

Brief reports

Open Access

## Analysis by *siRNA\_profile* program displays novel thermodynamic characteristics of highly functional siRNA molecules

Pirkko Muhonen\*<sup>1</sup>, Ranga N Parthasarathy<sup>2</sup>, Anthony J Janckila<sup>2</sup>, Kalman G Büki<sup>1</sup> and H Kalervo Väänänen<sup>1</sup>

Address: <sup>1</sup>Institute of Biomedicine, Department of Anatomy, University of Turku, Kiinamylynkatu 10, FIN-20520 Turku, Finland and <sup>2</sup>Veterans Affairs Medical Center, and University of Louisville, Louisville, KY, USA

Email: Pirkko Muhonen\* - [pirkko.muhonen@utu.fi](mailto:pirkko.muhonen@utu.fi); Ranga N Parthasarathy - [ranga.parthasarathy@med.va.gov](mailto:ranga.parthasarathy@med.va.gov); Anthony J Janckila - [anthony.janckila@med.va.gov](mailto:anthony.janckila@med.va.gov); Kalman G Büki - [kalbuk@utu.fi](mailto:kalbuk@utu.fi); H Kalervo Väänänen - [kalervo.vaananen@utu.fi](mailto:kalervo.vaananen@utu.fi)

\* Corresponding author

Published: 21 May 2008

Received: 13 August 2007

Source Code for Biology and Medicine 2008, 3:8 doi:10.1186/1751-0473-3-8

Accepted: 21 May 2008

This article is available from: <http://www.scfbm.org/content/3/1/8>

© 2008 Muhonen et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Objective:** Here we report the improved results of a new siRNA design program and analysis tool called *siRNA\_profile* that reveals an additional criterion for bioinformatic search of highly functional siRNA sequences.

**Methods:** We retrospectively analysed over 2400 siRNA sequences from 34 genes and with known efficacies to categorize factors that differentiate highly, moderately and non-functional siRNA sequences in more detail. We tested the biological relevance of *siRNA\_profile* in CHO cells stably expressing human TRACP.

**Results:** The highly functional siRNA molecules exhibited lower overall stabilities than non-functional siRNAs after taking into consideration all the nucleotides from 5'-terminus to the 3'-terminus along the siRNA molecule, in addition to the 5'-section of the antisense strand and the region between 9–14 nucleotides as previously has been acknowledged. Comparison of the *siRNA\_profile* program to five other programs resulted in a wide range of selected siRNA sequences with diverse gene silencing capacities, even when the target was only 197 nucleotides long. Six siRNA design programs selected 24 different siRNA sequences, and only 6 of them were selected by two or more programs. The other 18 sequences were individually selected by these six programs.

**Conclusion:** Low general stability of dsRNA plays a significant role in the RNAi pathway and is a recommended criterion to consider, in addition to 5'-instability, internal instability, nucleotide preferences and target mRNA position, when designing highly efficient siRNAs.

### Introduction

RNA interference (RNAi) is a gene silencing mechanism where short interfering RNA (siRNAs) and microRNA (miRNAs) molecules inhibit the transcription and translation of target genes in a sequence-specific manner [1-3].

siRNAs are exogenously produced ~21 nucleotides long, double stranded RNA molecules with complete complementarity to the target sequence. miRNAs are a family of endogenously encoded small non-coding RNAs, derived by processing of short RNA hairpins, that can inhibit the

translation of mRNAs bearing partial complementarity to the target sequences. RNAi has been acknowledged as a practical tool for new drug target discovery and RNAi drug development in mammalian cells [4]. Therefore designing highly functional siRNA molecules has become an essential part of RNAi methodology. We have developed a novel and user-friendly siRNA design algorithm *siRNA\_profile* with multiple options for minimizing the identification of non-functional, unspecific and immunostimulatory siRNA molecules. The analysis of functional and non-functional siRNA molecules were done by the *siRNA\_profile* program to demonstrate in more detail the characteristics of highly functional siRNA molecules to help scientists in their search for theoretically and biologically efficient siRNAs.

### Conventional siRNA design criteria

The rapid development of RNAi applications has revealed the need for efficient and specific siRNA design and analysis tools to maximize the efficiency while minimizing possible side-effects [5]. Computational methods and neural networks are tools approaching ideal siRNA design; however, so far none of them are perfect. Currently, thermodynamic characteristics of functional siRNA molecules guide siRNA design strategy. First, it has been shown that thermodynamic differences in the base-pairing stabilities of the 5'-ends of both siRNA and miRNA molecules play a critical role in determining which strand initiates RNA induced silencing complex (RISC) activation [6,7]. To achieve an efficient RNAi effect, activated RISC should be able to silence multiple copies of the target mRNA. A second criterion for effective siRNAs is low internal stability in the cleavage region between 9 to 14 nucleotides (calculated from the 5'-terminus) of the antisense strand. This is believed to have a critical role in mRNA cleavage and it may also help to release RISC from the cleaved target [7]. Third, there are findings of the nucleotide preferences over the length of siRNA sequence [8,9].

### siRNA\_profile program design

The *siRNA\_profile* program and a full text including additional data and a printable help page are available in the *siRNA\_profile* program web page [10]. Briefly, the *siRNA\_profile* program is based on findings of the asymmetric differences between functional and non-functional siRNAs [6,7,11] and on our studies of positional nucleotide differences and average dsRNA stability along the siRNA antisense strand. We have also incorporated some recommendations of Elbashir, S.M., et al. [2] and developed a novel, interactive and user-friendly siRNA design algorithm with multiple options for minimizing unspecific siRNA design. Previously, it has been recognized that 5'-UGUGU-3' motifs have immunostimulatory potential in synthetic siRNA molecules, in addition to CpG motifs

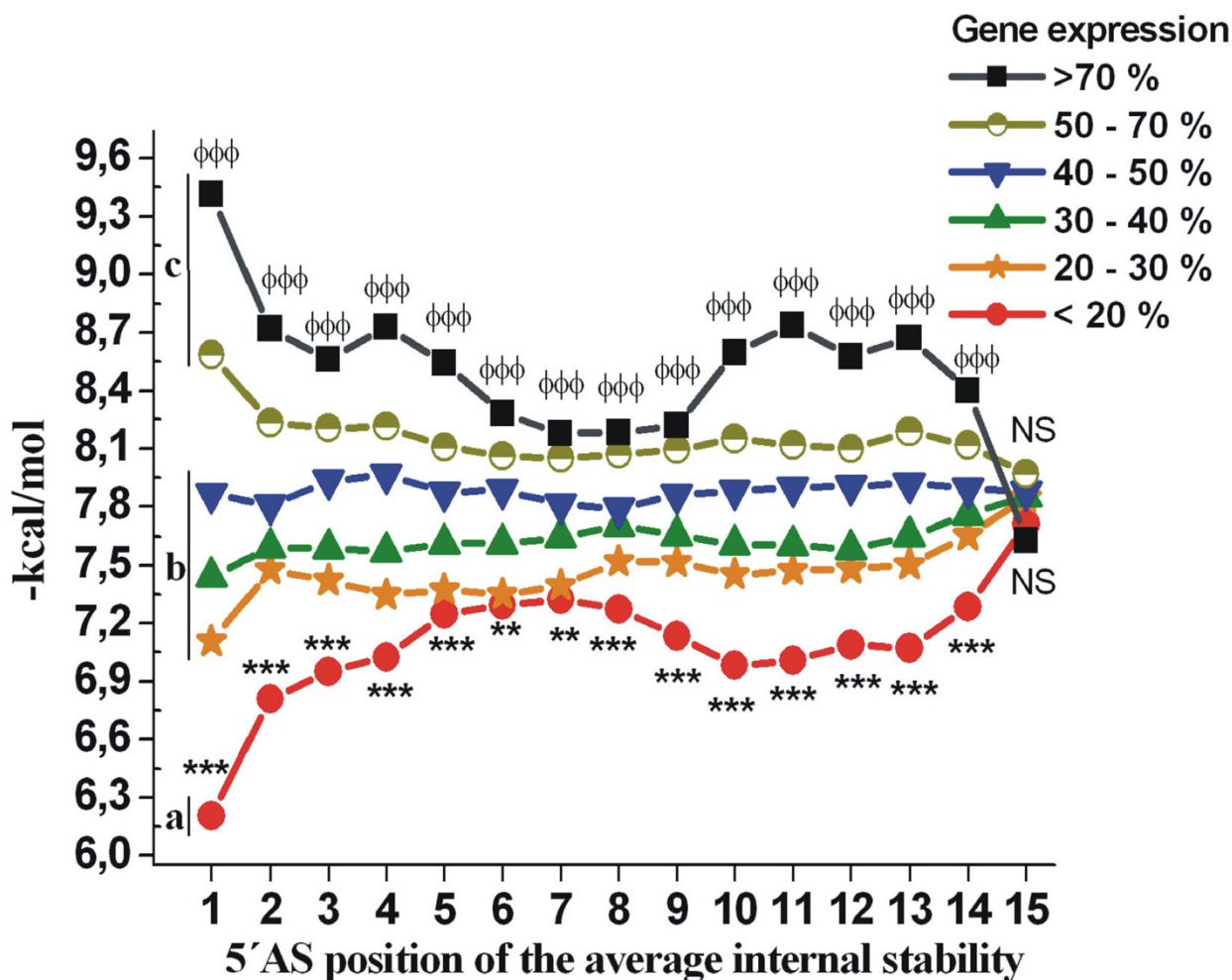
in single stranded RNAs and DNA oligomers [12-14]. We have incorporated a sensor able to recognize immunostimulatory motifs intending to avoid unnecessary false phenotypes by siRNA molecules and, additionally, by anti-miRNA oligomers. In addition, a scoring system was integrated into the *siRNA\_profile* program. It was adjusted based on our findings on nucleotide, purine and pyrimidine distribution along functional and non-functional siRNA sequences [the scoring system described in *siRNA\_profile* web page].

The *siRNA\_profile* program uses free energy values for calculation of average internal stability profiles. The average internal stability profiles were calculated as the sum of stability in five nucleotide windows by using the nearest-neighbour method as previously described [7,11]. Our program browses and calculates the target sequence from the antisense point of view, from the 5' - to the 3' - direction. However, it does not utilize nucleotides on the mRNA beyond the 3' - end of the siRNA as previously has been described [7], because this calculation method may easily lead to false bending of the profile. This unique and beneficial feature of the *siRNA\_profile* program allows the calculation of the factual, not the false dsRNA thermodynamic characters.

The *siRNA\_profile* program was originally developed on Red Hat Linux 9.0 and compiled to an executable file. The algorithm was written in C-programming language and the CGI application was written in Perl programming language. The application uses the user-given input values from the *siRNA\_profile* website to analyze and display the results.

### Findings

In order to understand the similarities in biologically effective siRNA sequences, previously published siRNA data were used [15]. 2431 siRNA sequences from 34 genes were fed into the *siRNA\_profile* program and siRNA profiles were divided into categories based on gene expression levels after siRNA silencing (gene expression level (GE) categories:  $\leq 20\%$ ,  $n = 280$ ;  $20.1\% - 30\%$ ,  $n = 455$ ;  $30.1\% - 40\%$ ,  $n = 445$ ;  $40.1\% - 50\%$ ,  $n = 405$ ;  $50.1\% - 70\%$ ,  $n = 736$  and  $>70\%$ ,  $n = 110$ ). The functional and non-functional siRNA sequences exhibited "the mirror energy profiles" (Figure 1). Low stability in the 5'-end of the antisense strand was a critical factor in determining the correct strand activating RISC and further catalysing mRNA degradation. This asymmetry property is in good agreement with previous results [6,7]. In addition, our results showed that highly functional siRNAs exhibited lower overall internal stabilities than non-functional siRNAs (Figure 1). The low overall stability properties of the most effective siRNAs were apparent at all the nucleotides from 5'-terminus to the 3'-terminus along the siRNA molecule

**Figure 1**

**siRNA stability analysis by *siRNA\_profile* and novel guidelines for highly efficient siRNA design.** The average energy profiles of 2431 siRNA sequences [15] were calculated and analyzed by *siRNA\_profile* program. The average energy profiles of siRNA sequences were divided into categories according to their efficiencies (GE categories:  $\leq 20\%$ , 20.1%–30%, 30.1%–40%, 40.1%–50%, 50.1%–70% and  $\leq 70\%$ ). The siRNA efficiency showed siRNA instability dependency: highly functional siRNA molecules (GE  $\leq 20\%$ , group a) illustrated lower overall stability in addition to asymmetric average energy profile than moderately (GE 20.1 – 50%, group b) or non-functional (GE  $>50\%$ , group c) siRNA molecules. Statistical analyses were performed with one-way ANOVA: (\*) average internal stabilities of highly functional siRNAs (a) compared to moderately functional ones (b);  $\Phi$  average internal stabilities of moderately functional siRNAs (b) compared to non-functional ones (c). \*\* $p < 0.01$ ; \*\*\*/ $\Phi\Phi\Phi p < 0.001$ .

excluding the last one. Noticeably, there was no position along the length of functional and non-functional siRNAs at which energy profiles crossed. This was true for both 5'-region and the region from nucleotides 9–14. These novel features were found to correlate siRNA energy profiles with their efficiencies. To demonstrate the differences of average internal stabilities of highly functional, moderately functional and non-functional siRNAs in more detail, these siRNAs were compared to each other. Average

internal stability of highly functional siRNAs (Figure 1, GE  $<20\%$ , marked as group a) showed significantly lower internal stability compared to moderately functional siRNAs (GE 20.1–50%, marked as group b). Non-functional siRNAs (GE 50–100%, marked as group c) exhibited significantly higher average internal stability than moderately functional ones. Here, our results obtained by *siRNA\_profile* program suggest that the thermodynamic characteristics focusing only on the 5'-end stability prop-

**Table 1: Comparison of the *siRNA\_profile* program and five other *siRNA* design programs**

|                       | <i>siRNA_profile</i> | Deqor | Ambion | siDesign | EMBOSS | SiDirect |
|-----------------------|----------------------|-------|--------|----------|--------|----------|
| <b>Number of hits</b> | 6 (4)                | 4(1)  | 3(0)   | 2(2)     | 6(4)   | 9(7)     |
| <b>Average GE%*</b>   | 10.4                 | 14.8  | 21.1   | 13.0     | 21.6   | 14.8     |
| <b>Median%*</b>       | 10.2                 | 13.1  | 26.9   | 13.0     | 16.1   | 14.9     |
| <b>Minimum GE%*</b>   | 5.4                  | 10.3  | 8.1    | 12.6     | 6.7    | 5.8      |
| <b>Maximum GE%*</b>   | 15.1                 | 22.8  | 28.3   | 13.3     | 54.1   | 32.2     |

197 nt human CyclophilinB sequence and data of known *siRNA* efficacies [8] were used for *siRNA* design program comparison. Number of hits that were not selected by any of five other programs is shown in brackets (). \* Gene expression (GE) % of the baseline

erties of the antisense strand are not enough to differentiate highly functional *siRNAs* from the moderately functional or non-functional ones.

To demonstrate the usefulness of *siRNA\_profile* program as a *siRNA* candidate search service, the biological efficacies of *siRNA* molecules targeting tartrate-resistant acid phosphatase (TRACP) experimentally designed by *siRNA\_profile* were validated in a Chinese Hamster Ovary (CHO) cell line stably over-expressing human TRACP under a CMV promoter [16] [data shown in *siRNA\_profile* web page].

The *siRNA\_profile* program functionality was compared to five other available *siRNA* design programs. Human CyclophilinB sequence region 193 – 390 [GenBank:M60857] [8] was fed into the *siRNA\_profile* program, Deqor [17], *siRNA* target finder (Ambion, USA), siDesign Center (Dharmacon, USA), EMBOSS explorer (gwilliam ° rfcgr.mrc.ac.uk) and siDirect [18], and optimal *siRNA* candidate search options were used. *siRNA* design results were compared to results obtained by *siRNA\_profile* by using the known efficacies of each *siRNA* sequences targeting the selected human Cyclophilin region (Table 1). The results showed that all tested programs selected functional *siRNA* sequences with significant variability. The novel *siRNA\_profile* program showed the highest selectivity. However, we noticed that all six programs tend to select different *siRNA* sequences even though the mRNA region selected for candidate search was only 197 nucleotides long. Programs selected altogether 24 different sequences, and only 6 of them were selected by two or more programs. The other 18 sequences were individually selected by these six programs. This result shows the dilemma, that even for programs that operate with related selection parameters, very different sequences were chosen with a range of *siRNA* efficacies.

In conclusion, our results suggest that low general stability of dsRNA plays a significant role in the RNAi pathway and is recommended to be used as an additional guideline for highly efficient *siRNA* design. Additionally, all known immunostimulatory motifs to date are highlighted in colour to prevent analysis of false phenotypes and the output

of the *siRNA\_profile* program displays the average internal energy profiles for each *siRNA* candidate. According to our results, the *siRNA* sequences exhibiting both favourable asymmetric properties and low overall stabilities according to Figure 1 are recommended for gene knock down experiments.

### Availability and requirements

*siRNA\_profile* program availability: <http://bonebiology.utu.fi/pimaki/main.html>

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

PM, KGB, HKV Study concept and design. PM, KGB Acquisition of the data. PM Analysis and interpretation of the data. AJJ, RNP CHO cell line stably expressing human tartrate-resistant acid phosphatase [additional data in the *siRNA\_profile* web page]. PM Drafting of the manuscript. HKV Study supervision. All authors read and approved the final manuscript.

### Acknowledgements

The authors thank Dr. Pentti Riikonen (Department of Information Technology, University of Turku), Anne Seppänen and Janne Harjunpää for providing valuable help during development of the *siRNA\_profile* algorithm and the web page application. The authors also thank Dr. Tiina Laitala-Leinonen (Bone Biology Research Consortium, University of Turku) for her assistance in formatting the text for publication.

### References

1. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T: **Identification of novel genes coding for small expressed RNAs.** *Science* 2001, **294(5543)**:853-858.
2. Elbashir SM, Lendeckel W, Tuschl T: **RNA interference is mediated by 21- and 22-nucleotide RNAs.** *Genes Dev* 2001, **15(2)**:188-200.
3. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC: **Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*.** *Nature* 1998, **391(6669)**:806-811.
4. Soutschek J, Akinc A, Bramlage B, Charisse K, Constien R, Donoghue M, Elbashir S, Geick A, Hadwiger P, Harborth J, John M, Kesavan V, Lavine G, Pandey RK, Racie T, Rajeev KG, Rohl I, Toudjarska I, Wang G, Wuschko S, Bumcrot D, Koteliensky V, Limmer S, Manoharan M, Vornlocher HP: **Therapeutic silencing of an endogenous gene by systemic administration of modified *siRNAs*.** *Nature* 2004, **432(7014)**:173-178.

5. Pei Y, Tuschl T: **On the art of identifying effective and specific siRNAs.** *Nat Methods* 2006, **3(9)**:670-676.
6. Schwarz DS, Hutvagner G, Du T, Xu Z, Aronin N, Zamore PD: **Asymmetry in the assembly of the RNAi enzyme complex.** *Cell* 2003, **115(2)**:199-208.
7. Khvorova A, Reynolds A, Jayasena SD: **Functional siRNAs and miRNAs exhibit strand bias.** *Cell* 2003, **115(2)**:209-216.
8. Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Khvorova A: **Rational siRNA design for RNA interference.** *Nat Biotechnol* 2004, **22(3)**:326-330.
9. Jagla B, Aulner N, Kelly PD, Song D, Volchuk A, Zatorski A, Shum D, Mayer T, De Angelis DA, Ouerfelli O, Rutishauser U, Rothman JE: **Sequence characteristics of functional siRNAs.** *Rna* 2005, **11(6)**:864-872.
10. **siRNA\_profile program** [<http://bonebiology.utu.fi/pimakil/main.html>].
11. Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MH, Neilson T, Turner DH: **Improved free-energy parameters for predictions of RNA duplex stability.** *Proc Natl Acad Sci U S A* 1986, **83(24)**:9373-9377.
12. Hemmi H, Takeuchi O, Kawai T, Kaisho T, Sato S, Sanjo H, Matsumoto M, Hoshino K, Wagner H, Takeda K, Akira S: **A Toll-like receptor recognizes bacterial DNA.** *Nature* 2000, **408(6813)**:740-745.
13. Judge AD, Sood V, Shaw JR, Fang D, McClintock K, MacLachlan I: **Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA.** *Nat Biotechnol* 2005, **23(4)**:457-462.
14. Sugiyama T, Gursel M, Takeshita F, Coban C, Conover J, Kaisho T, Akira S, Klinman DM, Ishii KJ: **CpG RNA: identification of novel single-stranded RNA that stimulates human CD14<sup>+</sup>CD11c<sup>+</sup> monocytes.** *J Immunol* 2005, **174(4)**:2273-2279.
15. Huesken D, Lange J, Mickanin C, Weiler J, Asselbergs F, Warner J, Meloon B, Engel S, Rosenberg A, Cohen D, Labow M, Reinhardt M, Natt F, Hall J: **Design of a genome-wide siRNA library using an artificial neural network.** *Nat Biotechnol* 2005, **23(8)**:995-1001.
16. Janckila AJ, Parthasarathy RN, Parthasarathy LK, Seelan RS, Yam LT: **Stable expression of human tartrate-resistant acid phosphatase isoforms by CHO cells.** *Clin Chim Acta* 2002, **326(1-2)**:113-122.
17. Henschel A, Buchholz F, Habermann B: **DEQOR: a web-based tool for the design and quality control of siRNAs.** *Nucleic Acids Res* 2004, **32(Web Server issue)**:W113-20.
18. Naito Y, Yamada T, Ui-Tei K, Morishita S, Saigo K: **siDirect: highly effective, target-specific siRNA design software for mammalian RNA interference.** *Nucleic Acids Res* 2004, **32(Web Server issue)**:W124-9.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

