Target site effects in the RNA interference and microRNA pathways

Gregor Obernosterer*1, Hakim Tafer† and Javier Martinez*

*Institute of Molecular Biotechnology (IMBA), Austrian Academy of Sciences, Dr. Bohr-Gasse 3, 1030 Vienna, Austria, and †Institute of Theoretical Biochemistry (TBI), University of Vienna, Währingerstrasse 17, 1090 Vienna, Austria

Abstract

In RNAi (RNA interference), siRNAs (small interfering RNAs) are loaded into the RISC (RNA-induced silencing complex), which then mediates endonucleolytic cleavage of complementary target RNAs. Although RNAi has become one of the most powerful tools in molecular biology to assess gene function, there remains a great number of ineffective siRNAs. It is already known that the assembly and activation of RISC is a crucial determinant of RNAi activity, but downstream effects such as target accessibility have not been analysed extensively. Therefore we assessed the effect of target site accessibility and found that it significantly improves the potency of siRNAs. Similarly, miRNAs (microRNAs) act by repressing protein synthesis through imperfect base-pairing to the 3′-UTR (untranslated region) of target mRNAs. We found that predicted target sites reside in regions of high accessibility and tested whether this criterion could be used in the search of functional miRNA targets. In addition, we performed reporter gene assays to test whether accessibility correlates with measured mRNA suppression levels. The results of our initial study suggest that secondary structures might add a so far underrepresented layer of complexity in the recognition of RNA targets by miRNAs.

Introduction

Over the past few years RNAi (RNA interference) has become an essential tool to investigate gene function by introducing siRNAs (small interfering RNAs) or shRNAs (short-hairpin RNAs) targeting homologous RNAs in cells [1–3]. During the activation of RISC (RNA-induced silencing complex), siRNAs are unwound and one of the strands assembles into RISC [4,5]. The small RNA molecule then acts as a template for the recognition and cleavage of cognate RNAs in the final step of the RNAi pathway [6,7]. Because of the wide applicability of RNAi in functional genomics, as well as the possibility to exploit it for therapeutics and biotechnology, consistent efforts have been made to design potent and specific siRNAs. It was shown that different siRNAs targeting the same gene greatly vary in their efficiency [8]. Insights from biochemical, structural and bioinformatics studies lead to the development of several RNAi design principles to improve the efficiency of siRNAs. Besides the exploration of various sequence-specific rules [9], the asymmetry criterion especially greatly improved the design of functional siRNAs [10,11]. The asymmetry rule determines which of the strands of a siRNA duplex is loaded into RISC by assessing either the type of base-pairs or the interaction energy of the terminal four bases.

As we know from in vivo and in vitro studies that RISC is unable to unfold and cleave structured RNA [12], we set out to investigate whether target site accessibility as the very downstream event in RNAi could be used as a design rule in creating siRNAs. Using experimental and computational methods entailing more than 900 siRNAs for 42 genes, we were able to demonstrate that taking accessibility into consideration significantly improves the design of siRNAs [13].

Another class of small RNAs that mediate post-transcriptional gene silencing by imperfectly binding to the 3′-UTR (untranslated region) of complementary mRNAs are miRNAs (microRNAs) [14–16]. It was previously shown that for a miRNA to be functional, nucleotides 2–7/8, the so-called ′seed′, must perfectly base-pair to the target mRNA [17]. It was suggested that the seed region alone is not the only reliable criterion for the identification of functional target sites [18–21]. Recent studies showed that the structure of the target RNA plays a crucial role in miRNA target recognition [22–24]. As target site accessibility plays an important role for the efficiency of siRNAs, we decided to investigate if this is also true for miRNAs. Using a reporter gene assay, we could indeed observe an effect on the miRNA-mediated target repression depending on the thermodynamics of the mRNA–miRNA interaction.

Results and discussion

It is well understood how RISC cleaves complementary miRNAs [25–27]. However, it was not known how RISC finds its target RNAs. Biochemical studies revealed that the accessibility of the target site correlates with the efficiency of cleavage. Using human purified RISC [28], Ameres et al. [12] have shown by in vitro cleavage assays that efficient...

Key words: cleavage, microRNA (miRNA), RNA-induced silencing complex (RISC), RNA cleavage, small interfering RNA (siRNA), target site accessibility.

Abbreviations used: miRNA, microRNA; RISC, RNA-induced silencing complex; RNAi, RNA interference; siRNA, small interfering RNA; UTR, untranslated region.

1To whom correspondence should be addressed (email gregor.obernosterer@imba.oeaw.ac.at).
target RNA cleavage is dependent on nucleotides 2–15 of the siRNA. They could also demonstrate that the 5′ part of the siRNA is essential for the initial target recognition. In order to investigate whether target site accessibility could be used as a design criterion to improve the efficiency of siRNAs and to study target RNA effects, we performed a comprehensive in silico study using the RNAplfold program [13]. Instead of assessing the MFE (minimum free energy structure) of mRNAs the RNAplfold program computes local base-pair probabilities and accessibilities within target RNAs [29]. This method differs from other approaches to evaluate target site accessibility as it considers the complete Boltzmann ensemble of all possible structures [30]. To test if target site accessibility can be used as a distinguishing feature for siRNAs of different silencing capacities, we applied the RNAplfold program to an extensive set of validated siRNAs. The targeting regions, 3′-UTRs and coding sequences of ~700 siRNAs were arbitrarily chosen without any rational design, giving a broad spectrum of silencing efficiencies. We generated a subset that contained only those siRNAs that resulted in either a clear target repression or did not show any silencing effects. This set of approx. 570 siRNAs was then used for optimizing the folding parameters of the RNAplfold program. We observed a significant separation between functional and non-functional siRNAs and found that the silencing efficiency correlated well with target site accessibility.

Which regions within the siRNA contribute to the separation of functional and non-functional siRNAs? We found that the six to eight consecutive nucleotides starting at the 3′-end of the target, which correspond to the ‘seed’ region within miRNAs, have to be accessible for efficient annealing. In addition, a second region was identified that contributed significantly to the separation of functional and non-functional siRNAs. This is in line with biochemical data, which demonstrated that a minimal stretch of ~16 nt must be accessible for efficient target RNA cleavage [12].

We also studied the influence of G/C content on RNAi since it was shown that G/C-rich sequences promote the formation of closed structures [9]. siRNAs were grouped according to their G/C content and the effect on accessibility was determined. For all G/C classes, even for G/C content >60%, we could significantly separate functional and non-functional siRNAs.

Clearly, target site accessibility is not the main criterion used in designing efficient siRNAs. It is the asymmetry, which is required for the correct strand selection, that most significantly improves the design of potent siRNAs [10,11]. To assess if accessibility is an applicable criterion, we compared it to sequence-specific rules, asymmetry and upstream effects (e.g. self-complementarity of siRNAs [31]). Amongst all the optimized criteria, only the asymmetry criterion performed better than accessibility. The stronger correlation might reflect the nature of the RNAi pathway where asymmetry acts upstream of target site accessibility effects. It also became evident when we compared each criterion that none of the parameters alone would be sufficient to design effective siRNAs. Therefore we combined accessibility, asymmetry and two additional rules describing the self-folding tendency of siRNAs into a freely available siRNA design tool called RNAxs (http://rna.tbi.univie.ac.at/cgi-bin/RNAxs). Importantly, all criteria used reflect important events in the RNAi pathway which affect distinct stages during siRNA maturation.

We validated the performance of RNAs in comparison with other design tools on an independent siRNA dataset consisting of 360 siRNAs whose silencing efficiencies were experimentally verified. The self-folding criteria performed poorly, whereas asymmetry and accessibility improved the siRNA selection significantly. In summary, more than 90% of the chosen siRNAs were functional giving >50% mRNA reduction and every third siRNA reduced gene expression by >90% [13].

miRNAs regulate gene expression in mammals by imperfect base-pairing to the 3′-UTR of target mRNAs, thereby mediating either target degradation or translational repression [32]. They control important events during development, differentiation, proliferation, and their deregulation leads to severe diseases such as cancer [33–36]. Experimental validation of target sites, mostly done in tissue-culture-based reporter gene assays, is difficult since miRNAs form intricate networks targeting more than a hundred 3′-UTRs [37,38]. The development of computational approaches such as PicTar or TargetScan improved the identification of functional target sites [39,40]. In essence, all of those tools rely on the evolutionary conservation of target sites containing ‘seed regions’. However, there is a constraint in that conservation cannot be applied to non-conserved miRNAs [41]. Some target prediction programs, e.g. RNAhybrid, exclusively compute the free energy between the miRNA and the target RNA [42].

Recently, it was shown that the incorporation of context determinants such as A/U (or G/C) content reliably improves the identification of specific sites [18]. Long et al. [23] proposed a structure-based model by combining known features of canonical miRNA target sites such as seed pairing with a two-step hybridization reaction. First, nucleation at the accessible target site takes place followed by miRNA annealing to disrupt closed secondary structures in order to establish a stable miRNA-target duplex. Kertesz et al. [22] have shown that prior to the binding of miRNAs any intramolecular base-pairings should be removed. They used conventional RNA folding algorithms to compute the energy cost necessary to remove local secondary structures within the target site. The total miRNA-binding site interaction energy is therefore the sum of the opening energy and the hybridization energy.

We assessed whether RNA sequences complementary to seed regions of known human miRNAs (mirBase release 11.0 [43]) are more accessible than expected. Using RNAplfold, we calculated the opening energy of the reverse complementary seed region and compared this energy with shuffled dinucleotide sequences. Importantly, the seed region was not shuffled, which ensures (i) that the location of the seed within the 3′-UTR remains the same as in the shuffled sequences, avoiding any border effects, (ii) the number of miRNA binding sites stays the same, and (iii) the G/C content is not altered, which might affect accessibility. This comparison was not...
limited to conserved seeds only. We found a clear difference in the shape of the distributions for the opening energy between miRNA seeds in real 3′-UTRs and seeds where the surrounding 3′-UTR sequences were randomized (Figure 1). Real miRNA seed sites have a significantly lower opening energy and are therefore more accessible than expected by chance (with a relative enrichment of 17.8%). This result indicates that enhanced accessibility is a feature of miRNA target sites and that it could help separate functional from non-functional target sites. The observation that miRNA targets reside in regions of higher accessibility is in line with other studies [22].

To experimentally test if accessibility affects miRNA-mediated translational repression, we cloned three let-7 binding sites behind a luciferase reporter gene [44]. This original construct was then modified, according to the RNAplfold, to contain either highly accessible or non-accessible target sites. In order to alter the accessibility, the surrounding regions around the target sequences were mutated. After transfection of the reporter construct into HeLa cells, we measured the repression efficiency mediated by endogenous let-7. Translational repression was almost completely abolished for the inaccessible construct, but remained unchanged for the accessible one (Figure 2). The accessibility and the energy cost to remove intramolecular structures in this small-scale study might explain the difference in the measured repression.

Various studies undermined the importance of the seed region to identify possible miRNA binding sites. Consideration of evolutionary conservation and context specific criteria significantly enhanced the specificity in miRNA target finding [18,19,39]. The addition of accessibility as suggested by Kertesz et al. [22] might improve our understanding of how miRNAs recognize their target sites [24]. It will be interesting to investigate whether folding kinetics, i.e. the formation, dissociation and shifting of base-pairs, might influence target recognition and translational repression [45]. Another layer of complexity for the formation of stable miRNA-binding duplexes might be the presence of RNA-binding proteins. It was shown that the binding of proteins to miRNA target sites affects miRNA binding efficiency [46].

In the first part of our study, we were able to demonstrate that accessibility is an important criterion in the siRNA design process. Accessibility combined with asymmetry and other non-sequence specific rules significantly improved the design of effective siRNAs. Optimizations of these criteria lead to the development of a siRNA design tool called RNAxs. In addition, we tried to assess if accessibility plays a role in miRNA target finding. We found miRNA seeds preferentially reside in regions of higher accessibility and initial cell-culture-based reporter gene assays indicate that accessibility of miRNA target sites influences translational repression.

H.T. is supported by Siemens and the Wiener Wissenschafts-, Technologie-, und Forschungsfonds. J.M. is Junior Group Leader at IMBA, the Institute of Molecular Biotechnology of the Austrian Academy of Sciences. We thank Philipp Leuschner for helpful discussions and comments on the manuscript.

References