AmiRNA Designer — new method of artificial miRNA design

Agnieszka Mickiewicz[1*], Agnieszka Rybarczyk[1,2*], Joanna Sarzynska[1], Marek Figlerowicz[1,2] and Jacek Blazewicz[1,2✉]

[1]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland; [2]Institute of Computing Science, Poznan University of Technology, Poznań, Poland

MicroRNAs (miRNAs) are small non-coding RNAs that have been found in most of the eukaryotic organisms. They are involved in the regulation of gene expression at the post-transcriptional level in a sequence specific manner. MiRNAs are produced from their precursors by Dicer-dependent small RNA biogenesis pathway. Involvement of miRNAs in a wide range of biological processes makes them excellent candidates for studying gene function or for therapeutic applications. For this purpose, different RNA-based gene silencing techniques have been developed. Artificially transformed miRNAs (amiRNAs) targeting one or several genes of interest represent one of such techniques being a potential tool in functional genomics. Here, we present a new approach to amiRNA design, implemented as AmiRNA Designer software. Our method is based on the thermodynamic analysis of the native miRNA/miRNA* and miRNA/target duplexes. In contrast to the available automated tools, our program allows the user to perform analysis of natural miRNAs for the organism of interest and to create customized constraints for the design stage. It also provides filtering of the amiRNA candidates for the potential off-targets. AmiRNA Designer is freely available at http://www.cs.put.poznan.pl/arybarczyk/AmiRNA/.

INTRODUCTION

RNA interference (RNAi) is a widespread mechanism responsible for regulation of gene expression in Eukaryotes. One of the key elements of the RNAi machinery are small regulatory RNAs (srRNAs) which serve as 20–30 nt probes targeting multienzymatic protein complexes to specific transcripts or genes (Obernosterer *et al.*, 2006). It has been suggested that 60–90% of eukaryotic genes are regulated by srRNA (Friedman *et al.*, 2009). srRNAs are involved in gene regulation at both transcriptional and post-transcriptional level (Nowacka *et al.*, 2012). Various classes of srRNAs differ in mode of biogenesis, precursor structure and in their function in gene regulation (Kim, 2005a, 2005b, Blazewicz *et al.*, 2011, Rybarczyk *et al.*, 2015, Szostak *et al.*, 2014). Among them, microRNAs (miRNAs) and small interfering RNAs (siRNAs) are found to be the most important for post-transcriptional gene silencing.

The miRNAs are endogenous in both plants and animals and originate from intergenic regions or introns, as primary precursors (pri-miRNAs) transcribed by RNA polymerase II (Lee *et al.*, 2004, Xie *et al.*, 2005). These precursors fold in stem-loop structures that contain a about 21-nucleotide (nt) miRNA in the 5' and/or 3' arm of the stem. However, there are differences both in the miRNA biogenesis and the structure of the precursor in plants and animals (Bologna *et al.*, 2013). Animal miRNA precursors are about 80 nt long and two different proteins, Drosha and Dicer, carry out the cleavage reaction. Drosha in the first nuclear step releases pre-miRNA (Han *et al.*, 2006) and then Dicer yields miRNA/miRNA* duplex in the cytoplasm. Plant miRNA precursors are often significantly longer, up to over 600 nt, display structural diversity and are cleaved in the nucleus by Dicer-like 1 protein (DCL1) (Kurihara and Watanabe, 2004) by several mechanisms, depending on the processing direction and the number of the cuts, eventually generating a about 21-nt miRNA/miRNA* duplex (Bologna *et al.*, 2013). Next, both in animals and plants, the miRNA (guide strand) is preferentially loaded onto the Argonaute protein (AGO) component of the RNA-induced silencing complex (RISC) while the miRNA* (passenger strand) is destroyed. The preferential loading of the guide strand is determined by thermodynamic asymmetry along the miRNA/miRNA* duplex (Eamens *et al.*, 2009). The characteristic, thermodynamically unstable 5' end of the miRNA strand allows this molecule to be loaded onto the RISC. The miRNA then directs the cleavage of the messenger RNA (mRNA) through the AGO in the RISC or represses translation of the transcript (Valencia-Sanchez *et al.*, 2006). Although the RNAi pathway is universal, there are differences in mode of its function in animals and plants. Plant miRNAs usually display perfect or near-perfect pairing with their mRNA target sites (Axtell *et al.*, 2011). Animal miRNAs are less complementary to their target sites. They harbour "seed" region: 2–8 nucleotides of the 5' ends complementary to ca. 7 nt of the target (Bartel, 2009). It was thought that the complementarity level is connected with the mechanism of inducing gene repression by miRNA. The most complement miRNAs were suggested to act through degradation of their target transcripts whereas the less paired — to repress translation. While this is true in the case of animal models, there is no obvious evidence for the same correlation in plants (Brodersen and Voinnet, 2009), and the mechanism still remains unclear. Another difference is in the location of miRNA target sites. In animals, miRNA target sites are located mostly within the 3' untranslated regions (3'

UTR) whereas in plants miRNAs bind both to the coding and untranslated regions. In plants miRNAs regulate mostly growth and developmental processes (Mallory & Vaucheret, 2004), and are involved in various stress responses (reviewed in Sunkar *et al.*, 2012). In animals, miRNAs are involved in proper functioning of every cell (Chekulaeva and Filipowicz, 2009), and their dysfunctions are associated with various diseases (reviewed in Meola *et al.*, 2009).

The second important class of molecules that trigger RNAi mechanism in plants and some animals are siRNAs (Carthew & Sontheimer, 2009). They originate from invading viruses, transgene inverted repeats, centromeres, transposons, and other repetitive sequences and they silence these sequences. The long double-stranded RNA (dsRNA) derived from a transgene construct is a common entry transcript in the siRNA pathway. It is diced sequentially into short (21–24 bp) fragments by Dicer2 (or Dicer1) in case of animals or Dicer-like 2–4 in plants. The siRNA is thought to be an RNA-based immune mechanism to fight off nucleic acid invaders or to maintain genome stability.

As the siRNA-based technology seems to be easy to implement, there were several applications based on long dsRNA to regulate a specific gene of interest (Chau & Lee, 2007). However, the generation of a pool of multiple short RNA may result in unspecific silencing (off-target effect), which has been especially carefully studied in animals (Petri & Meister, 2013). In contrast, the biogenesis of miRNA is a highly precise mechanism that results in arising of a homogeneous pool of short regulatory RNAs. For this reasons, the idea of miRNA-based method of both endo- and *exogenous* regulation has been developed. The main concept of the artificial miRNA (amiRNA) is to design a 21 nt-long RNA sequence that is able to regulate the target gene expression. This sequence is introduced into the miRNA precursor backbone to replace the native miRNA. Such structure forms a fold similar to natural miRNA precursor structure and will be processed in the miRNA biogenesis, despite the absence of the amiRNA sequence in the wild type organism and therefore it is called "artificial". Artificial miRNA seems to be a promising technique for both functional genetic studies and plant antiviral protection or plant genetic improvements.

For the time being, there are several bioinformatics tools capable of designing amiRNA for specific organism or a group of organisms. As the biogenesis and the mode of action of animal and plant miRNAs are significantly different, separate tools have been developed for amiRNA design in animal and plant species. One of the mostly used tools dedicated to plants is web-based software Web MicroRNA Designer (WMD3) (Ossowski *et al.*, 2008, Schwab *et al.*, 2006). Recently, a novel web-based tool for automated design of the amiRNA and synthetic trans-acting siRNAs sequences has been developed (Carbonell *et al.*, 2014, Fahlgren *et al.*, 2015). Another tools such as miR-Synth (Laganà *et al.*, 2014), AmiR-zyn (Baby *et al.*, 2012) are designed for human amiRNA prediction only. Recent biological study of amiRNA designed using WMD3 software has revealed that the rules for creating effective amiRNA are not yet fully understood (Li *et al.*, 2013). Some amiRNA sequences determined by the WMD3 and designated as recommended have shown less silencing efficiency than those less recommended, which have reached the efficiency of 90%. Most available tools are based on sequence complementarity and other parameters such as calculation of the overall hybridization free energy of the miRNA/target

duplex or/and the miRNA precursor secondary structure (Yu *et al.*, 2014). However, statistical analysis of the thermodynamic profiles of published miRNA molecules has revealed that the specific thermodynamic features play a crucial role in determining the functionality of these molecules. It has been shown, that functional miRNAs exhibit distinguishable thermodynamic characteristics, different than non-functional miRNAs (Khvorova *et al.*, 2003, Krol *et al.*, 2004).

In this work we present a new method of plant amiRNA design. Our method takes into consideration free energy decomposition of known miRNA/miRNA* and miRNA/target duplexes. On this basis, the thermodynamic profiles are generated, to which the free energy decomposition for the newly designed corresponding duplexes is compared. Both the free energy value and its trend at each position of the designed amiRNA is taken into the consideration.

Here, using our AmiRNA Designer, we generated a thermodynamic profile for miRNA/target and miRNA/miRNA* for *Arabidopsis thaliana*. Based on the obtained data we designed two independent sets of amiRNA candidates to silence two exemplary *A. thaliana* genes. The designed sequences of amiRNA and amiRNA* were analyzed with our Perl scripts for the off-target effect.

The AmiRNA Designer provides a flexible tool to semi-automated design of amiRNA for the organism of interest. The user can apply either pre-defined parameters, established based on the provided thermodynamic profiles for *A. thaliana,* or generate new profiles for their own set of miRNAs and create customized constrains that can be used during designing of amiRNA.

MATERIALS AND METHODS

Data sources. Sequences for miRNA precursors were downloaded from miRBase rel18. (Griffiths-Jones *et al.*, 2008) (truncated to an organism of interest — *A. thaliana* — in our case). Sequences of miRNA and their targets were taken from the ASRP database (Gustafson *et al.*, 2005). For the off-target analysis we used a file from PlantGDB (Dong *et al.*, 2004, Duvick *et al.*, 2008) as a source of all expressed genes. Detailed information about the input data formats are provided in the Software Manual (http://www.cs.put.poznan.pl/arybarczyk/AmiRNA/).

The target gene sequences were downloaded from TAIR database (http://www.arabidopsis.org/).

RNA secondary structure prediction. To predict the RNA secondary structure and calculate the free energy (ΔG — expressed in kcal/mol) decomposition of generated structures, we applied the UNAFold software package (Markham & Zuker, 2008). UNAFold version 3.8 was used with the following parameters: energy rules set to RNA(2.3) (Walter *et al.*, 1994) and temperature fixed to 22°C. Calculations for miRNA/target duplexes and miRNA precursors were performed using the following UNAFold scripts: hybrid-2s.pl (for duplexes) and hybrid-ss-2s.pl (for single sequences).

RNA thermodynamic profiling. Based on the free energy decomposition obtained from the secondary structure prediction, the following thermodynamic profiles were calculated (in the same manner): miRNA/target and miRNA/miRNA*.

The profiles were generated by assigning ΔG values to each nucleotide position in the miRNA sequence. The ΔG values were assigned based on the free energy decomposition provided by UNAFold for each of the ana-

Analysis stage:
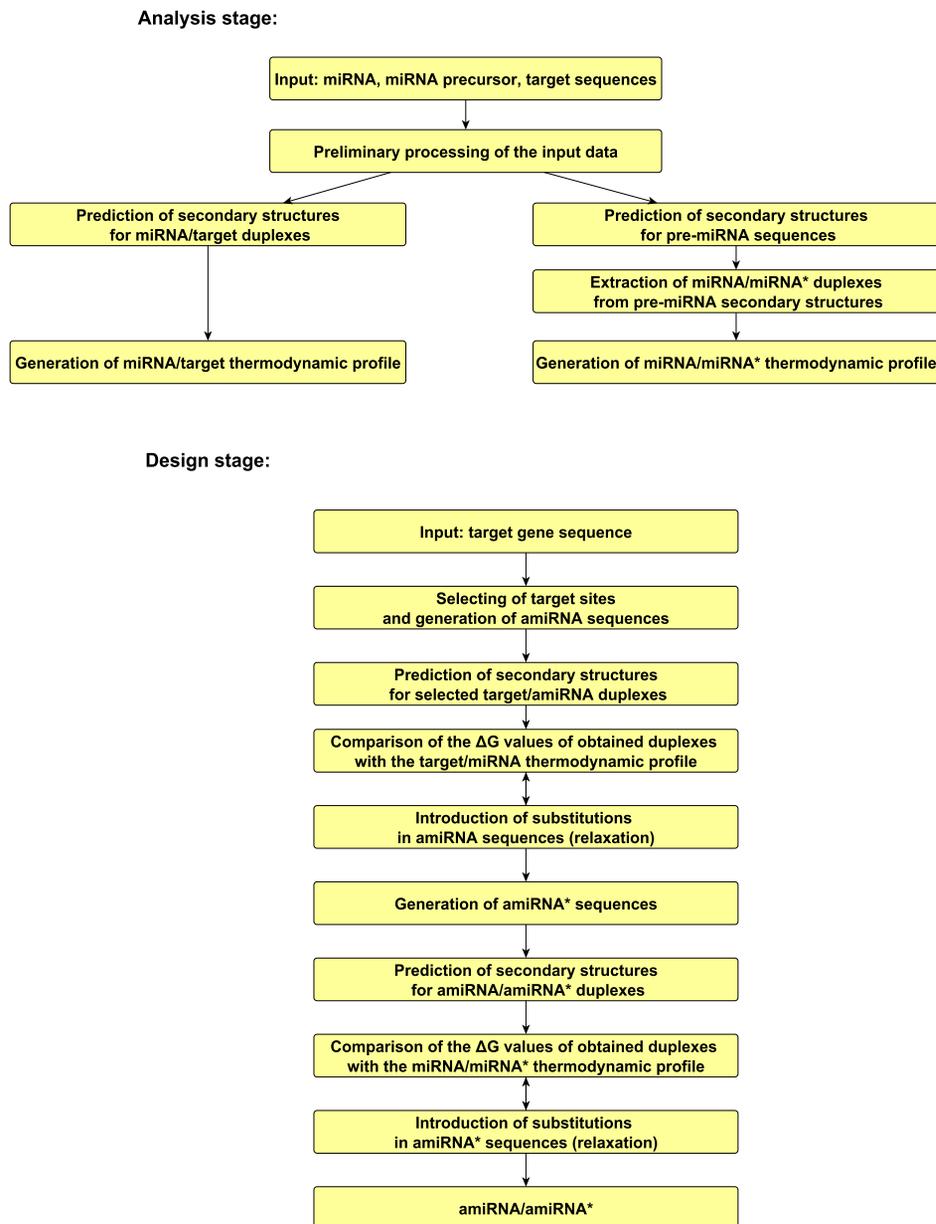
Design stage:

Figure 1. AmiRNA Designer workflow.
The scheme presents the most important steps of the analysis performed in AmiRNA Designer. It is divided into the following stages: "Analysis stage" and "Design stage" which can be performed independently. As a result of the first stage the thermodynamic profiles are created. The second step allows to design amiRNA and amiRNA* sequences basing on the results obtained in the "Analysis stage".

lysed duplexes (details of the profile generation are in Supplementary Materials). Next, statistical analysis of the thermodynamic data obtained for each miRNA was performed. The final profile provides free energy distribution associated with each miRNA position in miRNA/target and miRNA/miRNA* duplexes, respectively. The thermodynamic profiles are automatically created by AmiRNA Designer for the input data provided by the user and the results are saved as a statistic sheet in the csv format for each profile respectively.

Relaxation rules. During amiRNA design stage, it is checked whether the changes of the ΔG values at the individual positions of the amiRNA candidate show the same overall trend as the medians in natural miRNA/target thermodynamic profile. If a median value at a given position "i" in the miRNA/target profile is low-

er than at the position "i-1"(median decreases), the actual value of ΔG at the corresponding position "i" in amiRNA (ΔG_i) should be lower or equal to the median value at the position "i-1" in the miRNA/target profile and also lower than ΔG_{i-1}. The amiRNA sequences that do not satisfy these conditions at any position are immediately rejected. Otherwise, if median value at the given position "i" in the profile is higher than at the position "i-1" (median increases), the value of the ΔG at the corresponding position "i" in amiRNA should be greater or equal to the median value at the position "i-1" in the profile and the $\Delta G_i > \Delta G_{i-1}$. In this case, if the ΔG_i at the given position in amiRNA is in allowed region but its value is lower than required (the second condition is not fulfilled — the binding is too strong as compared to the profile), the nucleotide at this position can be
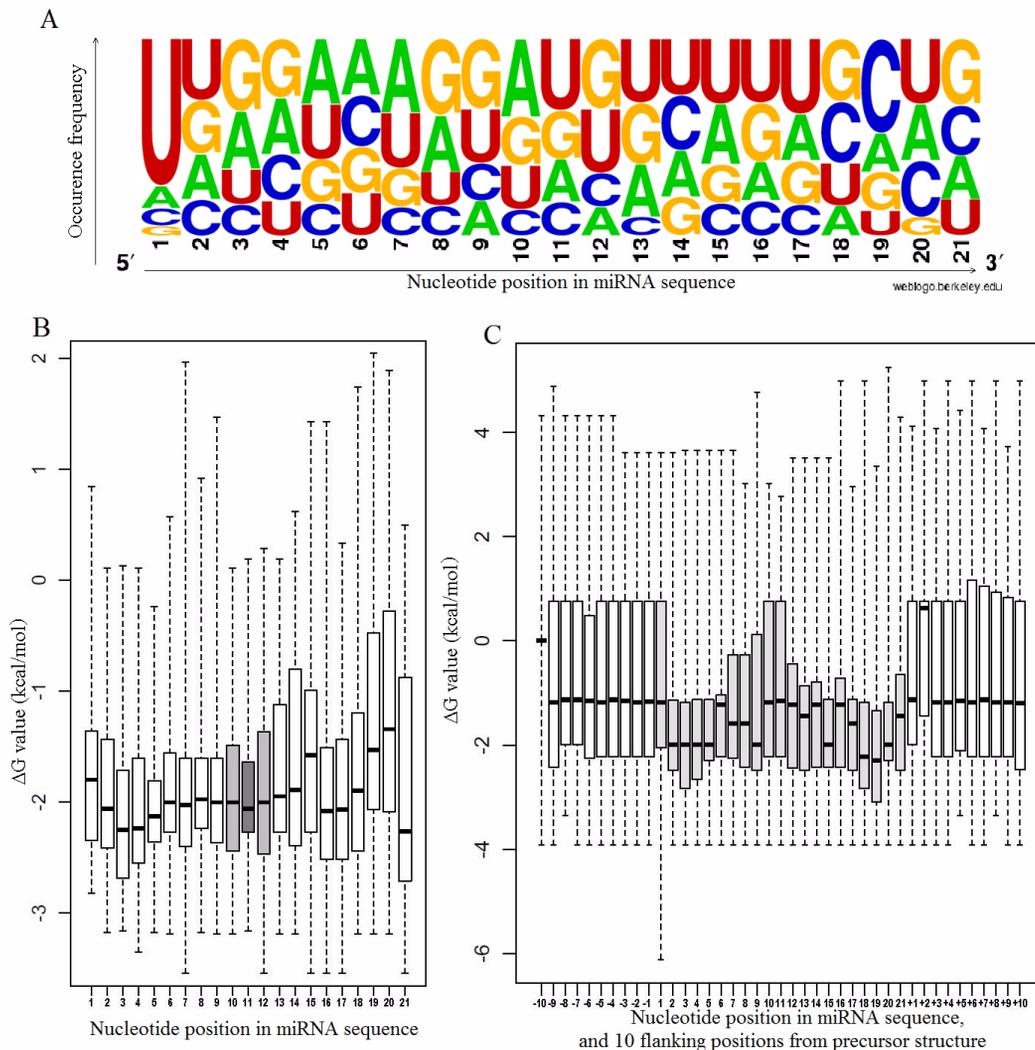
Figure 2. Analysis of *Arabidopsis thaliana* miRNA sequences.
(**A**). The statistical occurrence of nucleotides at the specific miRNA positions for a set of 21 nt miRNAs from *A. thaliana* (analysis prepared with http://weblogo.berkeley.edu/). (**B, C**). Thermodynamic profiles for *A. thaliana* generated by AmiRNA Designer. (**B**). The thermodynamic profile of miRNA/target duplex. (**C**). The thermodynamic profile of miRNA/miRNA* duplex. The bottom and the top of the box correspond to the lower and upper quartiles. Inside the box the median is located. Ends of the whiskers refer to the minimal and maximal values.

substituted. We have developed the term "relaxation" to describe this situation. Relaxation allows introducing non-canonical GU pairs, which are often found in RNA secondary structures (Mallory *et al.*, 2004). Exchange of the canonical base pair for the GU pair destabilizes the RNA duplex less than the introduction of mismatched nucleotide. Thus, the allowed substitutions in the amiRNA sequence are the following: A → G and C → U which result in GU and UG pairs in the amiRNA/target duplex, respectively. If the substitution in amiRNA at the given position is not possible then the next base pair is checked to determine whether the relaxation can be introduced. The substitutions are not allowed at the positions 10–12. Both the number of the allowed GU pairs (default 3) and the minimal distance between them (default 5) are user-defined.

Computational analysis of the off-targets. The results obtained with AmiRNA Designer were analysed for the off-target effect, using Perl scripts written for this purpose. The method to determine the direct interaction between the mRNA and amiRNA (or amiRNA*) is based on the Perl regular expressions. For each miRNA

and miRNA* the search for the binding site was performed along the entire transcripts with the shift of one nucleotide. We defined (in case of the Perl script) the following rules for the amiRNA/target binding: only one mismatch could occur in the nucleotides 3–9 and only two non-consecutive mismatches were allowed in the nucleotides 13–21 of the miRNA sequence.

RESULTS AND DISCUSSION

The AmiRNA Designer is a tool dedicated for generating artificial miRNAs based on the features of cellular native miRNAs. It consists of two stages: (i) analysis of the endogenous miRNAs and deriving parameters for filtering of the potential amiRNA, (ii) designing the amiRNA (and amiRNA*) and filtering using the parameters established in the previous stage. The workflow of AmiRNA Designer is shown in Fig 1. Each of the predicted amiRNA (and amiRNA*) was examined to determine whether it is likely to bind to any of the *A. thaliana* transcripts, to detect potential off-target interaction, with additional Perl scripts.

Table 1. The sequences of amiRNA and amiRNA* designed using AmiRNA Designer, where the relaxation rules have been applied.

id amiRNA	seq amiRNA	id amiRNA*	seq amiRNA*
amiR_Imm_410_**5U_18G**	UGUC**U**CUGAAGAGUACG**G**CAA	antyamiR_Imm_410_5U_18G_**15G**	UUG**CC**GUACUCUUC**GG**AGACA
amiR_Imm_410_**5U**	UGUC**U**CUGAAGAGUACGACAA	antyamiR_Imm_410_5U_**15G**	UUGUCGUACUCUUC**GG**AGACA
amiR_Imm_410_**18G**	UGUCCCUGAAGAGUACG**G**CAA	antyamiR_Imm_410_18G_**15G**	UUG**C**CGUACUCUUC**G**GGGACA

Substitutions introduced during the design stage in the amiRNA sequence are marked in red while those introduced in the amiRNA* sequence are marked in blue. Both sequences are presented in the 5'-3' direction. The name of the amiRNA sequence consists of: the target gene alias name (e.g. Imm), start position of the amiRNA/target binding site and the list of substitutions inserted during the design stage separated by underscore.

Analysis stage

157 sequences of 21 nt miRNA from *A. thaliana* were used to calculate nucleotide distribution. The obtained results showed that in over 70% of all miRNAs the first nucleotide in the sequence is U and in almost 40% the tenth is A (Fig. 2A, Fig. 1S and Table S1 at www.actabp.pl). Moreover, in the position 13 the C and in the position 20 the G is underrepresented (about 10%). Based on the obtained statistical data, we proposed some sequence characteristics that can be considered in artificial miRNA design. AmiRNA Designer uses filter parameters for the target sites, where the specific residues correspond to the *reverse complements of preferred* residues in the miRNA sequence (target site position = 22 — miRNA position). The above observations for miRNAs from *A. thaliana* allowed us to set up the following filter parameters for potential target sites: A at the position 21, U at the position 12, not G at the position 9 and not C at the position 2. We proposed the filter for A at the position 10 in amiRNA (that corresponds to U at the position 12 at the target site), which was additionally motivated by the observation that between the residue 10 and 11 of the miRNA a site-specific cleavage within the target RNA strand occurs. These parameters are optional and the user can decide which of them will be applied.

In the next step, two thermodynamic profiles were obtained: one for the miRNA/target duplex and second for the miRNA/miRNA*. These profiles showed the distribution of the ΔG value assigned to each position of miRNA, graphically depicted as boxplots. Calculated thermodynamic profiles served as the reference for filtering of the newly designed amiRNA sequences in the next stage.

From the 570 pairs of miRNA and their targets available in the ASRP database, 210 unique 21 nt sequences were used to generate the miRNA/target thermodynamic profile (Fig. 2B). The analyzed set of miRNA/target interactions contained both validated and computationally predicted pairs of miRNA/targets for *A. thaliana*, where the computational prediction was based on the homology with experimentally validated miRNA/mRNA pairs for another plants. There were several distinctive features of this profile. The most significant seems to be the central region of miRNA, between residue 10 and 11, where the AGO cleavage occurs in the mRNA strand. In this region the thermodynamic profile showed the local minimum of the ΔG value which implies the strongest base pairing at the cleavage site. There were also two other local minima, associated with positions 3–4 and positions 16–17 of miRNA, respectively. Moreover, our miRNA/target thermodynamic profile for *A. thaliana* revealed a discrepancy between the "5' half" and the "3' half" of miRNA. In case of "5' half" the median value was equal to –2.04 kcal/mol, whereas the corresponding value for the "3' half" was –1.83 kcal/mol, which can suggest a week analogy to the animal seed region, where about a

7 nt fragment from the miRNA 5' end showed almost perfect pairing whereas the rest showed weak complementarity (Lim *et al.*, 2003).

For 157 miRNA precursor sequences, secondary structures were predicted, miRNA/miRNA* pairs with 10 flaking bases from both sides were extracted and the miRNA/miRNA* thermodynamic profile was created. The obtained profile of *A. thaliana* miRNA/miRNA* (Fig. 2C) showed higher free energy associated with the first position of miRNA than with the last position, which is consistent with previously described features of animal miRNA precursors (Khvorova *et al.*, 2003; Krol *et al.*, 2004). In the central region, the local maximum occurred, which is in contradiction with the miRNA/target thermodynamic profile. We also observed two local minima: the first in the "5' half" of miRNA and the second one in the "3' half". These features, derived from the miRNA/miRNA* thermodynamic profile, were implemented in the design stage of amiRNA*.

Design stage

In the first step of amiRNA design stage, the sequence of the target gene (mRNA) to be silenced was fragmented into 21 nucleotide-long (or 20 nt) parts with a shift of one nucleotide. These 21 nucleotide-long sequences were considered as potential target sites for amiRNA. Then, for each of the potential target site, reverse complementary sequence was generated, to obtain a primary amiRNA/target duplex. The number of generated amiRNA/target duplexes can be reduced by applying sequence filters obtained in the analysis stage. Usually, amiRNA candidates are designed to have U at position 1. To achieve this, the user can choose the filter for A at the position 21 of the target, or reduce potential targets sites to the length of 20 nt and add the A nucleotide at the end, that introduces U at the first position of amiRNA (U-force option). In principle, the program works without any sequence filter. The U-force option additionally increases the number of potential amiRNA by allowing the mismatch at the first position of amiRNA/target. Next, the secondary structure of amiRNA/target duplex is predicted to obtain free energy decomposition, and recalculated according to the rules for the profile generation. The results are compared to the thermodynamic profile of miRNA/target generated previously. There are mainly two prerequisites that have to be fulfilled for amiRNA sequence to be accepted. First, ΔG values at each position in the given amiRNA sequence have to fit in the range of the lower and the upper quartile at corresponding positions in miRNA/target thermodynamic profile. The amiRNA sequences that do not satisfy this condition are immediately rejected. Next, it is checked whether the changes in the ΔG values at the individual positions in amiRNA show the same overall trend as the medians in miRNA/target thermodynamic profile. At this stage, if at some positions the ΔG

is too low (the biding is too strong), relaxation can be applied by introducing GU base pairs. When all necessary and permitted relaxations have been introduced, the thermodynamic profile of the modified amiRNA/target duplex is recalculated and compared with the reference miRNA/target profile. If the above mentioned prerequisites are fulfilled, then the modified amiRNA is accepted. In the next step, the amiRNA* sequence is designed. It is done similarly as in the case of amiRNA design. First, a reverse complementary sequence for amiRNA is created. Next, the thermodynamic profile for individual amiRNA/amiRNA* duplex is calculated and compared with the previously generated miRNA/miRNA* thermodynamic profile. The rules of the relaxations are the same as previously described and at this step only the amiRNA* sequence is modified whereas the amiRNA sequence is set as fixed. Finally, if the thermodynamic profile for the amiRNA/amiRNA* duplex fulfils all the defined rules, the amiRNA* sequence is accepted. The duplex amiRNA/amiRNA* is the final result of the design process.

The high complementarity to the target site makes amiRNA sequences almost unique. However, the introduction of substitutions in the amiRNA and amiRNA* sequence (that result in GU and UG base pairs), requires checking whether such modified amiRNA is able to silence genes other than the gene of interest. Our Perl scripts search all designed amiRNA and amiRNA* sequences against the off-target genes. As a result of this analysis, the user obtains the file that contains a list of artificial miRNA candidates that bind to the genes other than the target gene, together with their annotations (hypothetical off-targets), target sequences and mismatched positions.

Examples

As an application of AmiRNA Designer we generated amiRNA candidates for two target genes: AGB1 (ID: AT4G34460, alias names: ATAGB1, ELK4, ERECTA-LIKE 4, GTP BINDING PROTEIN BETA 1), and Immutans (ID: AT4G22260, alias names: IM, IM1, IMMUTANS, PLASTID TERMINAL OXIDASE, PTOX). The analysis performed with psRNATarget (Dai & Zhao, 2011) software showed the lack of the natural miRNAs for these genes. This analysis ensured us that the set of miRNA sequences used to create the thermodynamic reference profile did not contain the miRNA directed for the studied sequences. The design procedure was performed with the following parameters: the U residue on the position 12 within target site and the U forced option. In the case of ABG1 target, AmiRNA Designer revealed 22 amiRNA/amiRNA* duplexes whereas in the case of Immutans 20 amiRNA/amiRNA* candidate duplexes were generated. Exemplary miRNA/miRNA* sequences are shown in Table 1 (all designed sequences are presented in Table 2S and 4S at www.actabp.pl). Analysis of the off-target effect for AGB1 showed that only one amiRNA had a potential off-target effect on hypothetical protein (gi:110740355, gb:ak230271) and none of amiRNA* was complement to any of *A. thaliana* mRNA sequences (Table 3S at www.actabp.pl). In case of Immutans, there were at least 3 amiRNA/amiRNA* duplexes with no hypothetical off-target effect, however, the majority of amiRNA sequences were complementary to the hypothetical, putative or unknown proteins (Table 5S at www.actabp.pl). A list of the potential off-targets and miRNA mismatch positions for a given target gene were presented to the user. For the additional

verification of our software, we applied psRNATarget server (Dai *et al.*, 2011) to find targets for the designed amiRNA sequences. Using default parameters, the server generated the list of potential targets for our amiRNA candidates but the originally launched target sequence had significantly better scores than the remaining genes.

CONCLUSIONS

AmiRNA-mediated gene silencing has emerged as an effective and promising technique and became one of the most important tools widely explored in genetic engineering. By using dedicated programs, amiRNA sequences to target genes of interest can be designed, but there are still many problems concerning their efficiency and specificity that need to be solved.

We have demonstrated a novel, efficient approach for the semi-automated design of artificial microRNAs, which we have implemented as AmiRNA Designer software. The aim of the AmiRNA Designer is to design the amiRNA sequences for the target gene of interest that exhibit similar thermodynamic properties as the native ones, to increase the probability that the designed sequences will be processed such as those naturally occurring and will be target-specific. The thermodynamic properties of the native miRNA molecules are calculated based on the secondary structures of the input sequences. Based on the performed analysis the customized constraints are created for the design stage. The off-target silencing is verified using the algorithm implemented as Perl scripts that filters out the amiRNA candidates targeting the unwanted genes.

REFERENCES

Axtell MJ, Westholm JO, Lai EC (2011) Vive la différence: biogenesis and evolution of microRNAs in plants and animals. *Genome Biology* **12**: 221. doi: 10.1186/gb-2011-12-4-221.

Baby J, Vrundha N, Chandy S (2012) AmiRzyn: PERL Centered Artificial MicroRNA Designing Aid. *Int Res J Biol Sci* **1**: 18–23.

Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. *Cell* **136**: 215–233. doi: 10.1016/j.cell.2009.01.002.

Blazewicz J, Figlerowicz M, Kasprzak M, Nowacka M, Rybarczyk A (2011) RNA partial degradation problem: motivation, complexity, algorithm. *J Comp Biol* **18**: 821–834. doi: 10.1089/cmb.2010.0153.

Bologna NG, Schapire AL, Zhai J, Chorostecki U, Boisbouvier J, Meyers BC, Palatnik JF (2013) Multiple RNA recognition patterns during microRNA biogenesis in plants. *Genome Res* **23**: 1675–1689. doi: 10.1101/gr.153387.112.

Brodersen P, Voinnet O (2009) Revisiting the principles of microRNA target recognition and mode of action. *Nature Rev Mol Cell Biol* **10**: 141–148. doi: 10.1038/nrm2619.

Carbonell A, Takeda A, Fahlgren N, Johnson SC, Cuperus JT, Carrington JC (2014) New generation of artificial microRNA and synthetic trans-acting small interfering RNA vectors for efficient gene silencing in *Arabidopsis*. *Plant Physiol* **165**: 15-29. doi: 10.1104/pp.113.234989.

Carthew RW, Sontheimer EJ (2009) Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**: 642–655. doi: 10.1016/j.cell.2009.01.035.

Chau BL, Lee KA (2007) Function and anatomy of plant siRNA pools derived from hairpin transgenes. *Plant Methods* **3**: 13.

Chekulaeva M, Filipowicz W (2009) Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. *Curr Opin Cell Biol* **21**: 452–460. doi: 10.1016/j.ceb.2009.04.009.

Dai X, Zhao PX (2011) psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res* **39**: W155–W159. doi: 10.1093/nar/gkr319.

Dong Q, Schlueter SD, Brendel V (2004) PlantGDB, plant genome database and analysis tools. *Nucleic Acids Res* **32**: D354–D359.

Duvick J, Fu A, Muppirala U, Sabharwal M, Wilkerson MD, Lawrence CJ, Lushbough C, Brendel V (2008) PlantGDB: a resource for comparative plant genomics. *Nucleic Acids Res* **36**: D959–D965.

Eamens AL, Smith NA, Curtin SJ, Wang MB, Waterhouse PM (2009) The Arabidopsis thaliana double-stranded RNA binding protein DRB1 directs guide strand selection from microRNA duplexes. *RNA* **15**: 2219–2235. doi: 10.1261/rna.1646909.

Friedman RC, Farh KK-H, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19**: 92–105. doi: 10.1101/gr.082701.108.

Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miR-Base: tools for microRNA genomics. *Nucleic Acids Res* **36**: D154–D158.

Gustafson AM, Allen E, Givan S, Smith D, Carrington JC, Kasschau KD (2005) ASRP: the *Arabidopsis* small RNA project database. *Nucleic Acids Res* **33**: D637–D640.

Fahlgren N., Hill ST, Carrington JC, Carbonell A (2015) P-SAMS: a web suite for plant artificial microRNA and synthetic trans-acting small interfering RNA. *Bioinformatics* **32**: 157–158. doi: 10.1093/bioinformatics/btv534.

Han J, Lee Y, Yeom K-H, Nam J-W, Heo I, Rhee J-K, Sohn SY, Cho Y, Zhang B-T, Kim VN (2006) Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125**: 887–901.

Khvorova A, Reynolds A, Jayasena SD (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**: 209–216.

Kim JY, Kwak KJ, Jung HJ, Lee HJ, Kang H (2010) MicroRNA402 affects seed germination of *Arabidopsis thaliana* under stress conditions via targeting DEMETER-LIKE Protein3 mRNA. *Plant Cell Physiol* **51**: 1079–1083. doi: 10.1093/pcp/pcq072.

Kim VN (2005a) MicroRNA biogenesis: coordinated cropping and dicing. *Nature Rev Mol Cell Biol* **6**: 376–385.

Kim VN (2005b) Small RNAs: classification, biogenesis, and function. *Mol Cells* **19**: 1–15.

Krol J, Sobczak K, Wilczynska U, Drath M, Jasinska A, Kaczynska D, Krzyzosiak WJ (2004) Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. *J Biol Chem* **279**: 42230–42239.

Kurihara Y, Watanabe Y (2004) *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. *Proc Natl Acad Sci USA* **101**: 12753–12758.

Laganà A, Acunzo M, Romano G, Pulvirenti A, Veneziano D, Cascione L, Giugno R, Gasparini P, Shasha D, Ferro A *et al* (2014) miR-Synth: a computational resource for the design of multi-site multi-target synthetic miRNAs. *Nucleic Acids Res* **42**: 5416–5425. doi: 10.1093/nar/gku202.

Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN (2004) MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**: 4051–4060. doi: 10.1038/sj.emboj.7600385.

Li J-F, Chung HS, Niu Y, Bush J, McCormack M, Sheen J (2013) Comprehensive protein-based artificial microRNA screens for effective gene silencing in plants. *Plant Cell* **25**: 1507–1522. doi: 10.1105/tpc.113.112235.

Lim LP, Lau NC, Weinstein EG, Abdelhakim A, Yekta S, Rhoades MW, Burge CB, Bartel DP (2003) The microRNAs of Caenorhabditis elegans. *Genes Dev* **17**: 991–1008.

Mallory AC, Vaucheret H (2004) MicroRNAs: something important between the genes. *Curr Opin Plant Biol* **7**: 120–125.

Mallory AC, Reinhart BJ, Jones-Rhoades MW, Tang G, Zamore PD, Barton MK, Bartel DP (2004) MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *EMBO J* **23**: 3356–3364.

Markham NR, Zuker M (2008) UNAFold: software for nucleic acid folding and hybridization. *Bioinformatics* **453**: 3–31. doi: 10.1007/978-1-60327-429-6_1.

Meola N, Gennarino VA, Banfi S (2009) microRNAs and genetic diseases. *Pathogenetics* **2**: 7. doi: 10.1186/1755-8417-2-7.

Nowacka M, Jackowiak P, Rybarczyk A, Magacz T, Strozycki PM, Barciszewski J, Figlerowicz M (2012) 2D-PAGE as an effective method of RNA degradome analysis. *Mol Biol Reports* **39**: 139–146. doi:10.1007/s11033-011-0718-1.

Obernosterer G, Meister G, Poy MN, Kuras A (2006) The impact of small RNAs. Microsymposium on Small RNAs. *EMBO Rep* **8**: 23–27. doi: 10.1038/sj.embor.7400874.

Ossowski S, Schwab R, Weigel D (2008) Gene silencing in plants using artificial microRNAs and other small RNAs. *Plant J* **53**: 674–690. doi: 10.1111/j.1365-313X.2007.03328.x.

Petri S, Meister G (2013) siRNA design principles and off-target effects. *Methods Mol Biol* **986**: 59–71. doi: 10.1007/978-1-62703-311-4_4.

Rybarczyk A, Szostak N, Antczak M, Zok T, Popenda M, Adamiak R, Blazewicz J, Szachniuk M (2015) New *in silico* approach to assessing RNA secondary structures with non-canonical base pairs. *BMC Bioinformatics* **16**: 276. doi: 10.1186/s12859-015-0718-6.

Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D (2006) Highly specific gene silencing by artificial microRNAs in Arabidopsis. *Plant Cell* **18**: 1121–1133. doi: 10.1105/tpc.105.039834.

Sunkar Ramanjulu, Li Y-F, Jagadeeswaran G (2012) Functions of microRNAs in plant stress responses. *Trends Plant Sci* **17**: 196–203. doi: 10.1016/j.tplants.2012.01.010.

Sunkar R, Zhu JK (2004) Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell* **16**: 2001–2019.

Szostak N, Royo F, Rybarczyk A, Szachniuk M, Blazewicz J, del Sol A, Falcon-Perez JM (2014) Sorting signal targeting mRNA into hepatic extracellular vesicles. *RNA Biol* **11**: 836–844. doi: 10.4161/rna.29305.

Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev* **20**: 515-24.

Walter AE, Turner DH, Kim J, Lyttle MH, Müller P, Mathews DH, Zuker M (1994) Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc Natl Acad Sci* **91**: 9218–9222.

Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC (2005) Expression of Arabidopsis MIRNA genes. *Plant Physiol* **138**: 2145–2154. doi: http://dx.doi.org/10.1104/pp.105.062943.

Yu S, Pilot G (2014) Testing the efficiency of plant artificial microRNAs by transient expression in *Nicotiana benthamiana* reveals additional action at the translational level. *Front Plant Sci* **5**: 622. doi: 10.3389/fpls.2014.00622.