In silico designing of siRNA Molecule

Deepak Kapoor and Harpreet Kaur

University Institute of Biotechnology (UIBT), Chandigarh University, Mohali, India.

RNA interference (RNAi) is one of the finest technologies to be developed in the functional genomics to determine the role of genes. The phenomenon of RNAi was first reported by Napoli and Jorgenson in 1990. It was then used by Fire and Mello in 1997 for functional genomics studies involving the nematode *Caenorhabditis elegans* which reveals that dsRNA to be injected containing both sense and antisense RNA could specifically target the endogenous RNA before translation resulting in genetic interference (Fire *et al,* 1998). In 1999, certainly the antisense RNA leading to degradation of the co-suppressed RNA was discovered which may be due to reduction of nucleotide pairs to 25 nucleotides (Hamilton and Baulcombe, 1999). The link between RNAi and small interfering RNA (siRNA) is further proved in *Drosophila melanogaster*. This leads to the prediction that the short sized RNA activates a ribosomal complex called RNA Induced Silencing Complex (RISC) which binds with the target RNA in RNAi (Hammond *et al,* 2000).

RNA interference is a process of Post translation gene silencing carried out by double stranded RNA (dsRNA) similar to the gene undergoing silencing. The duplex of 19-21 bases nucleotide with 2nt overhangs, also called siRNA is specifically involved in initiation of gene silencing which is generated by ribosome III cleavage of longer dsRNAs (Zamore *et al,* 2000). siRNA basically induces post transcriptional deletion of gene products and silencing of the gene expression (Fire *et al,* 1998) by involving two steps-

a) generation of RNA induced silencing complex (RISC) and

b) recognition/ binding to target mRNA (Dykxhoorn *et al,* 2003).

Different genes require different siRNAs to suppress their expression. In fact, a siRNA is composed of DNA sequence that is formed by the small sequence of mRNA of the target gene (Tijsterman and Plasterk, 2002). The length of mRNA sequence is very large which leads to huge number of siRNAs. These huge numbers of siRNA are further tested to get an effective siRNA as every small sequence of mRNA is not an effective siRNA which is time consuming as well as laborious processes (Castanotto and Rossi, 2009).

Mechanism of RNA interference

RNAi is a process which involves the mechanism in which when a dsRNA is introduced, the cell activates the two step reaction. In the first step, the ribonuclease (RNase) III enzyme also called Dicer acts upon double stranded RNA(dsRNA) and processed it into small interfering RNA (siRNA) which subsequently leads to direct cleavage of homologous mRNA by RISC complex.(Hannon, 2002) .

Similarly, the siRNA that can silence gene activity are called micro-RNAs (miRNAs). Like siRNA, these single stranded RNA complexes are also the result of Dicer activity but it will only silence the gene containing hair-pin loop like structures. (Carrington and Ambros, 2003)

Screening for Dicers complexes in both animals and plants are then considered in for RNA interfering mechanism to take place. In animals, a pair of Dicer complex is involved named Dcr-1 and Dcr-2 whereas in humans and *C. elegans*, single Dicer complex is involved named Dcr-2 (Lee *et al,* 2004). The majority of siRNAs are produced by Dcr-2 and microRNAs by Dcr-1 essential for miRNA induced silencing. RNase activity signals Dcr-2 to bind siRNA with Dcr-2 associated protein complex called R2D2 which helps in the binding of siRNA with RISC complex (Liu *et al,* 2003). The two intermediate complexes including an initiating complex having Dicer associated with siRNA results into a large hollow complex containing dicers and R2D2 and functional RISC complex which converts double stranded (ds) siRNA into single stranded (ss) siRNA complex. The substrate specificity of Dicers leads to the formation of two different RISC complexes: Dcr-2 assembles siRISC onto siRNA and Dcr-1 assembles miRISC from miRNA. Since, miRNAs acts the same as siRNA when perfectly matched mRNA is present which leads to replacement of miRNA and thus imposing a blockage for translation as siRNA is functionally similar to miRNA (Doench *et al,* 2003).

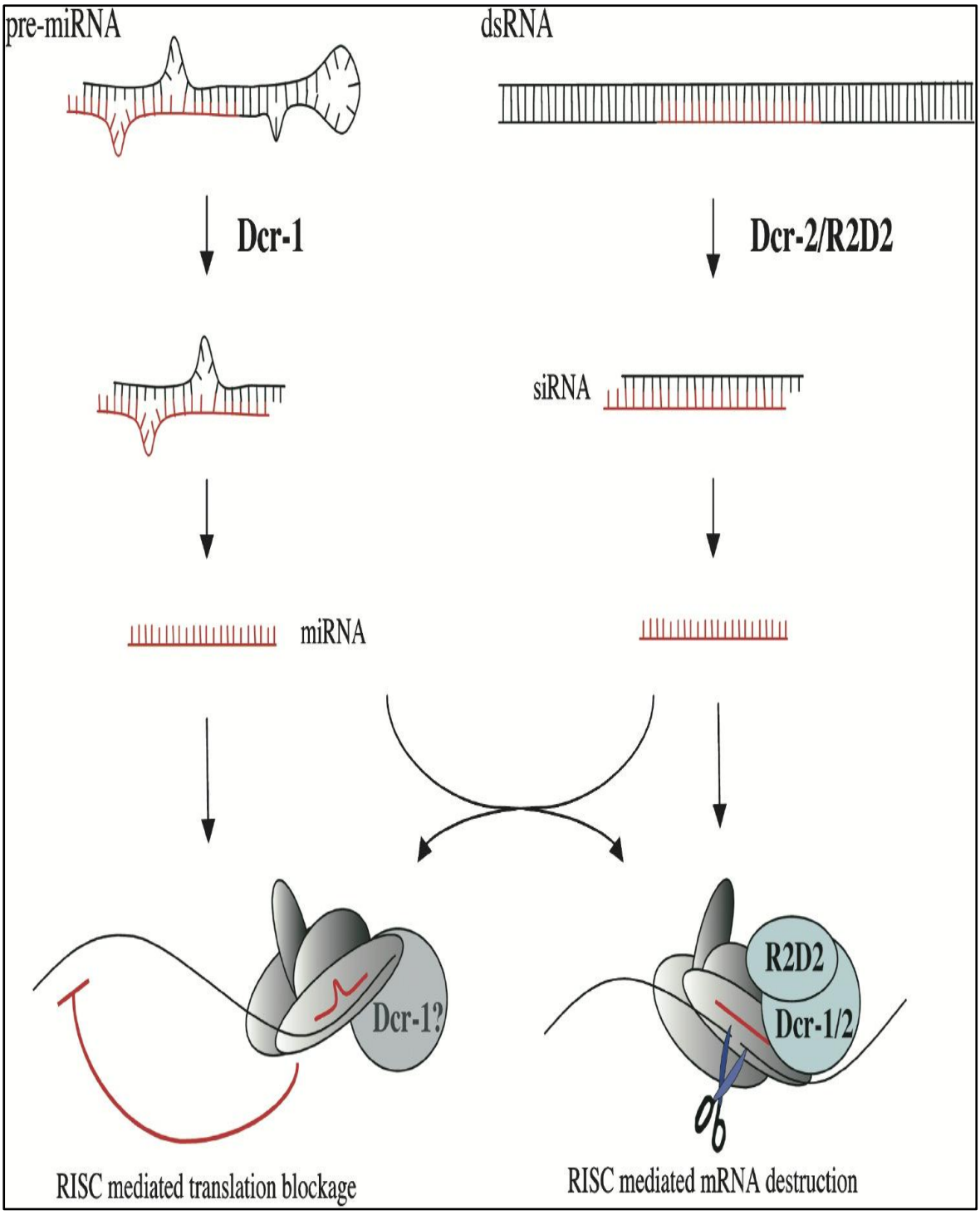

Figure1: Mechanism of RNA Silencing (Lee *et al,* 2004).

Rules for designing of an Efficient siRNA molecule

There are various guidelines/rules provided for efficient siRNA analysis, mainly by Ui-Tei, Amarzguioui and Reynolds. These rules include threshold score, GC content, Gibbs free energy, thermodynamics of free energy etc (Taxman *et al,* 2006). The rules are as follow:-

Ui-Tei Rules: - The efficiencies of siRNAs are exposed to be highly dependent on their sequences. The Ui-Tei rule is the empirical rule for the analysis of efficient siRNA. This criteria stress on 5' terminal nucleotides for determining siRNA efficiency. The efficient siRNA must satisfy the conditions simultaneously (Ui-Tei *et al,* 2004).

1) A or U at position 1 from 5' terminus of siRNA guide strand.
2) G or C at position 19
3) AU richness (AU≥4) in positions 1-7
4) No long GC stretch ≥10.

Amarzguioui Rules: - This rule set up some of the conditions that the 21 nucleotide pair must have to follow to be an efficient siRNA (Amarzguioui and Prydz, 2004). These conditions include:

1) Strong binding of 5' sense strand
2) No U at position 1
3) Presence of A at position 6
4) Weak binding of 3' sense strand
5) No G at position 19

Reynolds Rules: - In 2004, Reynold and coworkers formulated some rules that help in the designing & selection of efficient siRNA molecules. The rule is based on the scoring; if the molecule meets certain criteria then positive score is assigned and negative score is assigned only if G or C is present at 19^{th} position and G at 13^{th} position. The score of 6 or greater than 6 is considered to be optimal for the generation of efficient siRNA molecule.

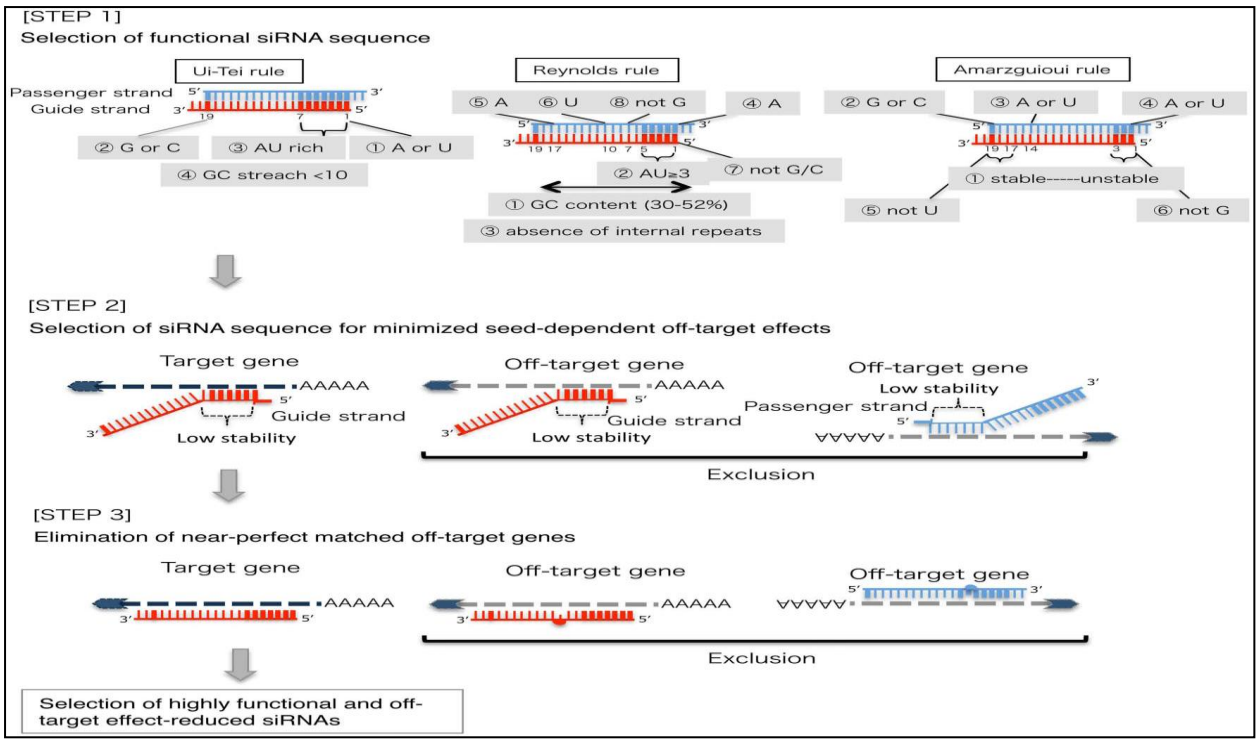

Figure 2: Representation of rules applied during siRNA design (Ui-Tei *et al,* 2004)

Other Set of Rules:- Other rules for selection and designing of siRNA includes

Tuschl Rules: - It is an early but still used attempt to design working siRNA duplexes leads to the following empirical rules. These rules include

1) Target sequence should contain 50% of GC content.
2) The target sequence complementary to mRNA must start from 50-100 nucleotide downstream
3) 23 nucleotide long target sequence with AA at position 19.

Hohjoh Rules: - These rules represent newly designed siRNA duplexes as Fork siRNA duplexes that can increase the RNAi activity.

1) The fork siRNA must mismatches at the 3' sense-strand.
2) G or C at position 1
3) A or U at position 8 and 19

Hsieh Rules: - The condition for siRNA to be efficient includes

1) No C at position 6

2) C or G at position 11
3) A at position 13
4) G at position 16
5) U at position 19
6) No G at position 19

Khvorova Rules: - This rule is based on the thermodynamic properties of sequences which are based on duplex unwinding and strand retention by RISC. (Gupta *et al,* 2010). These rules are

1) low stability at 3' end is less that 8.5 kcal/mol
2) low stability at 6-11 base pairs in strand greater than -7 kcal/mol

Takasaki rules (Elbashir *et al,* 2001): The rules include

1) No A/U at position 1
2) G at position 1 and 7
3) A at position 6 and 8
4) No U at position 7
5) No G at position 8 9 and 19
6) U at position 9 and 15

Stockholm rules (Chalk *et al,* 2004):

1) Total hairpin energy < 1
2) Antisense 5' end binding energy < 9
3) Sense 5' end binding energy in range 5-9 exclusive
4) GC content between 36% and 53%
5) Middle (7-12) binding energy < 13
6) 5' sense strand -5' antisense strand Binding Energy Difference < 0
7) Energy Difference within -1 and 0

.

References:

- Napoli, C., Lemieux, C. and Jorgensen, R. (1990) Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell.* **2(4):**279-289.

- Fire, A., Xu ,S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, **391(6669):**806–811.
- Hamilton, A.J. and Baulcombe, D.C. (1999) A Species of Small Antisense RNA in Posttranscriptional Gene Silencing in Plants. *Science,* **286(5441):**950-952.
- Hammond, S.M., Bernstein, E., Beach, D. and Hannon, G.J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature*, **404(6775):**293-296.
- Zamore, P.D., Tuschl, T., Sharp, P.A. and Bartel, D.P. (2000) RNAi: double stranded RNA directs the ATP-dependent cleavage of mRNA at 21-23 nucleotide intervals. *Cell*, 31;**101(1):**21-33.
- Dykxhoorn, D. M., Novina, C. D. and Sharp, P. A. (2003) Killing the messenger: short RNAs that silence gene expression. *Nat Mol Cell Biol* **4(6):**457-467.
- Castanotto, D. and Rossi, J.J. (2009) The promises and pitfalls of RNA-interference-based therapeutics. *Nature*, **457(7228):**426-433.
- Tijsterman, M. and Plasterk, R.H. (2002) Dicers at RISC: The mechanism of RNAi. Nature , 418, pp 244-251
- Carrington, J.C. and Ambros, V. (2003) Role of microRNAs in plant and animal development. *Science*, **301(5631):**336-338.
- Hannon, G.J. (2002) RNA interference. *Nature*, **418:**244-251.
- Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J. and Carthew, RW. (2004) Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell,* **117(1):**69-81.
- Liu, Q., Rand, T.A., Kalidas, S., Du, F., Kim, H.E., Smith, D.P. and Wang, X. (2003) R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. *Science,* **301:**1921–1925.
- Doench, J.G., Petersen, C.P. and Sharp, P.A. (2003) siRNAs can function as miRNAs. *Genes Dev,* **17:**438-442.
- Taxman, D.J., Livingstone, L.R., Zhang, J., Conti, B.J., Iocca, H.A., Williams, K.L., Lich, J.D., Ting, J.P. and Reed, W. (2006) Criteria for effective design, construction, and gene knockdown by shRNA vectors. *BMC Biotechnology,* **6:**7.
- Ui-Tei, K., Naito, Y., Takahashi, F., Haraguchi, T., Ohki-Hamazaki, H., Juni, A., Ueda, R. and Saigo, K. (2004) Guidelines for the selection of highly effective siRNA

sequences for mammalian and chick RNA interference. *Nucleic Acids Research*, **32(3):**936–948.

- Amarzguioui, M. and Prydz, H. (2004) An algorithm for selection of functional siRNA sequences. *Biochem Biophys Res Commun,* **316:**1050–1058.
- Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W.S. and Khvorova, A. (2004). Rational siRNA design for RNA interference. *Nat Biotechnol*, **22:**326–330.
- Hsieh, A.C., Bo, R., Manola, J., Vazquez, F., Bare, O., Khvorova, A., Scaringe, S. and Sellers, W.R. (2004) A library of siRNA duplexes targeting the phosphoinositide 3-kinase pathway: determinants of gene silencing for use in cell-based screens. *Nucleic Acids Res,* **32:**893–901.
- Gupta, S.K., Akhoon, B.A., Srivastava, M. and Gupta, S.K. (2010) A Novel Algorithm to Design an Efficient siRNA by Combining the Pre Proposed Rules of siRNA Designing. *J Comput Sci Syst Biol*, **3:**005-009.
- Elbashir, S.M., Lendeckel, W. and Tuschl, T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev,* **15:**188–200.
- Chalk, A.M., Wahlestedt, C. and Sonnhammer, E.L. (2004) Improved and automated prediction of effective siRNA. *Biochem Biophys Res Commun,* **319:**264–274.