

# AI-Enabled High-Throughput Platform for Accelerated Bispecific Antibody Discovery

Cassandra Koole, Crystal Richardson, Andrea O'Hara, Andrew Tsao, Prassanna Rao, Ming Yang, Riley Graham and Sumit Kumar

GENEWIZ from Azenta Life Sciences, South Plainfield, NJ 07080

## Abstract

**Introduction:** The rapid evolution of antibody discovery, driven by advances in automation, next-generation sequencing (NGS), and machine learning, has created new opportunities to accelerate identification and development of high-quality therapeutic candidates. Here, we present an integrated high-throughput platform that combines gene synthesis, antibody production, automated screening workflows, and AI-driven codon optimization to streamline antibody discovery.

Our platform enables the rapid identification and scalable production of thousands of antibody candidates by integrating *in vivo* (e.g., B cells, PBMCs) and *in vitro* (e.g., phage display) workflows with NGS. Compared to traditional approaches such as hybridoma technology or random colony screening using Sanger sequencing, NGS enhances sequence diversity and identification efficiency, uncovering more novel candidates. In addition to IgG antibodies, this platform supports the generation of complex formats, including bispecific and multispecific antibodies (BsAbs and MsAbs), which can simultaneously target multiple antigens or epitopes. These modalities are particularly valuable in applications such as cancer immunotherapy, where dual-target engagement can improve specificity and efficacy. Lastly, to address a critical bottleneck in antibody development - achieving high and consistent protein expression - our platform incorporates novel AI/ML-driven codon optimization. Traditional codon optimization approaches rely on static metrics such as codon adaptation index (CAI), GC content, and minimum free energy (MFE), which do not fully capture the complex, non-linear determinants of gene expression. Our AI-driven approach integrates a heuristic design framework with deep learning models trained on large-scale genomic and proprietary high-expression datasets. This enables the identification of sequence patterns associated with enhanced protein production and more accurately predicts expression outcomes across diverse constructs.

**Results:** The results of the application of this integrated platform demonstrated significant improvements across both discovery and expression stages. NGS-enabled workflows consistently yielded 5–50x more unique antibody sequences compared to conventional methods, substantially expanding the candidate pool. High-throughput screening pipelines successfully identified functional BsAbs candidates with desired binding profiles in reduced timelines.

In expression studies, AI-optimized sequences outperformed traditionally optimized constructs, resulting in higher antibody yields and improved consistency across multiple formats, including bispecific constructs. Wet-lab validation confirmed increased expression levels and reduced variability, enabling more efficient scale-up and downstream development.

**Conclusion:** These results highlight the power of combining high-throughput discovery technologies with AI-driven sequence optimization to transform antibody development. By addressing key bottlenecks in both candidate identification and protein expression, this integrated platform enables faster, more efficient generation of high-quality antibody therapeutics. The ability to rapidly discover and express complex modalities, such as BsAbs, positions this approach as a robust and scalable solution for next-generation biologics development.

## Machine Learning Enabled Antibody Production

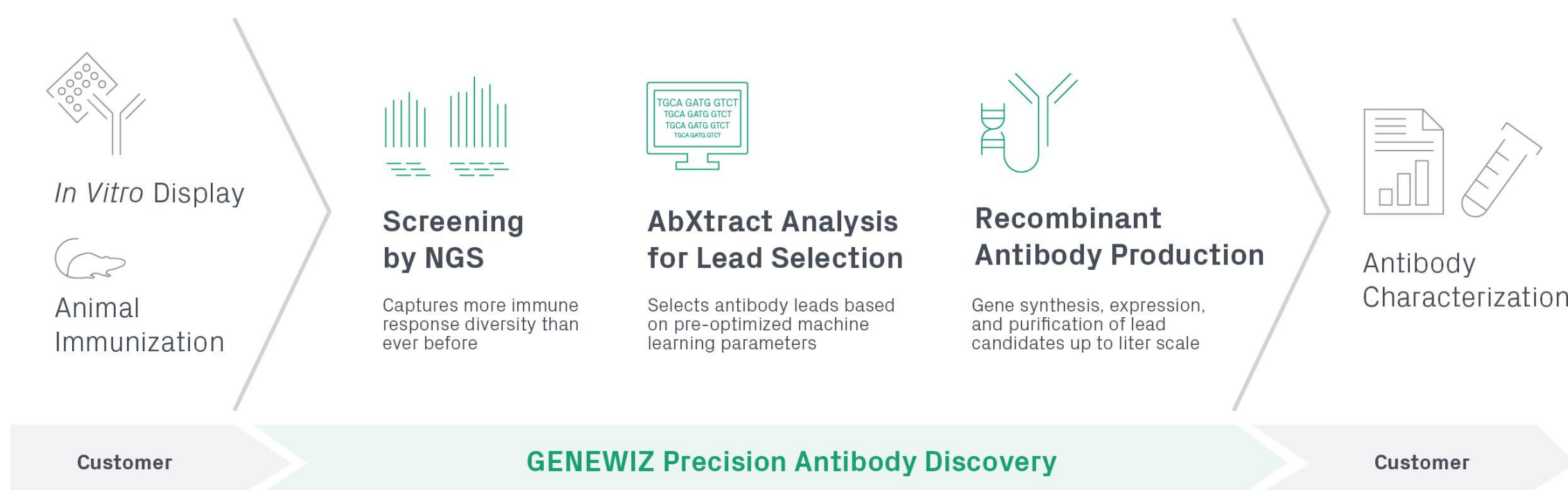
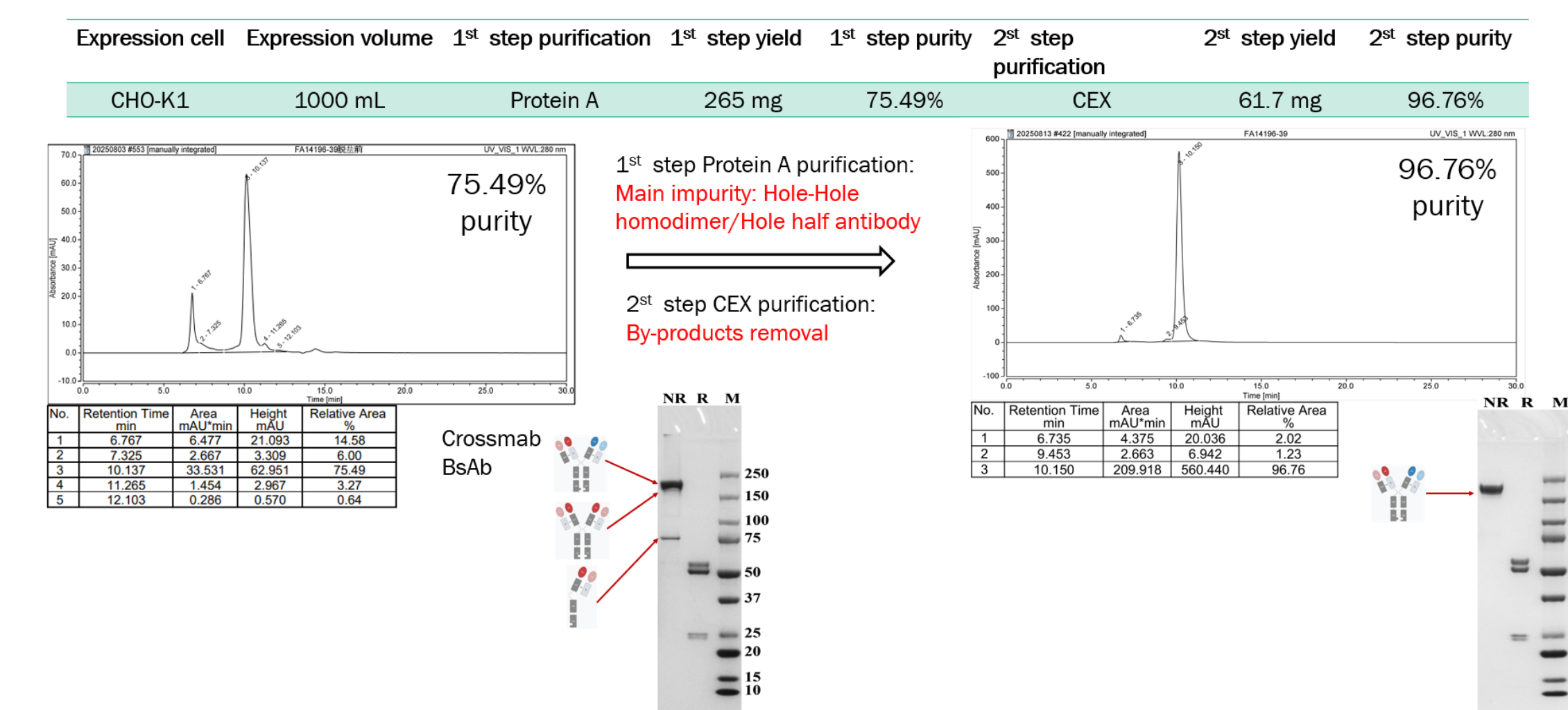


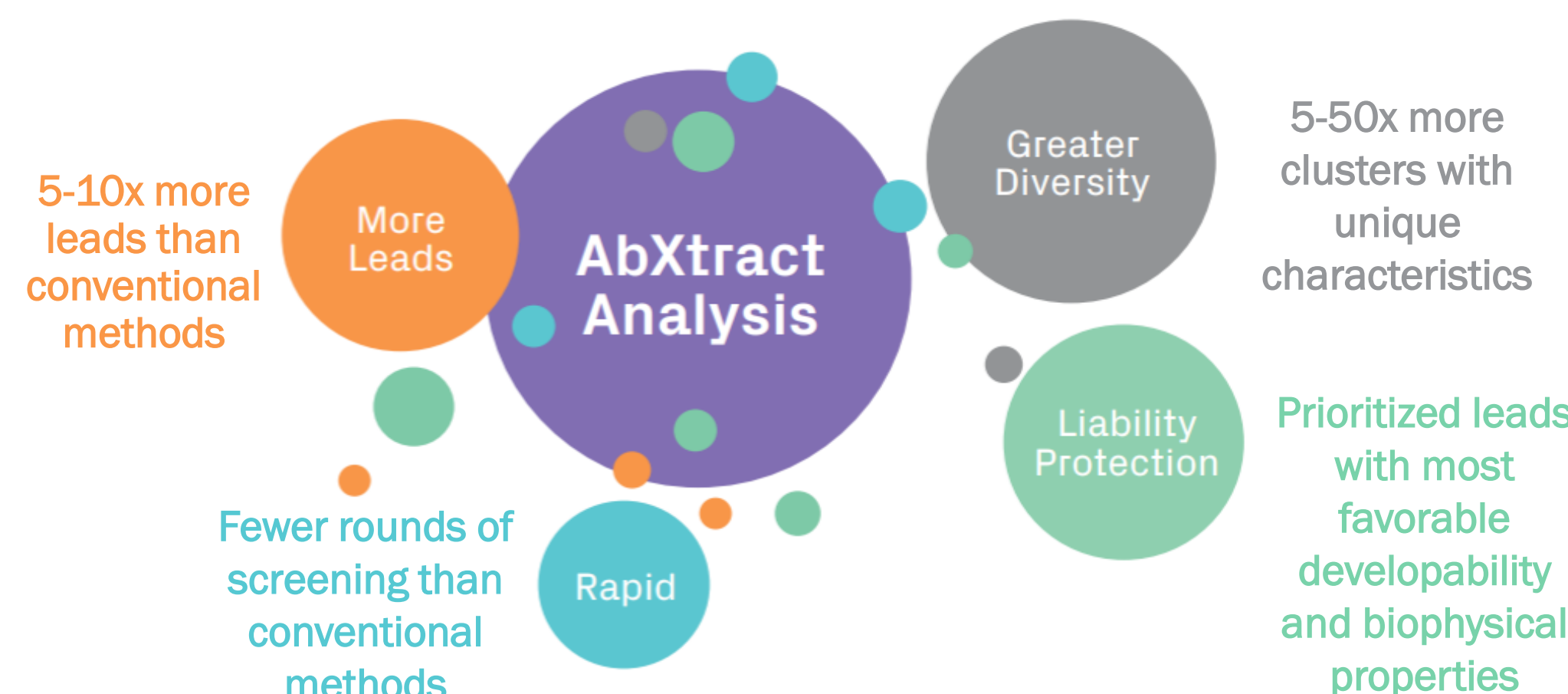
**Figure 1. AbXtract Workflow.** The bioinformatics pipeline utilizes machine learning for prioritization of optimal leads starting with quality filtering of input NGS data followed by annotating the sequences to identify regions of interest (ROI) and extract features. Relative Abundance and enrichment based on ROI can be calculated, the module then identifies abundant and rare clusters with unsupervised machine learning based on density-based clustering and quantifies sequence-based biophysical liabilities. The platform prioritizes the leads based on favorable NGS metrics to provide a gene synthesis-ready output.

## Application: Purification of Bispecific Antibodies



Two-step purification of a CHO-expressed BsAb using Protein A followed by cation exchange (CEX) chromatography. Initial purification shows heterogeneous species with mispaired and truncated byproducts, while the second step removes impurities and yields a highly pure, homogeneous antibody, confirmed by chromatographic profiles and SDS-PAGE.

## Advantages



## More Information



**Key Technical Contacts**  
Dr. Crystal M. Richardson  
[Crystal.Richardson@azenta.com](mailto:Crystal.Richardson@azenta.com)  
GENEWIZ Project Management  
[GS@azenta.com](mailto:GS@azenta.com)

## Optimizing Expression of Bispecific Antibodies

Asymmetric BsAbs are a specialized class of engineered antibodies designed to simultaneously bind two different antigens or epitopes. These antibodies are particularly valuable in therapeutic applications, as they can engage multiple targets with high specificity. For example, they can simultaneously bind a tumor-associated antigen and a T-cell receptor, effectively recruiting immune cells to attack cancer cells. This dual-targeting capability enhances therapeutic efficacy while minimizing off-target effects.

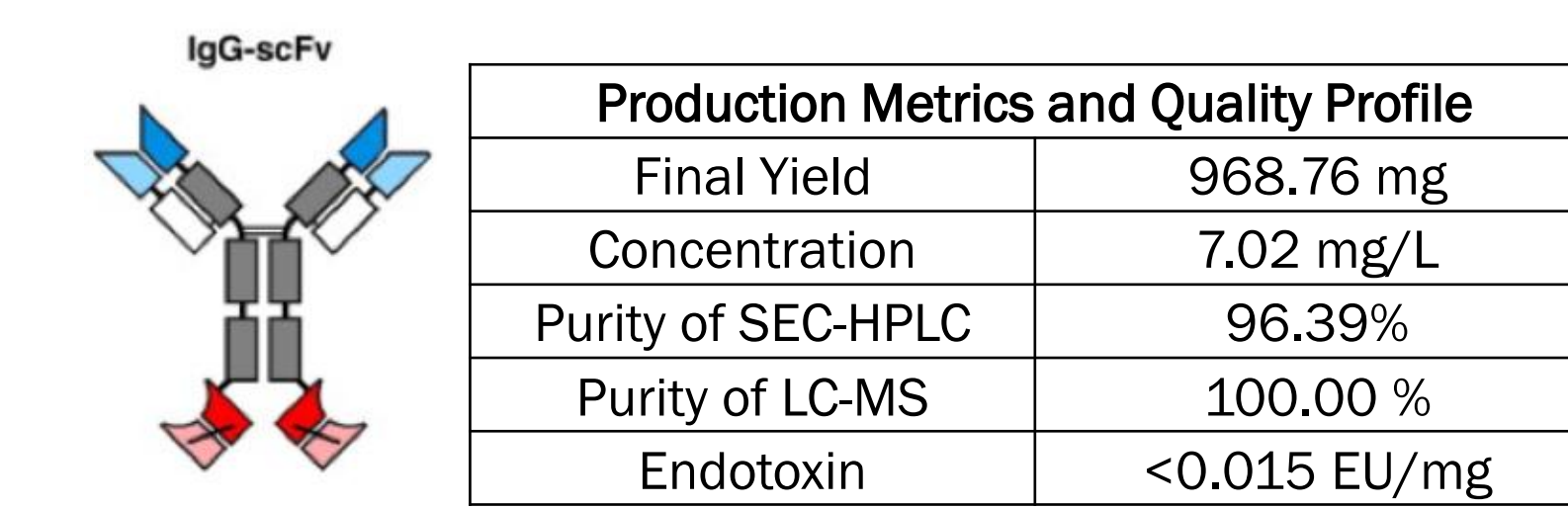


**Table 1. Production Metrics and Quality Profile for IgG-scFv.** Protein yield, concentration and purity evaluation. This approach enabled the production of a highly purified BsAb with strong yield and low endotoxin levels, making it suitable for *in vivo* applications. Structural integrity and purity were further confirmed using orthogonal analytical methods, including LC-MS and SEC.

**Figure 2. SDS-PAGE and SEC-HPLC results for IgG-scFv.**

### Overview

A BsAb required for *in vivo* studies could not be successfully produced using standard approaches.

### Synthesis and Cloning

The DNA sequence was re-engineered and optimized for CHO expression using advanced codon optimization tool to improve expression efficiency and overall construct performance.

### Expression and Purification

Transient expression was performed in CHO-K1 cells at large scale volume, followed by a custom multi-step purification workflow.

Despite their potential, the structural complexity of BsAbs introduces significant challenges in expression, including chain mispairing, formation of undesired byproducts, and lower overall yields compared to monoclonal antibodies. Optimizing expression is therefore critical for successful development. Our team uses strategic approaches including codon optimization, rational sequence engineering, and the use of robust expression systems like CHO cells to improve protein folding, assembly and expression efficiency.

## Production of Asymmetric IgG-like Bispecific Antibodies

Unlike symmetric BsAbs, which feature identical heavy and light chains, asymmetric BsAbs require the assembly of distinct chains, making their production significantly more challenging. Leveraging advanced design approaches, robust expression systems, and customized multi-step purification processes, we consistently generated high-quality BsAbs with exceptional reproducibility. Our tailored purification strategies effectively eliminate impurities such as homodimers and half-antibody species, ensuring superior purity, consistency and performance.

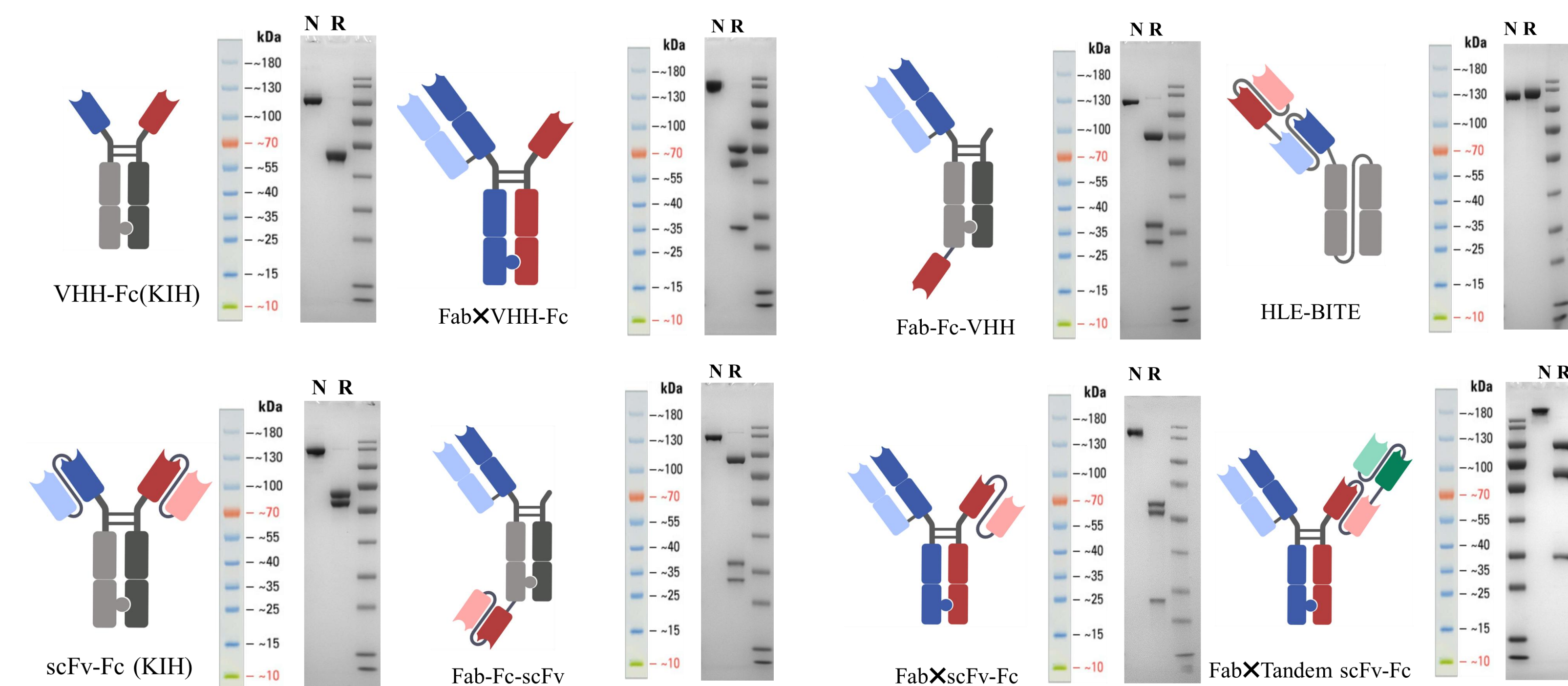


**Figure 3. Optimizing Recombinant Formats.** A variety of asymmetric IgG-like bispecific KIH antibodies are shown, with the corresponding representative SDS-PAGE.

## Summary

BsAbs have become highly valuable in areas such as oncology, immunotherapy and targeted drug delivery. However, their structural complexity introduces challenges in expression, including chain mispairing and formation of unwanted byproducts. Optimized design, expression and purification strategies are therefore essential for achieving consistent, high-quality production suitable for research and clinical development. By integrating machine learning, optimized sequence design, CHO-based expression, and tailored purification strategies, this approach overcomes key bispecific production challenges and enables reliable generation of high-quality candidates for downstream applications. This integrated solution accelerates the identification and production of high-quality therapeutic candidates.

## References

- Erasmus, M.F., Ferrara, F., D'Angelo, S. et al. Insights into next-generation sequencing guided antibody selection strategies. *Sci Rep*, 13: 18370 (2023).
- Ferrara, F., Erasmus, M.F., et al. A pandemic-enabled comparison of antibody discovery platforms: a naïve library directly delivering antibodies as potent as immune sources. *Nature Comm.* 13:462 (2022).
- Song, E., Zhu, P., Lee, SK, et al. Antibody-mediated in vivo delivery of small interfering RNAs via cell-surface receptors. *Nat Biotechnol*, 23: 709 (2005).
- Zhang X, Roth MJ. Antibody-directed lentiviral gene transduction in early immature hematopoietic progenitor cells. *J Gene Med*, 12: 945 (2010).

**Acknowledgements:** Suhani Nagpal<sup>1</sup>, Amy Migliori<sup>1</sup>, M. Frank Erasmus<sup>2</sup>, Laura Spector<sup>2</sup>, Kathryn Perea-Schmittle<sup>2</sup>

<sup>1</sup>OpenEye, Cadence Molecular Sciences, Santa Fe, NM; <sup>2</sup>Specifica, Santa Fe, NM

