Health Econ.* **2016**, *47*, 20.

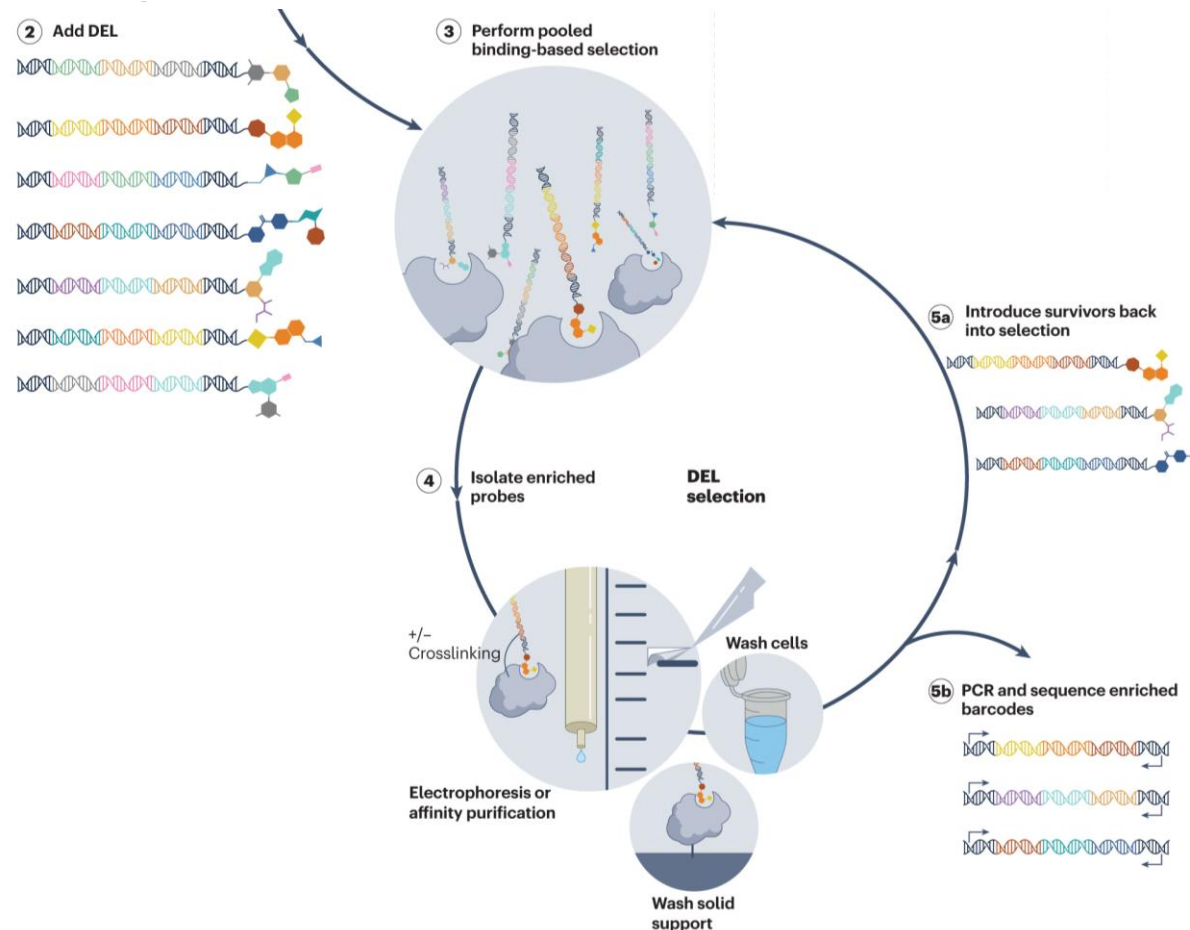
# DNA-Encoded Library

## General architecture of a DEL

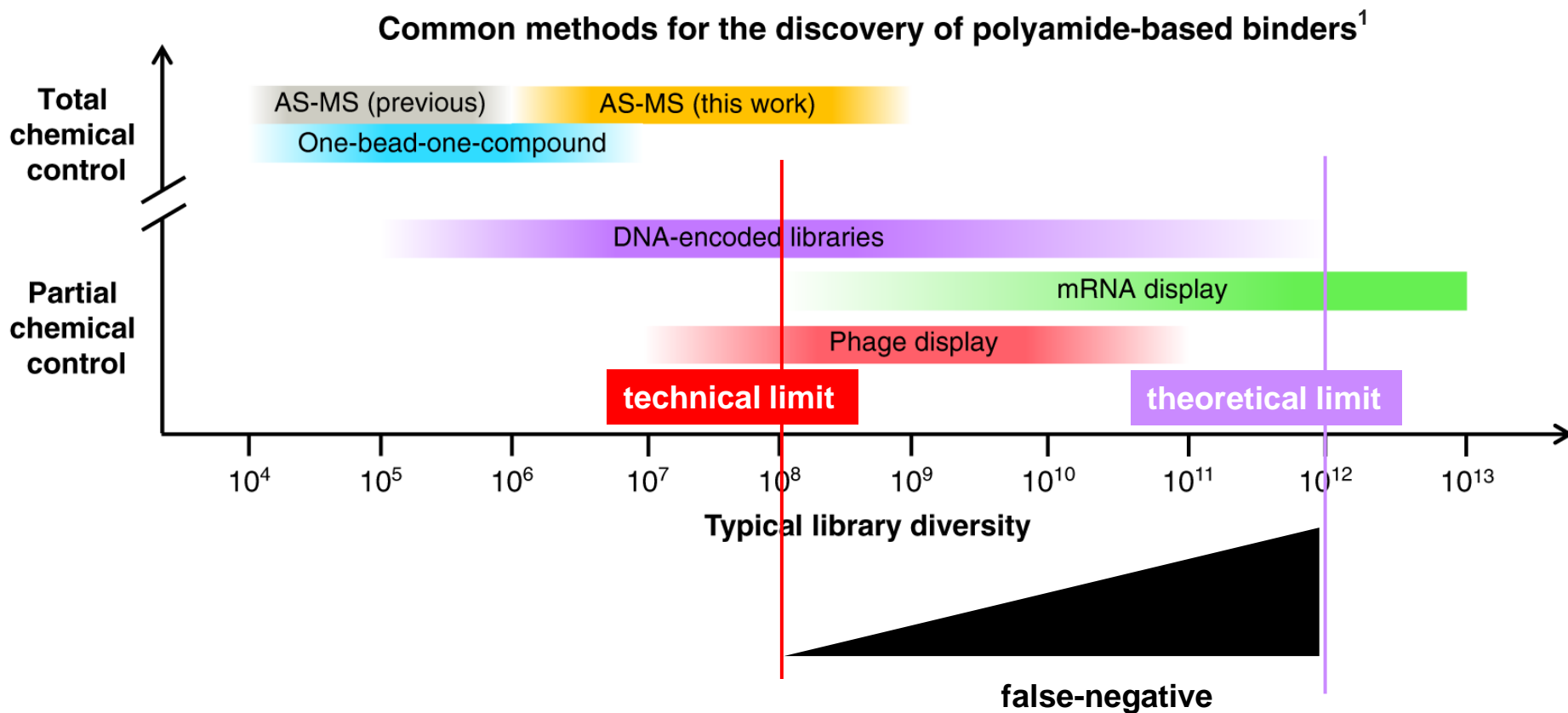


**DNA-encoded library (DEL):**  
conjugation of building blocks to  
DNA sequences which encode the  
structural information.

- ✓ high-throughput
- ✓ low concentration



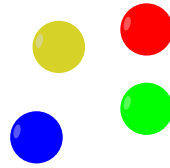
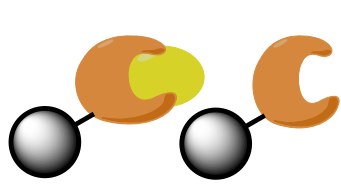
# Diversity Scale of DNA-Encoded Library



The upper limit of library size for DEL is estimated to be  $10^8$  with current technologies<sup>2</sup>.

1. Quartararo, A. J.; Gates, Z. P.; Somsen, B. A.; Hartrampf, N.; Ye, X.; Shimada, A.; Kajihara, Y.; Ottmann, C.; Pentelute, B. L. *Nat. Commun.* **2020**, *11*, 3183.
2. Satz, A. L.; Hochstrasser, R.; Petersen, A. C. *ACS Comb. Sci.* **2017**, *19*, 234.

# Challenges in Expanding the Diversity Scale



$$C_{\text{total}} = 10 \mu\text{M}$$

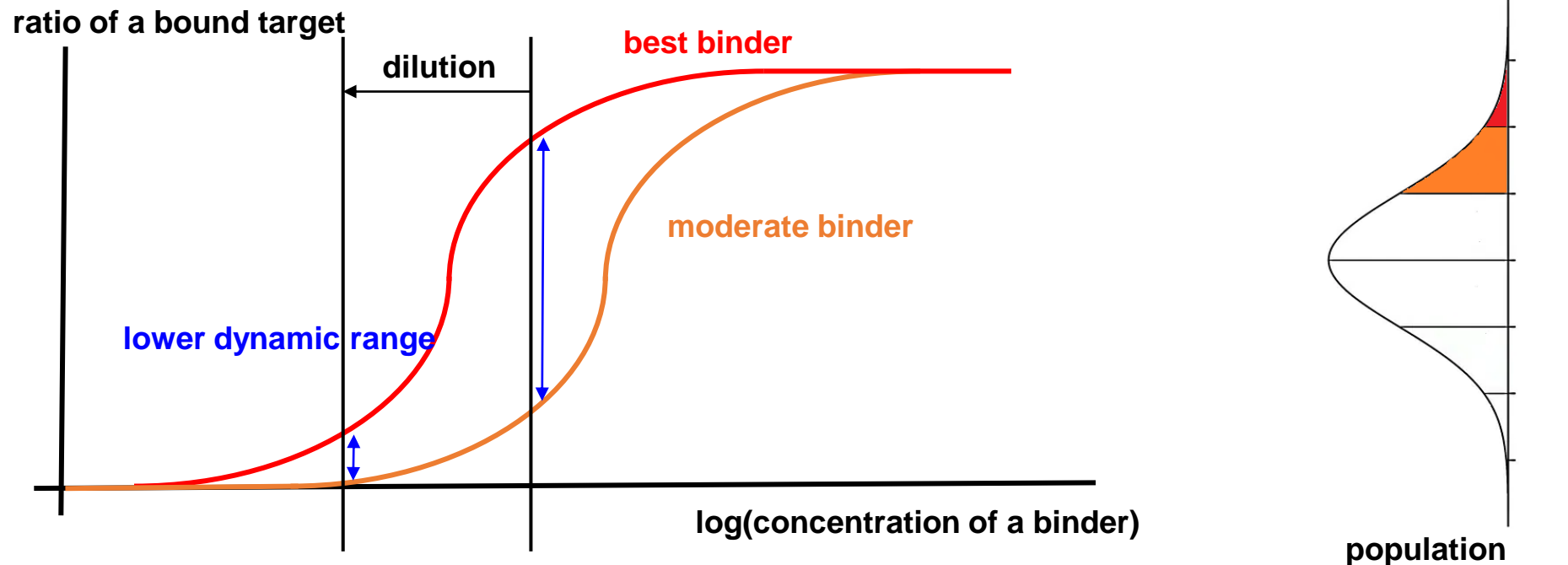
$$N = 10^6$$

$$c_{\text{mean}} = 10 \text{ pM}$$

$$N = 10^8$$

$$c_{\text{mean}} = 0.1 \text{ pM}$$

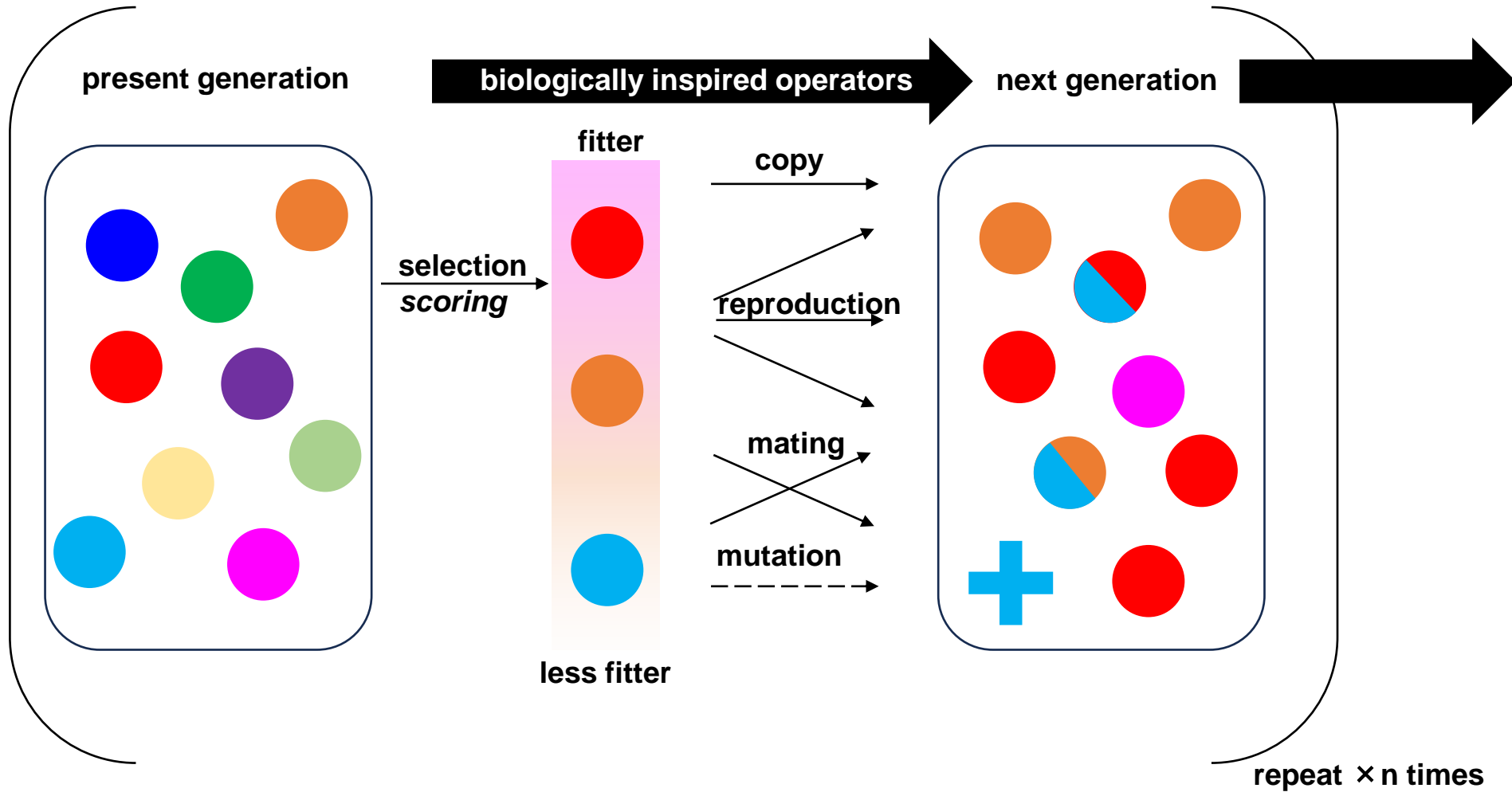
Expanding the library scale = Dilution of each binder



More abundant members with poorer affinity may overshadow the fitter binders.

# Genetic Algorithm

Genetic algorithm: a metaheuristic used to generate high-quality solutions to optimization featuring biologically inspired operators









# Enrichment of the Fittest in the Library

Analogies among Darwinian theory, genetic algorithm, and DNA-encoded library

	Darwinian theory	Genetic algorithm	DNA-encoded library
purpose	Survival of the fittest	Find the optimum solution	<b>Enrichment of the fittest binder</b>
operators		selection	binding affinity
		reproduction	PCR amplification
		copy	
		mating	chemical ligation hybridization DNA shuffling
		mutation	error-prone PCR

gray: not demonstrated in the main paper

Implementation of genetic algorithm on DEL could overcome a challenge in large libraries.

# Author's Profile

## Nicolas Winssinger

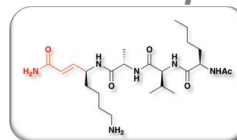
- B.Sc.: Tufts University
- Ph.D.: Scripps Research Institute (1995-2000)  
advised by Prof. K. C. Nicolaou
- Postdoctoral fellow: Scripps Research Institute (2000-02)  
advised by Prof. P. G. Schultz
- Assistant professor: Strasbourg University (2002-05)
- Professor: Strasbourg University (2005-12)
- Professor: University of Geneva (2012-)



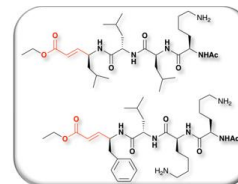
### Representative examples of hits identified with PNA-ended libraries

#### Research interests

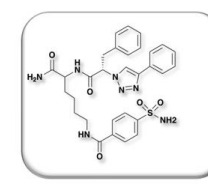
- Encoded library technology
- Biosupramolecular interaction
- Chemical biology probes



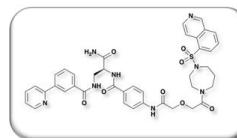
Cysteine protease inhibitor (Der P1)  
*Chem. Biol.* **2004**, *11*, 1361-1372  
*Tetrahedron* **2004**, *60*, 8677-8690



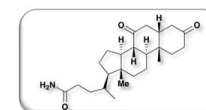
Orthogonal cathepsin K & F inhibitors  
*ChemBioChem*, **2006**, *7*, 1790-1797  
*Tetrahedron* **2007**, *63*, 6577-6586



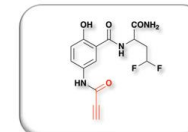
Carbonic anhydrase II inhibitor  
*Chem. Sci.*, **2011**, *2*, 770-775



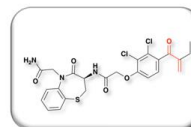
HSP70 inhibitor  
*Chem. Sci.*, **2015**, *6*, 739-744



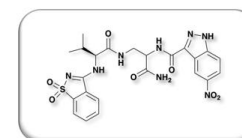
Streptavidin binder  
*OBC* **2012**, *10*, 1502-1505



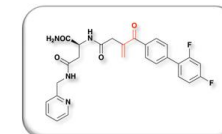
HERBB2 inhibitor  
*Chem. Sci.* **2013**, *4*, 2088-2092  
*Med. Chem. Commun.*, **2016**, *7*, 1340-1351



Covalent bromodomain inhibitor  
(BRD2-4, EP300, PHIP)  
*Angew. Chem. Int. Ed.* **2015**, *54*, 6057-61

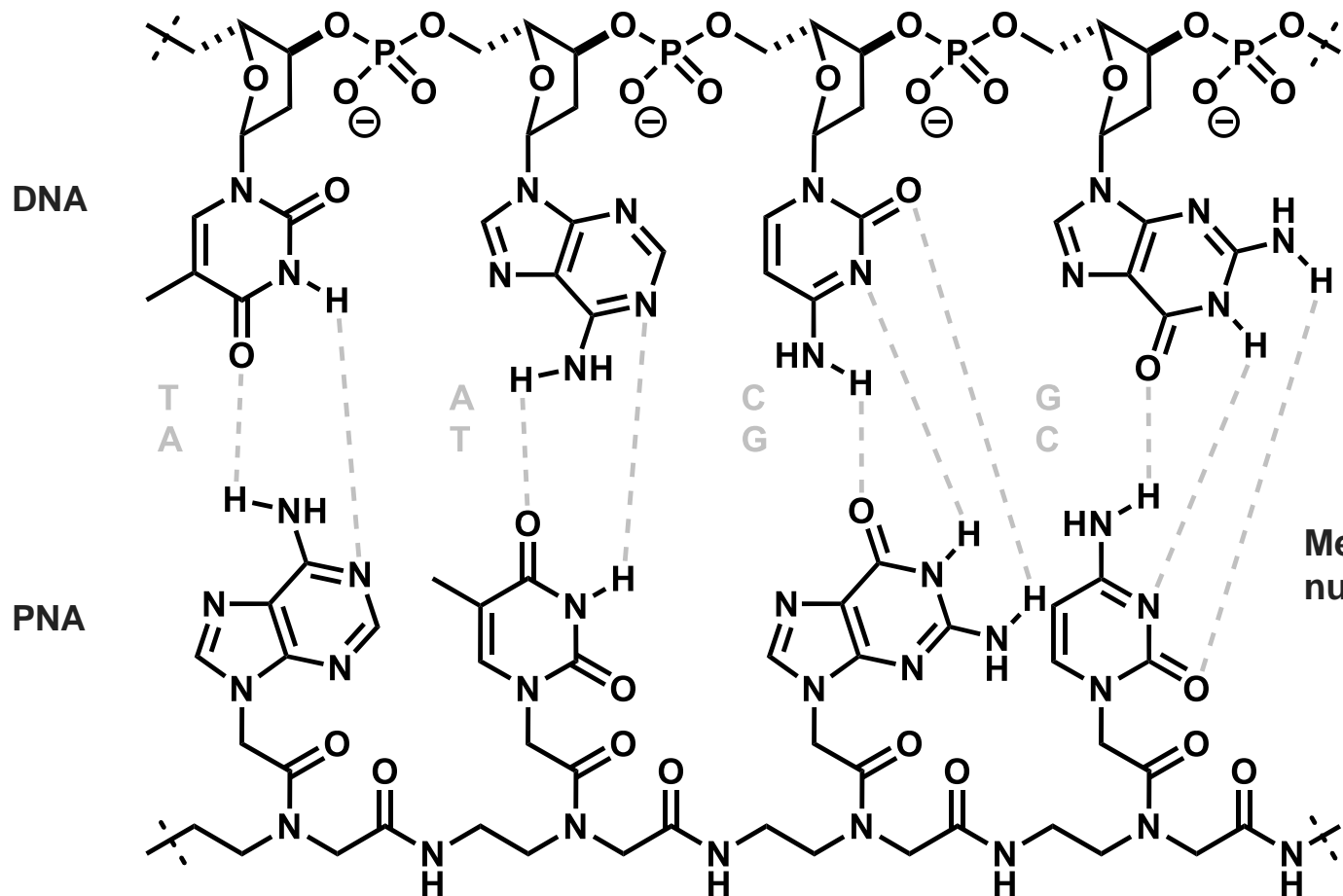


PTB1B inhibitor (selective vs TCPTP)  
*Bioorg. Med. Chem. Lett.* **2016**, *26*, 1080-1085



MEK inhibitor  
*Med. Chem. Commun.*, **2016**, *7*, 1340-1351

# Peptide Nucleic Acid (PNA)



Melting temperature of nucleic acid duplexes (15mer)

duplex	$T_m$
PNA-DNA	69 °C
DNA-DNA	54 °C

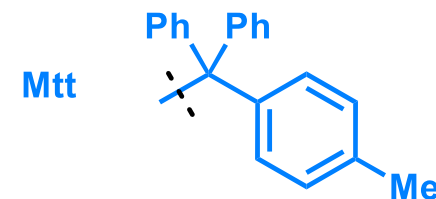
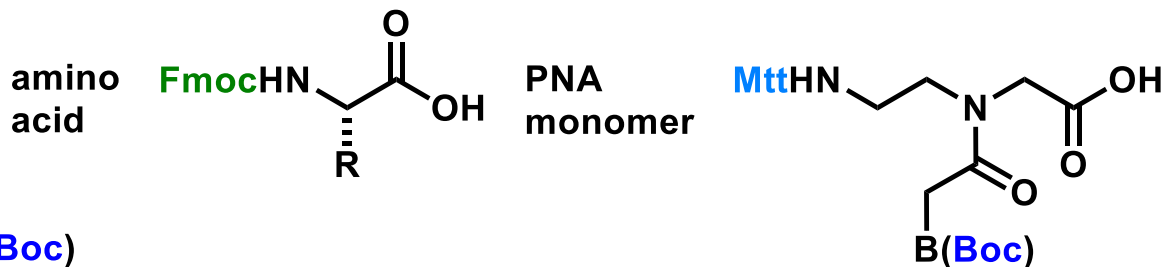
## PNA:

Artificial nucleic acid whose sugar units are substituted to amide and chain consists of aminoethylglycine instead of phosphodiester

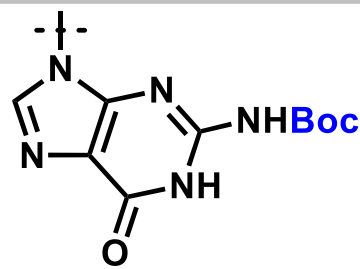
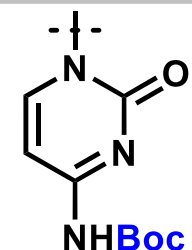
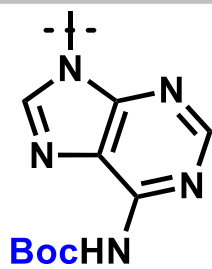
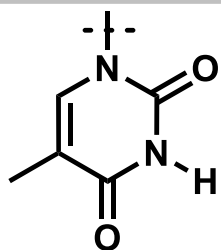
- Wide compatibility with standard solid-phase synthesis
- High stability (temperature, pH, enzymatic degradation, duplex)

# Solid-Phase Synthesis of PNA-Peptide Conjugate

12

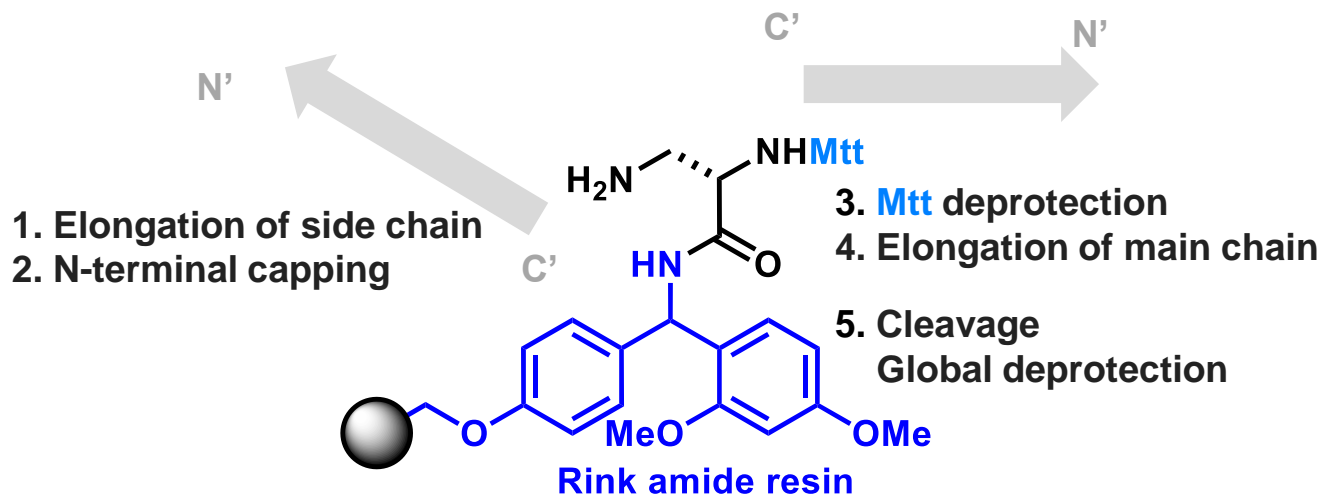


B(Boc)



**Mtt**: monomethyltrityl  
DAP: 2,3-diaminopropionic acid

**Mtt/Fmoc**-based solid-phase synthesis starting from **Mtt**-DAP-OH attached to Rink amide resin



## Elongation

- a. peptide
  - i. amide coupling
  - ii. **Fmoc** deprotection
- b. PNA
  - i. amide coupling
  - ii. **Mtt** deprotection

## **Fmoc** deprotection

piperidine/DMF (1/4)

## **Mtt** deprotection

(CF<sub>3</sub>)<sub>2</sub>CHOH/(CH<sub>2</sub>Cl)<sub>2</sub> (1/1)

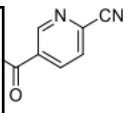
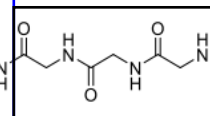
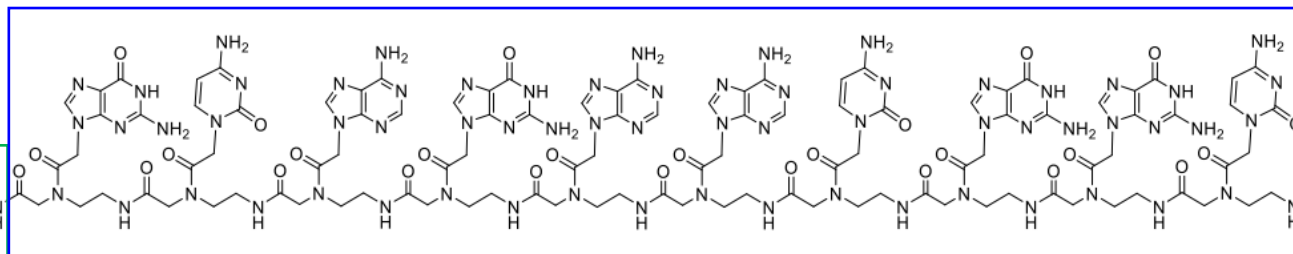
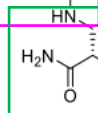
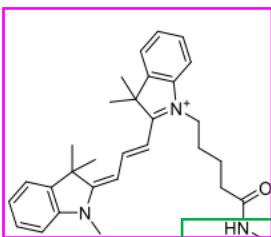
## Cleavage

TFA/PhOH/H<sub>2</sub>O/(*i*-Pr)<sub>3</sub>SiH  
(88/5/5/2)

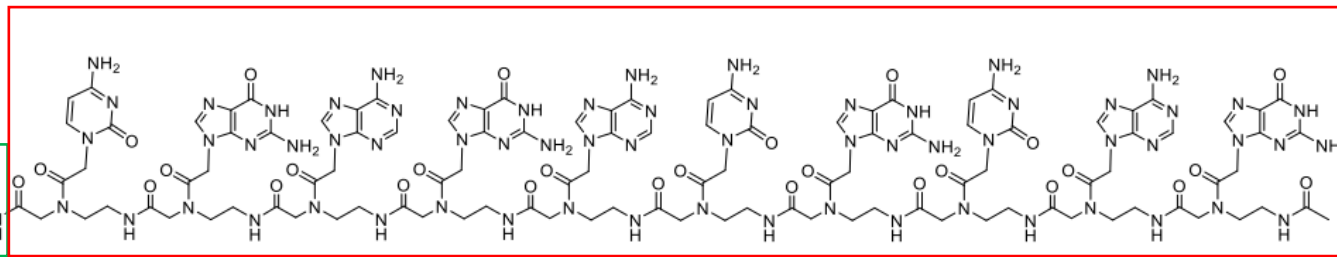
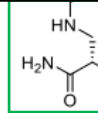
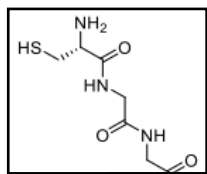
# Prepared PNA-Peptide Conjugate in Main Paper

13

C' **DAP(Cy3)-GCAGAACGGC**-Gly-Gly-Gly-CN N'



C' **DAP(Gly-Gly-Cys)-CGAAGACGCAG**-Ac N'



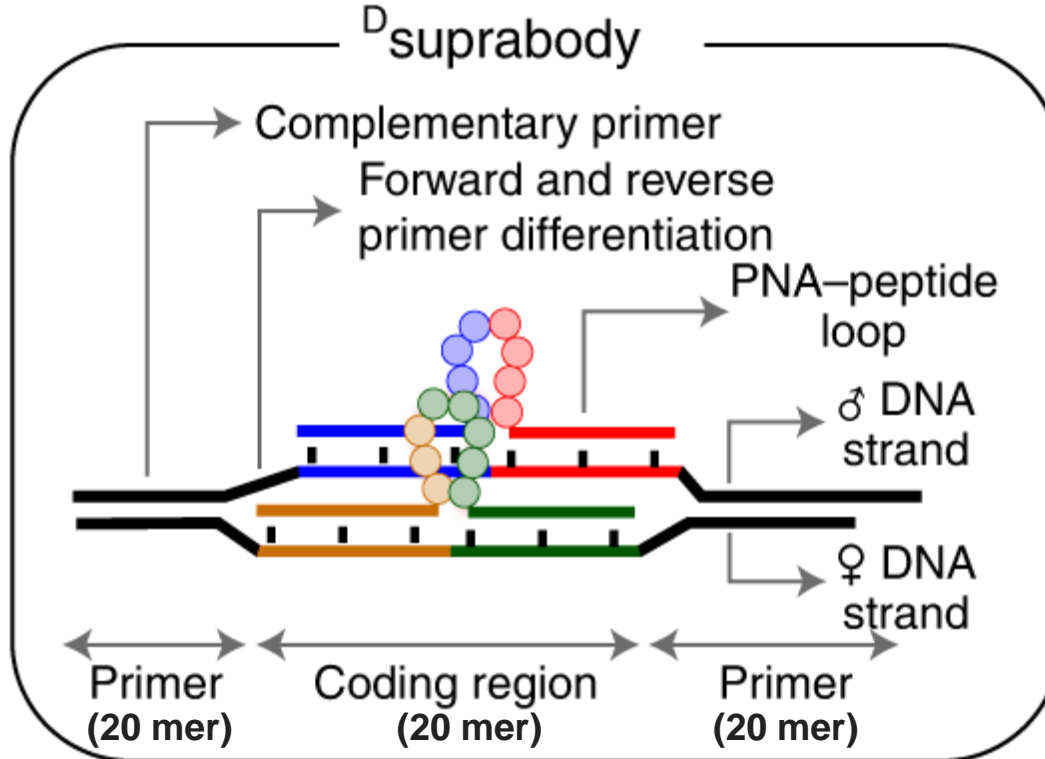


- Introduction
  - DNA-encoded library
  - Genetic algorithm
  - PNA
- **A mating mechanism to generate diversity for the Darwinian selection of DNA-encoded synthetic molecules**

Vummidi, B. R.; Farrera-Soler, L.; Daguer, J.-P.; Dockerill, M.; Barluenga, S.; Winssinger, N. *Nat. Chem.* **2022**, *14*, 141.

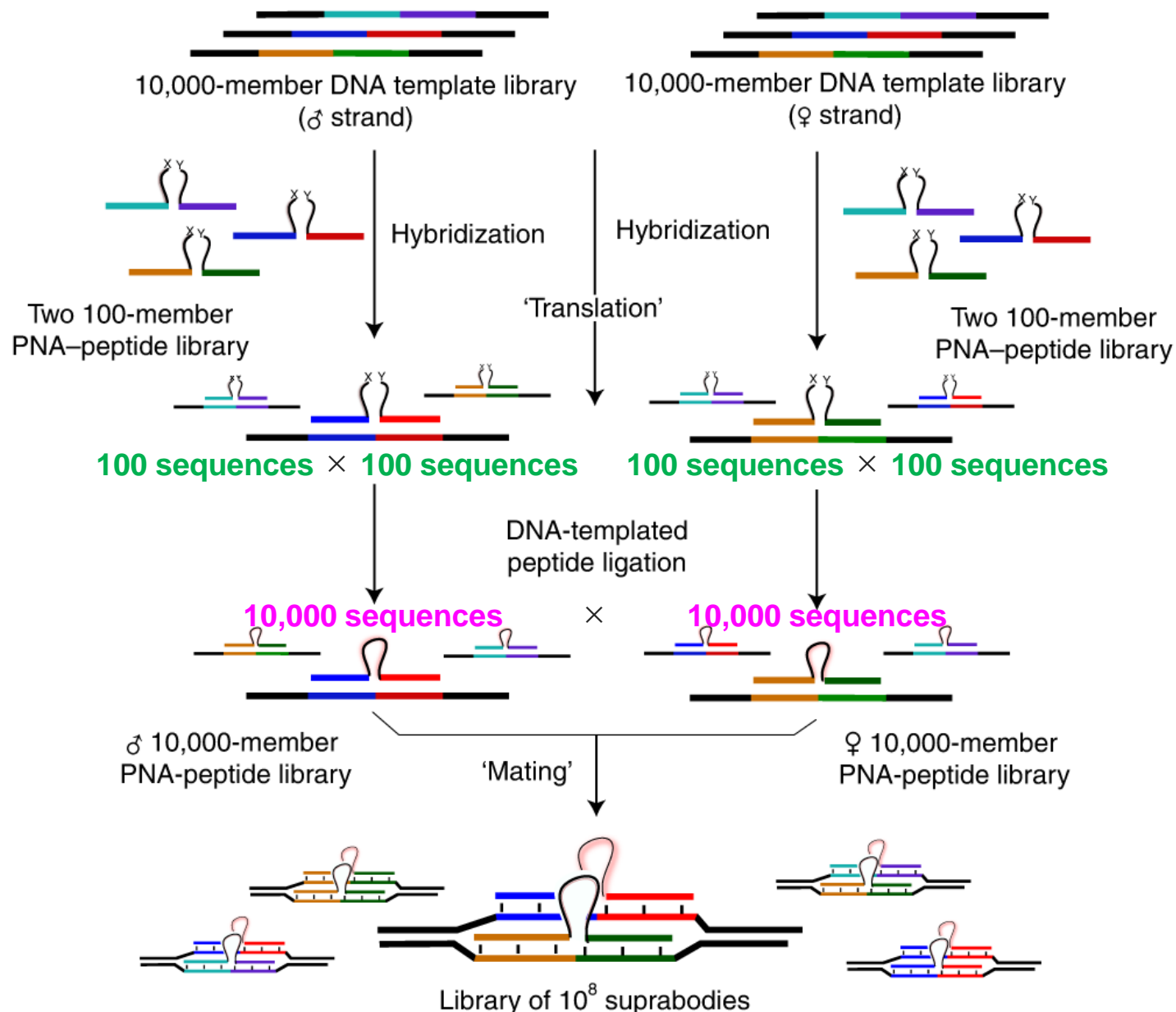
# <sup>D</sup>Suprabody

<sup>D</sup>suprabody (DNA suprabody): designed supramolecular composed of DNA, PNA, and peptide which would mimic structural features of an antibody



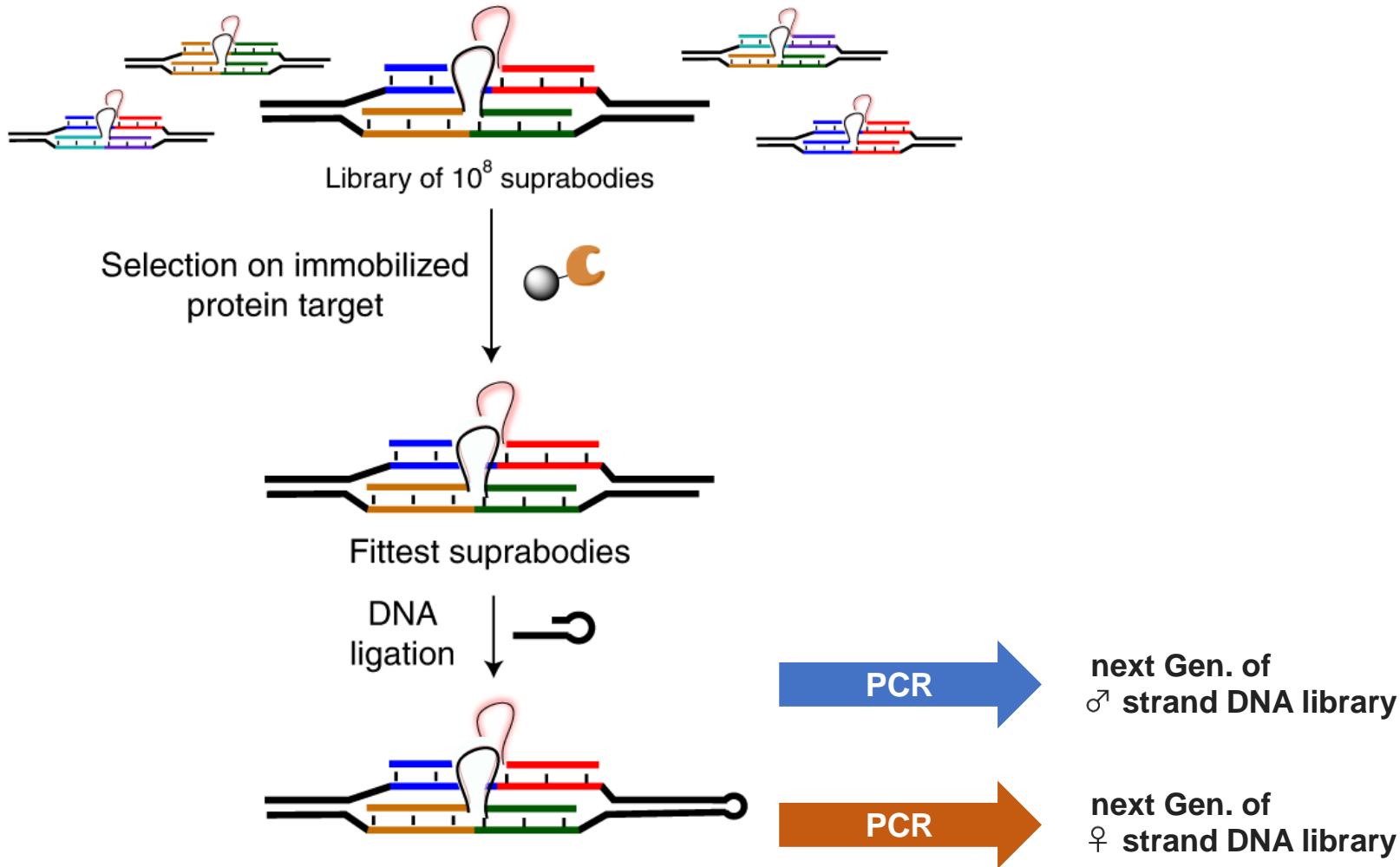
	peptide	PNA (10 mer)	DNA (60 mer)
function	two macrocyclic-like structure	encode the sequence of the peptide	complementarily retain the encoded information
features	restrained conformational profile	simple SPPS of PNA-peptide sequences high PNA-DNA duplex stability	amplification by PCR squaring the diversity by hybridization ( $\sigma \times \rho$ )

# Diversification of <sup>D</sup>Suprabody Library



**Combination of chemical ligation and hybridization realizes  $10^8$ -member diversity scale from 400-member PNA-peptide conjugates.**

# Selection and Amplification of <sup>D</sup>Suprabody Library 17



Reiterative rounds are carried out to let the DNA library be converged to the fittest.

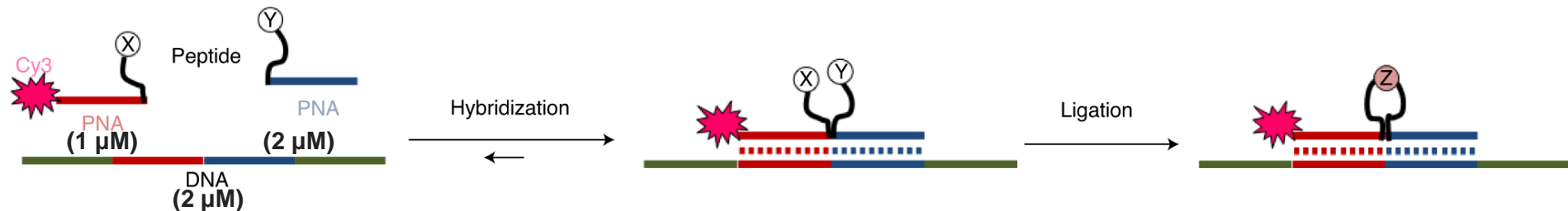
## Proof of Concept (PoC)

1. Diversity: chemical ligation
2. Selection/Amplification: enrichment and convergence to the fittest

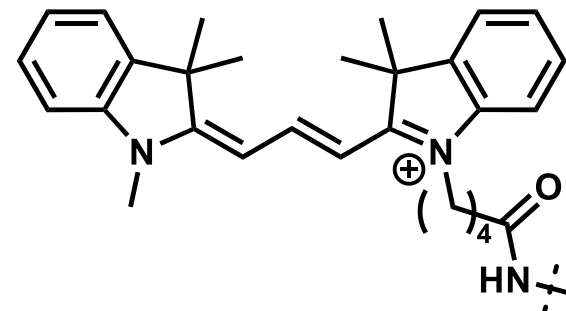
# PoC 1. Requirements for the Templated Ligations

18

DNA-templated peptide ligation amenable to suprabody library generation



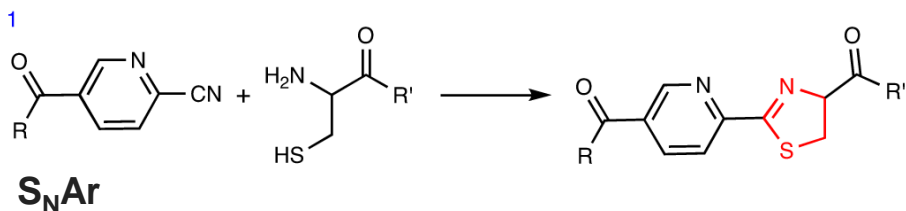
Cy3: Cyanine3



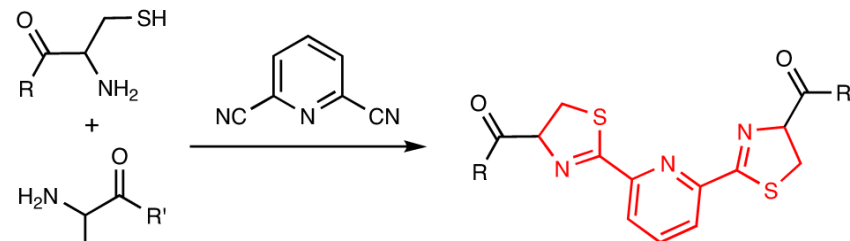
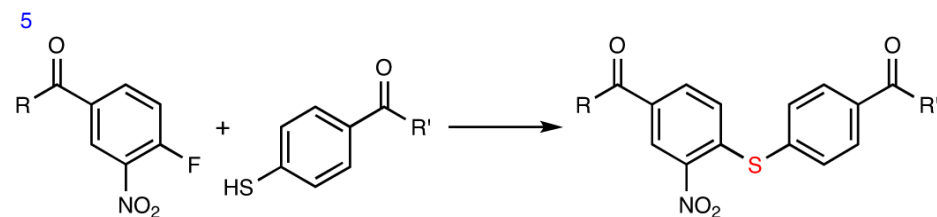
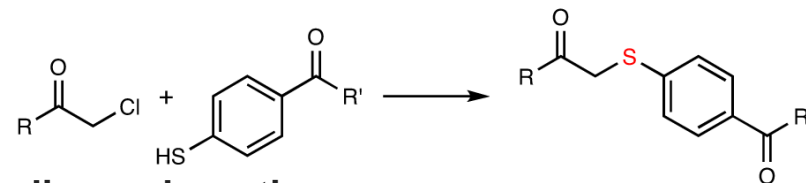
## requirements

- ✓ Compatible in an aqueous environment
- ✓ Faster than background bimolecular reactions

tested reactions (incubated in PBSt<sup>1</sup> at rt, 3 h)  
cysteine condensation



$\text{S}_{\text{N}}2$  6

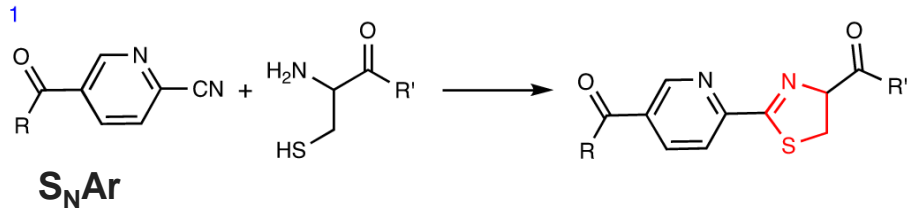


1. 0.01 M phosphate buffer, 2.7 mM KCl and 0.137 M NaCl, 0.05% Tween-20, pH 7.4

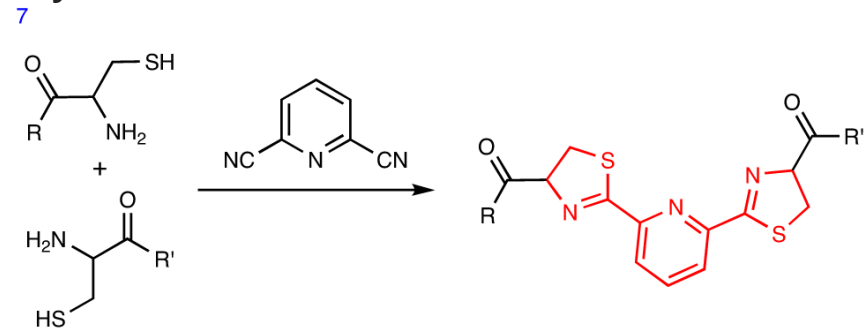
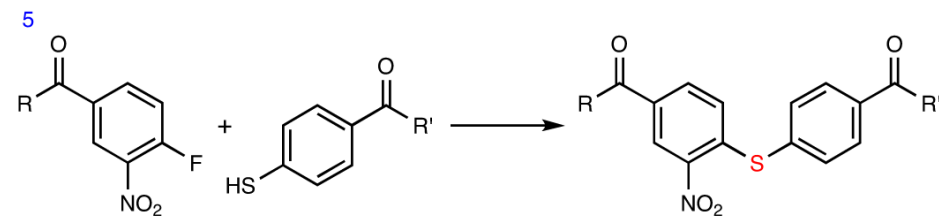
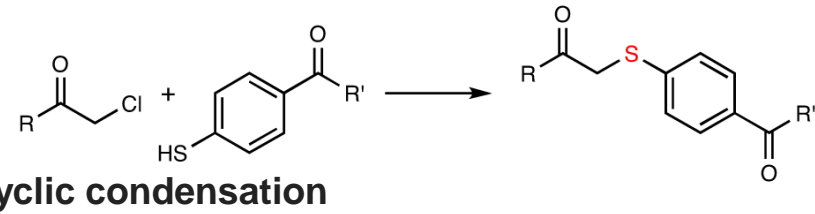


# PoC 1. Monitoring the Templated Ligations

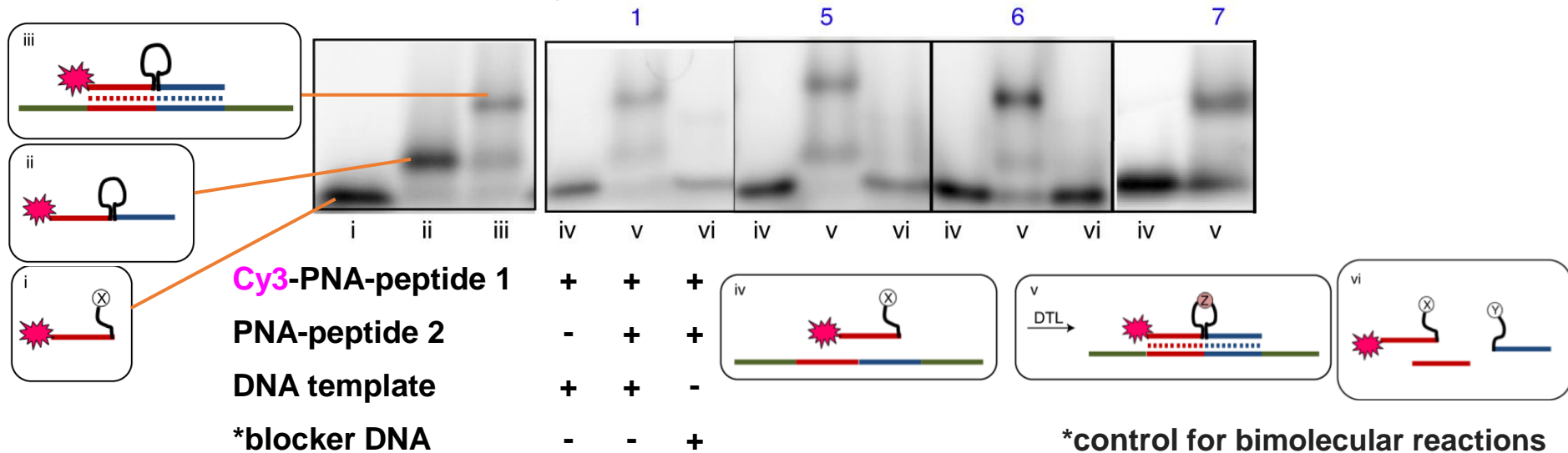
tested reactions  
cysteine condensation



$S_N2$  6



SDS-PAGE analysis with **Cy3** scanning



- Various reactions could be carried out to ligate peptidic fragments in a templated fashion.
- Cysteine condensation was employed for the construction of Dsuprabody library.

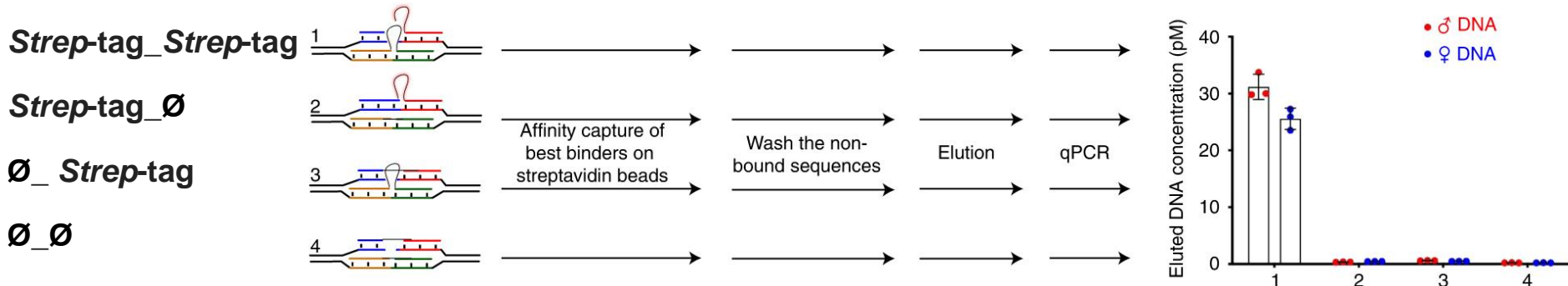
# PoC 2. Selection of *Strep*-Tag<sup>D</sup>Suprabody

## *Strep*-tag II (WSHPQFEK)

- Moderate binding affinity to streptavidin ( $K_D = 72 \mu\text{M}$ )<sup>1</sup>
- Dimeric *Strep*-tag showed higher avidity<sup>2</sup>

selection/amplification process of *Strep*-tag<sup>D</sup>suprabody

♂ PNA\_ ♀ PNA

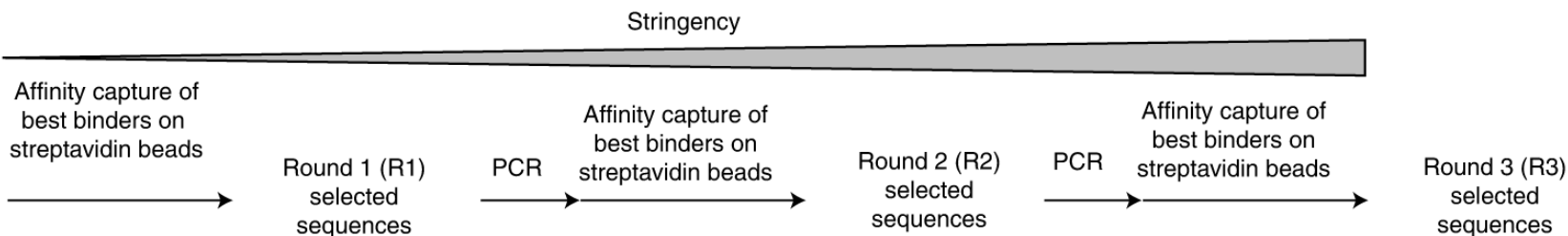


Only the dimeric *Strep*-tag assembly was successfully selected by binding affinity and amplified.

1. Schmidt, T. G. M.; Koepke, J.; Frank, R.; Skerra, A. *J. Mol. Biol.* **1996**, *255*, 753.
2. Busby, M. Stadler, L. K. J.; Ferrigno, P. K.; Davis, J. J. *Biophys. Chem.* **2010**, *152*, 170.

# PoC 2. Enrichment of the Fittest <sup>D</sup>Suprabody

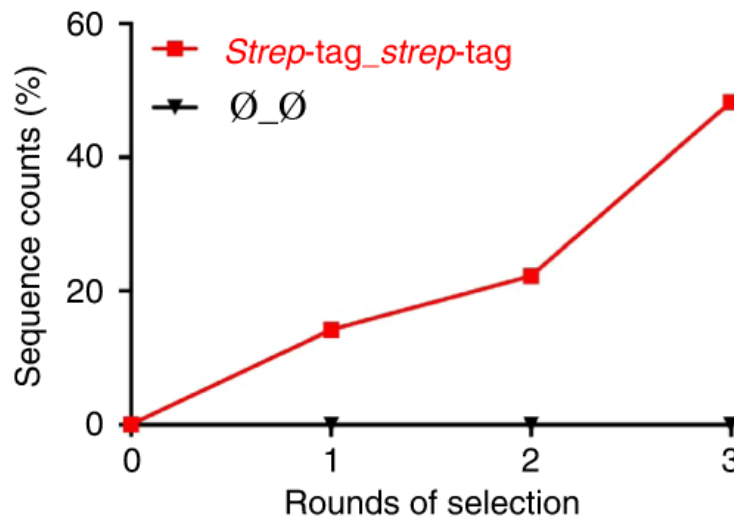
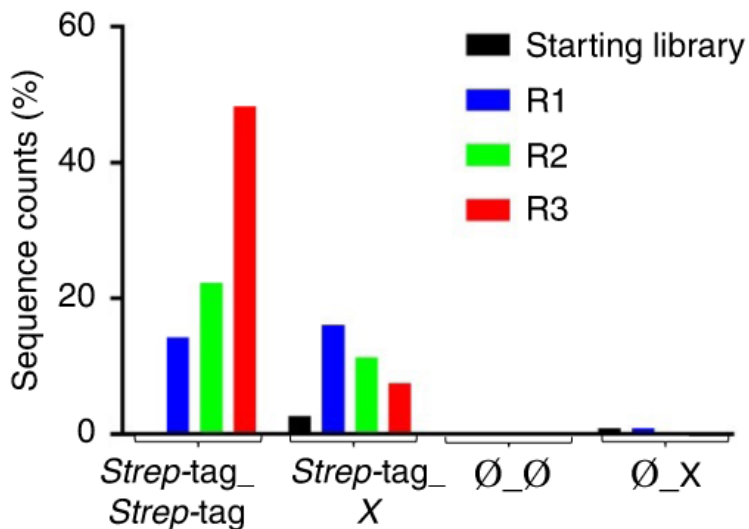
Starting library of 10<sup>8</sup> suprabodies



**Binders in the starting library**  
**strong binder (Strep-tag\_Strep-tag)**

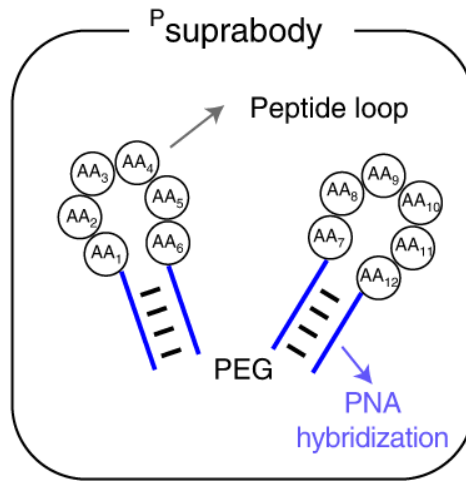
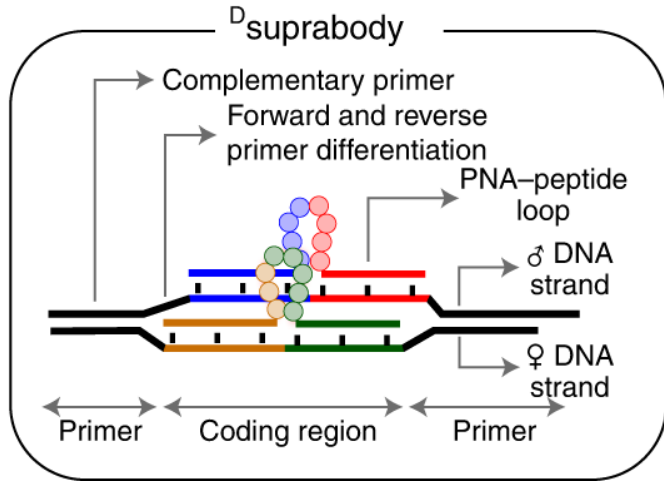
**moderate binder (Strep-tag\_X)**  
 × 10,000 more abundant than Strep-tag\_Strep-tag

**Strep-tag** positive control  
 ∅ negative control  
 X other sequences



Reiterative rounds of selection, amplification and mating enhanced the convergence to the fittest <sup>D</sup>suprabody.

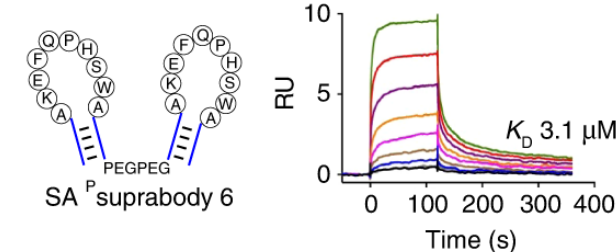
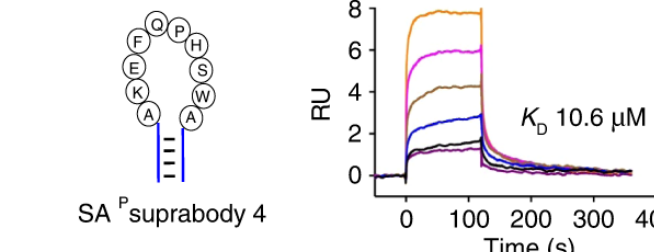
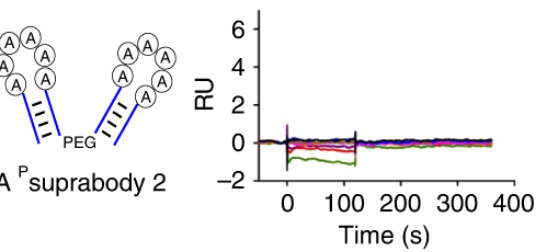
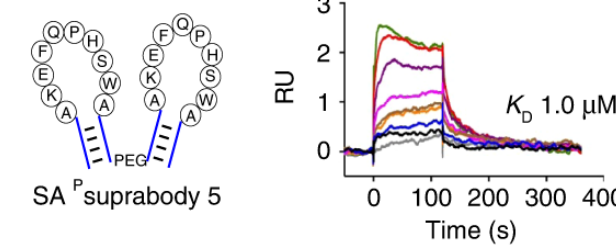
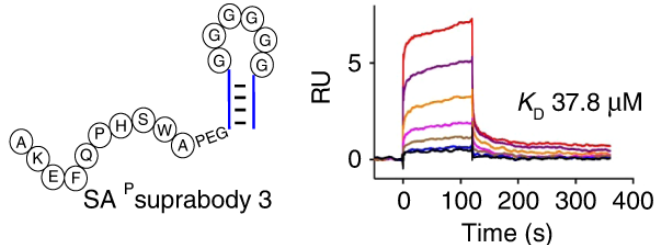
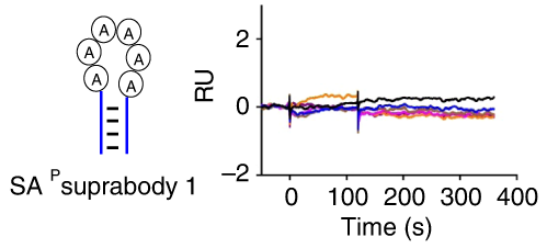
# PoC 2. Evaluation of the Converged Sequence



DNA suprabody (<sup>D</sup>suprabody)

- simplify the structure (~7 kDa)
- retain conformational state

PNA suprabody (<sup>P</sup>suprabody)



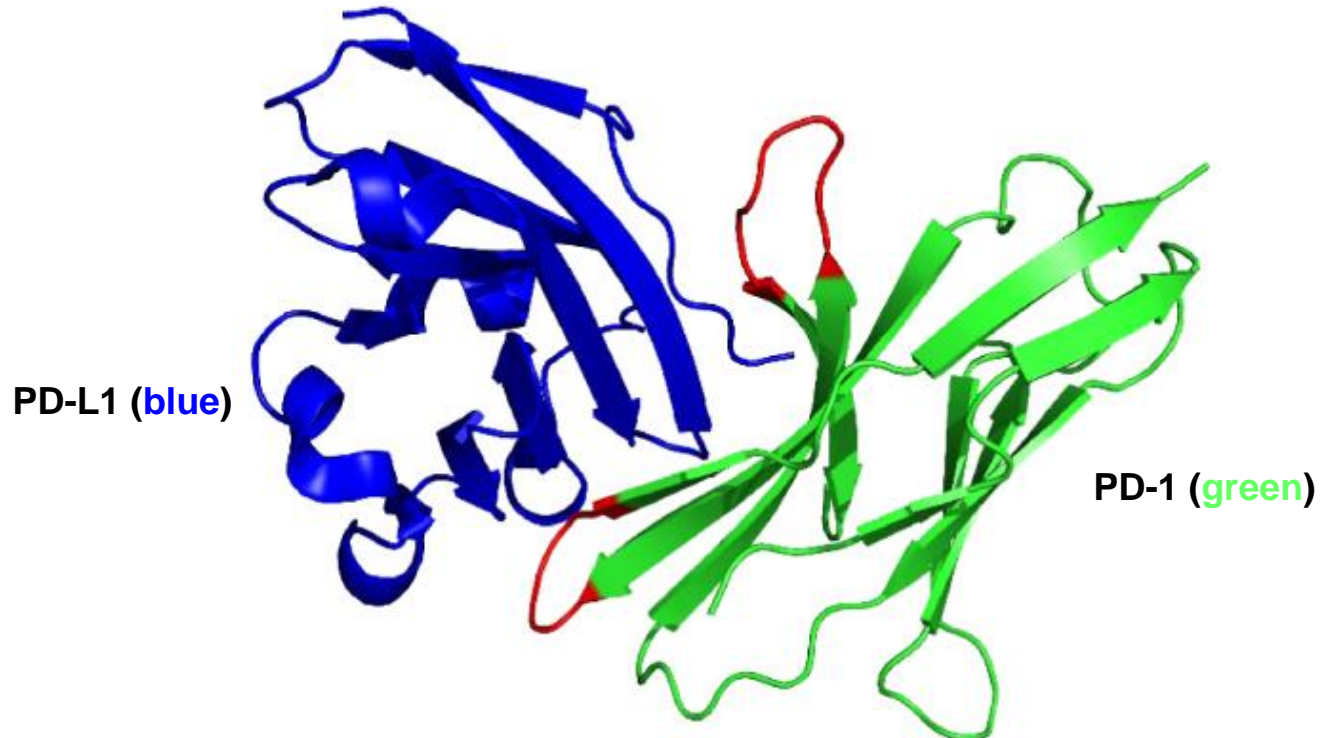
SA: streptavidin

- The affinity of the enriched PNA-peptide hairpin loops was evaluated with corresponding <sup>P</sup>suprabodies.
- The length of the linker in a <sup>P</sup>suprabody influenced the affinity.

# Ligand Discovery for PD-L1

programmed death ligand-1 (PD-L1) <sup>1,2</sup>

- immunological checkpoint which binds to PD-1 on T-cell
- overexpression of PD-L1 on cancer cells promotes the evasion from immune system surveillance.



PD-L1 (blue)

PD-1 (green)

two peptide loops (red) of PD-1 responsible for interaction of PD-L1 with PD-1<sup>3</sup>.

**Hypothesis: The geometry of the peptide loop can be mimicked by a <sup>D</sup>suprabody architecture.**

1. PDB ID: 4ZQK 2. Freeman, G. J.; Long, A. J.; Iwai, Y.; Bourque, K.; Chernova, T.; Nishimura, H.; Fitz, L. J.; Malenkovich, N.; Okazaki, T.; Byrne, M. C.; Horton, H. F.; Fouser, L.; Carter, L.; Ling, V.; Bowman, M. R.; Carreno, B. M.; Collins, M.; Wood, C. R.; Honjo, T. *J. Exp. Med.* **2000**, *192*, 1027. 3. Zak, K. M.; Kitel, R.; Przetocka, S.; Golik, P.; Guzik, K.; Musielak, B.; Doemling, A.; Dubin, G. Holak, T. A. *Structure* **2015**, *23*, 2341.



# Hypothesis Verification: Pilot <sup>D</sup>Suprabody Library

25 + 25 PNA-peptide conjugates

♂ strand 1-25

×

♀ strand 26-50

(wildtype sequence: 1-26)

Starting library of 625 <sup>D</sup>suprabodies derived from PD-1 loops

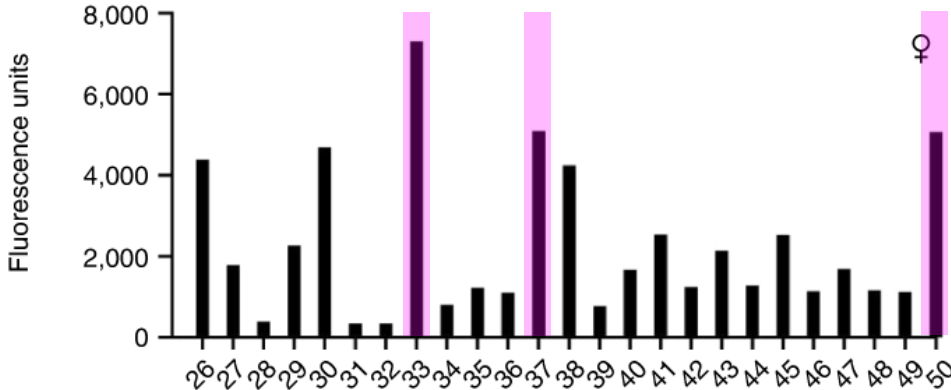
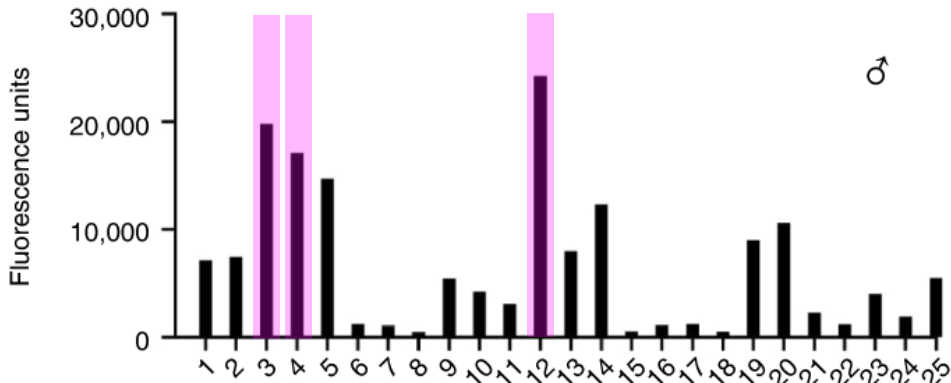
Affinity capture of best binders on PD-L1-loaded beads

single round

Identification of selected individual loops by microarray screening

pick up top3 strands (♂, ♀)

Fluorescence intensity on the microarray



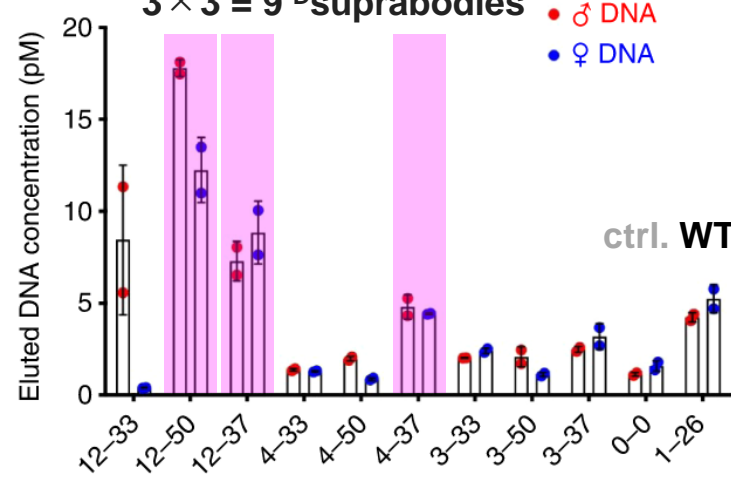
♂ selected loops

- 12 - VLNWYRMSAA<sub>11</sub>SNQTDKLA AF
- 4 - VLNWWRMSPSNQTDKLA AF
- 3 - VLNWAA<sub>2</sub>RMSPSNQTDKLA AF

♀ selected loops

- 33 - SLAPKAQIKESL
- 50 - LCGAMSVAPAA<sub>27</sub>LQFKESL
- 37 - LCGAVSIAPKAQFKESL

3 × 3 = 9 <sup>D</sup>suprabodies

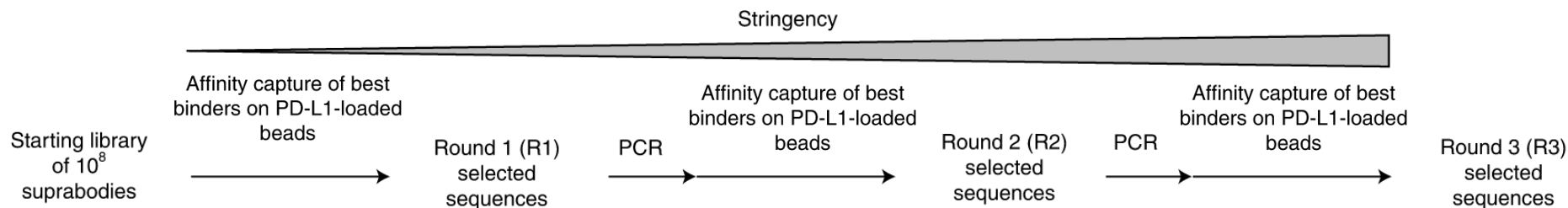


♂ - ♀	$K_D$ of <sup>P</sup> suprabody*
4-37	264 nM
1-26	640 nM

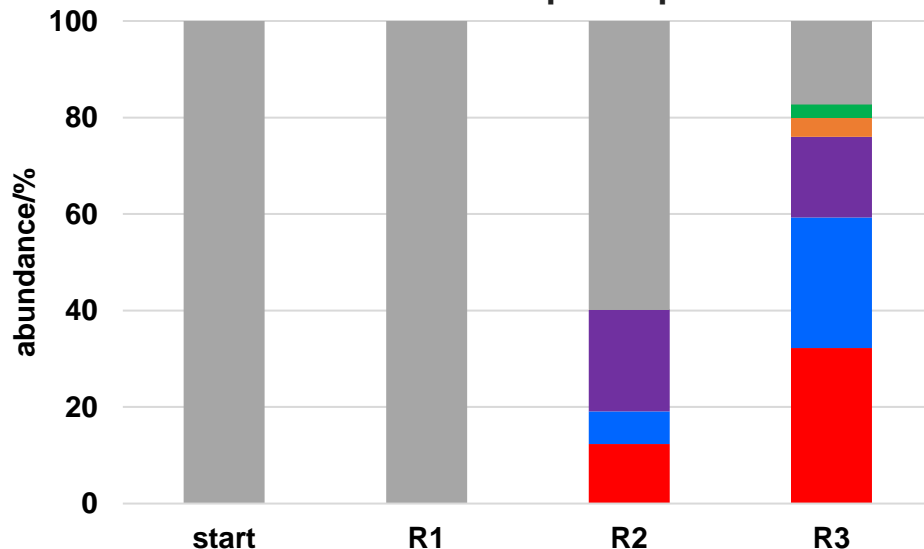
\*measured by BLItz

PD-L1 is a reasonable target for the ligand discovery based on <sup>D</sup>suprabody library.

# PD-L1 Selection with the 10<sup>8</sup>-Member Library



Transition of abundance of the top5 <sup>D</sup>suprabodies at R3

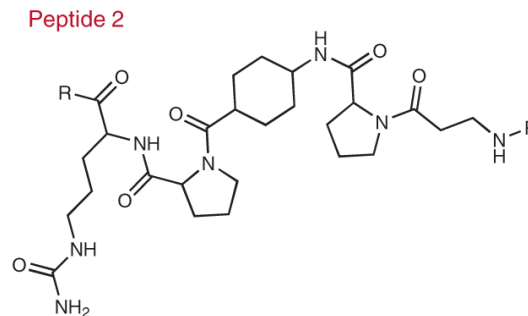
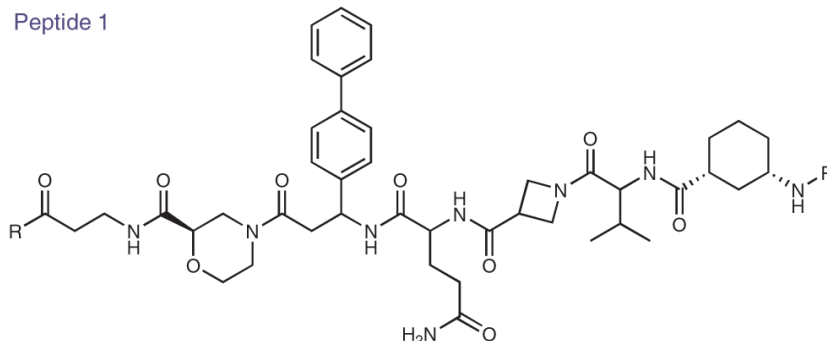
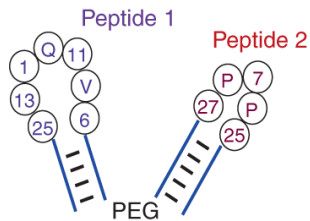


code (♂\_♀)  
 69.40\_47.94  
 94.74\_47.94  
 32.74\_76.38  
 32.74\_47.94  
 32.74\_68.37

K<sub>D</sub> of <sup>P</sup>suprabody\*  
 n.d.  
 70 nM  
 n.d.  
 9,600 nM  
 74 nM

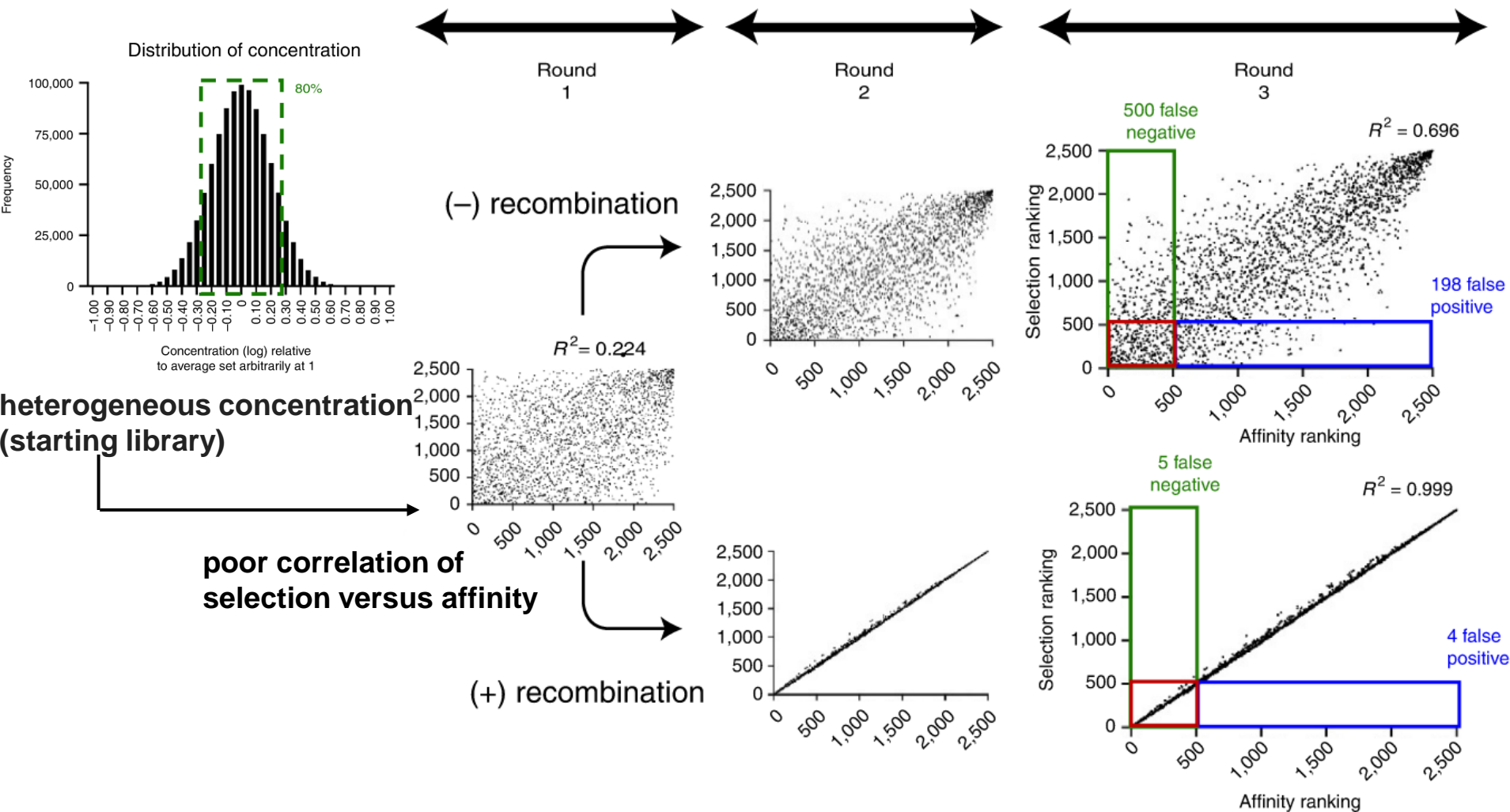
\*measured by BLItz

PD-L1 <sup>P</sup>suprabody 94.74\_47.94



Reiterative evolution of an unbiased library converged to a surpassing binder to PD-L1.

# Computational Simulation for Recombination



Recombination contributed the improvement of sensitivity/specificity of the selection.

# Summary

