



Deep Mutational Scanning for Paratope Mapping: A High-Throughput Approach to Antibody Optimization Using Anti-HER2 VHH

Background

Rational antibody optimization depends on understanding how individual amino acids within the paratope contribute to target binding. Traditional approaches to paratope mapping involve sequential site-directed mutagenesis campaigns where each variant is expressed separately, purified, and then characterized through individual binding assays. This workflow presents significant bottlenecks in both time and cost, particularly when comprehensive mapping requires testing dozens or hundreds of variants. The SPOC platform addresses these limitations by enabling parallel characterization of entire mutational libraries on a single SPR biosensor chip, with each variant synthesized directly from DNA through cell-free protein expression and immediately captured for kinetic analysis.

The anti-HER2 VHH (clone 2Rs15d¹) serves as an ideal model system for demonstrating this approach. HER2-targeting therapeutics represent a major class of cancer biologics, and nanobodies offer advantages in tissue penetration and manufacturing compared to full-length antibodies. Understanding which residues within the VHH paratope drive binding affinity provides a foundation for subsequent optimization efforts aimed at improving therapeutic properties while maintaining target specificity.

Experimental Design

The mutational library design focused on comprehensive coverage of all three complementarity-determining regions (CDRs) identified through combined manual assignment using Mitchell and Colwell criteria and confirmation via INDI.²⁻³ Each amino acid position within CDR1, CDR2, and CDR3 was individually substituted with four chemically diverse amino acids: alanine (neutral, non-polar), aspartate (negatively charged), lysine (positively charged), and serine (neutral, polar). This substitution strategy samples different chemical properties at each position while maintaining consistent codon usage across variants to avoid confounding effects from expression differences due to codon bias.

The resulting 92 variants were produced as C-terminal HaloTag fusions to enable covalent capture onto SPR biosensor surfaces. DNA encoding each variant was printed into individual nanowells on silicon substrates, with cell-free protein expression occurring in isolated nanoliter-volume chambers formed when the nanowell

slide is press-sealed against an SPR biosensor functionalized with HaloTag chloro-alkane linker. This approach generates a functional antibody array where each spot contains a unique variant captured at defined positions across the sensor surface (Figure 1). Recombinant HER2 extracellular domain (ECD) was then flowed across the entire array in a concentration series (1.4, 4.1, 12, 37, and 110 nM) with regeneration steps in between, allowing kinetic parameter determination for all variants in a single experimental run.

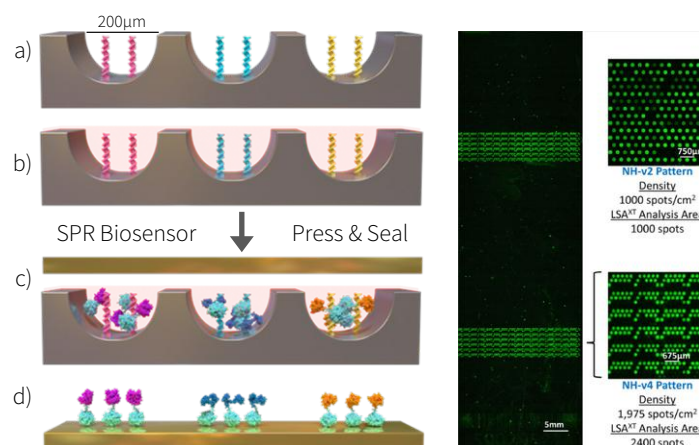


Figure 1: SPOC Platform Workflow and High-Density Nanowell Array Configurations. Left: a) A unique gene is printed in each nanoliter-volume well. b) Cell-free lysate is added to nanowells, for protein expression. c) SPR biosensor coated with capture ligand is pressed and sealed on the nanowells. d) Full-length proteins are captured on the SPR biosensor as pure-protein spots, with 192-1152 proteins on a single biosensor chip. Right: Anti-HaloTag fluorescence staining of expressed proteins on nanowell slides. Right: Full slide (5mm scale) showing two protein subarrays. Comparison of NH-v2 (1,000 spots/cm², 1,000 spots per LSA^{XT} analysis area) and NH-v4 (1,975 spots/cm², 2,400 spots per analysis area) nanowell patterns.

CDR1 Analysis: Positions 26-33

CDR1 showed a mix of permissive and critical positions. For example, position Y28 demonstrated complete substitution tolerance, with all four variants (Y28A, Y28D, Y28K, Y28S) maintaining affinities within the wildtype range of 13.7 ± 0.745 nM, while I29 variant achieved 3.1 ± 0.50 nM affinity, representing approximately 4-fold improvement over wildtype (Figure 3). However, this variant exhibited non-canonical kinetic behavior during dissociation, with a steep initial drop followed by a plateau where dissociation essentially stalls, suggesting the formation of a heterogeneous binding interface where one interaction mode becomes highly stable (Figure 4).

Position S32 proved sensitive to substitution chemistry, with S32K producing the lowest affinity variant in the entire study at 31 ± 2.8 nM, more than 2-fold weaker than wildtype. The lysine substitution at this position likely introduces electrostatic clashes or steric interference with the HER2 surface. Position C33 emerged as absolutely critical for binding - substitution with any of the four amino acids completely abolished detectable binding. Although the reported crystal structure for this nanobody does not predict disulfide formation, the complete loss of function upon C33 mutation suggests either direct paratope involvement or a structural role stabilizing the CDR1 conformation through an uncharacterized disulfide linkage.

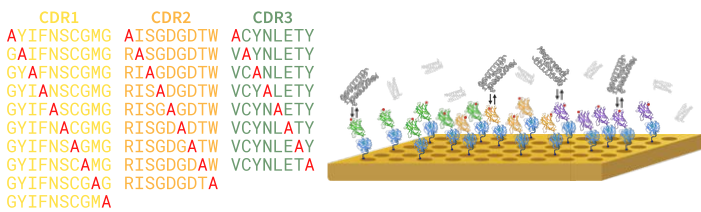


Figure 2: Anti-HER2 VHH Mutagenesis Library Design and SPR Biosensor Array Format. Left: CDR1, CDR2, and CDR3 sequences showing the 92-variant library. Each position within the CDRs (highlighted in color) was individually substituted with alanine (A), aspartate (D), lysine (K), or serine (S) to sample different chemical properties while maintaining consistent codon usage. Right: Schematic representation of the SPOC biosensor chip with HaloTag-captured VHH variants arrayed at discrete spots on the gold sensor surface for parallel kinetic analysis with HER2 ECD.

CDR2 Analysis: Positions 50-58

CDR2 contained both affinity-enhancing and structurally critical positions. Position I52 behaved similarly to C33, with all four substitutions (I52A, I52D, I52K, I52S) eliminating binding below detection limits. The hydrophobic nature of isoleucine at this position appears essential, possibly forming direct contacts with a hydrophobic patch on HER2 or maintaining CDR2 loop conformation. Position G54 showed context-dependent tolerance: G54S produced high-affinity binding with minimal signal loss, though overall binding levels were lower than wildtype, indicating potential effects on expression level or capture efficiency rather than intrinsic binding affinity.

Position T58 demonstrated remarkable permissiveness similar to Y28, with all four substitutions maintaining wildtype-like affinity. This suggests T58 sits outside the core binding interface or participates in contacts that can accommodate diverse side-chain chemistries without penalty. The pattern across CDR2 reveals a region where peripheral positions tolerate substitution while central positions (I52) play critical structural or contact roles.

CDR3 Analysis: Positions 95-102

CDR3, typically the most important region for antibody specificity, showed the most dramatic effects from mutagenesis. Position C99 matched the behavior of C33, with all substitutions abolishing

binding, further supporting the hypothesis of a critical disulfide bridge or essential structural cysteine. Position N101 produced interesting results: N101D maintained binding with higher overall signal levels compared to wildtype, indicating improved expression or capture while retaining function.

Position L102 yielded two of the three highest-affinity variants in the study: L102A at 3.1 ± 0.85 nM and L102S at 3.6 ± 0.64 nM, both representing approximately 4-fold improvements over the 13.7 nM wildtype affinity. Like I29D, these variants displayed non-canonical kinetics with steep dissociation followed by stable plateau phases. The leucine-to-alanine substitution likely removes steric bulk that constrains the binding interface, allowing the antibody to adopt a slightly different orientation that increases contact surface area or optimizes electrostatic complementarity. The success of both alanine and serine substitutions at position 102 suggests the improvement derives from removing hydrophobic bulk rather than introducing specific polar interactions.

Kinetic Parameter Patterns Across Substitution Types

Analysis of kinetic trends across all substitution types revealed clear patterns in how different amino acid properties affect binding. Lysine substitutions proved broadly detrimental to affinity, with lysine variants showing reduced binding at more positions than any other substitution type. The introduction of positive charge at paratope positions that normally contain neutral or hydrophobic residues likely creates electrostatic repulsion with basic patches on HER2 or disrupts the carefully balanced charge distribution required for optimal binding. A few positions tolerated lysine substitution without major affinity loss, with N31K showing modest improvement over wildtype (Figure 3).

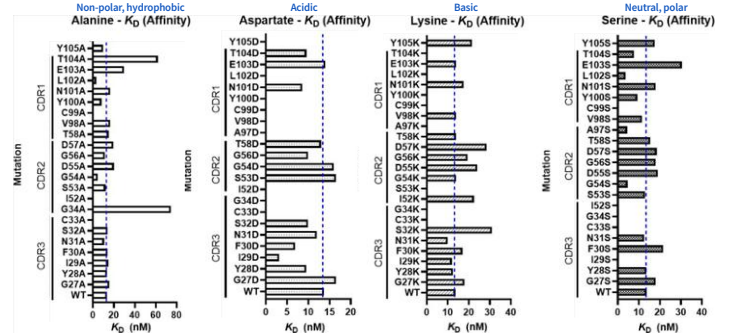


Figure 3: Affinity Profiles of HER2 VHH CDR Variants Across Four Substitution Types. Bar graphs showing KD values for all 92 variants organized by CDR region (CDR1, CDR2, CDR3) and substitution chemistry. Blue dashed line indicates wildtype affinity (13.7 nM). High-affinity variants achieving ~4-fold improvement include I29D (aspartate panel), L102A (alanine panel), and L102S (serine panel). Lysine and aspartate substitutions proved broadly detrimental, with aspartate eliminating or significantly reducing binding at over one-third of positions. Missing bars indicate complete loss of binding below detection limits (C33, C99, and I52 across all substitution types). Variants with bars extending beyond 100 nM showed substantially weakened binding compared to wildtype.

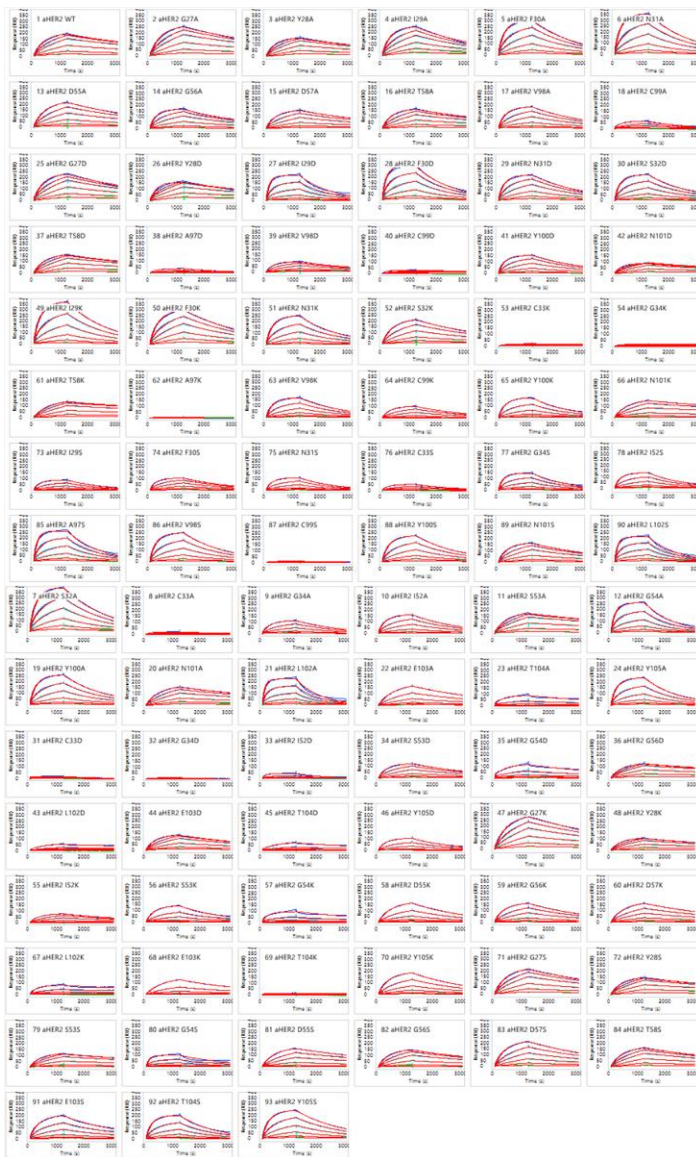


Figure 4: Sensorgrams for All 92 HER2 VHH CDR Variants Binding to HER2 ECD. SPR binding curves showing HER2 ECD titration (1.4, 4.1, 12, 37, 110 nM) across all variants under regenerative conditions. Red lines indicate 1:1 Langmuir binding model fits to the data. Wildtype (panel 1) shows $K_D = 13.7$ nM. Variants with minimal or no binding signal (e.g., C33A, C99A, I52A and their substitution series) indicate critical positions for binding or structural integrity. High-affinity variants (I29D, L102A, L102S) display characteristic non-canonical dissociation kinetics with initial rapid drop followed by stable plateau. Each sensorgram represents a single spot on the SPOC biosensor array, enabling direct comparison of all variants measured under identical experimental conditions.

Aspartate substitutions showed position-dependent effects: while improving affinity at 7 positions, aspartate eliminated or reduced binding at over one-third of positions tested. Introducing negative charge can be beneficial when it complements the electrostatic landscape of the HER2 surface, but is disruptive at other positions due to repulsion from acidic residues or disruption of hydrogen bonding networks. Alanine and serine substitutions were better tolerated, with alanine yielding affinity improvements at positions L102 and I29, and serine providing improvements at L102 and acceptable binding at several positions.

The three high-affinity variants (I29D, L102A, L102S) all exhibited similar kinetic characteristics: modest binding signals during the association phase followed by unusual dissociation behavior where the complex initially dissociates rapidly before reaching a stable baseline with minimal further dissociation. This kinetic signature suggests these variants form two populations of binding interactions - one transient mode that dissociates quickly, and a second extremely stable mode that resists dissociation under regeneration-free conditions (Figure 4). Understanding whether this reflects conformational heterogeneity in the antibody, multiple binding modes on the HER2 surface, or cooperative effects would require additional structural and biophysical characterization.

Methodological Controls and Reproducibility

To validate that measured affinities reflect intrinsic binding properties rather than artifacts from variable expression levels, the platform was tested with DNA concentration titrations for wildtype HER2 VHH. DNA encoding wildtype was printed at concentrations from 11 to 100 ng/ μ L, generating spots with different protein capture densities. Despite significant variation in binding signal magnitude (R_{max}) corresponding to different amounts of captured VHH, the calculated K_D values remained essentially constant across all DNA concentrations: 14, 15, 15, 17, 15, and 18 nM (average 16 ± 1.4 nM)⁴. This demonstrates that the kinetic fitting procedure correctly extracts affinity independently of protein density on the sensor surface, validating comparisons between variants that may express at different levels (Figure 5).

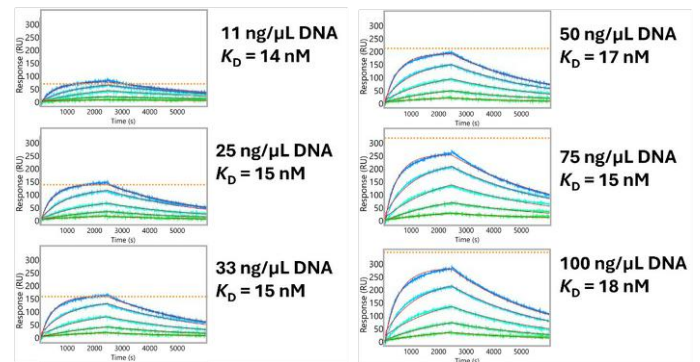


Figure 5: DNA Concentration Independence of Kinetic Measurements. DNA titration experiment showing that varying printed DNA concentrations (11-100 ng/ μ L) produces different protein capture levels but consistent affinity measurements ($K_D = 14$ -18 nM, average 16 ± 1.4 nM). Red lines indicate 1:1 Langmuir binding model fits. Orange dotted lines show maximum VHH capture levels measured by anti-HaloTag antibody binding. These results demonstrate that kinetic parameters are extracted independently of protein density on the sensor surface, validating comparisons between variants with different expression levels.

Applications to Antibody Engineering Workflows

The complete kinetic dataset generated from this single-chip experiment provides multiple entry points for rational optimization. For affinity maturation campaigns, the three variants showing 4-fold improvement (I29D, L102A, L102S)

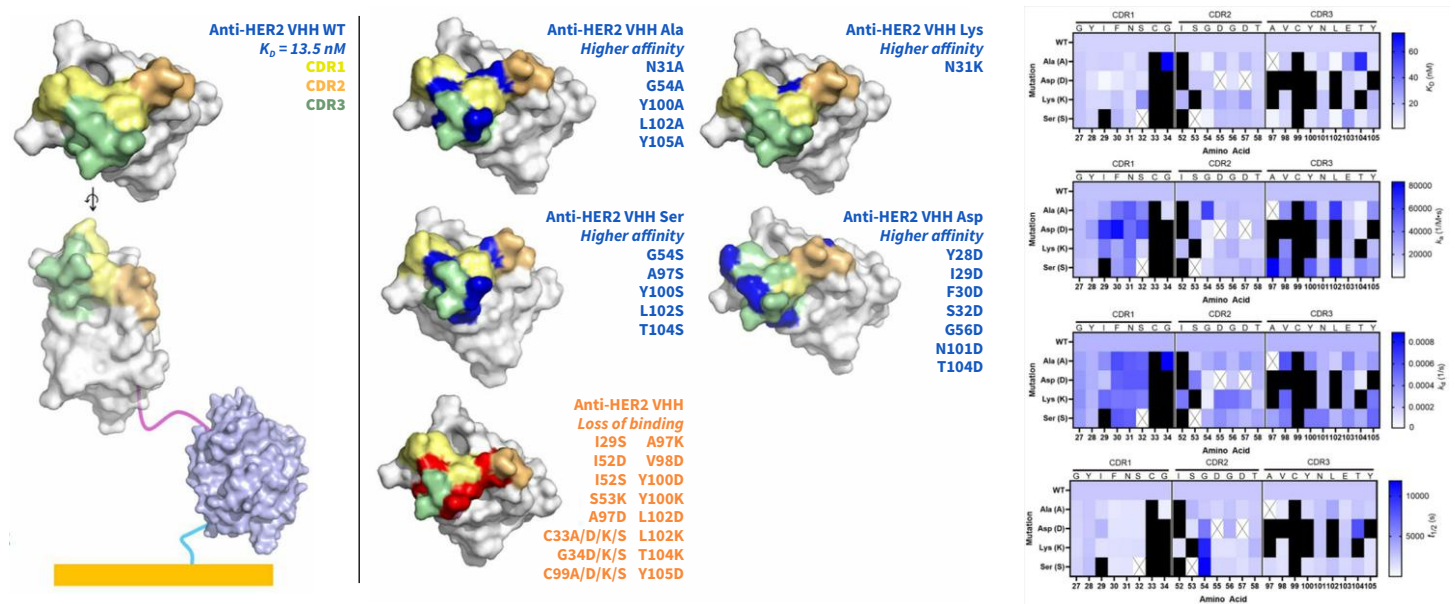


Figure 6: Structural Mapping of CDR Mutations and Kinetic Parameter Heatmaps. Left: Surface representation of anti-HER2 VHH showing CDR locations (CDR1 yellow, CDR2 orange, CDR3 green) with HER2 ECD binding partner shown below. Variants are grouped by substitution type, with higher-affinity variants (blue) and loss-of-binding variants (red) mapped onto the structure. Right: Heatmaps of kinetic parameters (K_D , R_{max} , k_d , $t_{1/2}$) organized by amino acid position and substitution type. Black boxes indicate complete loss of binding, X marks indicate wildtype residues not mutated. Color scales show parameter values with positions C33, C99, and I52 displaying loss of function across all substitution types.

represent validated starting points for combinatorial approaches where multiple beneficial mutations might be stacked. However, the unusual kinetic behavior of these variants requires further investigation before therapeutic advancement, as the heterogeneous binding may indicate stability concerns or unexpected interaction modes that could affect in vivo performance. For epitope mapping applications, the patterns of substitution tolerance and sensitivity create a functional map of the binding interface. Positions where only specific substitution types ablate binding (such as positions sensitive to charge introduction) likely make electrostatic contacts with the antigen, while positions sensitive to all substitutions (C33, C99, I52) either make critical direct contacts or maintain structural integrity of the paratope.

Developability screening represents another application for this data. Loss of binding from lysine and aspartate substitutions at most positions indicates this paratope is intolerant to charge modifications, with implications for chemical conjugation strategies or pH sensitivity. Permissive positions like Y28 and T58 could accommodate engineering modifications without compromising binding.

Beyond single-round characterization, this dataset enables iterative affinity maturation cycles where beneficial mutations can be combined and tested. The comprehensive kinetic profiles also provide high-quality training data for supervised learning models that predict how sequence changes affect binding kinetics, enabling computational approaches to guide subsequent optimization rounds.

Comparison to Traditional Sequential Mutagenesis

Traditional paratope mapping through sequential mutagenesis would require expressing and purifying 92 individual variants, followed by separate SPR experiments for each. With modern cloning and expression workflows, this sequential approach could be completed in approximately 2 months of continuous work.

The SPOC platform compressed this timeline to approximately three weeks from DNA receipt: one week for DNA printing and chip preparation, one week for cell-free expression and SPR data collection, and 2-4 days for SPR analysis of the complete library. The parallel format ensures all variants are compared under identical experimental conditions (same buffer, same HER2 preparation, same temperature, same analysis parameters) eliminating batch-to-batch variation that complicates sequential experiments. The platform scales to 1,152 variants per chip, enabling complete saturation mutagenesis where every CDR position is substituted with all 18 alternative amino acids (excluding self and cysteine), providing comprehensive coverage that would be impractical with sequential methods.

References

- Vaneycken, I. et al. Preclinical screening of anti-HER2 nanobodies for molecular imaging of breast cancer. *FASEB J.*, 25, 2433–2446 (2011)
- Mitchell, L. S. & Colwell, L. J. Comparative analysis of nanobody sequence and structure data. *Proteins*, 86, 697–706 (2018)
- Natural Antibody - INDI - an integrated nanobody database for immunoinformatics. <http://research.naturalantibody.com/nanobodies> (2024).
- Cook, R. et al. An approach to produce thousands of single-chain antibody variants on a SPR biosensor chip for measuring target binding kinetics and for deep characterization of antibody paratopes. *bioRxiv*, 2025.01.11.632576 (2025)