This AI-powered pipeline enables the computational design and validation of proteins containing specific domains of interest. The workflow begins with ESM3 generating novel protein sequences using a template structure while optionally incorporating functional constraints through targeted prompts. AlphaFold3 then validates the structural integrity of these sequences, confirming proper folding of the target domains and providing confidence metrics. Finally, DeepGO predicts biological functions to verify whether the designed proteins maintain desired functional properties. This integrated approach combines sequence generation, structural evaluation, and functional prediction into an efficient computational pipeline for domain-specific protein engineering, with applications in therapeutic development and synthetic biology.

For example，this is what we do with Ef-hand domain

### Experiment 1: ESM3-Based EF-hand Protein Sequence Generation

(Input specific functional tags to guide sequence generation) (1) Materials

Software Tools:

ESM3 pretrained models (open-source + 98B parameter versions) Source: <https://www.evolutionaryscale.ai/>

Database: Protein Data Bank (PDB)

(2) Methods

On the ESM3 interface, configure:

Temperature: 0.5

Model: esm-3-medium-2024-08

Output: Sequence

Structure Template: 1CLL (Human calmodulin, 4 EF-hand domains)

Functional Tags: Optional input of desired functional keywords (e.g., "calcium-binding," "thermostable")

Click "Generate."

### Experiment 2: AlphaFold3-Based Protein Structure Validation

(Validate generated sequences for structural integrity) (1) Materials

Software Tool: AlphaFold3 Source: <https://alphafoldserver.com/welcome>

(2) Methods

Input the sequence from Experiment 1.

Click "Generate."

Validation Metrics:

Confirm EF-hand domain folding (helix-loop-helix motifs).

Record per-residue confidence scores (pLDDT or predicted TM-score).

### Experiment 3: DeepGO-Based Functional Validation

(Verify if predicted functions match design goals) (1) Materials

Software Tool: DeepGO web server Source: <https://deepgo.cbrc.kaust.edu.sa/deepgo/>

(2) Methods

Input the generated sequence.

Output Analysis:

Record predicted Gene Ontology (GO) terms (e.g., "GO:0005509: calcium ion binding").

Compare with target functions (e.g., validate calcium-binding capability).

The generated sequence exhibits only 78.4% similarity to the natural sequence, yet retains functional EF-hand characteristics.