

REPORT



Biologics developability data analysis using hierarchical clustering accelerates candidate lead selection, optimization, and preformulation screening

Kevin James Metcalf^a, Galen Wo^b, Jan Paulo Zaragoza^a, Fahimeh Raoufi^a, Jeanne Baker^a, Daoyang Chen^a, Mehabaw Derebe^a, Jason Hogan^a, Amy Hsu^a, Esther Kofman^a, David Leigh^b, Mandy Li^a, Dan Malashock^a, Cate Mann^b, Soha Motlagh^a, Jihea Park^a, Karthik Sathiyamoorthy^a, Madhura Shidhore^a, Yinyan Tang^a, Kevin Teng^a, Katharine Williams^b, Andrew Waight^a, Sultan Yilmaz^a, Fan Zhang^a, Huimin Zhong^a, Laurence Fayadat-Dilman^a, and Marc Bailly^a

^aDiscovery Biologics, Merck & Co. Inc, Rahway, NJ, USA; ^bIT, Merck & Co. Inc, Rahway, NJ, USA

ABSTRACT

Identification of an optimal single protein sequence at the discovery stage for preclinical and clinical development is critical to the rapid development and overall success of a biologic drug. High throughput developability assessments at the discovery stage are used to rank potent molecules by their biophysical properties, deprioritize suboptimal molecules, or trigger additional rounds of protein engineering. Due to the amount of data acquired for these molecules, manual analysis methods to rank molecules are error prone and time-consuming. Here, we present applications of hierarchical clustering analysis for data-driven lead selection of biologics and preformulation screening using high throughput developability data. Hierarchical clustering analysis was applied here for prioritization of three different antibody modalities, including format and chain pairing of bispecific antibodies, sequence-optimized monoclonal antibodies from affinity maturation, preformulation screening of bispecific scFv-Fab fusion molecules, and monoclonal antibodies from an immunization campaign. This high-throughput method for ranking molecules by their developability characteristics and preformulation properties can substantially simplify, streamline, and accelerate biologics discovery and early development.

ARTICLE HISTORY

Received 11 March 2025
Revised 30 April 2025
Accepted 1 May 2025

KEYWORDS



Antibody discovery; antibody screening; biologics; biophysical properties; CMC; data-driven decision making; developability; hierarchical-clustering analysis; lead selection; manufacturability; monoclonal antibodies; preformulation; protein engineering; statistical analysis


Introduction

The goal of biologics preclinical research in the discovery and early development stages is to identify lead sequences that yield a single protein sequence that can progress through the clinical development as rapidly as feasible and will have the greatest chance of success during clinical trials.¹ Properties of an ideal biologic involve tuning the target binding, specificity, potency, pharmacokinetic properties, *in vivo* efficacy, and ease of manufacture. However, the antibody sequence space is vast² and our understanding of the relationship between sequence, biophysical properties, and function is still incomplete, making the identification of an optimal sequence difficult.³ As an example, a typical rodent immunization campaign can identify hundreds of monoclonal antibody (mAb) sequences that bind the target antigen with high affinity, but these antibodies may differ in their therapeutic properties, including affinity, specificity, potency, epitope, biophysical properties, and developability profile. Identifying a single sequence with optimal therapeutic and developability profiles for further development is thus an expensive and time-consuming process.^{4,5} Nevertheless, the clinical success of a biologic requires it to have an optimal potency, pharmacokinetic, efficacy, safety, and developability profile to enable a fast and successful path through clinical development.

Here, we focus on the developability properties and preformulation screening of sets of molecules from several

antibody modalities evaluated during the pre-developability stage.³ Pre-developability is a high-throughput screening workflow that is used to quickly reduce the number of potent candidate molecules, often 10–100 molecules, to a set of 1–5 lead molecules for further in-depth characterization. To measure these properties, experiments are performed in a high-throughput fashion on each purified protein to characterize biophysical properties, including resistance to stress, hydrophobicity, polyreactivity, thermal stability, aggregation, self-interaction, and heterogeneity.⁶ These assessments are routine for biologics selection and optimization at the discovery stage.^{3,4} The candidate molecules can be derived from *in vivo* immunization or *in vitro* campaigns (e.g., yeast or phage display), sequence optimization efforts involving several point mutations made to modify the properties of a parent protein sequence, or protein engineering efforts, including changing domain order and linker identity. At all stages of development, potency drives lead selection. In addition, when potency is equivalent, developability properties are used as additional criteria for lead selection. During the lead identification stage, for complex modalities like bispecific antibodies (BsAbs) the focus can be on the ability to produce high-quality molecules to enable testing potency (Figure 1a), while for less complex modalities the focus at this stage is on potency. During lead optimization and lead selection stages,

CONTACT Marc Bailly  marc.bailly@merck.com  Discovery Biologics, Merck & Co. Inc, 33 Avenue Louis Pasteur, Boston, MA 02115-5727, USA

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/19420862.2025.2502127>

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

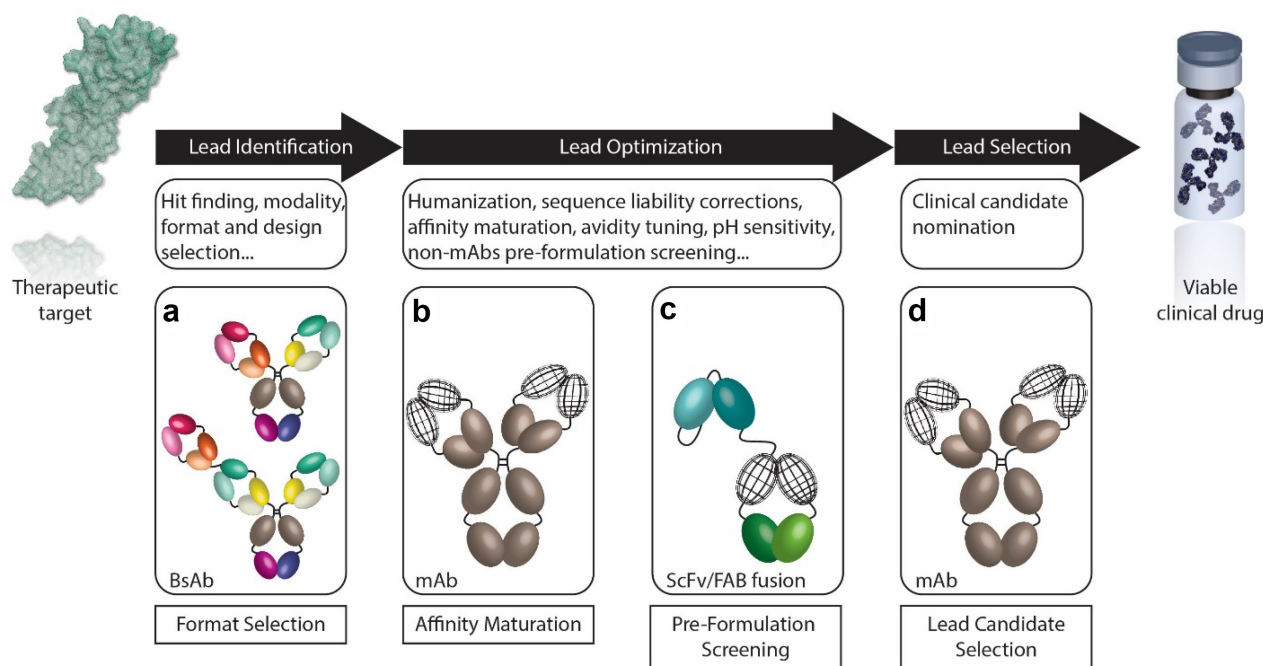


Figure 1. Preclinical biologics drug discovery process during hit-to-lead selection. Herein are presented case studies that span this process. (a) Case study 1: format and chain pairing selection of BsAb. (b) Case study 2: mAb optimization using yeast surface display. (c) Case study 3: preformulation screening of bispecific scFv-Fab fusion molecules. (d) Case study 4: mAb selection from immunization of a humanized mouse.

these molecules are considered to have equal and optimal potency and in vivo activity, such that molecules are differentiated by developability properties during affinity maturation (Figure 1b), preformulation screening (Figure 1c), and lead selection (Figure 1d). In addition to experimental measurements, computational prediction of biophysical properties based on clinical-stage and approved mAb properties can be used to triage sequences.^{7–10} The best sequences are then selected based on the combination of developability assessment data and pharmacology data for further testing toward single candidate selection, early development, and eventually clinical studies.

One challenge to overcome is that typical pre-developability assessments do not always readily identify optimal leads with the best properties across all developability parameters. Manual analysis is the current standard practice to analyze tens of molecules across tens of assay endpoints that yield a high-dimensional space of biophysical properties where manual analysis is time-consuming and prone to error. Establishing a clear set of data used for decision-making can clarify priorities for ranking molecules. These molecules are evaluated based on integrated data from multiple experiments. Structured data capture enables data integration and advanced data science methods, including clustering analysis to simplify and speed-up comparison within the set of molecules.

Hierarchical clustering analysis (HCA) uses high-dimensional data to cluster individual molecules or preformulations based on similar endpoint values and has been previously used to reveal relationships between developability assay endpoints of clinical mAbs⁸ and mAb isotypes.³ This analysis is well suited for high-throughput screening data, as it can be performed on high-dimensional data. The result of HCA is a dendrogram where molecules or preformulations

with more similar endpoint values are more closely connected. The closeness of the connections in the dendrogram is used to create clusters with similar properties. We hypothesized that HCA could be used to identify clusters of leads with optimal properties that could advance past the pre-developability stage. Identification of optimal clusters of molecules or formulations could streamline and standardize biologics pipeline decision-making toward lead selection with the ultimate goal of accelerating drug development.

Here, we describe how HCA of developability data can enable rapid, data-driven decision-making during the various stages of drug development to prioritize molecules and formulations with optimal properties. This work extends the previous use of hierarchical clustering of developability data^{3,8,11} to a new application of pipeline program decision-making for selecting optimal molecules and preformulations. HCA enables reproducible and systematic prioritization of lead molecules and preformulations based on developability endpoints. To illustrate key points, we discuss case studies for: 1) format and lead selection for BsAbs, 2) optimization of mAb sequences during affinity maturation, 3) preformulation screening for bispecific scFv-Fab fusion molecules, and 4) assessment of mAb sequences from an immunization campaign.

Results

During pre-developability assessment, multiple assays are run to assess and compare the biophysical properties, which at this stage are typically binding affinity and specificity, of ~10–100 molecules that have the desired biological properties. The pre-developability assessment of diverse biologic modalities follows a similar workflow and can occur multiple times during

the early discovery stage. To demonstrate that HCA of pre-developability data is generalizable to different modalities and at different stages of early discovery, we describe four case studies of production optimization of BsAb formats, mAb sequence optimization, buffer screening of bispecific scFv-Fab fusion molecules, and pre-developability assessment of mAbs from immunization (Figure 1a–d). Structured assay data were recorded and processed in a database to enable integration and analyses. Accurate data and metadata recording enabled the HCA presented here.

Case study 1: simultaneous BsAb production optimization and lead selection

Development of a BsAb can involve comparison of different formats, linkages, and chain pairings¹² that require high-throughput determination of production parameters. In case study 1, three different BsAb formats were considered: 1 + 1, 2 + 1, and 2 + 2 (Figure 2a). The Fab domains used to create these BsAbs were validated for target binding by Biacore surface plasmon resonance (SPR). This set of 40 BsAbs included eight 1 + 1 constructs, 24 2 + 1 constructs, and eight 2 + 2 constructs. These formats were considered to evaluate the effect of valency on potency. Within a given format, the connectivity of the binding domains was varied. These molecules were assessed for titer and purity/aggregation by size-exclusion chromatography (SEC) and capillary electrophoresis sodium dodecyl sulfate (CE-SDS) (Figures 2b and S1), as well as full complex formation by intact mass analysis by mass spectrometry. The goal at this stage was to identify constructs with acceptable titer and purity to enable in vitro potency assays.

A wide range of titer and purity values were observed. The key assay endpoints for purity were %Main for both SEC and non-reduced (NR) CE-SDS to include both high and low molecular weight impurities, which are expected result from aggregation and incomplete chain pairing in BsAbs. The main peak was confirmed to be the full BsAb complex by intact mass analysis by mass spectrometry (data not shown). These three assay endpoints were clustered, and 10 clusters were identified (Figures 2c–d and S1). Clusters 1–7 sorted into the top branch of the dendrogram and these molecules had purities greater than 50% by SEC and CE-SDS NR. Within these seven clusters, Clusters 1–3 were characterized by higher titer. Both 1 + 1 and 2 + 1 BsAb topologies were found in these three clusters. Clusters 4–7 had lower titer and both 2 + 1 and all 2 + 2 BsAb topologies were found in these four clusters. Clusters 8–10 sorted into the bottom branch of the dendrogram and were characterized by purity lower than 50% by SEC and CE-SDS NR. Only 2 + 1 BsAb topologies were found in these three clusters.

Cluster 1 was identified as having optimal titer and purity by SEC and CE-SDS NR (Figure 2e) and was composed of five BsAbs. Four of the five members of Cluster 1 were 1 + 1 BsAbs, and the remaining member of Cluster 1 was BsAb 2 + 1 16. The remaining four 1 + 1 BsAbs were found in Clusters 2 and 3, which had lower titer and purity by SEC and CE-SDS when compared with Cluster 1. Only four of the 24 2 + 1 BsAbs were found in Clusters 1–3. In Clusters 4–10 were 20 of the 2 + 1 BsAbs and all eight of the 2 + 2 BsAbs. Clusters 4 and 5 had low

titer and intermediate to good purity, and these clusters included all the 2 + 2 BsAbs and four 2 + 1 BsAbs. Cluster 6–10 had very low purity and were composed of 16 2 + 1 BsAbs. Taken together, these data revealed the 1 + 1 BsAb format as having the best titer and purity generally and identified BsAb 2 + 1 16 as having better titer and purity to all other 2 + 1 BsAb molecules. Although no 2 + 2 BsAbs were found in Cluster 1, potency studies to evaluate the effect of valency could be enabled by increasing the scale of production of the six 2 + 2 BsAbs found in Cluster 5, which had high purity and very low titer.

Case study 2: mAb lead optimization from affinity maturation

In case study 2, six parent mAbs were identified from a yeast display library as having desired target binding specificity and good developability properties. However, none of the six mAbs met the desired lead profile due to low-binding affinity in the range of 50–1000 nM. Affinity maturation was performed on these six parent mAbs and yeast surface display was used to identify variant mAbs (child mAbs) with tighter binding affinity. The six parent mAbs gave rise to 20 child mAb variants. To ensure that the affinity-matured child mAbs had good developability properties and specificity to the target, these 26 molecules were assessed for purity/aggregation by (SEC), predicted pI, hydrophobicity (by hydrophobic interaction chromatography (HIC)), thermostability (nano differential scanning fluorimetry (nanoDSF)), non-specific binding (NSB) by ELISA to off-target antigens (baculovirus particles (BVP), DNA, insulin), and polyspecificity by binding to CHO lysate proteins using the polyspecificity reagent (PSR) assay, self-interaction by affinity-capture self-interaction nanoparticle spectroscopy (AC-SINS), size distribution by dynamic light scattering (DLS), binding affinity by Biacore SPR and selectivity (Figure 3a). Selectivity measured the binding of the mAb to the target and to related antigens by Biacore and cellular binding assays. All assays were prioritized, as several of the parent mAbs had multiple suboptimal properties and affinity maturation can negatively impact developability properties.¹³

The mAb sequences were clustered, and eight clusters were identified (Figures 3b–c and S2). Cluster 6, which was composed of mAb Parent 2, mAb Child 2 A, and mAb Child 4 E, had optimal properties. The average value was optimal for all endpoints in this cluster (Figure 3d). The nearest cluster by distance measure to Cluster 6 was Cluster 7, which was differentiated by lower purity by SEC and higher self-interaction in acetate buffer by AC-SINS. Cluster 7 was composed of mAb Child 1 A and mAb Child 1 B. Clusters 1–5 were highly divergent from the optimal Cluster 6 and featured suboptimal values for AC-SINS in acetate and phosphate-buffered saline (PBS) buffers, PSR, NSB Average ELISA, and NSB BVP. Further, Clusters 1–5 trended with lower selectivity. From this clustering analysis, the three mAbs in Cluster 6 are recommended as having optimal properties, and the two mAbs in Cluster 7 as having the next best properties.

The parent and child mAbs had high sequence identity, and some parent and children mAbs had similar developability properties. For example, family 2 had both mAb Parent 2

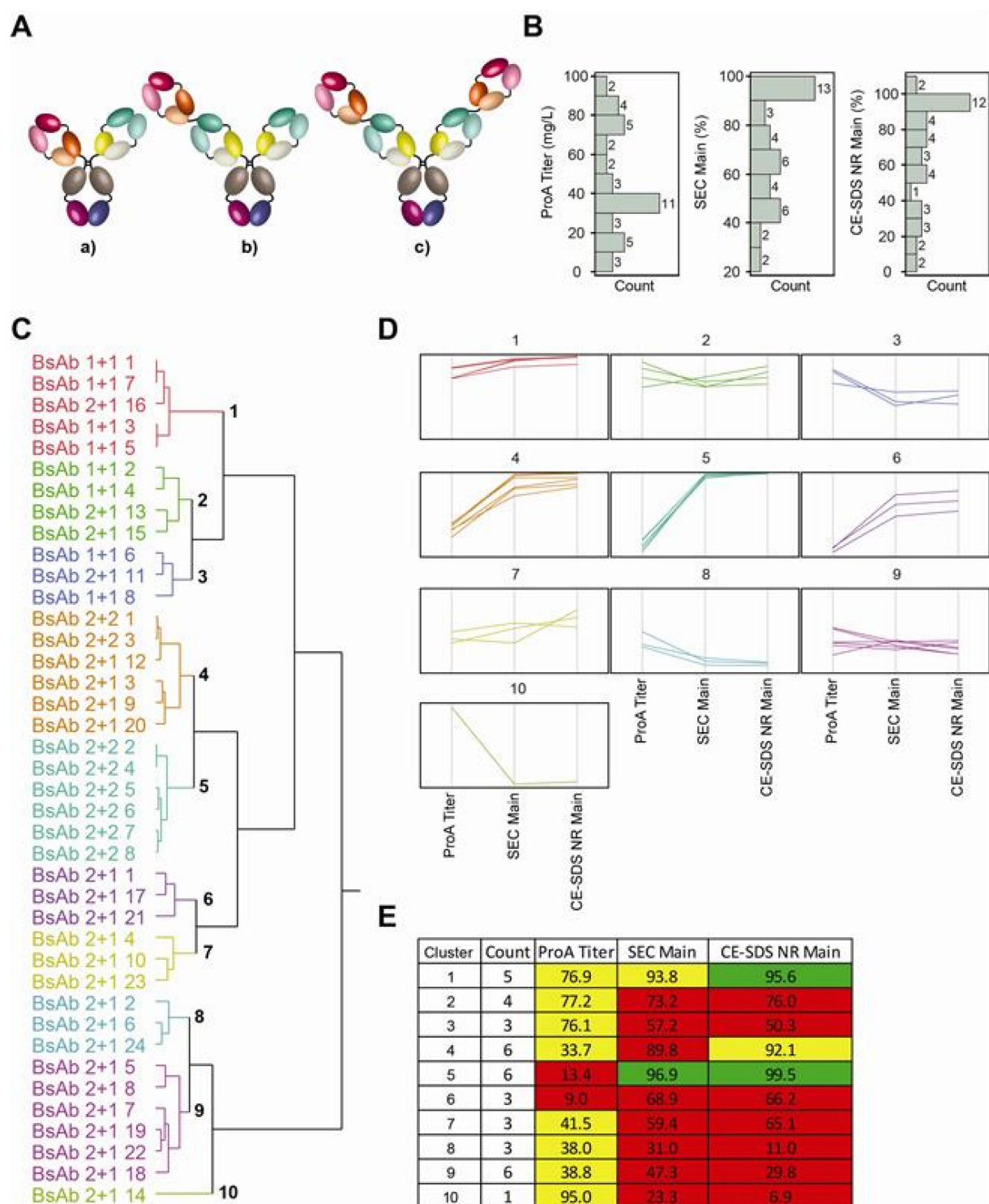


Figure 2. Developability data and HCA for BsAb production optimization and lead selection. A. Schematic of three BsAb topologies studied: a) 1+1, b) 2+1, and c) 2+2. B. Histograms of key assay endpoints. Each bar is labeled with the number of BsAbs in each bin. C. Dendrogram of 40 BsAbs in three different formats clustered using developability assay endpoints. Numbers in black indicate cluster number. D. Parallel coordinate plot. Each panel is a separate cluster, and the color of the lines in the panel matches the color of the clusters in the dendrogram in panel C. For each panel, each line corresponds to a different BsAb. Y-axis is normalized to minimum and maximum value for each assay endpoint. E. Cluster means table. The mean value for each assay endpoint used for clustering is reported for each cluster. Mean values are color coded to indicate different properties: green (optimal) indicates ProA Titer >100 mg/L, SEC Main >95%, and CE-SDS NR Main >95%; yellow (intermediate) indicates ProA Titer >25 mg/L, SEC Main >90%, and CE-SDS NR Main >90%; and red (suboptimal) indicates ProA Titer <25 mg/L, SEC Main <90%, and CE-SDS NR Main <90%.

and mAb Child 2 A – the only child mAb from family 2 – in the optimal Cluster 6. There were, however, notable exceptions. In mAb family 4, mAb Child 4 E was found in the optimal Cluster 6, while the other members were found in Clusters 2, 4, and 8, which had suboptimal properties, showing the diversity of properties that were achieved through affinity maturation. Clusters 2, 4, and 8 had suboptimal properties in HIC retention time (RT) (clusters 4 and 8), PSR (clusters 2, 4, and 8),

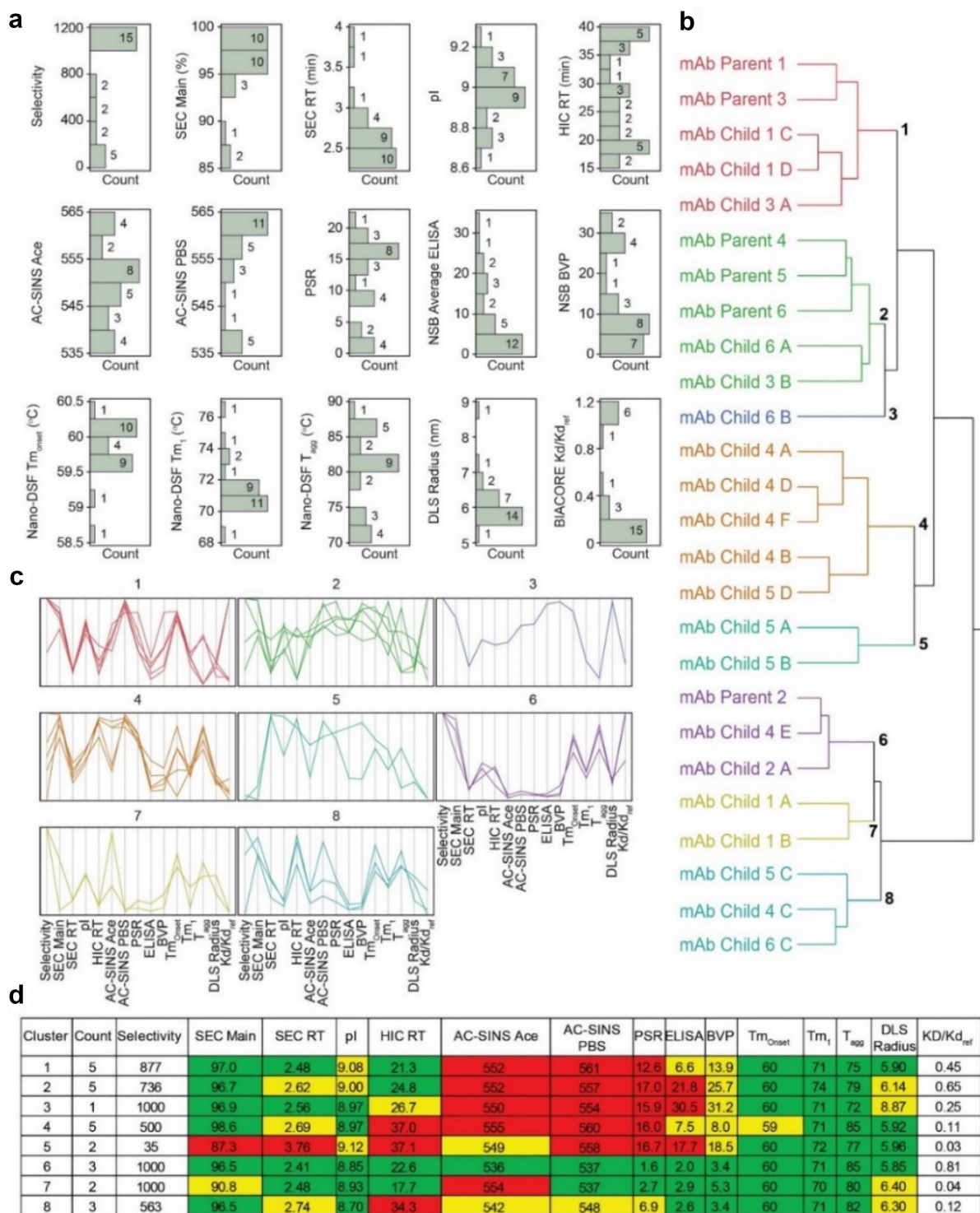


Figure 3. Developability data and HCA for mAb lead selection from sequence optimization. (a) Histograms of key assay endpoints. Each bar is labeled with the number of mAbs in each bin. (b) Dendrogram of 26 mAbs clustered using developability assay endpoints. Numbers in black indicate cluster number. (c) Parallel coordinate plot. Each panel is a separate cluster, and the color of the lines in the panel matches the color of the clusters in the dendrogram in panel B. For each panel, each line corresponds to a different mAb. Y-axis is normalized to minimum and maximum value for each assay endpoint. D. Cluster means table. The mean value for each assay endpoint used for clustering is reported for each cluster. Mean values are color coded to indicate different properties: green (optimal) indicates SEC Main >95%, SEC RT <2.5 min, pI <9.0, HIC RT <25 min, AC-SINS (Ace and PBS) <540 nm, PSR <4, NSB ELISA <3, NSB BVP <6, T_monset >60°C, T_m1 >70°C, T_{agg} >60°C, and DLS radius <6 nm; yellow (intermediate) indicates SEC Main >90%, SEC RT <3 min, pI <9.5, HIC RT <35 min, AC-SINS (Ace and PBS) <550 nm, PSR <8, NSB ELISA <15, NSB BVP <35, T_monset >50°C, T_m1 >60°C, T_{agg} >50°C, and DLS radius <10 nm; and red (suboptimal) indicates SEC Main <90%, SEC RT >3 min, pI >9.5, HIC RT >35 min, AC-SINS (Ace and PBS) >550 nm, PSR >8, NSB ELISA >15, NSB BVP >35, T_monset <50°C, T_m1 <60°C, T_{agg} <50°C, and DLS radius >10 nm.

AC-SINS (clusters 2, 4, and 8), NSB Average ELISA (clusters 2 and 4), and NSB BVP (clusters 2 and 4). These differences between mAb Child 4 E and other members of Family 4 were remarkable due to the high sequence identity within the family.

Case study 3: bispecific scFv-fab fusion preformulation screening for optimal buffer and early screening of leads

In case study 3, a preformulation screening assessment was performed on bispecific scFv-Fab fusion molecules (Figure 4a) to identify the optimal buffer formulation for developability assessment. These bispecific scFv-Fab fusion molecules were validated for target binding and potency before reaching the pre-developability assessment. Prior to this assessment, modalities such as bispecific scFv-Fab antibody fusions, antibody-drug conjugates, and other protein fusions undergo a preformulation screening assessment to optimize the formulation that is then used for the pre-developability assessment. These modalities have biophysical properties that can be sensitive to formulation, and it is crucial that they be screened in their optimal buffer. This set of molecules included nine bispecific scFv-Fab fusion molecules that represented two formats that differed in their connectivity: BsAb01 and BsAb02 were Format 1, and BsAb03-BsAb09 were Format 2. The two formats differed based on their linkers between binding domains. These nine BsAbs were assayed in four formulation buffers (histidine, citrate, glutamate, and acetate), giving 36 unique BsAb-buffer combinations. The molecules were assessed for purity/aggregation (SEC and CE-SDS), thermostability (nanoDSF), and temperature stress stability (40°C for 10 days with SEC and CE-SDS readouts) (Figure 4b and S3). Based on the program goal of optimizing stress stability to aggregation and thermostability, SEC and nanoDSF assay endpoints were prioritized as key assays, and CE-SDS was deprioritized.

The buffer screening assay endpoint data for the 36 molecule-buffer combinations were clustered and 12 clusters were identified (Figures 4c–d and S3). Clusters 1–6 are divergent from Clusters 7–12, with Clusters 1–6 having lower Tm1 and higher main peak monomer after 40°C 10 days stress by SEC. Interestingly, Clusters 1–6 were composed of BsAb01 and BsAb02, which shared the same Format 1, while Clusters 7–12 were composed of BsAb03-BsAb09, which shared the same Format 2. Thus, the clustering analysis of the buffer screening was also able to discriminate different BsAb formats based on preformulation screening data.

According to the program goal, Cluster 1 had optimal properties for BsAb01 (Format 1) formulation, Cluster 2 had optimal properties for BsAb02 (Format 1) formulation, and Cluster 10 had optimal properties for BsAb03-BsAb09 (Format 2) formulation, which was composed of 9 BsAb-buffer combinations (Figure 4e). The nearest clusters by distance measure to Cluster 10 were Clusters 11 and 12, which were composed of 1 BsAb-buffer and 5 BsAb-Buffer combinations, respectively. Compared to Cluster 10, Cluster 11 had lower Tm-onset and Tagg compared to the combinations in Buffer, and Cluster 12 had lower main peak monomer after 40°C for 10 days stress by SEC. For BsAb03-BsAb09 (Format 2), this analysis recommends formulations that are

found in Cluster 10, and all seven molecules formulated in 20 mM glutamate, pH 5.0 are found within this cluster. For BsAb01 and BsAb02 (Format 1), this analysis recommends formulations that are found in Clusters 1 and 2, which is 20 mM citrate, pH 6.0. The HCA analysis identified preformulations that were indeed superior during subsequent targeted stress studies, where BsAb03-BsAb09 (Format 2) had greater stability during 1 month temperature stress in 20 mM glutamate, pH 5.0 over the formulations in 10 mM histidine, pH 6.0 (data not shown), that was more pronounced than in the 10-day study reported here.

Case study 4: mAb lead selection from immunization

Case study 4 describes a pre-developability assessment of mAbs that came from a humanized mouse immunization campaign (Figure 5d). The mAbs were selected for the pre-developability study based on confirmed binding affinity and specificity to the target antigen and were predicted to have no primary post-translational modification liabilities within complementarity-determining region (CDR) residues. This set of 39 molecules included 37 mAbs that were confirmed binders (mAb01-mAb36) and two mAbs of clinical-stage molecules that were used as comparators (mAb38 and mAb39). The set of 37 confirmed binders had high sequence diversity, including 18 clades based on CDRH3 and both lambda and kappa isotypes. These molecules were assessed for purity/aggregation (SEC and CE-SDS), hydrophobicity (HIC), thermostability (nanoDSF), and non-specificity (NSB and PSR). Purity and aggregation were deprioritized because for traditional mAbs these properties can often be optimized during typical later-stage development. The distribution of each assay endpoint was assessed (Figure 5a and S4), and the hydrophobicity (HIC RT), thermostability (Tmonset, Tm1, Tagg), and non-specificity (NSB Average ELISA, NSB BVP, and PSR) were prioritized as key assays.

Next, the developability assay endpoint data for the 39 molecules were clustered, and 11 clusters were identified (Figures 5b and S4). The number of clusters was set manually by minimizing the within-cluster variance using the parallel coordinate plots while also maintaining a small number of clusters (Figure 5c). The cluster means table revealed that Cluster 10 had optimal properties based on the program criteria, which was composed of 7 mAbs (Figure 5d). Interestingly, both clinical mAb comparators were members of Cluster 10, in addition to the five mAbs identified from immunization. The cluster with the next best properties was Cluster 11, which was composed of four mAbs and was the closest cluster by distance measure to Cluster 10. Here, the most noticeable difference between Clusters 10 and 11 was a higher average value of HIC RT. Clusters 1–9 and 10–11 were highly divergent, with molecules in Clusters 1–9 having suboptimal properties in at least one of the following endpoints: HIC RT, NSB Average ELISA, NSB BVP, or PSR. Towards the goal of lead selection, this analysis recommends the five mAbs in Cluster 10 as having the best properties, and the four mAbs in Cluster 11 as having acceptable properties. The remaining 28 mAbs in Clusters 1–9 can be easily deprioritized due to multiple suboptimal properties.

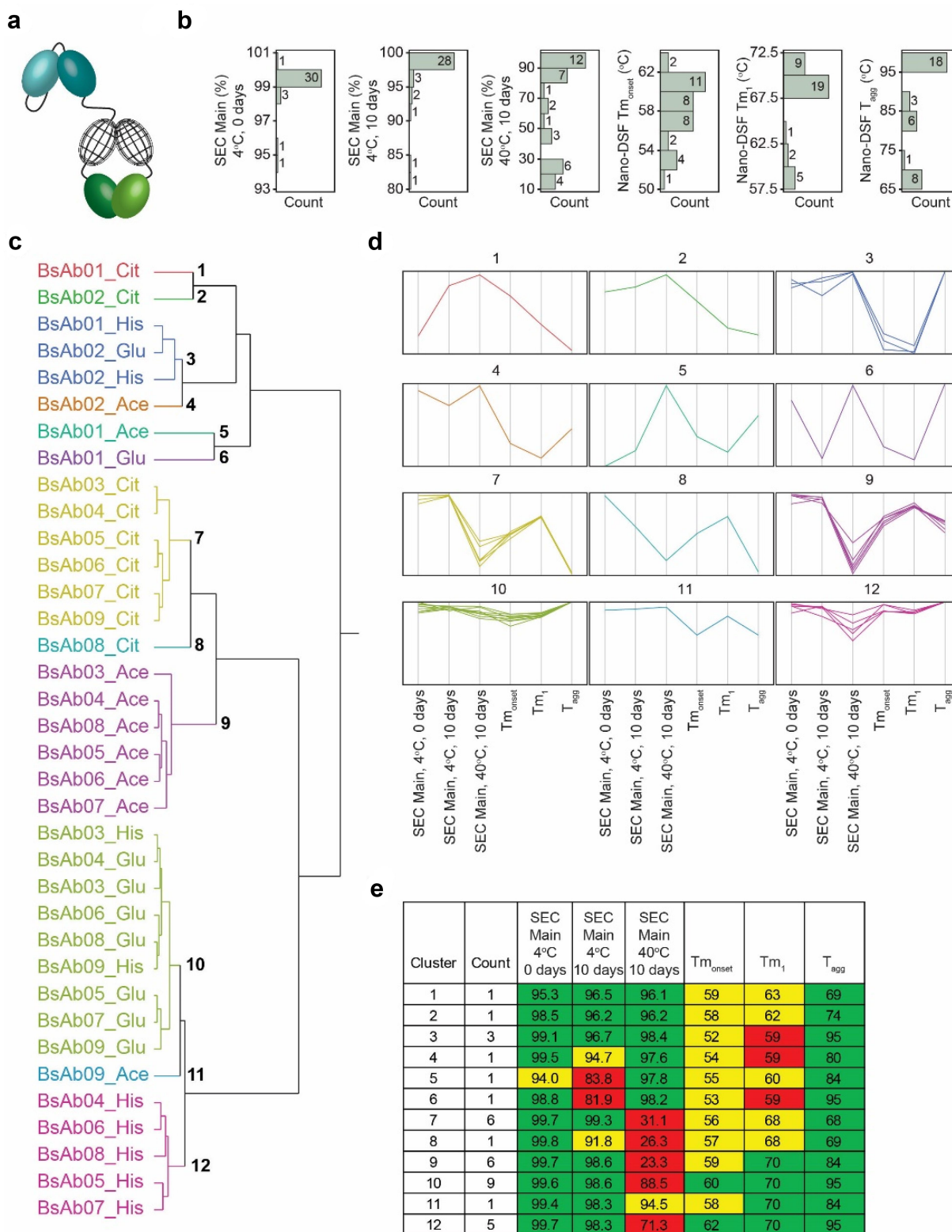


Figure 4. Developability data and HCA for bispecific scFv-Fab fusion preformulation screening and early screening of leads. A. Schematic of scFv-Fab fusion that forms core structure of lead molecules. B. Histograms of key assay endpoints. Each bar is labeled with the number of bispecific scFv-Fab fusion:buffer combinations in each bin. C. Dendrogram of 36 bispecific scFv-Fab fusion:buffer combinations clustered using developability assay endpoints. Numbers in black indicate cluster number. D. Parallel coordinate plot. Each panel is a separate cluster, and the color of the lines in the panel matches the color of the clusters in the dendrogram in panel B. For each panel, each line corresponds to a different bispecific scFv-fab fusion:buffer combination. Y-axis is normalized to minimum and maximum value for each assay endpoint. E. Cluster means table. The mean value for each assay endpoint used for clustering is reported for each cluster. Mean values are color coded to indicate different properties: green (optimal) indicates SEC Main >95%, T_monset >60°C, T_{m1} >70°C, and T_{agg} >60°C; yellow (intermediate) indicates SEC Main >90%, T_monset >50°C, T_{m1} >60°C, and T_{agg} >50°C; and red (suboptimal) indicates SEC Main <90%, T_monset <50°C, T_{m1} <60°C, and T_{agg} <50°C.

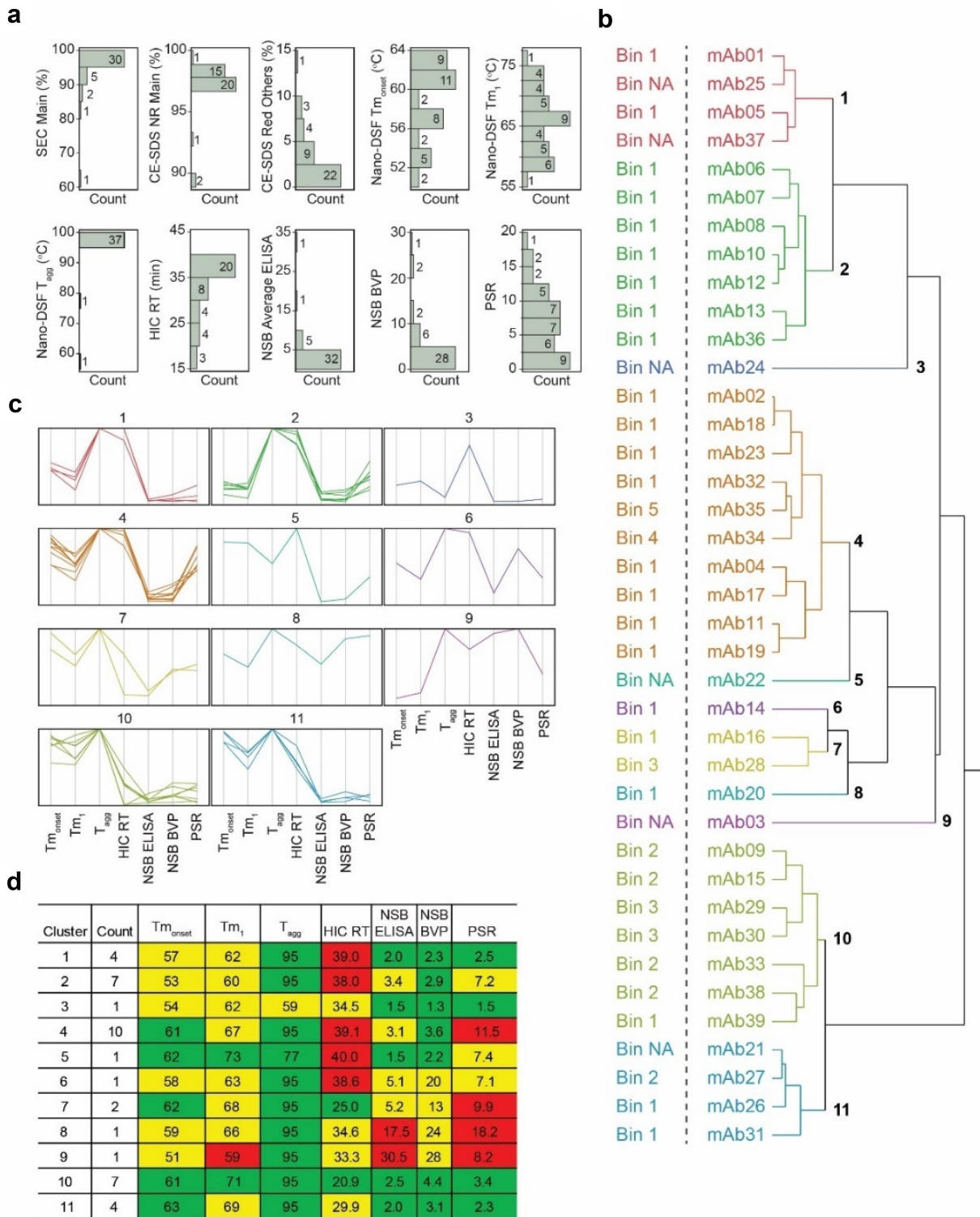


Figure 5. Developability data and HCA for mAb lead selection from mouse immunization. (a) Histograms of key assay endpoints. Each bar is labeled with the number of mAbs in each bin. (b) Dendrogram of 39 mAbs clustered using developability assay endpoints. mAb38 and mAb39 were clinical-stage competitors. Numbers in black indicate cluster number. Epitope binning results for each mAb are indicated. (c) Parallel coordinate plot. Each panel is a separate cluster, and the color of the lines in the panel matches the color of the clusters in the dendrogram in panel B. For each panel, each line corresponds to a different mAb. Y-axis is normalized to minimum and maximum value for each assay endpoint. (d) Cluster means table. The mean value for each assay endpoint used for clustering is reported for each cluster. Mean values are color coded to indicate different properties: green (optimal) indicates T_m_{onset} >60°C, T_m₁ >70°C, T_{agg} >60°C, HIC RT <25 min, NSB ELISA <3, NSB BVP <6, and PSR <4; yellow (intermediate) indicates T_m_{onset} >50°C, T_m₁ >60°C, T_{agg} >50°C, HIC RT <35 min, NSB ELISA <15, NSB BVP <35, and PSR <8; and red (suboptimal) indicates T_m_{onset} <50°C, T_m₁ <60°C, T_{agg} <50°C, HIC RT >35 min, NSB ELISA >15, NSB BVP >35, and PSR >8.

Comparison of the clustering analysis of mAbs based on developability properties with an epitope binning assessment is important to evaluate epitope diversity of lead molecules. Because epitopes can be partially overlapping, mAbs with similar epitopes are classified into communities. Cluster 10, which was identified to have the best developability properties, contained mAbs from 3 of 5 epitope communities, maintaining epitope diversity (Figure 5b and S4). This result shows that HCA identified mAbs with good developability while maintaining epitope diversity.

Discussion

HCA accelerates decision making based on pre-developability data

We applied HCA to prioritize molecules and preformulation selection for 26–40 molecules representing mAbs, BsAbs, and antibody fusions modalities from immunization, yeast display, and in silico design at different stages of early discovery. HCA is simple, straightforward, fast, and easy to implement within the current biologics drug pipeline, and is currently used for analyzing small molecule predicted properties¹⁴ and toxicology profiles.^{15,16} This method allows researchers to establish criteria based on desired developability properties and apply these criteria quantitatively to identify molecules with optimal properties from a set of biologics. A key aspect of HCA is the use of pre-developability assay data for clustering. The pre-developability workflow can be run in high throughput on many molecules, and assays in this workflow are selected based on their ability to de-risk candidate molecules during early-stage biologics drug discovery. All molecules in a pre-developability set are analyzed using identical methods for protein production and analytical assays, such that the assay endpoint values are comparable. Inclusion of data for molecules with dissimilar methods would likely bias the clustering along this axis. In addition, when the validated clinical molecules were available for comparative analysis, our data showed that the two clinical comparators were members of the optimal cluster when co-assayed with 37 mAbs identified from an immunization campaign (Figure 5), validating the use of HCA to identify optimal lead candidates.

Careful selection of which assays to include in HCA is critical to the success of this analysis. During biologics drug discovery, a team must first define the goal of the desired molecule. In addition to the desired potency and tolerability, other properties like pharmacokinetics, stability, and manufacturability must be defined and have priority and quantitative criteria set. This step identifies fit-for-purpose data, such that only assays with decision-making data are used for HCA. Additionally, decision-making assay endpoints also must show sufficient variability in the dataset, as was shown in the histogram analysis at the beginning of each case study. If the assay endpoint data have low variability and are not differentiated, then these data should be excluded from HCA to prevent clustering based on insignificant differences. The four case studies presented here come from different points in early-stage drug discovery (Figure 1), each with distinct goals and different analytical experiments included in the HCA. When

screening BsAb formats and topologies (case study 1), the focus was on titer and purity, and did not require analysis of other properties at this early stage, as the majority of molecules had suboptimal titer or purity. During sequence optimization of mAbs from yeast display (case study 2), we included all assay endpoints to ensure developability was not affected by affinity maturation.¹³ During preformulation screening of bispecific scFv-Fab fusion molecules (case study 3), temperature stress stability to aggregation and thermostability was prioritized as key parameters for these molecules, and clipping by CE-SDS was deprioritized. During lead selection of mAbs from immunization (case study 4), we prioritized hydrophobicity, thermostability, and non-specificity endpoints, and deprioritized purity and aggregation as these properties can be optimized during typical cell line development for traditional mAbs.

HCA has previously been used to cluster mAbs based on their isotype³ and developability assay endpoints based on their relatedness.^{8,11,17} These analyses looked at many mAbs across many target antigens, with the goal of understanding overall characteristics of mAbs with good developability properties or correlations between analytical assays. In contrast, the approach described here applies HCA to support lead candidate selection in biologics pipeline programs. By performing the clustering analysis on related molecules, the optimal region of the multidimensional space of developability properties can be identified to enable lead selection. While applying HCA for lead selection, it is important to remember that molecules within a given format (i.e., therapeutic modality), topology, or clonotype will often have similar developability properties, and will therefore often cluster together, as was seen in the separate clustering of IgG1 and IgG4 mAbs,³ the three different formats in the BsAb production optimization (Figure 2) and the two different bispecific scFv-Fab fusion formats in the buffer screening example (Figure 4). In addition, in another example, the two best clusters by developability properties maintained epitope diversity (Figure 5b and S4). This shows that mAbs with good developability properties could be found in each epitope community, unlike what was found for modality and topology (Figures 2 and 4).

Extending HCA of pre-developability data to accelerate decision-making in the pipeline of the future

Future biologics drug development will be shaped by advances in artificial intelligence and machine learning (AI/ML).^{1,18,19} HCA will be an important tool in supporting pipeline decision-making based on AI/ML models for biologics discovery. For small molecule drug discovery, commercial tools are currently deployed in the industry to use HCA on modeled property data, such as solubility, hydrophobicity (logP), blood–brain barrier penetration, and P-gp transport, to cluster potential lead molecules based on similarity over 10+ predicted properties.¹⁴ Here, we propose that HCA could be used for biologics model data that is rapidly growing to rival small molecule model data. The majority of current approaches focus on building models to predict a single endpoint. For example, there are AI/ML models with good predictive power of primary developability endpoints, including hydrophobicity of mAbs,⁹ polyspecificity of

mAbs⁹ and VHHs,²⁰ thermostability of scFvs,²¹ viscosity of mAbs,²² and deamidation²³ of CDR residues in mAbs. However, biologics drug discovery teams frequently seek to optimize multiple developability endpoints simultaneously. Multi-parameter optimization of model data in small molecule drug discovery is now routine.^{24,25} The predicted values from multiple assay endpoint models for hundreds of sequences could be analyzed by HCA to identify sequences with optimal predicted developability profiles across the high-dimensional space of predicted developability endpoints. In addition, the combination of experimental data with AI/ML predicted values could augment the analysis based on purely experimental data presented in the case studies.

Advances in high-throughput experimentation have also increased the need for analysis tools to interpret large datasets to support lead molecule selection.^{11,26–28} Greater numbers of molecules can be experimentally assessed due to improvements in recovery of clones from immunization and *in vitro* display,²⁹ miniaturization of protein production, and miniaturization and automation in analytical assays.⁵ Furthermore, modalities such as BsAbs and bispecific scFv-Fab fusions can be fused or paired in multiple ways, leading to a greater number of possible sequences that can be subjected to a developability assessment.¹² In the four case studies discussed here, HCA was applied to between 26 and 40 entities, and 3–15 assay endpoints. This HCA method scales to larger sets of molecules³⁰ and can allow for rapid and systematic ranking of the best molecules from a set of hundreds.

While the examples presented here apply HCA to pre-developability assessments, we expect that other evaluations during early-stage drug development could also benefit. For example, screening clones from an antibody discovery campaign for binding affinity, species cross-reactivity, and specificity often involves analysis of hundreds of molecules assessed over a high-dimensional assay space. In this example, incorporation with AI/ML-predicted endpoints could further accelerate decision making and molecule lead selection. Additionally, the impact of HCA for data-driven decision making during early-stage drug development can be increased through democratization of the method. Data availability and interactive tools like dashboards allow for the integration and sharing of data that often comes from cross-functional teams and the application of HCA.

Materials and methods

Hierarchical clustering analysis

Analyses were performed in JMP using Ward's method and each value was normalized using the Z-score. The number of clusters was set manually by minimizing the within-cluster variance using the parallel coordinate plots while also maintaining a small number of clusters. The data used for clustering can be found in Tables S1–S4.

Protein production

Proteins were produced using transient transfect of ExpiCHO cells followed by Protein A purification as previously

described.³ The protein titers were determined by measuring the protein recovered after protein A purification using an A280 measurement to determine concentration, multiplying by the total volume of purified protein, and dividing by the culture volume.

Size-exclusion chromatography

Proteins were analyzed by SEC using a Waters UP-LC system and a BEH200 column as previously described.³

Hydrophobic interaction chromatography

Proteins were analyzed by HIC using a Dionex Pro Pac HIC-10 column as previously described.³

Nano differential scanning fluorimetry

Proteins were analyzed by nanoDSF using a Nanotemper Prometheus NT.48 instrument as previously described.³

Capillary electrophoresis sodium dodecyl sulfate

Proteins were analyzed by CE-SDS using a LabChip GXII (Perkin Elmer) with an HT Protein Express Chip (Perkin Elmer) as previously described.³

Affinity-capture self-interaction nanoparticle spectroscopy

Proteins were analyzed by AC-SINS in 20 mm sodium acetate pH 5.5 and PBS 1× pH 7.4 as previously described.³

Dynamic light scattering

Proteins were analyzed by DLS using a DynaPro Plate Reader II (Wyatt, Santa Barbara, CA) as previously described.³

Biacore surface plasmon resonance

Proteins were analyzed on a Biacore T200 or Biacore 4000 (GE Healthcare) as previously described.³

Epitope binning and community analysis

A biosensor surface containing an array of covalently coupled mAbs was prepared on an HC30M sensor chip (Carterra) via amine-coupling using the LSA's surface array preparation protocol. The running buffer for chip preparation was 25 mM MES pH 5.5, 150 mM NaCl, and 0.05% (v/v) Tween-20. The entire sensor surface was first activated with a mixture of 40 mM EDC and 10 mM sulfo-NHS in 100 mM MES buffer at pH 5.5 for 5 min. mAbs diluted to approximately 10 µg/mL in 10 mM sodium acetate pH 4.5, 0.05% Tween-20 in duplicate were coupled to unique spots on the EDC-sulfo-NHS-modified surface for 10 min, followed by a 5 min injection of 1 M ethanolamine-HCl (pH 8.5) to quench any remaining reactive esters. The resulting array of covalently coupled mAbs on the sensor surface was used for subsequent epitope

binning measurements. Binning in this study was performed in a “classical sandwich format” at 25 °C in 1×HBS-TE +0.5 mg/ml bovine serum albumin (BSA) using the single flow cell mode of the LSA instrument. Each binning experiment cycle involved first a 3 min injection of the antigen at 100 nM followed by a 3 min injection of the secondary antibody at 50 µg/mL. The array surface was regenerated with two 15 s pulses of 10 mM glycine pH2.2 after each binning cycle. The raw high-throughput epitope binning results were processed using the LSA epitope software to generate a heat map as well as other visualizations such as a combined dendrogram and network/community plots.

Nonspecific binding assays

The NSB Average ELISA assay followed the methods of Jain et al.⁸ Briefly, ELISA assays were run with six different antigens: cardiolipin (50 µg/mL, Sigma #C0563), dsDNA (1 µg/mL, Sigma #D4522), ssDNA (1 µg/mL, Sigma #D8899), insulin (5 µg/mL, Sigma #I9278), KLH (5 µg/mL, Sigma #H8283), and LPS (10 µg/mL, InvivoGen #tlrl-ebbps). Each antigen was coated on to 96-well half-well plates (Corning #3690) by incubating 25 µL per well at 4°C overnight. Then, plates were blocked with 0.5% BSA in PBS at RT for 1 hr. Next, 100 nM of each antibody in 0.5% BSA in PBS was incubated at RT for 1 hr. The wells were washed three times with 0.1% Tween-20 in PBS, and anti-Hu H/L HRP Peroxidase AffiniPure Goat Anti-Human IgG (H+L), (Jackson ImmunoResearch #109-035-088) at 1:5000 dilution in 0.1% Tween-20 in PBS was incubated at RT for 1 hr. The wells were washed three times with 0.1% Tween-20 in PBS, and the wells were developed with 1-step TMB (Thermo #34028). After 15 min, 1 N HCl was added to each well and the plates were read at 450 nm. The value reported is fold over zero, where the antibody well absorbance value is divided by the absorbance value of the well without any antibody. The fold over zero value for cardiolipin, dsDNA, ssDNA, insulin, KLH, and LPS was then averaged and reported as NSB Average ELISA.

The NSB baculovirus particle (BVP) assay was like the NSB Average ELISA assay, except BVP (2.8 µg/ml in 50 mM sodium carbonate, pH 9.6, Lake Pharma) was used to coat the plates, and Tween-20 was not included in the buffers. The fold over zero value for BVP was reported as NSB BVP.

Polyspecificity reagent assay

The PSR assay was performed as previously described.⁹ Values are reported as fold over zero, where the fluorescence value is divided by the fluorescence value of the sample without any antibody.

Selectivity assay

Selectivity was calculated from the binding affinity of the mAb to the target that has high allelic diversity in the human population by Biacore and cellular assays that include binding to cells that express the target allele and binding to cells that express the non-target allele. The selectivity data represents results from multiple assays: (1) Biacore binding to target

allele, (2) binding to cell lines that express either the target allele or a non-target allele of the same gene, and (3) binding to human PBMCs with and without endogenous target allele expression. The selectivity score value was set where 1000 was the highest specificity, and 0 was no specificity.

Preformulation screening

Samples were prepared for buffer screening by dialysis against 4 mM sodium acetate pH 5.5, concentration to 5 mg/mL, and jump dilution in one of four buffers to 1 mg/mL final protein concentration. The four buffers are: 20 mM sodium acetate, pH 5.5; 20 mM glutamate, pH 5.0; 10 mM histidine, pH 6.0; and 20 mM citrate, pH 6.0. Samples were then subjected to SEC, CE-SDS, nanoDSF assays. Stress studies were performed on aliquots that were incubated at 40°C for 10 days and 4°C for 10 days. These stress study samples were then subjected to SEC, CE-SDS, and DLS assays.

Acknowledgments

We would like to thank Yao Yu, Nestor Gutierrez, Karin Vroom, Huan Yang, Susmita Borthakur, Hyelin Ha, and Ming-Tang Chen for their contributions advancing these pipeline programs through generating molecules and data, performing analyses, and program leadership. We would also like to thank Sandro Vivona, for fruitful discussions and reviewing our manuscript and Alan Cheng for discussions on modeling and data science in small molecule drug discovery.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported that there is no funding associated with the work featured in this article.

References

1. Carter PJ, Rajpal A. Designing antibodies as therapeutics. *Cell*. 2022;185(15):2789–2805. doi: [10.1016/j.cell.2022.05.029](https://doi.org/10.1016/j.cell.2022.05.029).
2. Olsen TH, Boyles F, Deane CM. Observed antibody space: a diverse database of cleaned, annotated, and translated unpaired and paired antibody sequences. *Protein Sci: A Publ Protein Soc*. 2022;31(1):141–146. doi: [10.1002/pro.4205](https://doi.org/10.1002/pro.4205).
3. Bailly M, Mieczkowski C, Juan V, Metwally E, Tomazela D, Baker J, Uchida M, Kofman E, Raoufi F, Motlagh S, et al. Predicting antibody developability profiles through early stage discovery screening. *Mabs-austin*. 2020;12(1):1743053. doi: [10.1080/19420862.2020.1743053](https://doi.org/10.1080/19420862.2020.1743053).
4. Mieczkowski C, Zhang X, Lee D, Nguyen K, Lv W, Wang Y, Zhang Y, Way J, Gries J-M. Blueprint for antibody biologics developability. *Mabs-austin*. 2023;15(1):2185924. doi: [10.1080/19420862.2023.2185924](https://doi.org/10.1080/19420862.2023.2185924).
5. Makowski EK, Wu L, Gupta P, Tessier PM. Discovery-stage identification of drug-like antibodies using emerging experimental and computational methods. *Mabs-austin*. 2021;13(1):1895540. doi: [10.1080/19420862.2021.1895540](https://doi.org/10.1080/19420862.2021.1895540).
6. Xu Y, Wang D, Mason B, Rossomando T, Li N, Liu D, Cheung JK, Xu W, Raghava S, Katiyar A, et al. Structure, heterogeneity and developability assessment of therapeutic antibodies. *Mabs-austin*. 2019;11(2):239–264. doi: [10.1080/19420862.2018.1553476](https://doi.org/10.1080/19420862.2018.1553476).

7. Raybould MIJ, Marks C, Krawczyk K, Taddese B, Nowak J, Lewis AP, Bujotzek A, Shi J, Deane CM. Five computational developability guidelines for therapeutic antibody profiling. *Proc Natl Acad Sci USA*. 2019;116(10):4025–4030. doi: [10.1073/pnas.1810576116](https://doi.org/10.1073/pnas.1810576116).
8. Jain T, Sun T, Durand S, Hall A, Houston NR, Nett JH, Sharkey B, Bobrowicz B, Caffry I, Yu Y, et al. Biophysical properties of the clinical-stage antibody landscape. *Proc Natl Acad Sci USA*. 2017;114(5):944–949. doi: [10.1073/pnas.1616408114](https://doi.org/10.1073/pnas.1616408114).
9. Waight AB, Prihoda D, Shrestha R, Metcalf K, Bailly M, Ancona M, Widatalla T, Rollins Z, Cheng AC, Bitton DA, et al. A machine learning strategy for the identification of key in silico descriptors and prediction models for IgG monoclonal antibody developability properties. *MAbs*. 2023;15(1):2248671. doi: [10.1080/19420862.2023.2248671](https://doi.org/10.1080/19420862.2023.2248671).
10. Ahmed L, Gupta P, Martin KP, Scheer JM, Nixon AE, Kumar S. Intrinsic physicochemical profile of marketed antibody-based biotherapeutics. *Proc Natl Acad Sci USA*. 2021;118(37):118. doi: [10.1073/pnas.2020577118](https://doi.org/10.1073/pnas.2020577118).
11. Jain T, Boland T, Vásquez M. Identifying developability risks for clinical progression of antibodies using high-throughput in vitro and in silico approaches. *MAbs*. 2023;15(1):2200540. doi: [10.1080/19420862.2023.2200540](https://doi.org/10.1080/19420862.2023.2200540).
12. Labrijn AF, Janmaat ML, Reichert JM, Parren PWHI. Bispecific antibodies: a mechanistic review of the pipeline. *Nat Rev Drug Discov*. 2019;18(8):585–608. doi: [10.1038/s41573-019-0028-1](https://doi.org/10.1038/s41573-019-0028-1).
13. Teixeira AAR, D'Angelo S, Erasmus MF, Leal-Lopes C, Ferrara F, Spector LP, Naranjo L, Molina E, Max T, DeAgüero A, et al. Simultaneous affinity maturation and developability enhancement using natural liability-free CDRs. *MAbs*. 2022;14(1):2115200. doi: [10.1080/19420862.2022.2115200](https://doi.org/10.1080/19420862.2022.2115200).
14. StarDrop reference guide v7.4, 2023, Optibrium Ltd. <https://optibrium.com/products/product-installation/>
15. Yang ZY, He JH, Lu AP, Hou TJ, Cao DS. Application of negative design to design a more desirable virtual screening library. *J Med Chem*. 2020;63:4411–4429. doi: [10.1021/acs.jmedchem.9b01476](https://doi.org/10.1021/acs.jmedchem.9b01476).
16. Marvel SW, To K, Grimm FA, Wright FA, Rusyn I, Reif DM. ToxPi graphical user interface 2.0: dynamic exploration, visualization, and sharing of integrated data models. *BMC Bioinf*. 2018;19(1):80. doi: [10.1186/s12859-018-2089-2](https://doi.org/10.1186/s12859-018-2089-2).
17. Jain T, Prinz B, Marker A, Michel A, Reichel K, Czepczor V, Klieber S, Sun W, Kathuria S, Oezguer Bruederle S, et al. Assessment and incorporation of in vitro correlates to pharmacokinetic outcomes in antibody developability workflows. *MAbs*. 2024;16(1):2384104. doi: [10.1080/19420862.2024.2384104](https://doi.org/10.1080/19420862.2024.2384104).
18. Prihoda D, Maamary J, Waight A, Juan V, Fayadat-Dilman L, Svozil D, Bitton DA. BioPhi: a platform for antibody design, humanization, and humanness evaluation based on natural antibody repertoires and deep learning. *MAbs*. 2022;14(1):2020203. doi: [10.1080/19420862.2021.2020203](https://doi.org/10.1080/19420862.2021.2020203).
19. Notin P, Rollins N, Gal Y, Sander C, Marks D. Machine learning for functional protein design. *Nat Biotechnol*. 2024;42(2):216–228. doi: [10.1038/s41587-024-02127-0](https://doi.org/10.1038/s41587-024-02127-0).
20. Harvey EP, Shin JE, Skiba MA, Nemeth GR, Hurley JD, Wellner A, Shaw AY, Miranda VG, Min JK, Liu CC, et al. An in silico method to assess antibody fragment polyreactivity. *Nat Commun*. 2022;13(1):7554. doi: [10.1038/s41467-022-35276-4](https://doi.org/10.1038/s41467-022-35276-4).
21. Harmalkar A, Rao R, Richard Xie Y, Honer J, Deisting W, Anlahr J, Hoenig A, Czwikla J, Sienz-Widmann E, Rau D, et al. Toward generalizable prediction of antibody thermostability using machine learning on sequence and structure features. *MAbs*. 2023;15(1):2163584. doi: [10.1080/19420862.2022.2163584](https://doi.org/10.1080/19420862.2022.2163584).
22. Lai PK, Gallegos A, Mody N, Sathish HA, Trout BL. Machine learning prediction of antibody aggregation and viscosity for high concentration formulation development of protein therapeutics. *MAbs*. 2022;14(1):2026208. doi: [10.1080/19420862.2022.2026208](https://doi.org/10.1080/19420862.2022.2026208).
23. Delmar JA, Wang J, Choi SW, Martins JA, Mikhail JP. Machine learning enables accurate prediction of asparagine deamidation probability and rate. *Mol Ther Methods Clin Dev*. 2019;15:264–274. doi: [10.1016/j.omtm.2019.09.008](https://doi.org/10.1016/j.omtm.2019.09.008).
24. Koscher BA, Canty RB, McDonald MA, Greenman KP, McGill CJ, Bilodeau CL, Jin W, Wu H, Vermeire FH, Jin B, et al. Autonomous, multiproperty-driven molecular discovery: from predictions to measurements and back. *Science*. 2023;382(6677):eadi1407. doi: [10.1126/science.adi1407](https://doi.org/10.1126/science.adi1407).
25. Segall M. Multi-parameter optimization: identifying high quality compounds with a balance of properties. *Curr Pharm Des*. 2012;18(9):1292–1310. doi: [10.2174/138161212799436430](https://doi.org/10.2174/138161212799436430).
26. Liu Y, Caffry I, Wu J, Geng SB, Jain T, Sun T, Reid F, Cao Y, Estep P, Yu Y, et al. High-throughput screening for developability during early-stage antibody discovery using self-interaction nanoparticle spectroscopy. *MAbs*. 2014;6(2):483–492. doi: [10.4161/mabs.27431](https://doi.org/10.4161/mabs.27431).
27. Phan S, Walmer A, Shaw EW, Chai Q. High-throughput profiling of antibody self-association in multiple formulation conditions by PEG stabilized self-interaction nanoparticle spectroscopy. *Mabs-austin*. 2022;14(1):2094750. doi: [10.1080/19420862.2022.2094750](https://doi.org/10.1080/19420862.2022.2094750).
28. Chai Q, Shih J, Weldon C, Phan S, Jones BE. Development of a high-throughput solubility screening assay for use in antibody discovery. *MAbs*. 2019;11(4):747–756. doi: [10.1080/19420862.2019.1589851](https://doi.org/10.1080/19420862.2019.1589851).
29. Georgiou G, Ippolito GC, Beausang J, Busse CE, Wardemann H, Quake SR. The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat Biotechnol*. 2014;32(2):158–168. doi: [10.1038/nbt.2782](https://doi.org/10.1038/nbt.2782).
30. Ward JH. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc*. 1963;58(301):236–244. doi: [10.1080/01621459.1963.10500845](https://doi.org/10.1080/01621459.1963.10500845).