VHH_Affinity_Maturation_Wet-Lab_Workflow

This guide helps you understand how to carry out downstream library construction, display screening, sequencing analysis, and candidate validation after the VHH affinity maturation computational design is completed. Specific experimental conditions may be adjusted according to the parental VHH sequence, antigen format, affinity range, and the available experimental platform.

> **Note:** The wet-lab steps described in this guide—library construction, display screening, NGS sequencing, and candidate validation—are performed by the user; the affinity maturation service delivers the computational design results, the library construction scheme, and the NGS-based selection scheme, together with scheme-level consulting support. The library size, sequencing depth, antigen concentration, and similar figures given below are typical reference ranges and should be adjusted to your own platform and reagents.

Computational Workflow

The affinity maturation computational design proceeds as follows:

1. Based on the input parental VHH sequence and position design parameters, design mutation-position combinations (sublibraries);
2. Predict structure and binding, and score the position combinations;
3. Output the recommended mutant sublibrary design table by maximizing the overall score across all sublibraries (global optimum).

Overview

1.1 Objective

Starting from the parental VHH, select variants with improved affinity (binding) while maintaining specificity, favorable expression properties, and developability.

1.2 Strategy

Build libraries by NNK degeneracy at the model-recommended mutation-position combinations (sublibraries), screen by yeast or phage display with stepwise increasing

pressure, and use NGS to track enrichment throughout and rationally select candidates. Because a VHH is a single-domain antibody (with no light chain), all recommended positions lie on the same VHH sequence; there is no need to separate light- and heavy-chain libraries, and different recommended sublibraries may be constructed and screened separately or pooled in defined ratios into a single combined library. The display platform may be chosen as follows:

Scenario / need	Recommended platform	Notes
Quantitative discrimination of affinity gradients; normalization by display level	Yeast display	Two-color FACS quantitative sorting; reduces high-expression/ low-affinity false positives
Large library capacity; solid/ solution-phase panning or off-rate pressure	Phage display	Capacity up to 10^8–10^9; convenient for off-rate (dissociation-rate) selection

Workflow at a glance: computational design → library design & synthesis → display library construction → pre-screening pilot → display screening (yeast/phage) → NGS tracking & analysis → monoclonal validation → soluble expression & functional validation

1.3 Core Approach

The computational step provides "positions worth exploring by combinatorial mutagenesis," not specific amino acids; in the wet lab, NNK diversification at these positions together with display enrichment and NGS analysis focuses the search on the model-predicted beneficial-mutation space. All recommended sublibraries should be retained (global optimum), and candidate selection should be driven by NGS enrichment trajectories rather than final abundance alone.

Stage 1: Parental Sequence Construction and Parental Pilot

Before handling any mutant library, determine the parental sequence and establish a reliable screening benchmark.

2.1 Construction of the Parental VHH Sequence

A VHH is a single-domain antibody and contains no light chain. The downstream library design, display screening, and candidate validation are all based on the same parental VHH sequence.

First, confirm the amino acid sequence and reading-frame integrity of the parental VHH, including the N-terminus, the CDRs, the framework regions (FRs), and the terminal tag or cloning-adaptor design. If yeast display, phage display, or another display system is to be used, the signal peptide, display fusion protein, detection tag, and cloning sites should also be determined according to the requirements of the chosen vector.

Once the parental VHH amino acid sequence is determined, codon optimization may be performed according to the intended downstream display system. For example, Saccharomyces cerevisiae may be chosen as the optimization host for yeast display, whereas Escherichia coli may be chosen for phage display or E. coli expression systems. The resulting parental VHH DNA sequence serves as the basis for mutant library design, parental display-control construction, and subsequent experimental validation.

2.2 Parental Pilot Experiment Before Screening

Before formal library screening, a small-scale pilot experiment using the parental VHH display strain or parental phage is recommended. Testing the parental VHH across an antigen concentration gradient and comparing it with negative controls such as empty vector, irrelevant VHH, and no-antigen controls can provide an initial estimate of the antigen concentration range suitable for affinity maturation screening.

Because this is an affinity maturation library, the screening goal is not to find binders from a fully random library but to select variants with improved binding over the parent. Accordingly, the first-round concentration and gating strategy may be set with reference to the parental signal and made moderately more stringent than the parental benchmark, to preferentially enrich mutants that retain strong binding at lower antigen concentrations. With this pilot, the number of screening rounds can usually be limited to 2–3.

Decision point: If the parental VHH cannot reproducibly yield a detectable binding signal in the current experimental system, optimize the conditions first (e.g., antigen quality, labeling method, buffer system) before proceeding to library screening; do not screen the mutant library directly.

Stage output: the parental VHH DNA sequence, the parental display control, and the antigen concentration range and gating benchmark to be used as references in subsequent screening.

Stage 2: Affinity Maturation Library Design and Construction

3.1 Library Design and Synthesis

After the VHH affinity maturation computational design is completed, the recommended mutant library design table can be viewed and downloaded from the "Detailed Mutation Results and Data Download" section of the final report. The report typically presents recommendations by sublibrary, with each row corresponding to one recommended mutant sublibrary and including information such as sublibrary ID, recommended mutation positions, number of mutated residues, and model score. For example, the report may recommend a sublibrary containing four mutation positions such as G31, S52, R100, and Y102. Here, "mutation positions" denote a combination of positions that the model considers worth diversifying locally, and the "score" can be used as a reference for prioritizing sublibraries.

Note that the report recommends "which positions are worth exploring by combinatorial mutagenesis" rather than specifying the exact amino acid to introduce at each position. Therefore, during downstream library construction, the recommended amino acid positions may be mapped back to the corresponding codons in the parental VHH DNA sequence, and these codons replaced with NNK degenerate codons. NNK covers all 20 common amino acids while reducing the relative proportion of stop codons, making it suitable for affinity maturation library construction. In other words, a sublibrary with three recommended positions typically corresponds to three NNK codons, and one with four recommended positions to four NNK codons.

Each sublibrary should be designed from the same parental VHH sequence. Apart from replacing the recommended positions with NNK, the rest of the VHH sequence should remain identical to the parent, so that screening results can be unambiguously attributed to changes at the recommended position combinations. After design, each sublibrary should retain an intact reading frame, with no additional insertions, deletions, or premature stop codons. The VHH DNA sequences carrying NNK sites may

be submitted for synthesis as dsDNA fragments, or split into cloning-compatible fragments according to the specific platform.

It is strongly recommended to synthesize all sublibraries recommended in the report. The library design maximizes the model score over the full set of sublibraries as a whole, yielding a globally optimal solution rather than optimizing any single sublibrary in isolation; the sublibraries are complementary in their mutation combinations and jointly cover the beneficial-mutation space predicted by the model. Selecting only some sublibraries would break this global configuration, omit key beneficial combinations, and reduce the success rate of affinity maturation—hence all sublibraries are recommended. If budget or experimental scale is genuinely limited, sublibraries may be ranked by model score, mutation-region location, and structural rationale and constructed in batches, with the remaining sublibraries completed when feasible.

Because a VHH has only one variable chain, all recommended positions lie on the same VHH sequence. There is thus no need to build or screen separate light-chain and heavy-chain libraries. Different recommended sublibraries may be constructed and screened separately, or pooled in defined ratios into a single combined VHH affinity maturation library for screening.

From a sequencing standpoint, the VHH variable region is usually short, so the overall length is well suited to conventional short-read tracking. With paired-end strategies such as PE150 or PE250, a well-designed sequencing window can efficiently cover the VHH body or the main mutation region and thereby track abundance changes of different mutation combinations during screening. If the recommended positions are widely dispersed, the sequencing scheme can be designed according to the actual fragment length, read length, and primer positions to ensure that key positions are effectively resolved.

When pooling different sublibraries, the differences in theoretical diversity arising from different numbers of NNK positions should be considered. A sublibrary with three NNK positions has a theoretical diversity of about 32^3, and one with four NNK positions about 32^4—roughly 32-fold higher. Therefore, if individual variants are to be represented as evenly as possible in the pooled library, the input of a four-NNK sublibrary may be set to about 32 times that of a three-NNK sublibrary. In practice, this ratio can be adjusted according to the concentration and total amount of synthesized product and the scale of downstream library construction.

3.2 Parental Display Control and Library Construction

A parental VHH display control should be constructed in parallel with the mutant libraries. The parental display vector should use the same VHH scaffold, display tag, linkage, and vector backbone as the library, but without mutations. The parental control can be used to assess display level, antigen-binding signal, staining conditions, and an appropriate starting concentration for screening.

The mutant library may be constructed by yeast display, phage display, or another display system according to the platform. For yeast display, common vectors such as pYD1 or pCTcon2 may be selected; for phage display, standard phagemid systems may be used. During library construction, the actual library size should adequately cover the theoretical diversity: because each amino acid under NNK is encoded by 1–3 codons with uneven frequencies, the actual library size is recommended to reach roughly 5–10× the theoretical diversity (e.g., four NNK positions give a theoretical diversity of ~10^6; a yeast-display target of ⩾10^7 transformants is recommended, while phage display can typically reach 10^8–10^9), reducing the risk of losing beneficial variants to transformation bottlenecks. In addition, about 1/32 of each NNK position is a stop codon (TAG): this is read through by the amber-suppressor strains commonly used in phage display, but in yeast display TAG acts as a true stop, so a sublibrary with four NNK positions loses roughly 12% of clones to truncation—this loss should be factored into estimates of library size and effective diversity.

If different recommended sublibraries are constructed separately, it is advisable to number, store, and quality-control them separately from the construction stage onward, so that downstream screening and NGS analysis results remain clearly traceable. If several sublibraries are pooled into a combined library, each sublibrary should be assessed for concentration, integrity, and clone quality before pooling, and the mixing ratios recorded to aid later interpretation.

Stage output: two core experimental inputs—first, the recommended VHH mutant sublibrary design schemes; and second, the synthesizable VHH library sequences obtained by converting the recommended positions to NNK. These sequences can then be used for display-vector cloning, yeast or phage display screening, NGS (next-generation sequencing) enrichment analysis, and functional validation of candidate VHHs.

Stage 3: Display Screening

The mutant library may be screened by yeast display or phage display; in both cases the overall strategy is "fewer rounds with stepwise increasing pressure," usually 2–3 rounds.

4.1 Yeast Display Screening Strategy

If yeast surface display is used, the VHH mutant library should be induced, stained, and sorted as an independent sample. The antigen used for detection and sorting typically needs to be biotinylated and detected with a fluorophore-labeled streptavidin (or directly fluorophore-labeled), so that the display-tag signal and the antigen-binding signal can be read simultaneously by flow cytometry. The overall strategy is "fewer rounds with stepwise increasing pressure": the first round uses the appropriate antigen concentration determined from the parental pilot, and subsequent rounds gradually reduce the antigen concentration (typically by about 2–10-fold per round) while continuously recovering the fraction of cells with the strongest antigen-binding signal.

Both display-tag signal and antigen-binding signal should be analyzed during screening, comparing antigen-binding intensity within cell populations of comparable display level to reduce the risk of mistakenly selecting high-expressing but low-affinity clones. Sorting gates may be set with reference to the signal positions of the parental VHH and the negative controls, preferentially selecting populations above background and at or to the right of the parental signal.

Samples should be retained after each round for subsequent NGS analysis and, if needed, retrospective review. The final enriched population can proceed to monoclonal picking, Sanger sequencing, and functional validation.

4.2 Phage Display Screening Strategy

Besides yeast display, phage display can also be used for VHH affinity maturation screening. It is suitable for solution-phase or solid-phase panning and makes it convenient to increase selection pressure by lowering antigen concentration, increasing wash stringency, or introducing competitive conditions.

A parental VHH phage control can be used to gauge antigen coating concentration, solution-phase antigen concentration, wash stringency, and background binding. For affinity maturation projects, 2–3 rounds of panning are usually sufficient. Later rounds may gradually reduce the antigen concentration or add strategies such as competitive

selection or off-rate (dissociation-rate) selection, to preferentially enrich candidates with slower dissociation and more stable binding.

Samples should be retained from each round for NGS sequencing and subsequent candidate-sequence tracking. The final enriched population can be used for monoclonal picking, phage ELISA, Sanger sequencing, and further validation.

Stage output: the enriched populations after each round and the systematically retained samples (initial library, each-round product, final product), for the Stage 4 NGS analysis and monoclonal validation.

Stage 4: NGS Analysis and Candidate Validation

5.1 NGS Sequencing and Enrichment Analysis

The initial library, the enriched samples after each round, and the final screening product should be systematically retained for NGS analysis. For VHH libraries, sequencing and analysis usually focus on the VHH region containing the recommended positions; if read length allows, the full VHH sequence may be covered to assess both the recommended combinations and any unintended mutations. To keep the abundance trajectories reliable, the sequencing depth of each sample should adequately cover the expected unique combinations (typically ~1–5 M reads per sample, with effective reads exceeding 10× the expected number of combinations); the initial library may be sequenced more deeply to ensure that low-abundance combinations are also covered.

Key aspects of NGS analysis include coverage of mutation combinations in the initial library, abundance after each round, fold enrichment relative to the initial library or the previous round, and whether a continuous enrichment trend is present. Combinations that are covered initially, enrich stably across rounds, reach relatively high final abundance, and remain reasonable from sequence and structural perspectives are generally prioritized.

NGS results help determine which mutation combinations are genuinely enriched under selection pressure and provide a basis for subsequent monoclonal validation, combinatorial mutation design, and prioritization of candidate VHHs.

5.2 NGS-Based Candidate Analysis and Monoclonal Validation

Building on the population-level enrichment analysis in 5.1, this section focuses on selecting and validating candidate monoclonal clones from the enriched population.

Picking clones at random from the enriched library carries substantial randomness: the chance of picking a given clone depends on its abundance distribution in the library, and high abundance does not necessarily equate to high affinity (display/expression advantages and amplification bias can also inflate abundance). Candidate selection is therefore best driven by NGS data rather than by random picking and final abundance alone.

To this end, NGS sequencing is recommended for the initial pre-screening library and for the products of every screening round, so that the abundance trajectory of each mutation combination can be reconstructed across the entire process. Once sequencing is complete, all VHHs enriched in the final round should be assessed against the following criteria:

- **Enrichment relative to the parental sequence:** the abundance gain of the candidate mutation combination over the parental sequence.
- **Fold enrichment relative to the pre-screening library:** final abundance divided by the initial-library abundance.
- **Sustained enrichment across rounds:** continuous, monotonic enrichment is more reliable than a single-round spike.
- **Enrichment under low antigen concentration:** variants that continue to enrich as the antigen concentration is progressively lowered are more likely to have genuinely improved affinity.

After prioritizing candidate mutation combinations on the basis of these criteria, pick the corresponding candidate clones (the number relates to the parental VHH's affinity level and the NGS data-analysis results, such as the degree of enrichment) and perform Sanger sequencing to confirm the mutation combination and exclude frameshifts, stop codons, and other abnormalities; then compare candidate and parental display level and antigen-binding signal by flow cytometry (yeast display) or phage ELISA (phage display), and complete quantitative affinity validation (KD together with the kinetic parameters kon and koff).

After validation, confirmed beneficial mutations may be combined to pursue higher affinity. For combinations, preferentially select advantageous variants from different recommended sublibraries: because they typically act on distinct position regions, their beneficial effects are more likely to be independent and thus more readily additive or synergistic. A second round of combination validation can be designed based on the structural location of the mutations, the enrichment trends, and the monoclonal validation results.

5.3 Soluble VHH Expression and Functional Validation

The final candidate sequences may be converted into soluble VHH, Fc-fusion VHH, bivalent VHH, multispecific VHH, or other target formats for expression and functional validation according to project needs. If the intended application format is not monomeric VHH, it is advisable to validate the expression, stability, specificity, and affinity of candidate mutants in the target molecular format as early as possible.

Functional validation may use ELISA, BLI, SPR, flow-cytometry-based cell-binding assays, or other methods suited to the target antigen format. For affinity maturation projects, in addition to KD, kinetic parameters such as kon and koff should be monitored—particularly whether koff is reduced. The final candidate VHH should simultaneously meet the requirements of improved binding, retained specificity, favorable expression properties, and feasibility for downstream development.

Stage output: quantitatively validated candidate VHHs (combination mutants where appropriate) with improved affinity over the parent and good specificity and expression properties.

6. Delivery and Follow-up Support

Through the affinity maturation service, you obtain a computationally designed VHH mutant library scheme and recommended mutation combinations. Combined with downstream display screening, NGS enrichment analysis, and monoclonal validation, this forms a complete loop from "computational design" to "experimental screening" to "candidate VHH validation".

Depending on the scope of collaboration, the final deliverables may include the recommended mutation positions, detailed information on the different position combinations (recommended sublibraries and their position combinations), the library construction scheme, and the NGS-based selection scheme. Downstream validation may be carried out by yeast display, phage display, or another screening system according to the platform, and preferred VHHs may further be converted into soluble VHH, Fc-fusion VHH, or other target application formats for development assessment.

7. FAQ and Notes

7.1 No or weak enrichment after screening

- Antigen failure: re-validate antigen activity and labeling (biotinylation efficiency, conformation).
- Too much first-round pressure: using very low antigen or harsh washing already in R1 may lose rare variants; increase pressure progressively.
- Low display/infection efficiency: check induction and display level, helper-phage titer and ratio.
- Insufficient sequencing depth: increase reads per sample (see 5.1).

7.2 Single or highly biased enriched sequences

- "Biological-advantage" amplification may have occurred (clones that grow or display fast but do not necessarily bind better); increase the antigen pressure across rounds or switch the screening strategy.
- Judge by NGS enrichment trajectories (sustained enrichment across rounds and continued enrichment at low antigen concentration) rather than final abundance alone (see 5.2).

7.3 Insufficient library size / loss of beneficial variants

- Improve transformation efficiency and ensure roughly 5–10× oversampling (see 3.2).
- For yeast display, note that the NNK TAG codon is a true stop, so a sublibrary with four NNK positions loses about 12% of clones to truncation—account for this when estimating effective library size.

7.4 High-expression, low-affinity false positives

- Compare antigen-binding intensity within cell populations of comparable display level, using two-color gating for normalization (see 4.1).

7.5 NGS cannot accurately resolve mutation combinations

- Confirm that read length and primers cover all recommended positions; for dispersed positions, design a linked-read scheme according to the actual fragment length, read length, and primer positions (see 3.1).