# Rational design of a next-generation synthetic Fab library for antibody discovery

Nick Hollmer, Youssef Atef AbdelAlim, Chang Yang, Ngan Ho, Mina Abdollahi, Filmawit Belay, Zachary Anderson, Sophia Ulmer, Matt Salotto, Murali Anuganti, Roushu Zhang, Anita Ghosh, Mahesh Vangala, Satyendra Kumar, Haisun Zhu, Rob Meijers, Deepash Kothiwal

Institute for Protein Innovation, Harvard Institutes of Medicine, 4 Blackfan Circle, Boston, MA 02115

Institute for Protein Innovation

## Introduction

Fragment antigen binding (Fab) display on the yeast cell surface is a powerful method for antibody discovery. Synthetic libraries are popular tools but often are plagued with non-productive sequences, leading to low functional diversity and display. Developability is another major issue due to the presence of liability motifs, causing failures during antibody production and development. To overcome these obstacles, we first must understand the characteristics of naturally occurring complementary determining regions (CDRs) in the human antibody repertoire.

The heavy chain CDR3 (HCDR3) is particularly important due to its dominant role in defining antigen specificity.[1] Amino acids within HCDR3 are typically distributed non-randomly, exhibiting length and position specificity.[2] At the same time, not all HCDR3 lengths are equally represented, suggesting a bias towards certain lengths.[3] Incorporating these features becomes paramount to enhance the frequency of functional Fabs within synthetic libraries.

Similarly, many antibodies fail during production due to developability issues. Extensive research has been conducted in this area, leading to the identification of amino acid motifs detrimental for developability. Using this comprehensive knowledge, we have designed an antigen-agnostic Fab library incorporating liability-free synthetic HCDR3s, thereby establishing a robust platform for antibody discovery capable of targeting a wide range of antigens.

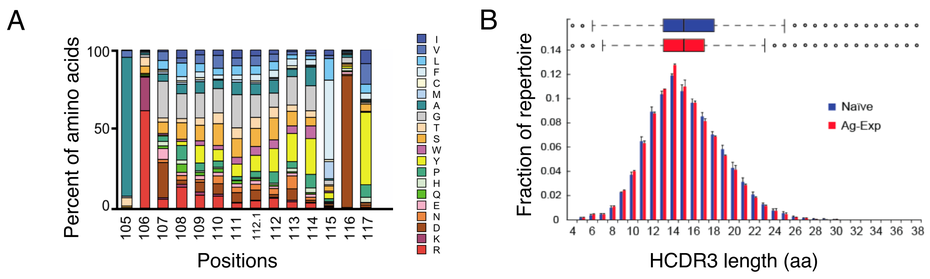


**Figure 1**. **Characteristics of human naïve HCDR3s**. **(A)** Human B cell sequencing data showing HCDR3 amino acid frequency[1] and **(B)** Length distribution.[2]

## Objectives

To design a high diversity, next-generation Fab library with desirable characteristics for yeast display :

- Position and length-specific amino acid distribution for HCDR3.
- Productive HCDR3 lengths.
- Liability-free CDRs.

## Library design

- Analyzed ~1.5 M unique naive HCDR3 sequences from the OAS database.
- 20 HC:LC combinations, 5 heavy chains (VH1-69, VH3-23, VH3-7, VH4-34 and VH5-51) and 4 light chains (VK1-39, VK3-15, VK3-20 and VK4-1).
- 8.55E+11 liability free HCDR3 sequences.
- Light chain barcodes to determine HC:LC pairing.

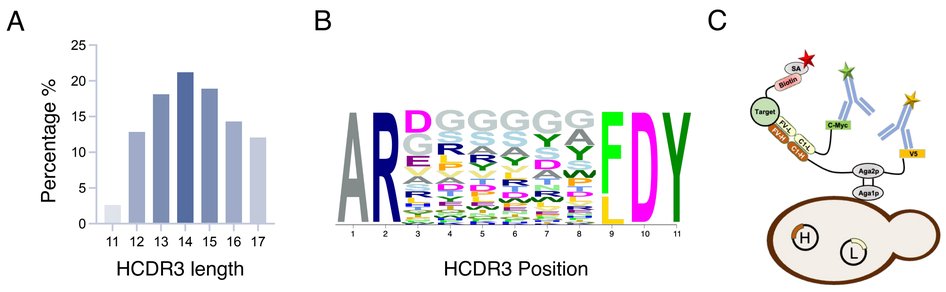


**Figure 2**. **IPI's design to create synthetic HCDR3s for the next-generation Fab library**. **(A)** Distribution of selected HCDR3 lengths. **(B)** Amino acid frequencies for an 11 AA long HCDR3. **(C)** IPI's yeast display and selection system.

## Results

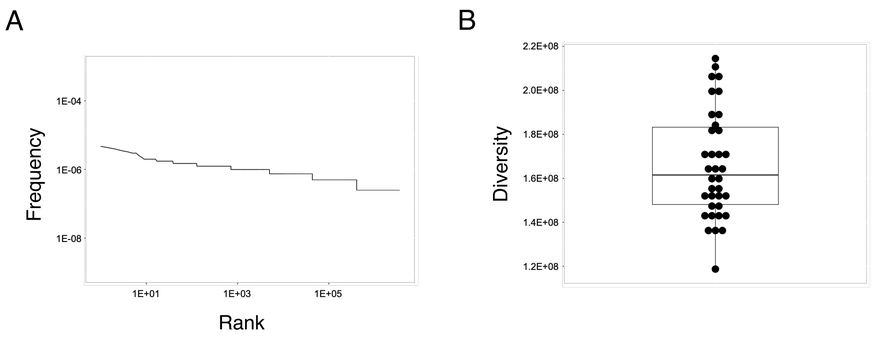


**Figure 3**. **NGS characterization of libraries built using heavy chain VH1-69**. **(A)** HCDR3 frequency and dominance plot. **(B)** Initial analysis of HCDR3 diversity using Capture-Recpature.[4]

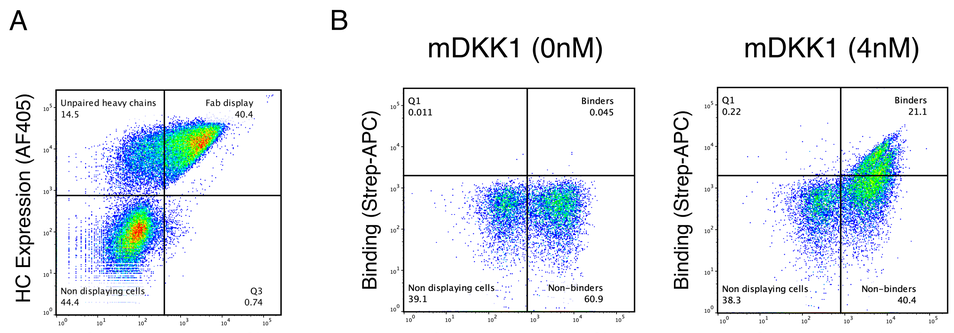


**Figure 4**. **Library display and selection**. **(A)** Flow cytometry analysis of Fab display on the yeast cell surface and **(B)** the final selected population against biotinylated mDKK1 at 4nM.

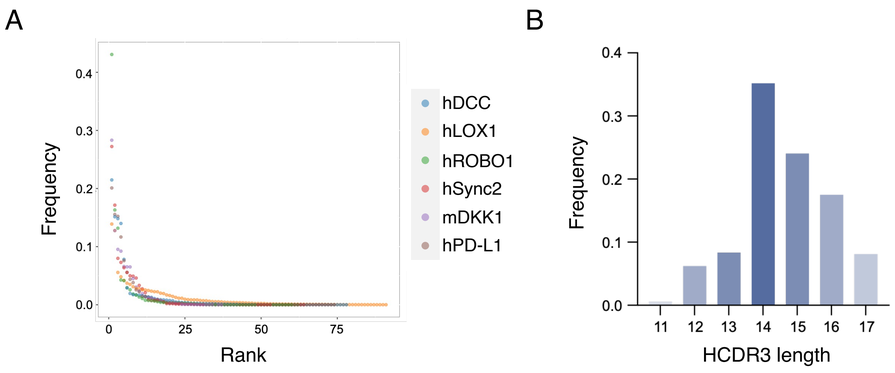


**Figure 5**. **NGS analysis of HCDR3 sequences after final selection**. **(A)** Frequency and dominance plot showing enriched but diverse output not dominated by one or a few sequences. **(B)** Bar graph showing HCDR3 length distribution of selected population.

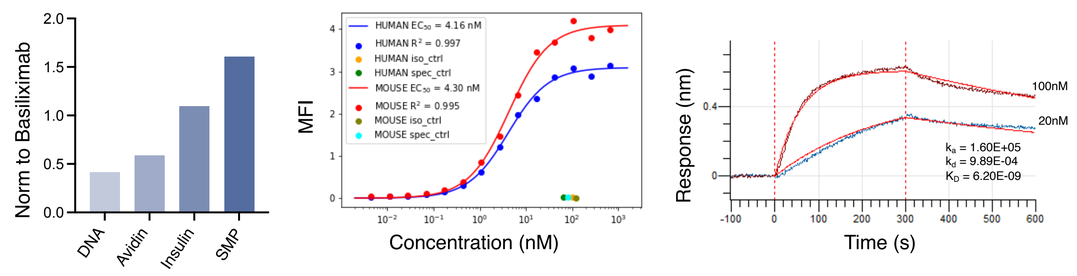


**Figure 6**. **Characterization of a mDKK1 antibody (Ab06)**. **(A)** Bar graph showing binding to various non-specific antigens. **(B)** Cell display showing cross-reactivity of Ab06 to hDKK1 and single-digit nanomolar binding to antigen displayed on mammalian cells**. (C)** Antigen-antibody interaction kinetics using Bio Layer Interferometry (BLI)**.**

## Summary

Here, we designed a Fab library with synthetic HCDR3 sequences, where the amino acid distribution mirrored that of naive human antibody repertoire. This approach resulted in higher Fab display and greater functional diversity. We also used multiple productive HCDR3 lengths, resulting in a very diverse CDR3 population. Initial NGS analysis suggests that our library contains approximately 1.6E+08 unique CDR3 sequences. When combined with 20 different germline heavy and light chain combinations, the diversity increases to around 3E+09 unique Fabs. Importantly, we purged all the CDRs of liability motifs, resulting in a library that produces highly developable antibodies. A discovery campaign against a panel of antigens resulted in a diverse Fab population that was easily converted into full IgGs and showed strong binding in cell display with single-digit nanomolar $K_D$ in BLI assay. In conclusion, we have created a next-generation, antigen-agnostic synthetic Fab library that is a robust platform for generating developable antibodies against a wide range of target antigens.

## References

1. Xu, J. L. & M. M. Davis 2000, Immunity, 13, 37.
2. Zemlin, M., et al. 2003, J Mol Biol, 334, 733.
3. DeKosky, B. J., et al. 2016, Proc Natl Acad Sci U S A, 113, E2636.
4. Weinstein, J. A., N. Jiang, R. A. White, 3rd, D. S. Fisher & S. R. Quake 2009, Science, 324, 807.

## Acknowledgments

We thank Debora Marks, Andrew Kruse, Joe Hurly, Aaron Kollasch and Nirakar Basnet for helpful discussions.

## Want To Collaborate With Us?

We have generated antibodies to more than 200 cell surface receptor targets.

We are open to collaborations from industry and academia.

Contact us at **https://proteininnovation.org/receptor-engineering/** or via **rob.meijers@proteininnovation.org**

Scan to connect with IPI Receptor Engineering