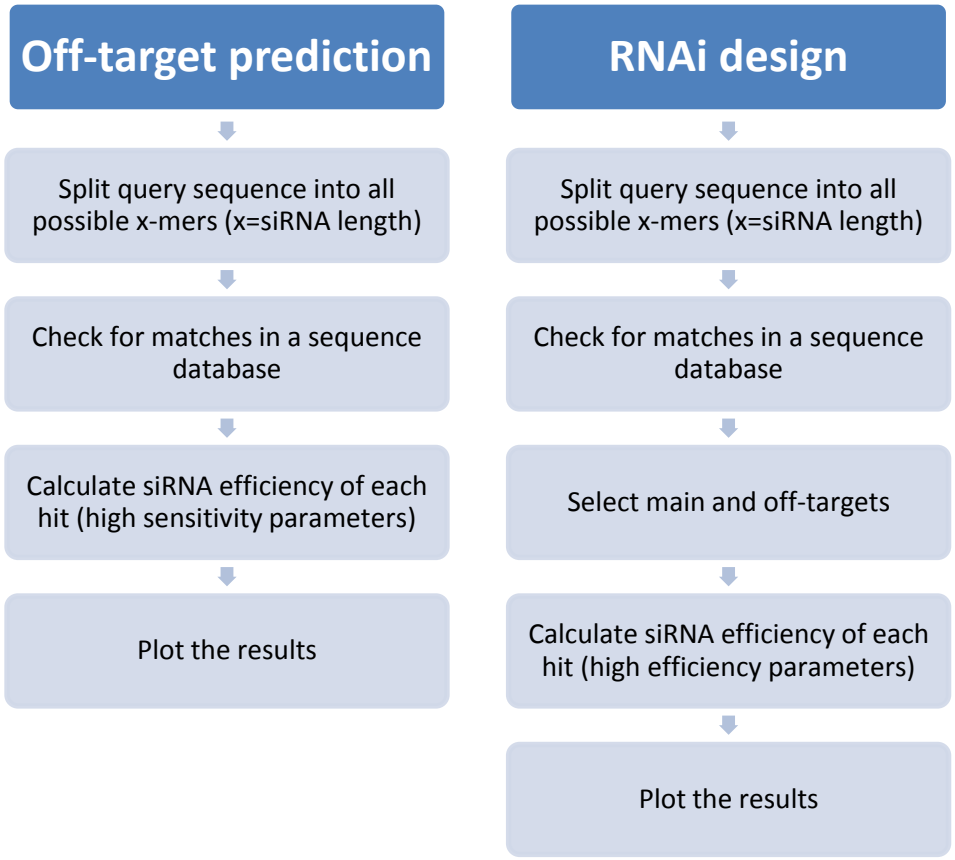
Help at a glance

Long double stranded RNAi construct design and RNAi off-target prediction

Stefanie Lück
Dimitar Douchkov
IPK-Gatersleben 2017

siFi21 Workflow



NOTE: For checking existing RNAi construct design. The query should be the RNAi trigger sequence.

NOTE: For selecting optimal sequence for RNAi construct design. The query should be the sequence of the target mRNA.

siRNA structure and strand selection mechanism

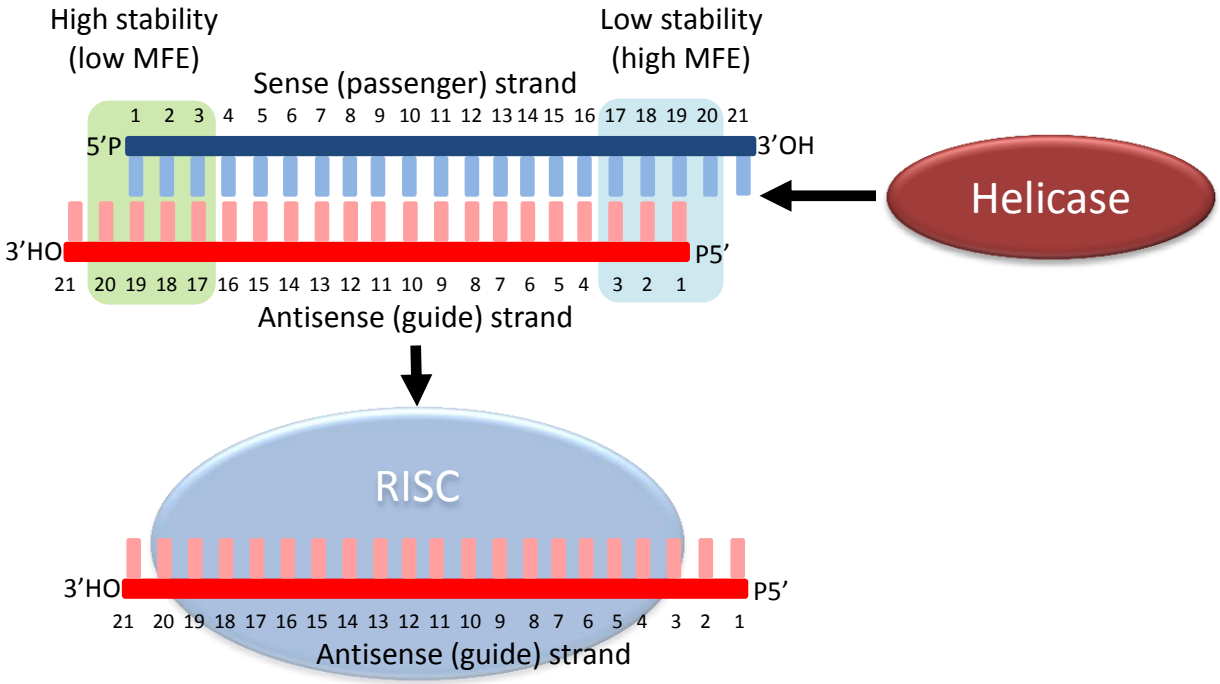


Figure 1. Thermodynamic and sequence based parameters

Thermodynamic parameters: The thermodynamic stability of the first few base pairs of either siRNA strand affects the ratio of RISCs containing the antisense (red) or sense (blue) strands of siRNAs. The relatively low thermodynamic stability in the 5' end of the antisense strand (blue shaded box) compared with the high thermodynamic stability in the 5' end of the sense strand (green shaded box) leads to a preference for the incorporation of the antisense strand into RISC and cause efficient silencing of the target. RISC - RNA-induced silencing complex; MFE - thermodynamic minimum free energy.

5' Terminal nucleotide rule: According several studies the 5'-terminal nucleotide of both strands plays a critical role in strand selection. Basically, strands with U or A at the 5'-end are incorporated more efficiently into RICS than strands with G or C on the same position. **siFi21** checks for the 5'-terminal nucleotides of both strands, then calculates the minimum free energy (MFE) of the first 3 nucleotides of the 5' end of each strand by taking into account also the 3'-dangling ends of the opposite strand.

Elbashir et al. Nature 2001; Schwartz et al. Cell 2003; Dorsett and Tuschl Nature 2004; Malefyt et al. FEBS 2014 etc.

Target site accessibility (mRNA)

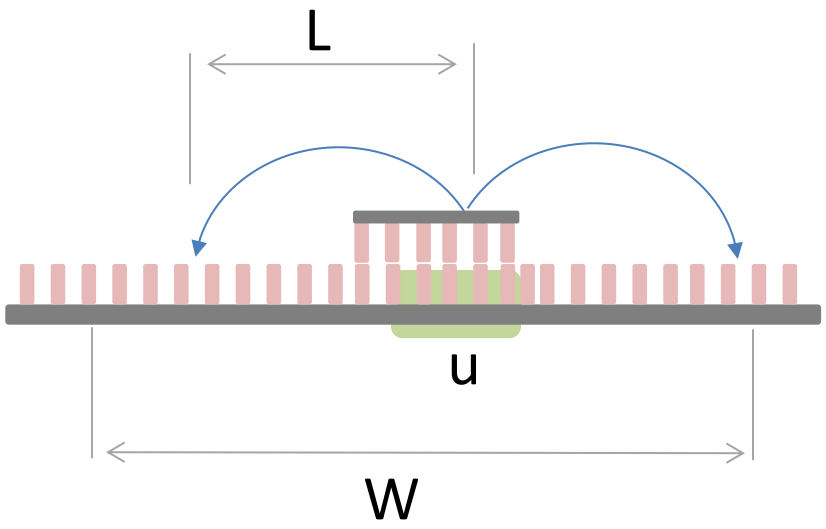
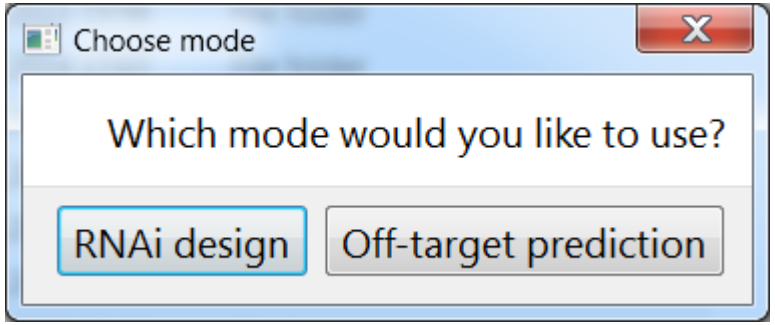


Figure 2. Local base pairing probability

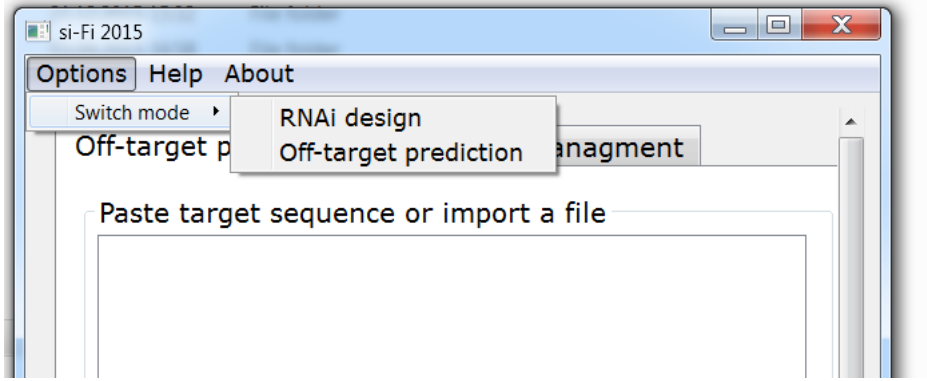
The secondary structure of RNA often includes internally paired regions of double stranded RNA. These regions are nor readily accessible for the RISC complex as the interaction requires pairing of the siRNA with the target. However, calculation of the complete secondary structure of an mRNA is usually inaccurate and very computational demanding. Moreover, large secondary structure of mRNA are usually destroyed in the translation process by the progressing ribosomes. Therefore **siFi21** calculates the local (in a small region) base-pairing probability of the target site by using RNAplfold (Vienna RNA package). It gives averaged pairing probability for a region of lengths **u** within a distance **L** and window **W**. Higher pairing probability of a region is associated with lower mRNA accessibility, which makes it a suboptimal RNAi target.

Bernhart et al. Bioinformatics 2006; Tafer et al. Nature Biotech 2008.

Getting started with **siFi21**

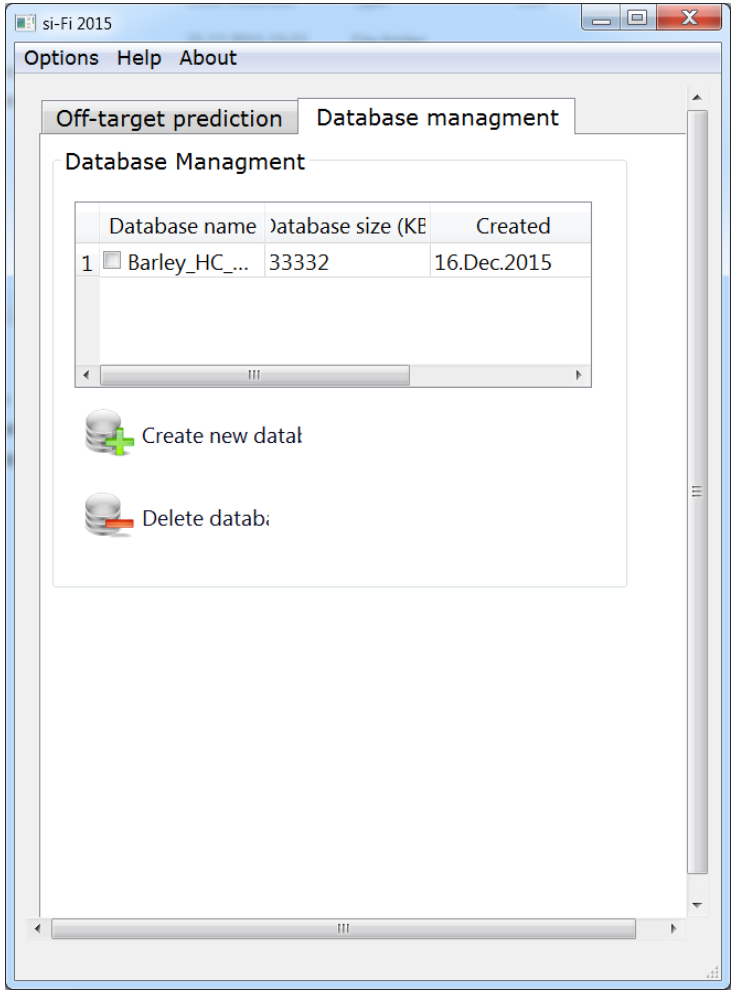


At the start you will be asked to select a mode:
RNAi design mode is for selecting optimal region to be targeted by RNAi
Off-target prediction mode is to test putative or existing RNAi construct for target specificity within a custom sequence database.



You can switch between the two modes at any time.

Database management tab

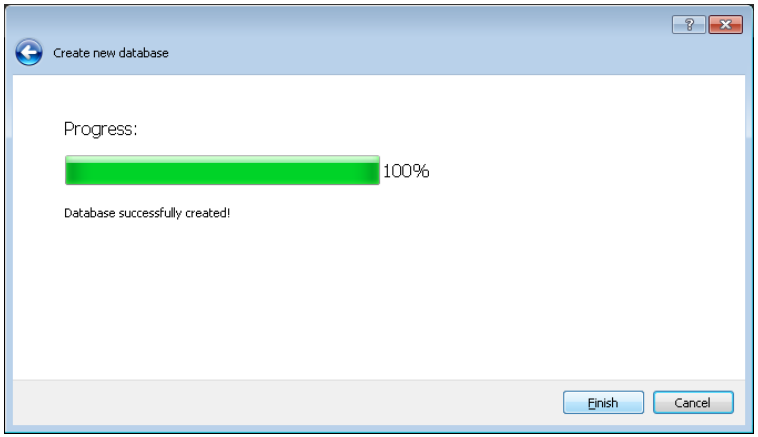


Database management tab

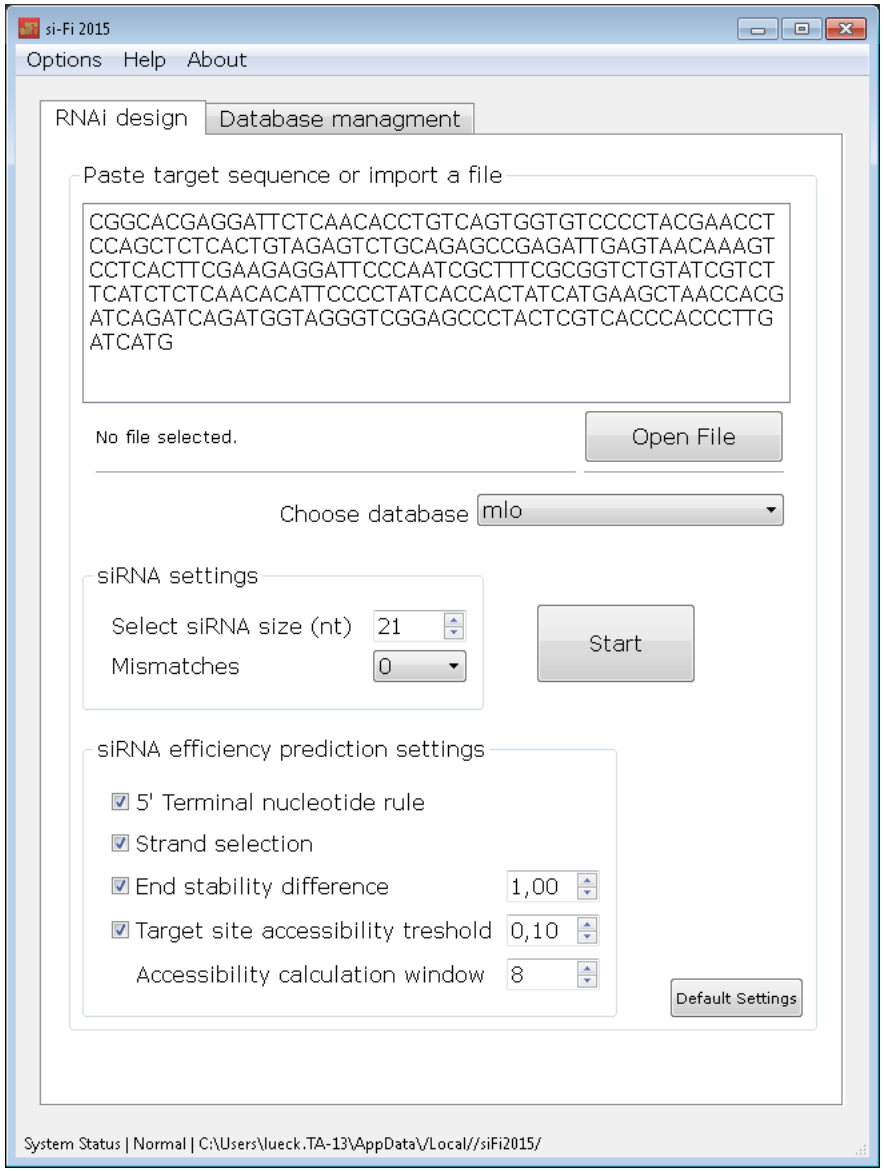
Prior to to start the analysis (regardless of RNAi design or Off-target search) a nucleotide sequence database must be created. It will be used to find targets and off-targets of the analyzed sequence. Typically this database will contain cDNA sequences from the organism of interest (ideally the complete transcriptome of the organism).

In **Database management** tap you can create sequence database from a multiple FASTA file or delete existing databases by selecting it and pressing **Delete database** button.

Important! Creation of very large databases may take up to few hours of time! Please be patient and wait for the accomplishing of the process confirmed by a message **Database successfully created** and pushing the **Finish** button:



RNAi design mode



The **RNAi design** mode is for selecting optimal region for designing RNAi constructs.

1. Paste the **mRNA target** sequence (in FASTA format) or load it from a file. This will be (as full as possible) sequence of the gene that you want to silence.

2. Select a nucleotide sequence database. Your RNAi construct will be tested against this database for main- and off-targets. See also the “Database management” tab.

3. Select a siRNA length (default is 21 nucleotides)

4. Choose the number of allowed mismatches of siRNA to the target (default is 0).

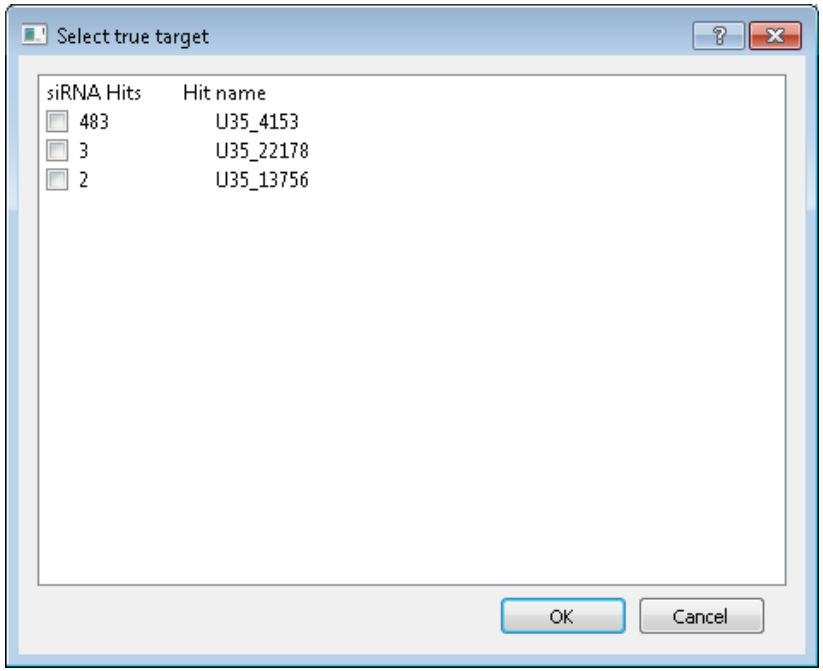
   NOTE: If mismatches are allowed the siRNA efficiency calculations are disabled since no reliable data is available.

5. Select criteria for definition of efficient siRNA:
   - Check the 5’-terminal nucleotides of both strands
   - Use MFE based strand selection (see Fig. 1)
   - Minimum MFE difference (in kcal/mol; higher is better, default 1.00)
   - Minimum target site accessibility (higher is better, default 0.1)
   - Length of region u (see Fig. 2; default is 8)

The default parameters are selected for higher efficiency of the siRNAs.

Press the **Start** button to start the analysis
Press the **Default Settings** button to restore the default values.

RNAi design mode



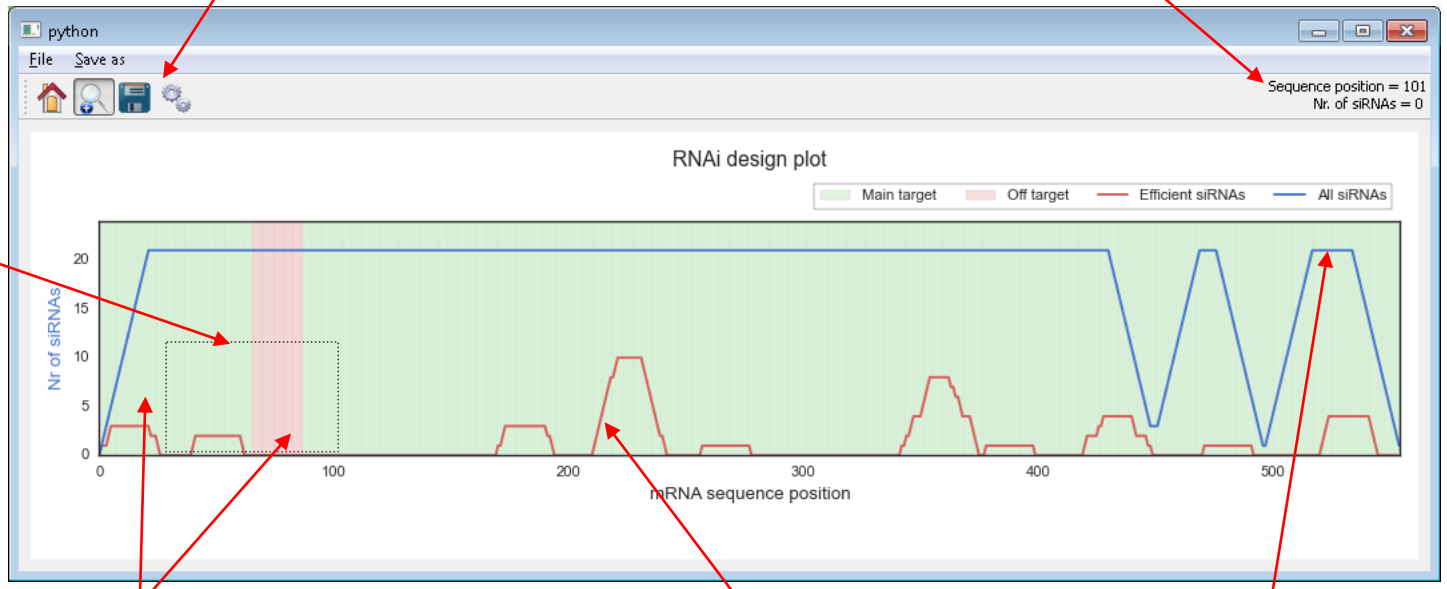
With the start of the analysis **siFi21** will generate all possible siRNAs that may originate from the sequence of interest (the query sequence) and will search with all of them against the provided sequence database. The ID of all database sequences where siRNA hits were found will be listed in the **Select true target** window. Select your main target gene(s). This is the gene or genes that you want to silence . All unselected sequences will be treated as off-targets.

You can select more then one main target.

Result plot for RNAi design

Toolbar with buttons for "Return to default view", "Zoom" and "Plot settings"

Coordinates of the pointer showing the sequence position and the siRNA counts on this position.



Zoom rectangle (see next page)

Green area – siRNAs generated from this region will hit only the selected main target
Red area – siRNA generated from this region may hit off-targets besides the main target

Number of putatively generated efficient siRNAs that will match the corresponding main target.

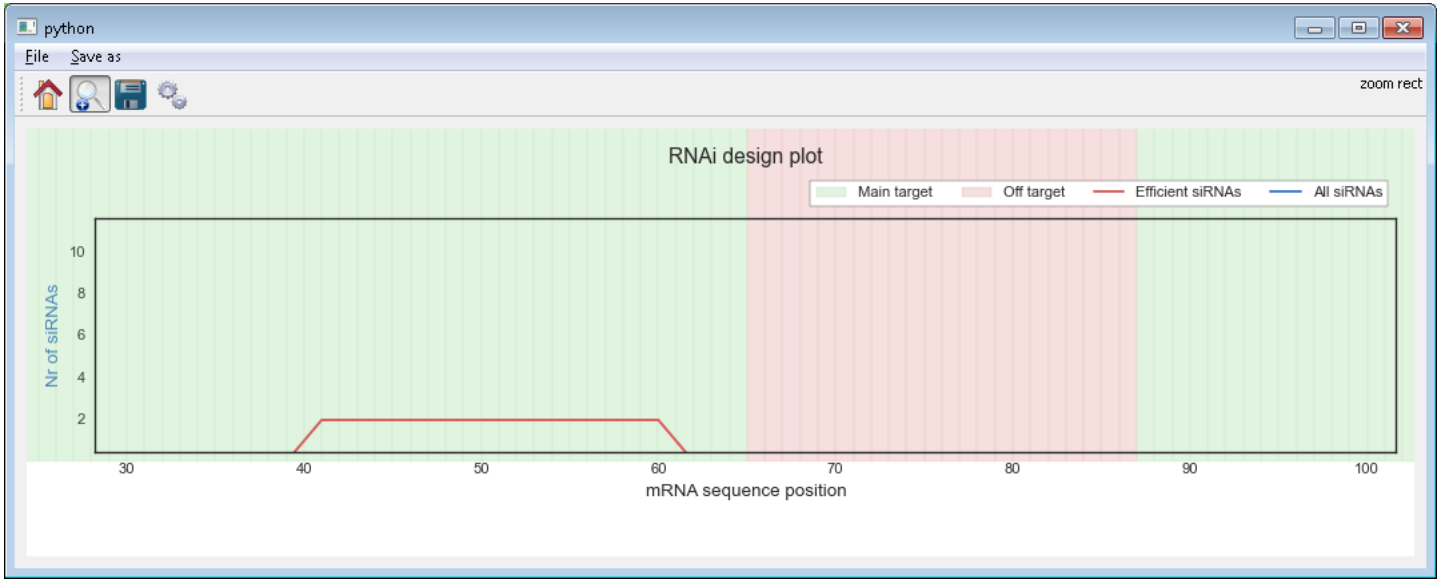
Total number of putatively generated siRNAs that will match the corresponding main target.

On X-axis is the nucleotide position of the mRNAi target sequence.
On Y-axis is the number of putatively generated siRNA per nucleotide position of the mRNA.

The results can be exported in a tabular form and as an image file.

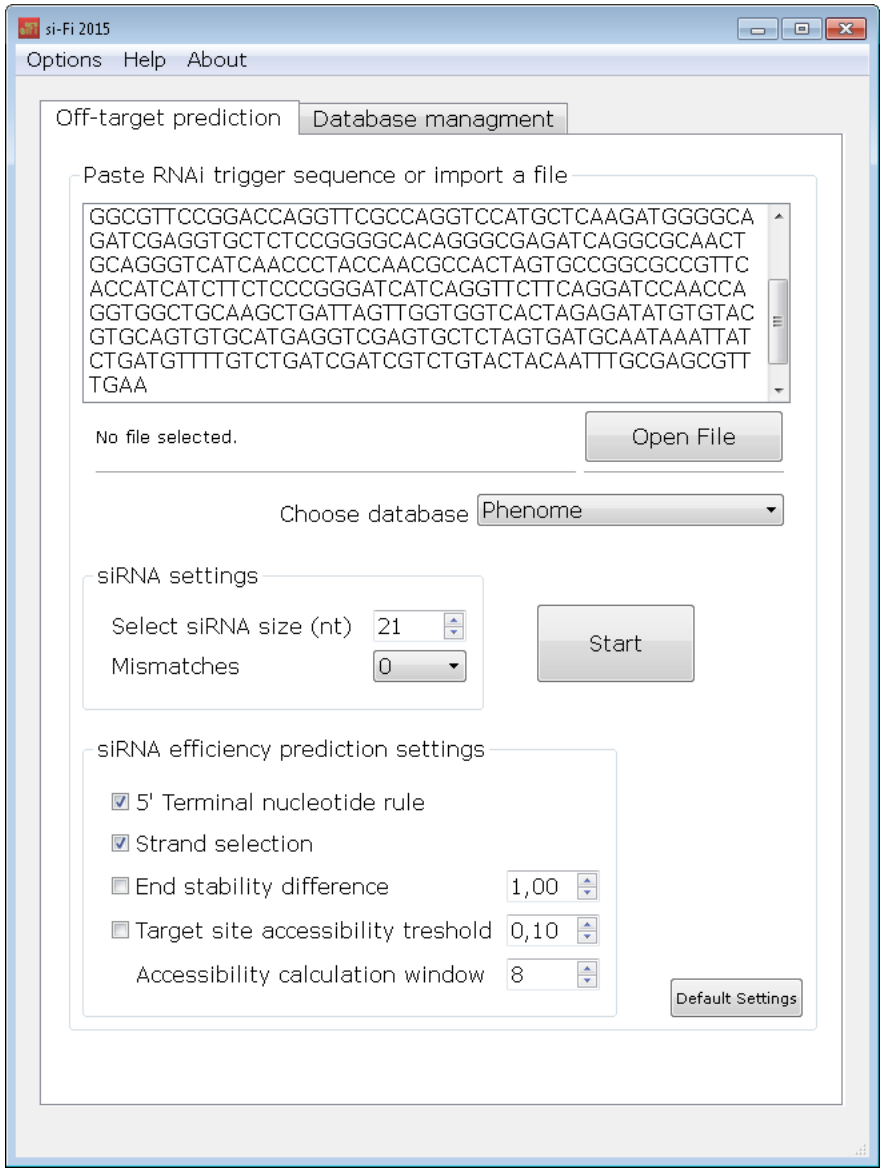
An ideal region for designing an RNAi construct will contain only green area (main-target specific, no off-targets) and as much area covered by red line (presence of efficient siRNAs) as possible.

Result plot for RNAi design



Zoomed area from the plot on the previous page with finer scaling of X- and Y-axis.

Off-target prediction mode



The **Off-target prediction** mode is for checking RNAi construct design for off-targets within the provided sequence database.

1. Paste the **RNAi trigger** sequence (in FASTA format) or load it from a file.

2. Select a nucleotide sequence database. Your RNAi construct will be tested against this database for main- and off-targets. See also the “Database management” tab.

3. Select a siRNA length (default is 21 nucleotides)

4. Choose the number of allowed mismatches of siRNA to the target (default is 0).

NOTE: If mismatches are allowed the siRNA efficiency calculations are disabled since no reliable data is available.

5. Select criteria for definition of efficient siRNA:
   - Check the 5’-terminal nucleotides of both strands
   - Use MFE based strand selection (see Fig. 1)
   - Minimum MFE difference (in kcal/mol; higher is better, default is “not selected”)
   - Minimum target site accessibility (higher is better, default “not selected”)
   - Length of region u (see Fig. 2; default is “not in use”)

The default parameters are selected for higher sensitivity for off-targets.
Press the **Start** button to start the analysis
Press the **Default Settings** button to restore the default values.

Result plots for **Off-target prediction**

NOTE: Maximum of 5 targets will be displayed in the plot. It there are more target hits the will be shown only in the Summary table.

Toolbar with buttons for "Return to default view", "Zoom" and "Plot settings"

Coordinates of the pointer showing the sequence position and the siRNA counts on this position.

Result summary table. Number of found siRNA hits (total and efficient) to the corresponding target.

Sequence IDs of the found putative targets of your query RNAi sequence. In this mode **siFi21** will show all found targets without determination of true- and off-targets.

Blue line - total number of generated siRNAs with match to the corresponding sequence.

Red line - number of found efficient siRNAs with match to the corresponding sequence.

On X-axis is the nucleotide position of the RNAi trigger (query) sequence.
On Y-axis is the number of putatively generated siRNA per nucleotide position of the RNAi trigger sequence.

Typically the sequence with the most hits will be the main-target of the tested RNAi construct and the rest are putative off-targets.

The results can be exported in a tabular form and as an image file.

siRNA **Fi**nder ver. siFi21_1.2.3-0008

Stefanie Lück (programming)
Dimitar Douchkov (program design)
Patrick Schweizer (supervisor)
IPK-Gatersleben 2017

python powered