User Guide for Atlas.Y

Version 1.0

Tongji-Software

Contents

1 Introduction

This user guide provides detailed instructions on how to use the Atlas.Y. The website allows users to design subcellular localization of proteins in yeast by selecting appropriate signal peptides, linkers, and target proteins.

1.1 Website Overview

Atlas.Y (Advanced Targeting and Localization Assistance System for Yeast) is a specialized web-based tool developed by Tongji-Software to facilitate the subcellular localization of proteins within yeast cells. This system is designed to assist researchers in creating fusion proteins that can be accurately targeted to specific subcellular compartments. By leveraging the vast knowledge of signal peptides and linkers, Atlas.Y helps users generate fusion proteins that carry the necessary signals for precise localization.

Upon entering the user's input, the system searches through its extensive signal peptide and linker databases to find the most suitable components for engineering the target protein. The result is a fusion protein with defined localization signals. The generated proteins undergo a comprehensive evaluation system that includes 3D structure prediction using ESM-fold, stability assessment through Rosetta, and functionality verification with the CAD-score model. This ensures the final protein design meets both structural and functional expectations.

For users needing dynamic control of protein localization in specific spatiotemporal contexts, Atlas.Y also provides an advanced dynamic version, utilizing optogenetics for precise temporal and spatial control over protein localization behavior.

1.2 Prerequisites

Atlas.Y is a web-based application, so no local installation is required. To ensure smooth operation, users should meet the following system requirements:

- **Basic Understanding of Protein Engineering:** Familiarity with concepts like signal peptides, linkers, and fusion protein design will help users navigate the platform and make informed decisions during the design process.

- **Input Data:** Users need to prepare the protein sequences they wish to modify. These sequences, along with their corresponding PDB (Protein Data Bank) files and FASTA format files, will be used as the target proteins for localization.

- **Internet Access and Compatible Browser:** Atlas.Y is a web-based tool, and users must have a stable internet connection. We recommend using the latest versions of Google Chrome, Mozilla Firefox, or Microsoft Edge for optimal performance.

- **User Account:** While some basic features may be accessible without an account, users will need to register for an account to save their designs, access advanced features, and manage their projects.

2   Website Navigation

2.1   Homepage Overview

Upon entering the homepage of Atlas.Y, users are greeted with a clean and intuitive interface. The homepage is divided into two main sections: the left-hand navigation menu and the workspace.

Navigation Menu: The navigation menu is located on the left side of the screen and provides access to the following key sections:

- **Tag Designer:** Users can switch between the **Basic Designer** and **Dynamic Designer** versions depending on their project needs.

- **Store:** In this section, users can apply for a membership or purchase additional algorithms to enhance their protein design capabilities.

- **Documentation:** The user guide and technical documentation are available here, providing comprehensive instructions and resources for using the website.

- **About:** This section provides information about the Tongji-Software team, the mission, and the values behind the project. Users can also find contact details for reaching out.

- **Support:** This section contains FAQs, video tutorials, and options for users to provide feedback or get assistance with any issues they may encounter.

Workspace: The workspace occupies the central part of the homepage. Here, users can:

- Click the **START** button to begin a new protein localization query by uploading their target protein's FASTA and PDB files.

- View and select from a list of historical queries displayed below the START button, allowing users to quickly restore and review previous searches.

2.2   Functionality Overview

The Atlas.Y website offers a variety of functional interfaces to assist users in designing and analyzing fusion proteins. The following is an overview of the main functional pages:

Protein Localization Request Submission Page: Users can input their target protein and relative requirements.

Query Results Page: Displays multiple fusion proteins generated by the system, sorted by stability score.

Entry Details Page: Provides detailed information on the selected fusion protein.

Functional Scoring Interface: Shows the overall and specific site CAD-score for fusion protein functionality.

Stability Scoring Interface: Displays the overall and specific stability score of the fusion protein.

Variant Protein Details Page: Shows the results of sequence optimization with multiple variant options.

Temporal Dynamic Localization Design System:The system generates matching fusion proteins and provides a design report with sequence details and optogenetic pair recommendations.

3 How to Use the Website

3.1 Protein Localization Request Submission

After clicking **START**, you will be directed to **the Protein Localization Request Submission page**, where you can begin by uploading the necessary files for localization editing. You are required to upload the `.fasta` file and `.pdb` file of the protein you wish to modify.

Once the files are uploaded, you can proceed to select your desired target localization from the options available on the page. If you are unfamiliar with the abbreviations used for localization, feel free to refer to the **demonstration animation** located in the left-side legend for clarification.

Finally, you will be prompted to choose the properties of the linker that best suit your design. Should you have any doubts regarding the appropriate linker selection, you can consult the User Guide via the provided link in the Atlas.Y interface for further assistance.

3.2 Query Results Page

After initiating the query, users will be directed to the **Query Results Page**, where the software generates and matches multiple localization fusion proteins. These results are displayed in order based on stability scores, which are derived from protein structure predictions and algorithmic analysis(Note: A lower stability score means a higher stability). Users can easily browse the list of entries, each providing a brief description, including information about the signal sequence, linker sequence, and stability score of the fusion protein.

Users can select specific entries according to their preferences. By clicking on any entry, they will be taken to a detailed page where more in-depth information about the selected fusion protein is available for further analysis and decision-making.

3.3 Entry Details Page

Once a user selects a specific entry from the query results, they are taken to the **Entry Details Page**. This page provides comprehensive information related to the chosen fusion protein, including:

- Detailed views of the **original protein sequence**, **signal sequence**, **linker sequence**, and **fusion protein sequence**.

- On the left side of the page, a **3D structure** of the fusion protein is displayed, allowing users to drag and rotate the model for an in-depth visualization of the protein's spatial conformation.

At the bottom of the page, users are presented with several action buttons to proceed with further analysis:

- **Functional Analysis:** Evaluate the functional retention of the fusion protein to ensure its performance in the designated subcellular environment.

- **Stability Analysis:** Further calculate the stability of the fusion protein in the new environment to ensure its long-term effectiveness.

- **Sequence Optimization:** Provides users with options to optimize the sequence, helping improve the performance of the fusion protein.

These features allow users to receive comprehensive feedback on their design and make necessary improvements to the protein structure as needed.

3.4   Functional Scoring Interface

Free version users can access an **overall functional analysis** at no cost. The software will return a total CAD-score, which provides feedback on the functional results of the designed protein. For users interested in specific sites of the protein, such as an enzyme's active site, they can choose the **advanced version** of the software, which includes the **Site-Specific CAD Scoring System**.

In the **Site-Specific CAD Scoring System**, users can view the functional scores for selected core sites of their protein. There are two methods to select a site:

1. Enter the site number directly.

2. Click on the site using the visual map on the left side of the page.

Once the user selects the desired site, the software will return two scores:

- **CAD A Score:** This score is ideal for understanding the overall structural stability and global adjustments of the protein.

- **CAD S Score:** This score focuses on local interactions within specific functional regions, such as active sites, making it particularly useful for users who are interested in specific site functionality.

Users who are more concerned with the overall protein structure should prioritize the CAD A Score, while those interested in site-specific functions, such as an enzyme's activity center, should focus on the CAD S Score for more detailed local interaction insights.

3.5   Stability Scoring Interface

In the **Stability Scoring** feature, users are able to review the overall stability score for their designed fusion protein. This score provides a comprehensive measure of how stable the protein is in its predicted conformation.Additionally, users can explore the contribution of specific forces to the protein's stability.

By providing insights into how each of these forces contributes to the overall stability, users can better understand the strengths and weaknesses of their fusion protein design. This information can guide further optimization of the protein for enhanced stability.

3.6   Sequence Optimization System

To improve the success rate of fusion protein localization, users can opt for the **Sequence Optimization System** feature. This feature enhances the localization capabilities of the fusion protein by recommending optimized variants.

Once the user selects this option, the software will calculate and generate a list of variant proteins. These variants are designed to improve localization efficiency, and the software provides multiple entries for users to choose from.

Upon selecting a variant, the user will be directed to the **Variant Protein Details Page**, where the following information is displayed:

- **Protein Sequence:** The complete sequence of the selected variant protein.

- **Functional Score:** The software will calculate and display the functional score of the variant, helping the user understand the protein's functionality in the context of its subcellular localization.

- **Stability Score:** The corresponding stability score for the variant protein is also provided, giving insights into the structural robustness of the protein.

By reviewing the sequence, functional score, and stability score, users can compare different variants and choose the most suitable one for further analysis or experimentation. The detailed feedback allows users to make data-driven decisions, ensuring the success of the designed fusion protein.

3.7   Temporal Dynamic Localization Design System

In addition to the basic version, users can switch to the **Dynamic Version** via the menu bar. Within the **Temporal Dynamic Localization Design System**, users have the ability to select pre-defined optogenetic conditions that control the temporal localization of fusion proteins. These conditions are supplemented with other matching requirements that are consistent with the basic version of the system.

Once the user sets their desired conditions, the software will generate multiple matching results based on the provided inputs. The preliminary steps for selecting and uploading data remain the same as in the basic version, ensuring a familiar workflow for users.

At the final step, the system will output a comprehensive Design Report tailored to the user's needs. This report will include:

- **Optogenetic Pair:** Detailed information about the recommended optogenetic pair for controlling temporal localization.

- **Component Sequences:** The complete sequence information of each component in the fusion protein, including the signal peptide, linker, and target protein.

This dynamic version of the localization design allows users to achieve precise spatiotemporal control over their fusion proteins, with clear guidance provided by the system's design report.

4   Core Features

4.1   Signal Peptide Selection

Users can choose appropriate signal peptides for their fusion protein design. The website interface provides a list of pre-defined signal peptides from a comprehensive database, allowing users to select one that best suits their target localization.

Signal Peptide	Localization
NLS	Cytoplasm to Nucleus
NES	Nucleus to Cytoplasm
SP	Cytoplasm to Endoplasmic Reticulum Lumen
SP_TM	Anchored on the Endoplasmic Reticulum Membrane
SP_GPI	Cytoplasm to Endoplasmic Reticulum Lumen, then to Cell Surface
GPI	Cell Surface
TM	Embedded in Cell Membrane
PTS	Peroxisome
MT	Mitochondria
LD	Lipid Droplets

Table 1: Signal Peptides and Their Corresponding Localization

4.2   Linker Selection

Users can choose linkers from the website's extensive linker database, which contains both rigid and flexible options. Each linker is categorized based on its length, flexibility, and sequence properties, making it easy for users to find the most suitable option for their fusion protein design.

When selecting the properties of linkers, users can refer to the following guidance:

Rigid Linker: If your design requires the two fused protein domains to maintain a fixed relative position, and structural stability with minimal movement is needed for functionality, a rigid linker is more appropriate. This type of linker is typically used to preserve a specific protein structure, ensuring precise interaction between functional domains while preventing unnecessary movement.

Flexible Linker: A flexible linker is better suited when the two fused protein domains need more freedom of movement, or when the function requires a certain degree of conformational flexibility. Flexible linkers allow for a wider range of motion, making them ideal for situations where the fusion protein needs to adjust its conformation freely without interfering with the functionality of each domain.

Hydrophilic Linker: If the fusion protein is intended to function in an aqueous environment, or if the target protein's function involves the cytosol or other hydrophilic environments, a hydrophilic linker is more suitable. Hydrophilic linkers have good water solubility and can maintain stability in polar solvents, reducing non-specific aggregation.

Hydrophobic Linker: If the fusion protein needs to operate in a hydrophobic environment (such as membrane-related structures or functions), or if the fused protein domains are located in the cell membrane or other hydrophobic regions, a hydrophobic linker is more appropriate. Hydrophobic linkers provide better stability in non-polar environments, helping to maintain the structural integrity of the fusion protein and avoiding interference from aqueous environments.

In summary, rigid linkers are suitable for scenarios where structural integrity is crucial, while flexible linkers are preferred for applications requiring greater flexibility in protein function. Hydrophilic linkers are ideal for proteins functioning in aqueous environments, while hydrophobic linkers are better suited for proteins that function in membrane or hydrophobic environments.

5 Membership Purchase and Algorithm Purchase

5.1 Membership Plans

Atlas.Y offers three membership tiers to cater to different user needs: Free, Professional, and Enterprise versions.

- **Free Version**: Users can access basic features such as protein localization design and overall functionality scoring. This version is ideal for students, beginners, and occasional researchers.

- **Professional Version**: This plan includes all features of the Free Version plus full access to the algorithm store. It is available as a monthly or annual subscription and is suited for active researchers or small to medium-sized labs requiring advanced tools.

- **Enterprise Version**: Tailored for large research institutions or companies, this version offers advanced team management features, custom algorithm support, and enhanced data security services. It is available on a monthly or annual subscription basis, with custom development options.

5.2 Algorithm Store

For users who need specific functionalities, Atlas.Y provides an Algorithm Store where individual algorithms can be purchased on a pay-per-use basis. The available algorithms include:

- **Site-Specific CAD Scoring System**: Offers detailed evaluation of specific protein sites, such as enzyme active sites.

- **Temporal Dynamic Localization Design System**: Allows users to implement optogenetic-controlled protein localization over time.

- **Optimized Protein Evolution Localization System**: Assists in optimizing protein evolution for specific localization within organelles.

- **Protein Localization Prediction System**: Predicts protein localization within cellular structures using advanced algorithms.

Users can purchase individual algorithms or opt for bundled packages for cost savings. These options provide flexibility, ensuring that both occasional and frequent users can access the tools they need.

6   Support and Feedback

6.1   Contact Support

For detailed troubleshooting guides and frequently asked questions, please visit the **Support** section of our website. This section contains FAQs, video tutorials, and options for users to provide feedback or get assistance with any issues they may encounter.

If you encounter any technical issues or require customer service assistance, please feel free to contact us at `tongji-china2019@163.com`.

6.2   Submitting Feedback

We highly value user feedback and encourage you to share your thoughts or report any bugs you encounter. To submit feedback, please visit the **Support** section of our website. Users can fill in the feedback form and submit their suggestions. We regularly review user feedback to prioritize future updates.

Your feedback helps us improve Atlas.Y, and we appreciate your contribution to making the platform better.