MESA Designer Complete Usage Guide

Distribution

Introduction

The MESA Designer transforms biological targets into production-ready genetic constructs through an intelligent design pipeline. This guide covers every feature and option available in the application, including validation checks and troubleshooting tips.

For installation instructions, please refer to the MESA Designer Installation Instructions. After starting the webapp locally or via Docker, access it at http://localhost:8501/.

Input-Output Pipeline Overview

Understanding the MESA Framework

MESA (Modular Extracellular Sensor Architecture) receptors are engineered cellular sensors that detect specific molecules and trigger intracellular responses. The design process starts with identifying your target molecule and ends with downloadable genetic constructs ready for synthesis.

Supported Input Types

Antibody Database Integration:

- Direct integration with SAbDab provides access to over 10,000 experimentally validated antibody structures
- Automatic retrieval of structural data with resolution and organism information
- Integrated affinity data from both SAbDab and SKEMPI v2.0
- Real-time 3D structure visualization for informed selection

Custom Binder Import:

- Supports *de-novo* designed sequences from BindCraft
- Compatible with experimentally determined binders from selection campaigns
- Natural receptor sequences can be directly imported
- Maximum sequence length: 5,000 amino acids per chain
- Accepts standard amino acid codes (ACDEFGHIKLMNPQRSTVWY) plus stop codons (*)

Component Libraries:

- Validated transmembrane domains from Edelstein et al., 2020
- Curated linker patterns with customizable repeat numbers
- Multiple TEV protease variants with different activities
- Auto-inhibitory peptides for background reduction
- Detection tags (FLAG, HA, Myc, custom)

Output Formats

The tool generates multiple file types packaged in a convenient ZIP archive:

GenBank Files:

- Industry-standard format compatible with Geneious Prime, SnapGene, and Benchling
- Each functional domain is properly annotated with feature tags
- Color-coded regions provide easy visualization
- Includes both DNA (if optimized) and protein sequences

Sequence Optimization Options:

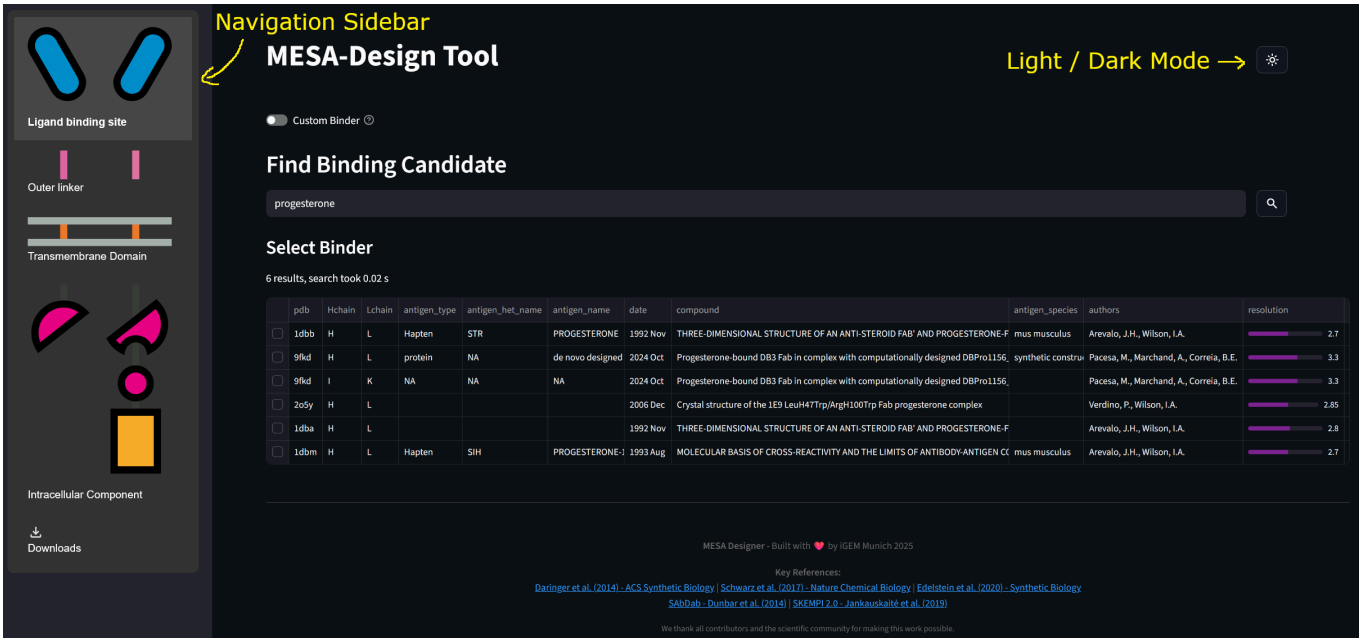
- Codon optimization for various target organisms (H. sapiens, M. musculus, etc.)
- Restriction site avoidance (RFC10, RFC12, RFC1000, BioBrick standards, custom). This is especially interesting for iGEM Teams as an iGEM Compatible site avoidance collection is also available
- Automatic detection of non-standard amino acids that prevent optimization

Additional Outputs:

- FRET validation constructs with mCerulean/mVenus fusions for dimerization confirmation evaluation experiments
- Original PDB files when using SAbDab search provided by RCSB PDB
- Comprehensive design summary with all parameters
- SAbDab search results in CSV format

Step-by-Step Usage Instructions

Initial Setup and Navigation



When you first open MESA Designer, you'll see the main interface with a navigation sidebar on the left. The sidebar dynamically updates to show your current position in the design workflow:

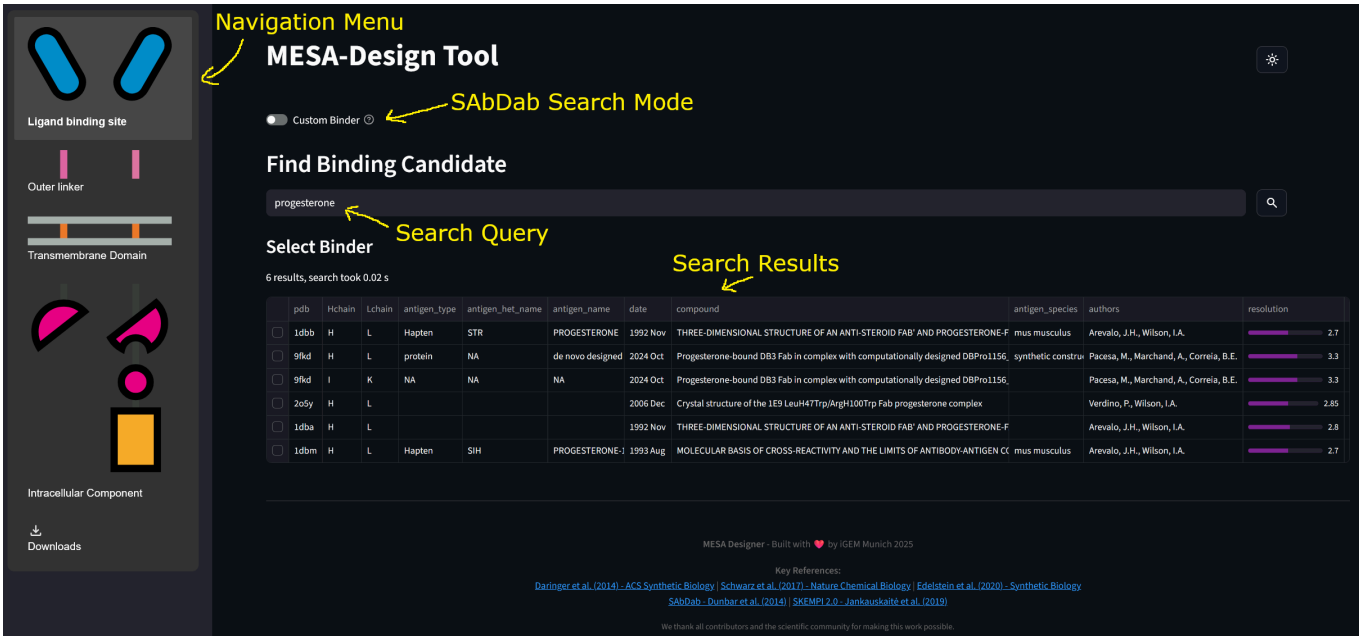
- **Ligand binding site** - Target selection and binder configuration
- **Outer linker** (or just "Linker" for non-transmembrane designs) - Flexible connector design
- **Transmembrane Domain** - TMD selection (only appears when enabled)
- **Intracellular Component** - Protease and cargo configuration
- **Downloads** - Final construct generation and export

The application supports both light and dark themes. Toggle between them using the sun/moon button in the top right corner.

Step 1: Target Selection

You have two pathways for defining your binding domain:

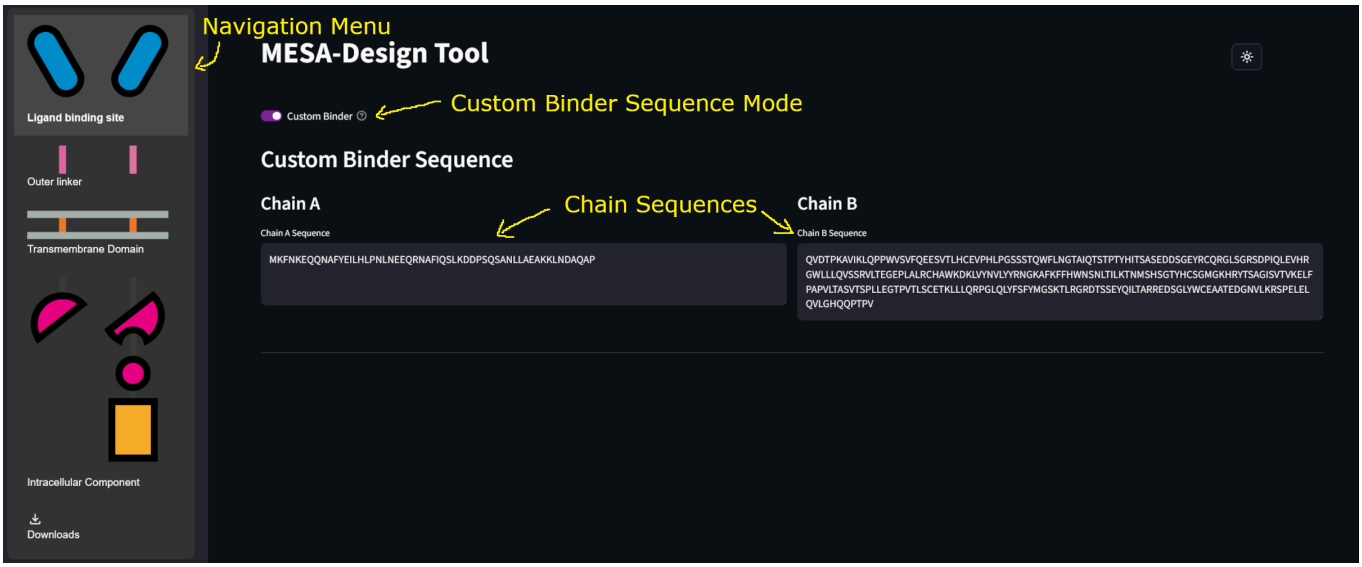
Option A: Database Search



The search interface connects directly uses SAbDab for validated antibody structures and associated affinity data:

1. **Enter your target antigen:** Type common names like "progesterone", "CD19", or "PD-L1"
    - Search queries are limited to 100 characters
    - Only alphanumeric characters, spaces, hyphens, underscores, and periods allowed
    - The system validates your input and displays warnings for invalid characters
2. **Review search results:** The results table shows:
    - PDB ID and structure date
    - Resolution (displayed as a progress bar, lower is better)
    - Source organism
    - Antigen information
    - Affinity data when available (Kd values from SKEMPI database)
    - Results are sortable and filterable
3. **Select your preferred binder:** Consider these factors:
    - Resolution quality (aim for < 3.0 Å)
    - Species compatibility with your expression system
    - Available affinity data
    - Publication date (newer structures may have better refinement)

Option B: Custom Binder Entry



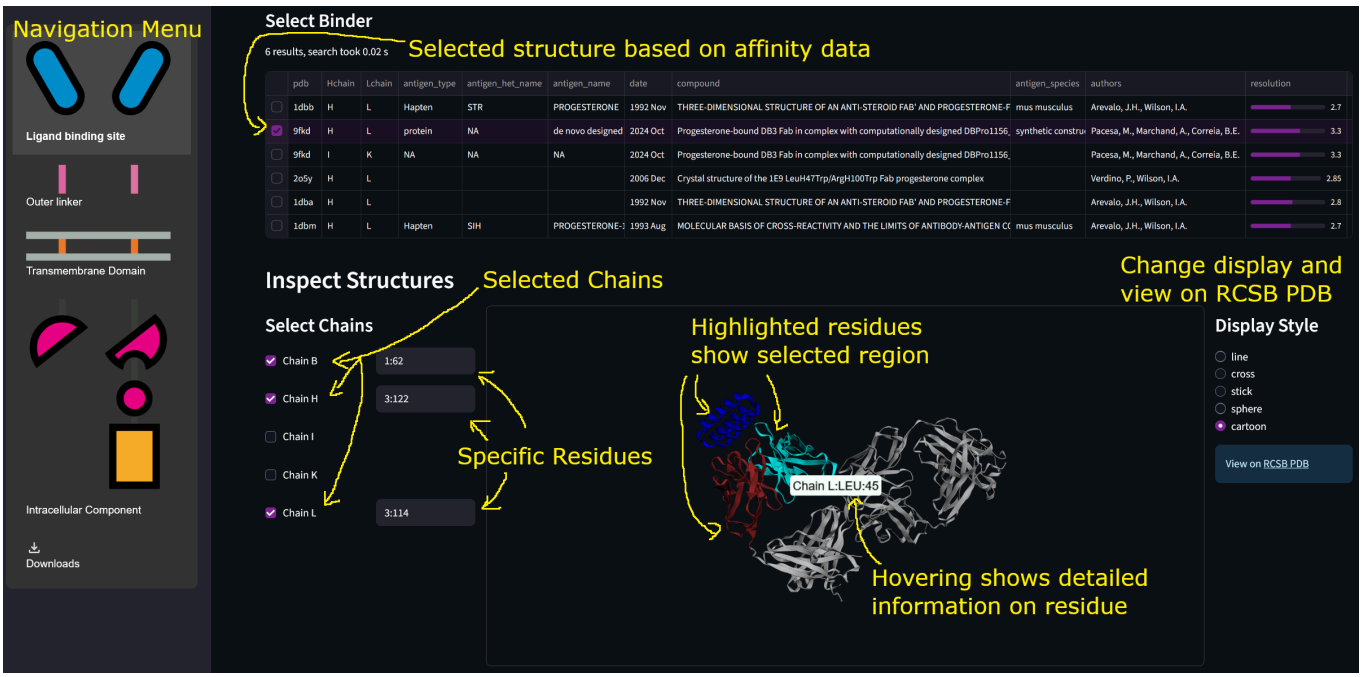
Toggle "Custom Binder" to manually enter sequences:

1. **Paste custom sequences:** Enter Chain A and/or Chain B sequences
   - Maximum 5,000 amino acids per chain
   - System automatically converts to uppercase and removes whitespace
   - Non-standard amino acids trigger warnings but don't prevent usage
2. **Validation checks performed:**
   - Sequence length verification
   - Non-standard amino acid detection
   - Internal stop codon warnings
   - Empty sequence handling (single chain designs allowed)

Common validation warnings you might encounter:

- "Contains non-standard amino acid codes" - May affect sequence optimization
- "Contains internal stop codon(s)" - Will terminate translation prematurely
- "Sequence exceeds maximum length" - **Automatically truncated to 5,000 AA**

Step 2.1: Structural Inspection and Chain Selection



Once you've selected a structure from the database, the 3D viewer activates:

3D Structure Viewer Controls

- **Display styles:** Choose from line, cross, stick, sphere, or cartoon (default)
- **Mouse controls:**
  - Left click + drag: Rotate structure
  - Right click + vertical drag: Zoom in/out
  - Scroll wheel: Zoom in/out
  - Scroll wheel click + drag: Move structure
  - Hover over residues: Display chain:residue:number information
- **Color coding:** Each chain gets a unique color for easy identification upon selection via the chain selection menu
- **Background:** Automatically adapts to your selected theme

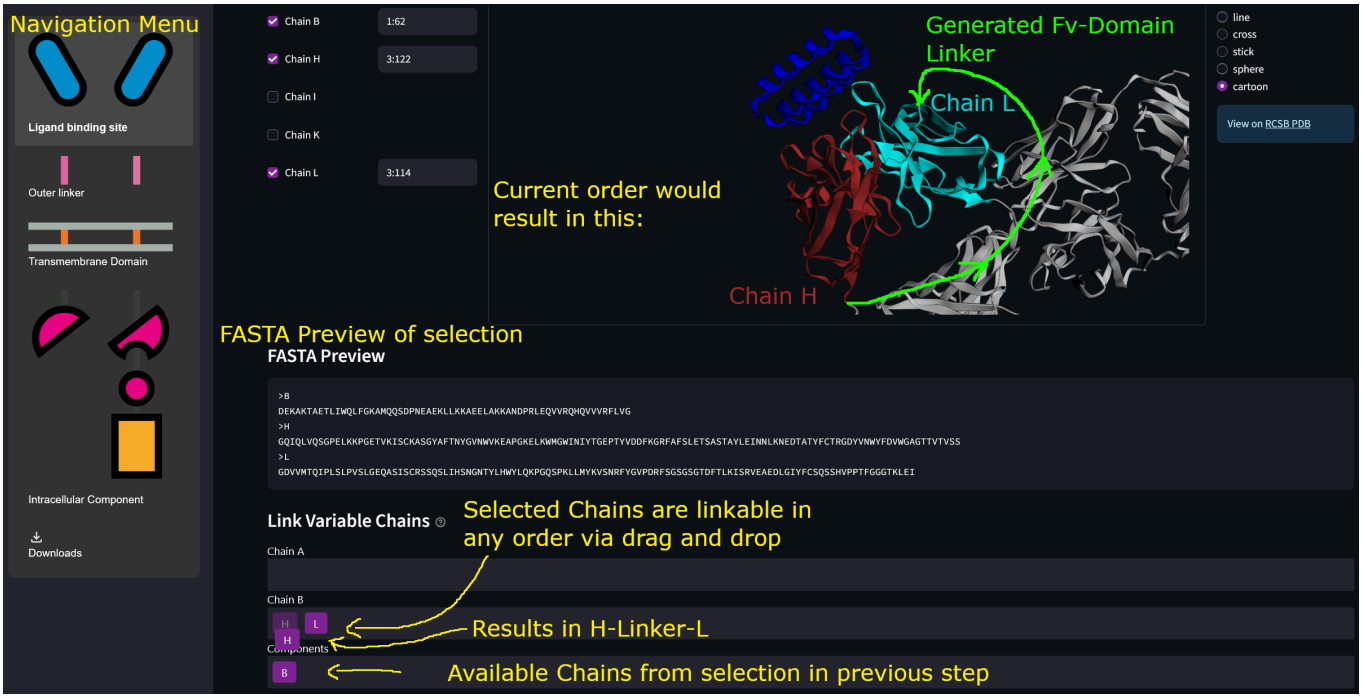
Chain and Residue Selection

The left panel shows all available chains in your selected structure:

1. **Select entire chains:** Check the box next to any chain to include it entirely
   - Selected chains are highlighted in the 3D viewer
   - Multiple chains can be selected simultaneously
2. **Specify residue ranges:** After selecting a chain, customize the range:
   - Format: `start:end` (e.g., `1:100`)
   - Uses 1-based indexing matching PDB numbering
   - Updates highlight in real-time on the structure
3. **Verification:** The FASTA preview below shows your exact selection
   - Each selected chain/range appears as a separate entry
   - Sequences are extracted directly from the PDB file
   - This preview helps confirm you've selected the intended regions

Tip: For antibody structures, you typically want to select the variable domains (VH and VL chains). These are usually the first ~120 residues of heavy and light chains.

Step 2.2: Assemble Desired Sequence



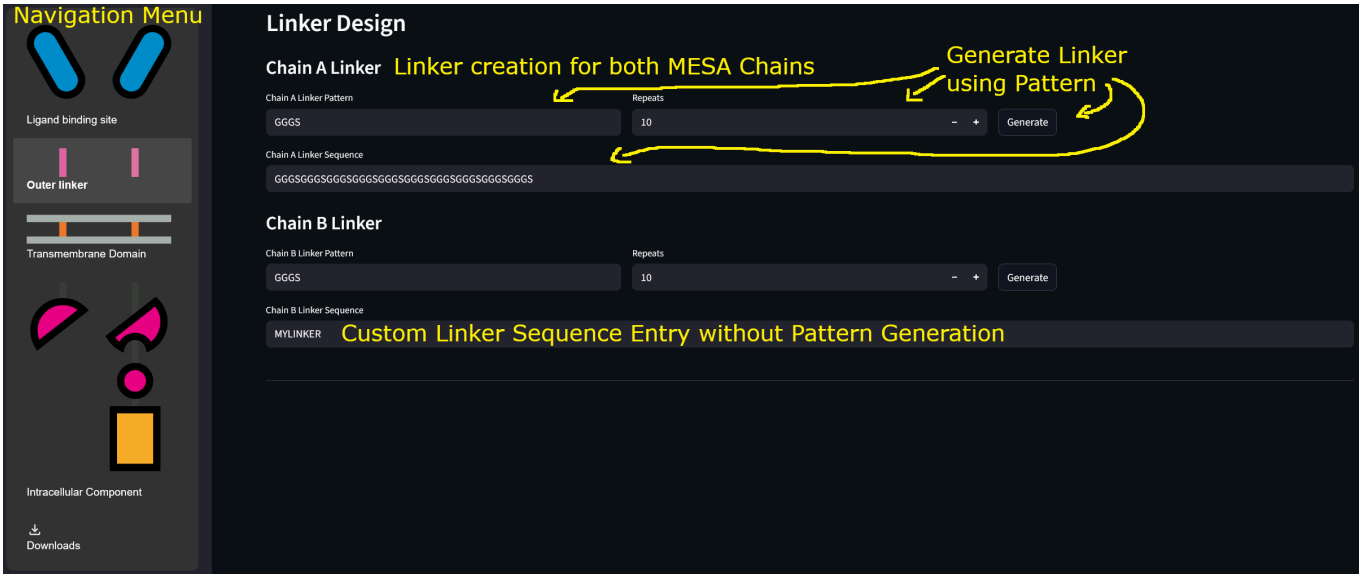
The drag-and-drop interface lets you create custom arrangements. These can be very useful for chaining antibody variable fragments and essentially extracting and linking the binding parts of an antibody into a MESA binder:

Creating Nanobodies and scFv Constructs

1. **Drag chains to target containers:**
   - **Chain A container:** Sequences dropped here become part of MESA Chain A
   - **Chain B container:** Sequences dropped here become part of MESA Chain B
   - **Components container:** Holds unused chain selections
2. **Automatic linker insertion:**
   - When multiple components are in one chain, they're connected with GGGGS×3 linkers
   - These inter-variable linkers are different from the TMD-binder linkers you'll design later
   - Order matters: drag components to arrange them N-terminus to C-terminus. Especially when you have already performed custom modeling, this can help estimate an optimal sequence, although both arrangements can work
3. **Common configurations:**
   - **Single domain antibody:** One VH or VL in Chain A, leave Chain B empty
   - **scFv:** VH and VL in the same chain, automatically linked
   - **Fab-like:** VH in Chain A, VL in Chain B
   - **Custom nanobody:** Any arrangement of selected domains

The assembled sequences appear in the Chain A/B sequence fields, ready for the next step.

Step 3: Linker Design



Linkers provide crucial flexibility between your binder and transmembrane domain or intracellular component when using intracellular MESA:

Linker Generation Options

For each active chain, you can:

1. **Use pattern-based generation:**
   - **Pattern field:** Enter a short amino acid sequence (e.g., GGGGS, EAAAK)
     - Maximum 100 amino acids per pattern
     - Validated for standard amino acids only
     - Common patterns:
       - GGGGS: Highly flexible, commonly used
       - EAAAK: More rigid, α-helical tendency
       - Custom: Any valid amino acid sequence
   - **Repeats field:** Number of times to repeat the pattern (1-1000)
     - Typical range: 5-15 repeats
     - Longer linkers (>15 repeats) provide more flexibility but may affect expression
     - Very long linkers (>30 repeats) might create structural instability
   - **Generate button:** Creates the full linker sequence
2. **Manual sequence entry:**
   - Directly type or paste a complete linker sequence
   - Maximum 1000 amino acids total
   - Overwrites any generated sequence

Linker Validation and Warnings

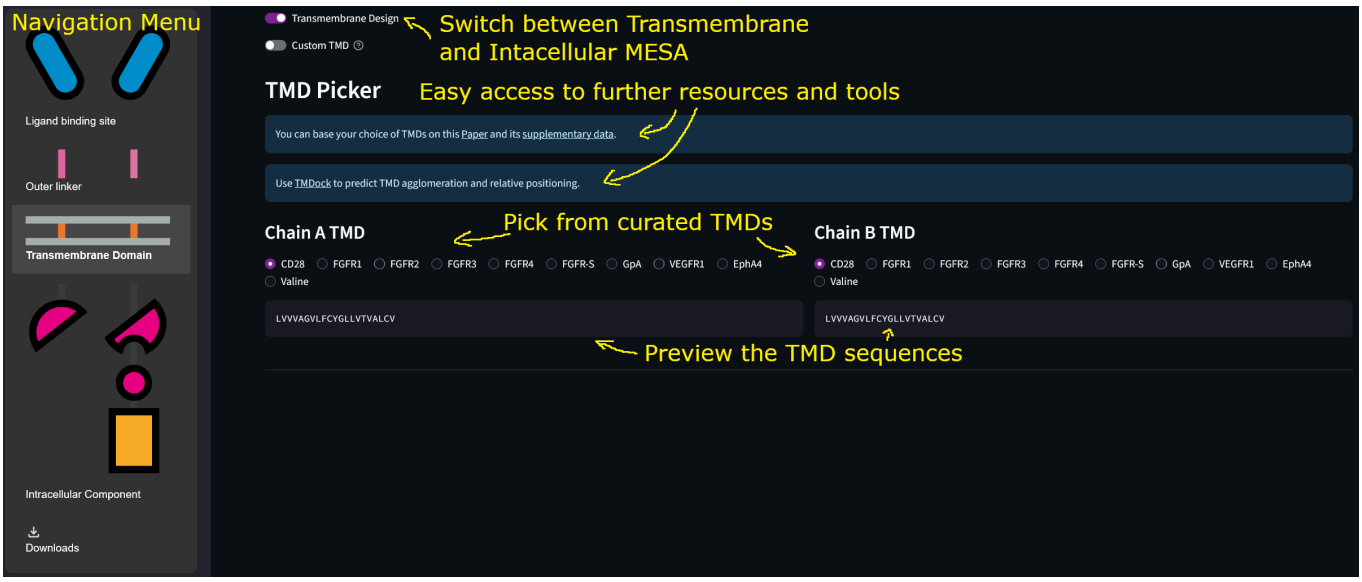
The system checks linkers for potential issues:

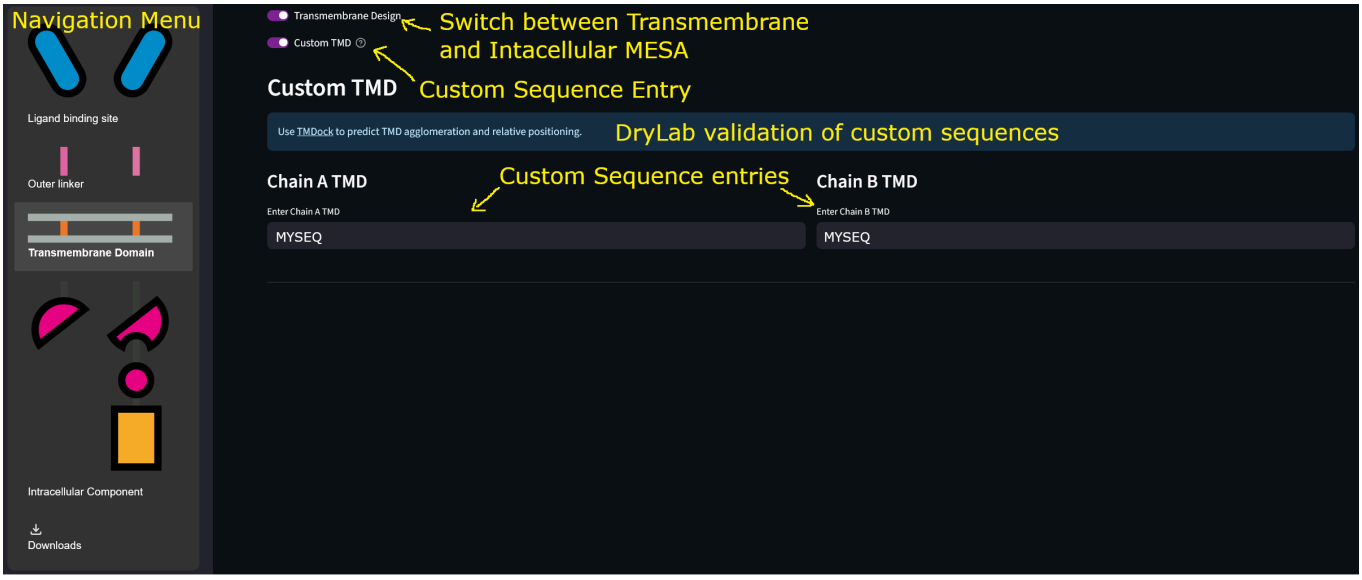
- **Length warnings:**
  - "Very short (<5 AA)" - May not provide sufficient flexibility
  - "Very long (>200 AA)" - Verify this is intentional
- **Composition warnings:**
  - ">50% hydrophobic" - May cause aggregation issues
  - "Contains stop codon(s)" - Will terminate translation
- **Non-standard amino acids:** Flagged but allowed

Design considerations:

- Flexible linkers (G/S-rich) allow more conformational freedom
- Rigid linkers (A/E-rich) maintain fixed spacing
- Hydrophobic residues should generally be avoided
- Length affects both flexibility and expression efficiency

Step 4: Transmembrane Domain Selection





The TMD anchors your receptor in the cell membrane and controls dimerization:

Transmembrane Design Toggle

First, decide if you need membrane localization:

- **ON (default):** Creates membrane-bound receptors
  - Includes signal peptide (CD4) for ER targeting
  - Adds TMD for membrane anchoring
  - Typical for cell-surface sensing applications
- **OFF:** Creates cytoplasmic constructs
  - No signal peptide or TMD
  - Starts with methionine

- For intracellular applications only

Pre-validated TMD Options

When using standard TMDs, choose based on dimerization properties. You can find a lot of data about different dimerization tendencies and fold-differences in Edelstein et. al paper's supplementary data.

Each TMD option shows the complete sequence and links to supporting literature.

Custom TMD Entry

Toggle "Custom TMD" for your own sequences:

1. **Enter TMD sequences:** One for each active chain
   - Typical length: 15-50 amino acids
   - Should be predominantly hydrophobic (>50%)
   - Low charged residue content (<15%)
2. **Validation performed:**
   - Length check (warns if outside 15-50 AA range)
   - Hydrophobicity calculation
   - Charged residue detection
   - Stop codon warning
3. **Use TMDock for validation:**
   - The tool links to TMDock for structure prediction
   - Verify your custom TMD will insert properly
   - Check predicted dimerization interfaces

TMD design tips:

- Start with validated sequences when possible
- Maintain high hydrophobicity for membrane insertion
- Avoid charged residues in the membrane-spanning region
- Consider homo- vs hetero-dimerization needs
- Check out existing data on the effect of different TMDs on dimerization
- A combination of FGFR4 and CD28 appears to be a good starting point by offering high fold-change upon dimerization Edelstein et. al

Step 5: Intracellular Component Configuration

The intracellular domain determines your receptor's output:

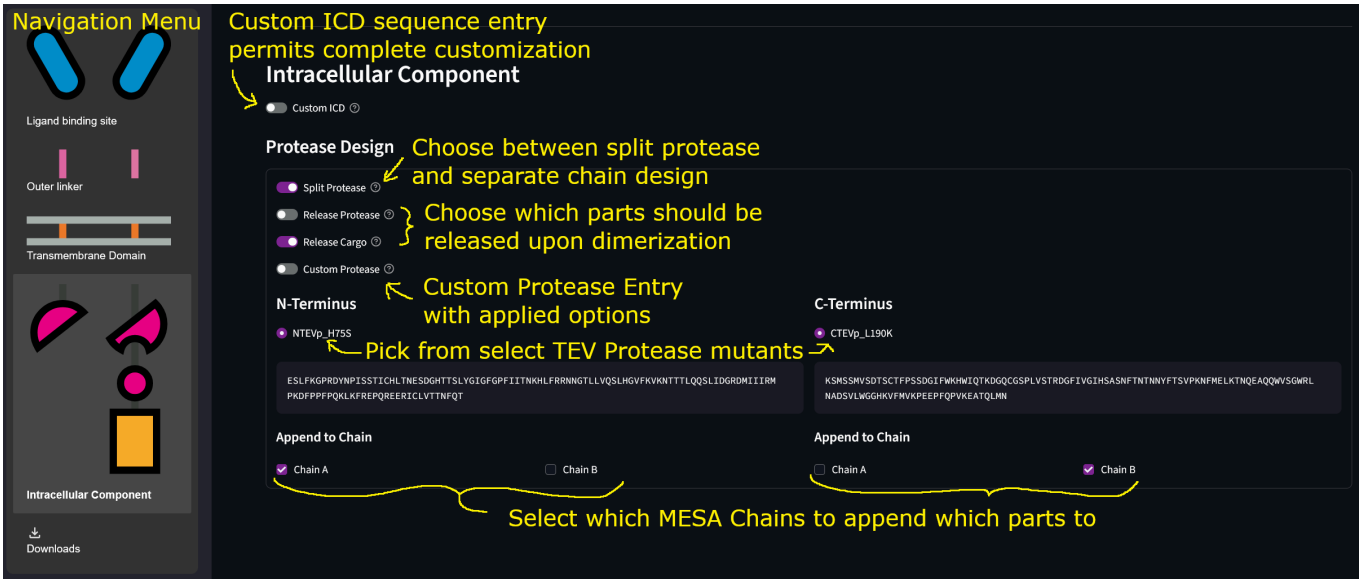
ICD Design Mode Selection

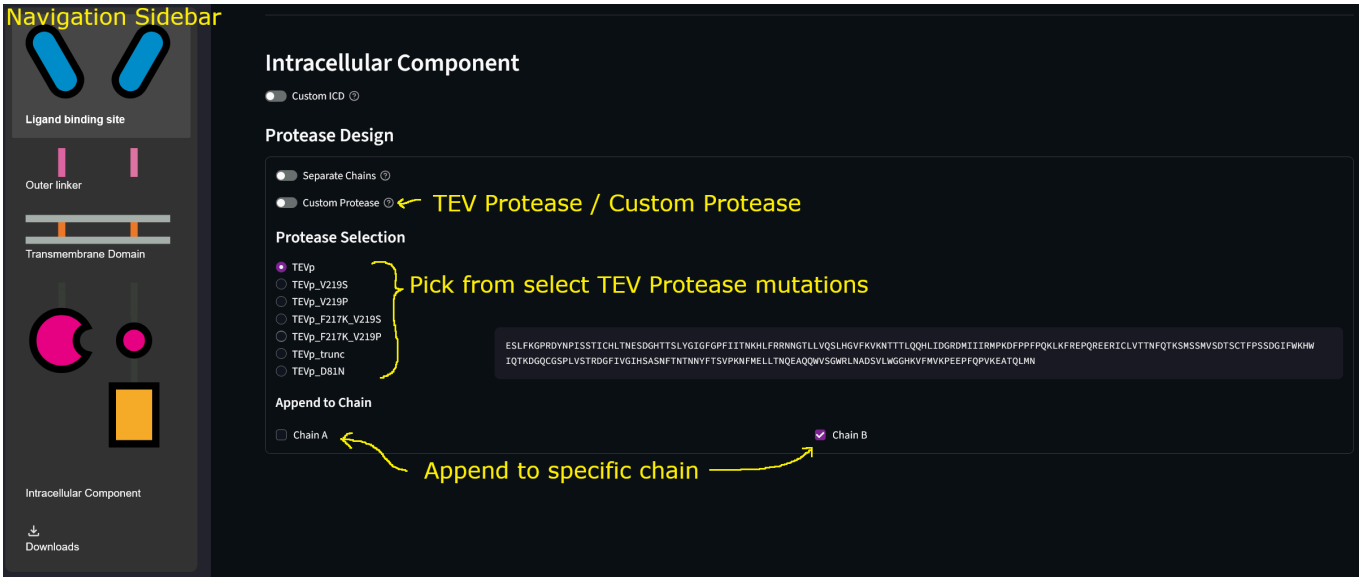
Choose between two approaches:

1. **Protease System (default):**
   - Uses protease for signal transduction via cleavage of recognition site
   - Highly modular and predictable
   - Multiple validated variants available
   - Supports split and complete configurations
2. **Custom ICD:**
   - Enter any intracellular sequence
   - Maximum 5,000 amino acids per chain
   - For non-protease signaling mechanisms or other custom needs
   - Validation checks for stop codons and length

Protease Configuration

Split vs. Separate Chains Toggle





This critical decision affects background signaling:

Split Protease (default):

- Protease divided into N-terminal and C-terminal fragments
- Fragments only active when brought together by receptor dimerization
- Lower background activity
- Requires careful chain assignment

Separate Chains:

- Complete protease on one chain

- PRS and cargo on the other chain
- Simpler design but potentially higher background
- Good for testing or when split sites unknown or hard to determine

Protease Selection

For Split Design:

- **N-terminus options:**
  - nTEVp: Standard N-terminal fragment
  - nTEVp_alt: Alternative split site
  - Each shows complete sequence
- **C-terminus options:**
  - cTEVp: Standard C-terminal fragment
  - cTEVp_alt: Matching alternative split
- **Assignment:** Check boxes to assign each fragment to Chain A or B
  - Can put both on same chain (defeats split purpose)
  - Typically one fragment per chain

For Complete Design:

- Multiple TEV variants with different activities

Custom Protease Option

When toggled on:

- Enter sequences for N-term, C-term (split) or complete protease
- Use SPELL to find optimal split sites when using split variant otherwise simply enter desired custom protease sequence
- Maximum 1,000 amino acids per fragment
- Validation checks for non-standard amino acids

Release Mechanisms

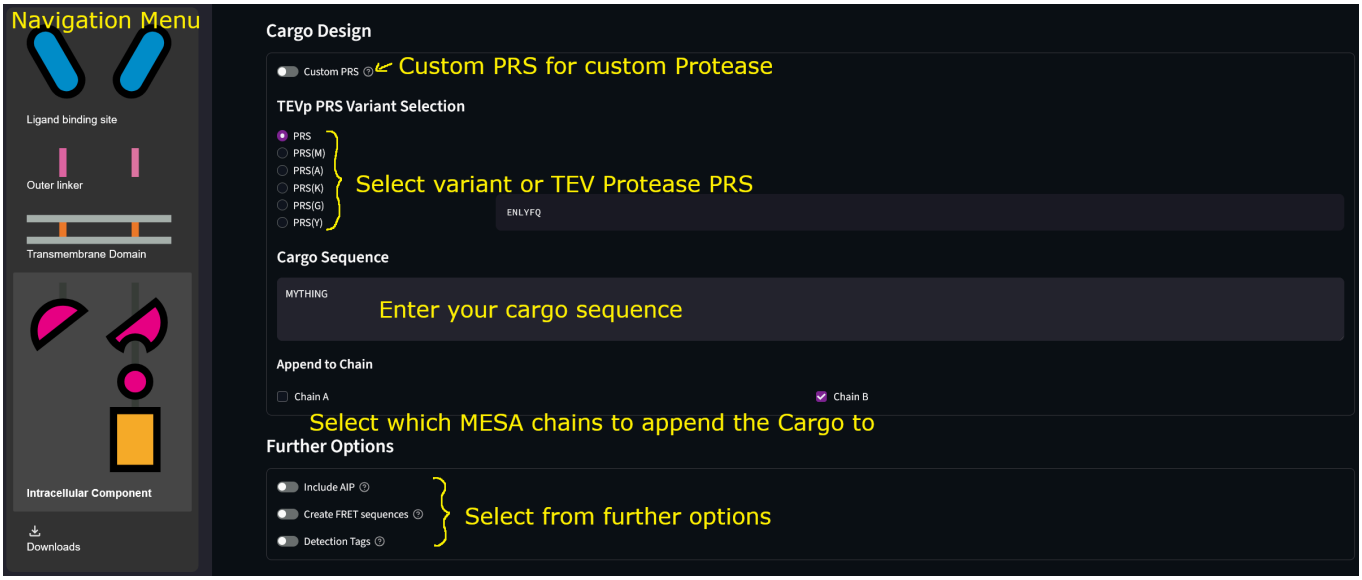
Release Protease toggle (split design only):

- ON: PRS placed before protease, releasing it upon cleavage
- OFF: Protease remains construct/membrane-tethered
- Affects signal dynamics and localization

Release Cargo toggle:

- ON: Cargo separated from protease after cleavage
- OFF: Cargo remains attached to protease
- Important for transcription factor applications

Cargo Configuration



Protease Recognition Sequence (PRS)

The PRS determines where TEV protease cuts:

Standard PRS variants:

- Each variant affects cleavage kinetics

Custom PRS:

- Enter any sequence up to 50 amino acids
- Typically 6-10 amino acids optimal
- Must match your protease specificity
- Validation warns about unusual lengths

Cargo Sequence

Your functional output protein:

1. **Enter cargo sequence:**
   - No length limit (up to 10,000 AA practical)
   - Can be transcription factors, reporters, enzymes
   - Common cargos:
     - tTA/rtTA: Inducible gene expression
     - Gal4/VP64: Synthetic transcription factors
     - GFP/mCherry: Fluorescent reporters
     - Cre/Cas9: Gene editing tools
2. **Validation checks:**
   - Large cargo warning (>2000 AA)
   - Internal stop codons flagged
   - Non-standard amino acids noted
3. **Chain assignment:**
   - Check boxes to add cargo to specific chains
   - Can be on same or opposite chain from protease
   - Consider split vs. complete protease configuration

Additional Options

Auto-Inhibitory Peptides (AIP)

AIPs reduce background by blocking protease when inactive:

Standard AIPs:

- Different mutations available
- Custom: Enter 10-30 AA sequence

Design considerations:

- Place after protease sequence
- Separated by GGGSGGGS linker
- More important for complete protease designs
- Can reduce sensitivity along with background

Detection Tags

For experimental validation and purification:

Available tags per chain:

- **FLAG (3x):** DYKDDDDKDYKDDDDKDYKDDDDK
  - Excellent for Western blots
  - Multiple commercial antibodies available
- **HA:** YPYDVPDYA
  - Smaller tag, good for immunofluorescence
- **Myc:** EQKLISEEDL
  - Compatible with many detection systems
- **Custom:** Any sequence you need

Tags are prepended after the signal sequence but before the binder.

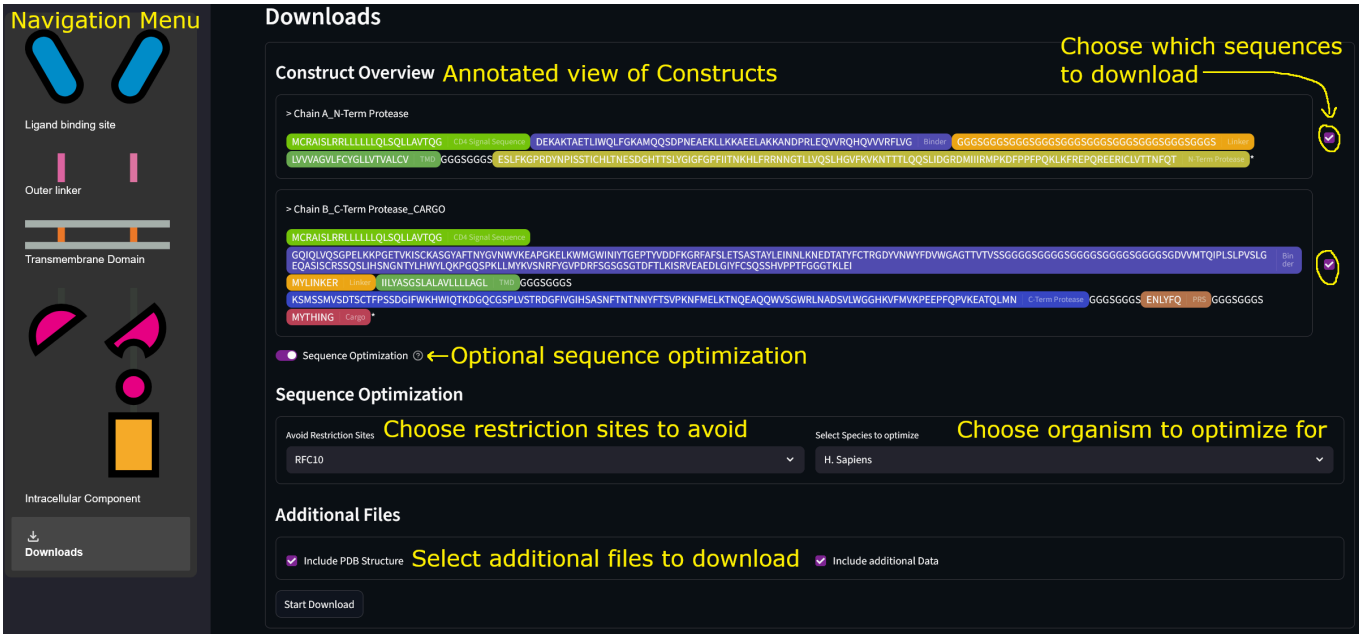
FRET Validation Constructs

Automatically generate constructs for testing:

- Creates mCerulean and mVenus fusion versions
- Tests binder dimerization without protease complications
- Each chain gets both FRET variants

- Energy transfer indicates successful dimerization

Step 6: Review and Download



Component Overview

Review all your selections:

- **Chain Tags:** Selected detection tags per chain
- **Signal Sequence:** CD4 sequence (if transmembrane)
- **Binder Overview:** Complete FASTA of selected sequences
- **Linkers:** Full sequences for each chain
- **TMD Overview:** Selected or custom TMD sequences
- **ICD Components:** Protease, AIP, PRS, and cargo configurations

Validation Warnings Summary

If any issues were detected, they're summarized here:

- Warnings grouped by component
- Severity indicated by icon color
- Won't prevent download but may affect function
- Common warnings and solutions:
  - Non-standard amino acids: Check sequence or disable optimization
  - Internal stop codons: Remove or verify intentional
  - Unusual lengths: Confirm design intent

Construct Assembly

The tool automatically assembles complete constructs:

- Each construct shown with color-coded annotations
- Domain boundaries clearly marked
- Automatic stop codon addition
- Checkboxes to select/deselect for download

Construct naming follows the pattern:

- `ChainID_Component_Modifier`
- Examples: "Chain_A_N-Term_Protease_CARGO", "Chain_B_FRET_mVenus"

Sequence Optimization

When available (no non-standard amino acids), configure:

Restriction Sites to Avoid:

- RFC10: EcoRI, NotI, PstI, SpeI, XbaI
- RFC12: Adds AvrII, NheI, PvuII, SapI, XhoI
- RFC1000: BsaI, SapI only (Type IIS)
- iGEM BioBrick Full: Comprehensive set for Registry compliance
- Custom: Select any enzymes from the database

Target Organism:

- H. sapiens: Human codon optimization
- M. musculus: Mouse expression systems
- D. melanogaster: Fly systems
- G. gallus: Chicken/avian systems

The optimizer uses DNAChisel for constraint satisfaction and codon optimization.

Download Package Contents

Click "Start Download" to generate:

1. **GenBank files:** One per selected construct
   - .gb extension
   - Full sequence annotations
   - Feature tags for each domain
   - Compatible with all major cloning software
2. **Selected PDB:** (if database search used)
   - Original structure file
   - Stored in `selected_pdb/` subdirectory
3. **Additional data:** (if selected)
   - `mesa_design_summary.txt` : Complete design parameters
   - `sabdab_data.csv` : Full search results with affinity data
   - Timestamp and configuration details
4. **References file:** Always included
   - Key publications
   - Database citations
   - Attribution requirements

All files are packaged in a ZIP archive named `mesa-design.zip`.

Troubleshooting Common Issues

Search Problems

- **"No targets were found":** Try synonyms or partial names
- **Search takes too long:** Check search query, very short queries may overload your system
- **Invalid characters warning:** Remove special characters, use only letters/numbers

Sequence Validation

- **Can't enable sequence optimization:** Check for non-standard amino acids (BJOUXZ)
- **Warning about hydrophobic linkers:** Redesign with more G/S residues
- **Stop codon warnings:** Remove asterisks unless at sequence end

Assembly Issues

- **Empty constructs:** Verify chain assignments for all components
- **Missing components:** Check that boxes are checked for protease/cargo assignment
- **Unexpected construct names:** Review chain associations in each section

Download Problems

- **Can't download:** Ensure at least one construct is selected
- **GenBank file errors:** Report specific error messages for debugging
- **Missing files:** Verify all checkboxes in download options

Best Practices

For Reliable Designs

1. Start with validated components when possible
2. Test with FRET constructs before full implementation
3. Use split protease for lower background
4. Include detection tags for troubleshooting
5. Keep linkers between 30-50 amino acids

For Optimization

1. Check all sequences for non-standard amino acids early
2. Choose restriction sites based on your cloning strategy
3. Match codon optimization to your expression system
4. Consider mRNA stability and expression levels

For Validation

1. Generate FRET constructs for new binders
2. Include both **positive and negative controls**
3. Test multiple TMD variants if unsure
4. Verify with both Western blot and functional assays

Support and Resources

For technical issues: team@igem-munich.com

For feature requests: Create an issue on our GitLab Page

Additional resources:

- MESA Original Paper
- SAbDab Database
- BindCraft Colab
- TMDock Server
- SPELL Tool

MESA Designer - Built with ❤️ by iGEM Munich 2025