Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 11 (2012) 139 – 148

Procedia Computer Science

Proceedings of the 3rd International Conference on Computational Systems-Biology and Bioinformatics (CSBio 2012)

Conceptual design of RNA-RNA interaction based devices

Jittrawan Thaiprasit [a], Supapon Cheevadhanarak[b], Dujduan Waraho[a], Asawin Meechai[a, c,*]

[a]Biological Engineering Program, Faculty of Engineering,
[b] *Department of Biotechnology, Faculty of Bioresources and Technology,* [c]*Department of Chemical Engineering, Faculty of Engineering, King Mongkut's University of Technology Thonburi, Bangkok, Thailand, 10140*

Abstract

A key goal of synthetic biology is to use biological molecules to create novel biological systems. Due to their role as transmitters in such systems, RNA molecules have gained much attention from synthetic biologists to design and construct novel RNA molecules with desirable functions and properties. In recent decades, the design of RNAs, however, has been limited to RNA architecture with primitive functions: aptamer and catalysis. To expand the paradigm of RNA-based design, we herein propose a conceptual design of RNA-RNA interaction based systems, considering domain-based structures of RNAs, as well as thermodynamic properties of RNA molecules and their interactions. Two evaluation scores, namely structural score (SS) and affinity score (AS), are used as criteria for selection of proper RNA sets. We employ this concept to design various RNA sets, each of which contains three RNA strands that altogether function like a comparator device. With these criteria, we show that four out of forty RNA sets would behave like a biological comparator since they have appropriate structure (SS = 1) and proper interaction order (AS > 1). The proposed scores are proven to be proper criteria for selection of RNA sets with required functions. This preliminary design offers an opportunity for synthetic biologists to expand the design of RNA sequence from a single strand to multiple strands that would behave in the same manner as enzymatic reactions.

Selection and/or peer-review under responsibility of the Program Committee of CSBio 2012.

Keywords: synthetic RNA, RNA-RNA interaction, thermodynamics, comparator

* Corresponding author. Tel.: +662-470-9221-34 Ext. 405; fax: +662-428-3534.
E-mail address: asawin.mee@kmutt.ac.th.

1877-0509 © 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.procs.2012.09.015

1. Introduction

In synthetic biology, molecular building blocks—i.e. DNA, RNA, and protein—have been used to make a number of biological parts and devices, including logic gates [1-3], oscillators [4-6], and sensors [7-9]. Within the next generation, synthetic biologists aim to integrate existing synthetic systems with more complex and advanced networks from an engineering perspective [10]. Since synthetic RNA devices can respond sensitively to alterations in the levels of endogenous signalling molecules, similar to a transmitter in an information processing system, they can potentially link signal-associated pathways to cell fate decisions. The integration of RNA transmitters with other devices would lead to more complex informative networks in a programmable cell. Thus, synthetic biologists have been attempting to construct newly designed RNAs to synthesize more complex cellular systems [11].

A crucial principle in RNA design is selecting the RNA sequence that will adopt a target secondary structure for a desired function [12]. Previously, synthetic RNAs have been selected with no prior information by primitive *in vitro* selection based methods [11]. With these methods, a set of random RNA sequences is used as an RNA pool that may contain many undesired RNAs. The target RNAs are isolated by either their binding or their catalytic activities, and are then amplified. These processes are repeated to isolate active RNAs. Later, the information about functional motifs and structures of RNAs have been accumulated [12-15], and several tools based on this information have been developed to help predict the secondary structure of RNAs before laboratory validation [16, 17]. The increasing availability of these tools has led to the development of various design approaches for synthesizing the RNA architecture with required functions. The parameters involving equilibrium stability of RNA structures, such as minimal free energy, and partition function are mainly utilized as selection criteria [12, 18, 19]. For example, the difference of the free energy between two states of RNA structures is efficiently used in the design of novel RNAs called allosteric ribozymes [20]. Allosteric ribozymes designed based on the equilibrium stability of RNA structures properly function both *in vitro* and *in vivo* [2, 20]. However, most RNA design approaches have focused on the design of a single RNA strand with primitive functions, such as catalysis and aptamer [11, 21]. There are few studies on the design of multiple RNA strands altogether interacting towards a given function. In 2004, Isaacs *et. al.* proposed a synthetic system of RNA-RNA interaction using a particular motif called **YUNR** (p**Y**rimidine-**U**racil-**N**ucleotide-pu**R**ine) for effective hybridization that has opened the way for the design of novel devices containing multiple RNA strands. In this system, there are two RNA strands, *cis*-repressed (*cr*RNA) and *trans*-acting RNAs (*ta*RNA), which can hybridize with each other. The *t*aRNA is transcribed only in the presence of an exogenous signal. Once transcribed, it then hybridizes with the existing *cr*RNA, triggering the conformational change of the *cr*RNA that in turn allows the translation of repressed genes. Because this work was an early study of synthesizing RNA-RNA interaction, the design of the multiple RNA strands was only based on the location of particular motifs and the structure of RNA strands. To provide a systematic design for an RNA-RNA interaction based system, it requires additional criteria for selecting proper RNAs that adopt the desired secondary structures, and behave as expected.

In this study, we have proposed a conceptual design for an RNA-RNA interaction based system, and attempted to create an RNA-based system that behaves as a biological comparator. The conceptual method consists of two main steps: the design of RNA strands from interacting domains and the selection of RNAs by two criteria, namely structural score (SS) and affinity score (AS). The proposed design is demonstrated using a comparator-like function as a model comprised of 3 pieces of RNA strands interacting in 2 different reactions. This design approach presented herein will enable the synthesis of new sets of multiple RNA strands with advanced functions.

2. The Strategy Of Conceptual Design For RNA-RNA Interaction Based System

We proposed a strategy of conceptual design for an RNA-RNA interacting based system comprised of two main steps: design of various sets of multiple RNAs from interacting domains and selection of proper RNA sets by two evaluation scores. The overall strategy is schematically shown in Fig. 1.

2.1 The specification of model reaction and interacting domains

The design strategy starts with the specification of a synthetic system of RNA-RNA interaction with desired functions and behaviors. In general, an RNA-RNA interaction involves the hybridization between an RNA strand with its cognate RNA (target strand) that subsequently causes either the conformational changes or the catalysis of such target strand. A starting model for RNA-RNA interaction based system can be described as a simple model of chemical reaction by

$$R_n + R_m \rightleftharpoons R_nR_m$$

where R_n and R_m are an RNA strand and its cognate RNA, respectively, and R_nR_m represents the hybridized RNA complex between the two strands.

To design a set of multiple RNA strands, all RNA strands in the desired system are first categorized into interacting domains. The domains are either units of hybridization between RNAs or particular RNA motifs. For example, the riboregulator of Isaacs *et al*. [22] have two RNA strands called *ta*RNA and *cr*RNA in such system. Fig. 2 shows a structure of the hybridized RNA between *ta*RNA and *cr*RNA that can be categorized into 4 domains consisting of 3 domains for RNA hybridization (domains 1, 3 and 4), and 1 domain for an RNA motif (domain 2).

2.2 Generation and concatenation of RNA sets based on interacting domains

In this study, the candidate RNA sequences are generated based on interacting domains by using the Domain Design (DD) software [23]. Specific regions or motifs, such as the ribosomal binding site (RBS), are indicated for no mutation in the program. Also, each domain is assigned the scores for mutation corresponding to the strength power of its binding. All RNA sets are concatenated into the initial designed RNA strands.

2.3 Determination of evaluation scores

To assess whether the generated RNA sets from the DD program are proper, we proposed two evaluation scores, namely SS and AS, that indicate the reaction sequence in a given RNA set. The available tools that help predict the secondary structures of both single stranded and hybridized RNAs are also employed to determine the evaluation scores. The consideration and calculation of the two scores is described as follows.

Structural score (SS)

SS represents the ratio of the number of required characters found in a designed RNA set to the total number of required characters, as described in Equation (1). The required characters such as loop, bulge, and closed motif of an RNA strand or a hybridized strand can be observed from their predicted secondary structures using available tools e.g. NUPACK [17], and the Vienna RNA package [24].

$$\text{Structural score (SS)} = \frac{\text{Number of required characters in designed RNA set}}{\text{Total number of required characters}} \quad (1)$$

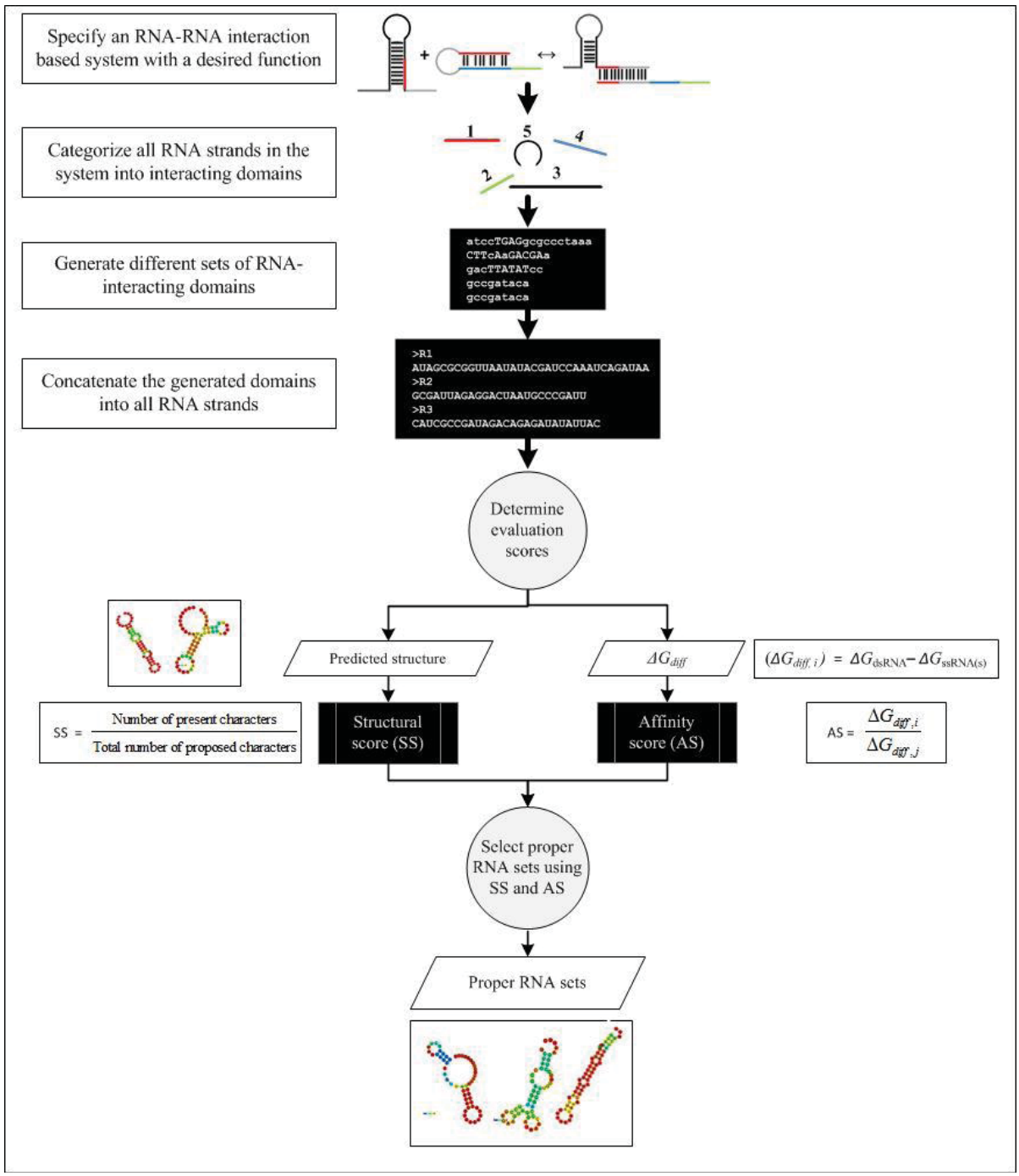


Fig. 1 The flowchart for conceptual designing of RNA-RNA interaction based systems.

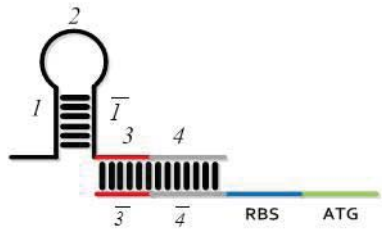


Fig. 2 A structure of the hybridized RNA between *ta*RNA and *cr*RNA containing 4 interacting domains (modified from [22]).

Affinity score (AS)

AS represents the ratio of the energy differences (ΔG_{diff}) of two reactions as can be described in Equation 2:

$$\text{Affinity score (AS)} = \frac{\Delta G_{diff,i}}{\Delta G_{diff,j}} \quad (2)$$

where $\Delta G_{diff,\ i}$ and $\Delta G_{diff,\ j}$ represent the energy differences of reaction *i* and *j* within a given RNA system, respectively.

In thermodynamics, the energy difference (with negative value) of a reaction indicates the spontaneity of the reaction as well as the strong binding affinity. As such, the lowest energy difference means the highest binding affinity and vice versa. In this study, AS which may be referred to the ratio of binding affinity between two reactions is used to predict the sequence of binding reactions within a given RNA system. Therefore, reactions *i* and *j* take place simultaneously if AS equals one. On the other hand, reaction *i* occurs before reaction *j* if AS is greater than one, and vice versa. The criterion of AS for selecting the proper set is assigned by the required sequence of reactions in the specified RNA system.

The energy difference of a reaction can be calculated by subtracting the free energy of the reactants from that of the product (s), as represented in Equation 3. In our RNA system, reactants and product are single stranded RNAs (ssRNA) and hybridized RNA (hdRNA), respectively.

$$\text{The energy difference of reaction } i\ (\Delta G_{diff,\ i}) = \Delta G_{\text{hdRNA}} - \Delta G_{\text{ssRNA(s)}} \quad (3)$$

The energy of structures can readily be determined using available tools e.g. NUPACK [17], and the Vienna RNA package [24].

3. Demonstration of the Conceptual Design

3.1 Proposed model for biological comparator based on RNA-RNA interaction

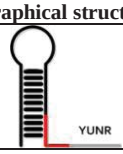
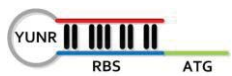
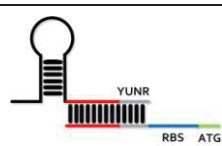
To demonstrate the conceptual design, we proposed a model of an RNA-RNA interaction based system that has a comparator-like function. In engineering, a comparator is a device that compares two input signals and returns an output to indicate which signal is greater. In general, one of the input signals is set as a reference, while the other is set as a test signal. The proposed model of RNA-based comparator comprises of 3 distinct RNAs; I_1, I_2, and M_1, that react in 2 reactions shown below.

$$I_1 + I_2 \rightleftharpoons I_1I_2 \quad \text{(First reaction)}$$

$$I_1 + M_1 \rightleftharpoons I_1M_1 \quad \text{(Second reaction)}$$

I_1 acts as the internal input of this RNA-based system whose transcription is triggered by an external signal via an inducible promoter. In this study, I_2 will be set as a reference signal for the system. The hybridization reaction of I_1 and I_2 is used for subtracting the quantity of the input signal with that of the reference signal. Thus, to make this model to act as a comparator, this reaction is desired to be the first reaction of the proposed model, followed by the hybridization of M_1 with the remaining I_1 that causes the I_1M_1 complex as the output of this model.

To realize the biological-based comparator, the structure of I_1 should contain the hanging region that can bind to the YUNR motif on M_1's loop. YUNR is the particular motif used to increase the effective initial binding of an RNA-RNA interaction [25]. M_1 is assigned to locate at the upstream of a reporter gene e.g. fluorescent protein. Its structure is folded to form a hairpin-loop that closes the ribosome binding site (RBS) to repress a translation of the reporter gene. The loop of M_1 also contains the YUNR motif. In addition, the hairpin-loop structure should provide 3 bulges on stems for the ease of its conformational change when binding to I_1 [22] . The complex I_1I_2 should have more than 24 contiguous base pairs in order for RNase III to degrade it [26]. The output of the proposed model, I_1M_1 complex can alter the hairpin-loop structure of M_1 and subsequently open the accessibility of RBS and the translation of the reporter protein. The descriptions of all considered characters are summarized in Fig. 3.

RNA type	Graphical structure	Description of required characteristics[a]
I_1 (ssRNA)		1. Hanging regions of YUNR motif for initial hybridization with M_n
M_1 (ssRNA)		1. YUNR motif on loop for effective initial binding 2. There are 3 bulges on the stem 3. Closed RBS when no binding to I_1
I_1I_2 (hdRNA)		1. More than 24 contiguous base pairings to provide RNase III degradation
I_1M_1 (hdRNA)		1. Opened RBS for ribosome accessibility 2. Contiguous base pairing for providing a strong binding

[a] *The required characteristics as previous work [22]*

Fig. 3. A summary of desired characteristics for the biological comparator based on RNA-RNA interaction.

3.2 Domain of interaction in comparator-like system

To generate an optimal sequence, 3 distinct RNAs are divided into 3 interacting domains based on the hybridization regions. Domain 1 is the region for I_1 and I_2 binding and the recognition site for RNase III with contiguous base pairings of more than 24 nucleotides placed on the sticky ends of both *Is.* Domain 2 is the region for the initial binding between the sticky end of I_1 and the YUNR motif of M_1. Domain 3 is the region for the subsequent binding between the internal strand of I_1 and the stem of M_1. The structures of the three domains are shown in Fig. 4.

To generate the desired RNA sequences, the template sequences of all domains are specified in the DD program to perform *in silico* mutation. Domain 1 is comprised of a partial sequence of RBS for increasing the binding strength after the initial hybridization between I_1 and M_1. The conserved sequence, YUNR motif is

specified in the generation of Domain 2. Domain 3 contains a complementary sequence to part of RBS. In this preliminary work, eight templates (Fig. 4) are used to generate the sequences of 40 candidate RNA sets. The scores for binding strength are assigned by the sequence of hybridization in which the scores for domain 1 is highest and that of domain 3 is lowest.

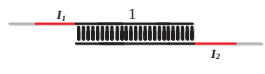
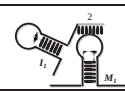
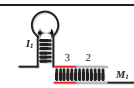
Template no.	Domain no. (5'→ 3')		
	1	2	3
			
T1	**GAGGATA***NNN***GG***NN***GG***NNN***TATCCTG**	*NNN***CT***N***G**	*NNN***AT***N***ACA**
T2	**GAGGATA***NNN***GG***NN***GG***NNN***TATCCTG**	*NNN***CT***N***A**	*NNN***AT***N***ACA**
T3	**GAGGATA***NNN***GG***NN***GG***NNN***TATCCTG**	*NNN***TT***N***A**	*NNN***AT***N***ACA**
T4	**GAGGATA***NNN***GG***NN***GG***NNN***TATCCTG**	*NNN***TT***N***G**	*NNN***AT***N***ACA**
T5	**CAGGATA***NNN***GG***NN***GG***NNN***TATCCTC**	*NNN***CT***N***A**	*NNN***AT***N***ACA**
T6	**CAGGATA***NNN***GG***NN***GG***NNN***TATCCTC**	*NNN***CT***N***G**	*NNN***AT***N***ACA**
T7	**CAGGATA***NNN***GG***NN***GG***NNN***TATCCTC**	*NNN***TT***N***A**	*NNN***AT***N***ACA**
T8	**CAGGATA***NNN***GG***NN***GG***NNN***TATCCTC**	*NNN***TT***N***G**	*NNN***AT***N***ACA**

Fig. 4. Eight template sets used in DD program (N is a position for sequence mutation).

All generated domain sequences are edited and concatenated into 3 designed RNA strands. Patterns of RNA concatenation in the direction from 5' to 3' are represented as follows:

$$I_1 = 2 + 3 + 1$$
$$I_2 = (\mathrm{N})_{16} + \bar{1}$$
$$M_1 = (\bar{1})_{10} + \bar{3} + \bar{2} + 3 + (1)_{10}$$

4. Selection Of Proper RNA Sets Using Two Evaluation Scores

In this study, we proposed the design strategy for an RNA-RNA interaction based system and employed it to design a biological comparator. The secondary structures of ssRNAs and hdRNAs from 40 designed RNA sets were predicted using the Vienna package [24]. The free energies of the ssRNAs and hdRNA used for the energy difference calculation were determined using NUPACK [17]. The SS and AS were calculated and used for selection of proper RNA sets. Herein, a designed RNA set would be considered a proper RNA set for the biological comparator only if the values of SS and AS equal to one and more than one, respectively. It was found that out of the forty designed RNA sets only five sets have SS of 1 while 37 sets have AS of more than one (Fig. 5). Since the structures of RNA strands in a given RNA set are crucial for the biological comparator to properly function, the cutoff of SS is set at the maximum, thus nearly 90 percent of the designed RNA sets are discarded. The five sets that pass the SS criterion include T5_S3, T7_S2, T7_S3, T7_S5, and T8_S3. Fig. 6 shows an example of the secondary structures of proper and improper RNA sets with SS of 1 (T5_S3) and SS of less than one (T5_S5), respectively. The highlighted sequences of I_1 strand indicate the opened and closed domain 2 for RNA set number T5_S3 and T5_S5, respectively. In addition, M_1 in the set number T5_S3 has a complete YUNR sequence (domain $\bar{2}$), while M_1 in the set number T5_S5 contains an incomplete part of YUNR sequence (Fig. 6). Due to the essence of the YUNR motif for initial binding, RNA set number T5_S5 with SS below 1 will not function as the biological comparator.

Interestingly, the 5 RNA sets that pass the SS criterion are not among the sets of high AS values, particularly T7_S3 that has an AS of less than 1. The AS of T7_S3 suggests that T7_S3 will unlikely function as the comparator. This is because the reaction of I_1I_2 hybridization in T7_S3 RNA set is not a first reaction for comparing the input signal with the reference. Therefore, the output signal of T7_S3 cannot tell us whether the input signal is larger than the reference signal. In summary, by using the criteria of the two scores, only four RNA sets (T5_S3, T7_S2, T7_S5, and T8_S3) are selected for further construction of the RNA based comparator.

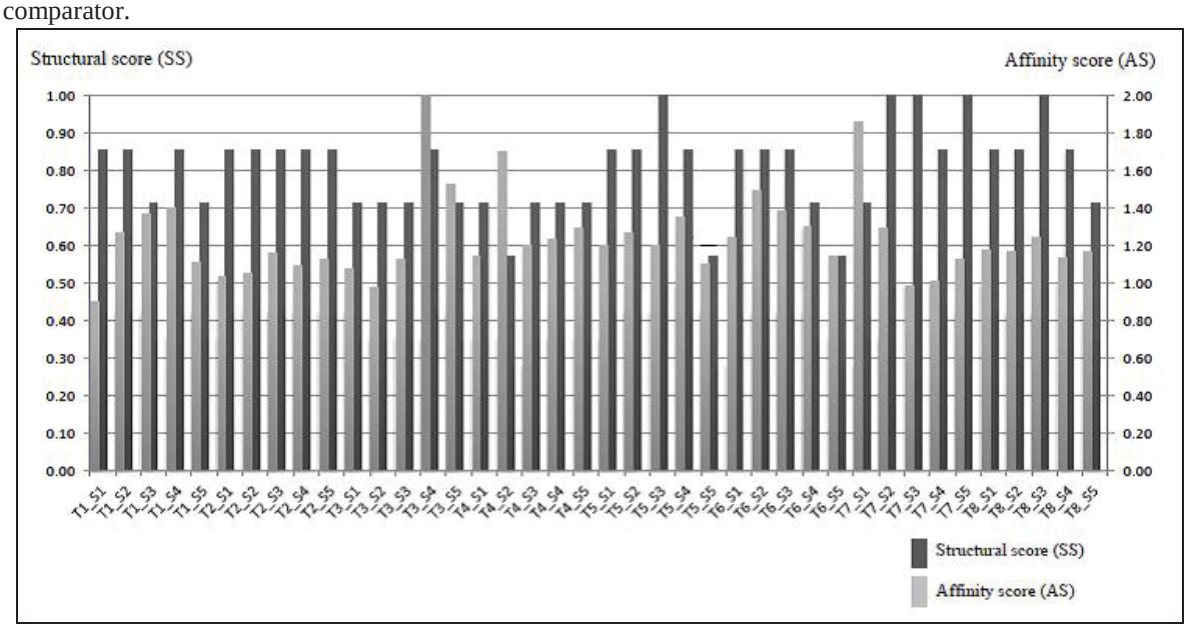


Fig. 5. The structural score (SS) and affinity score (AS) for each RNA set.

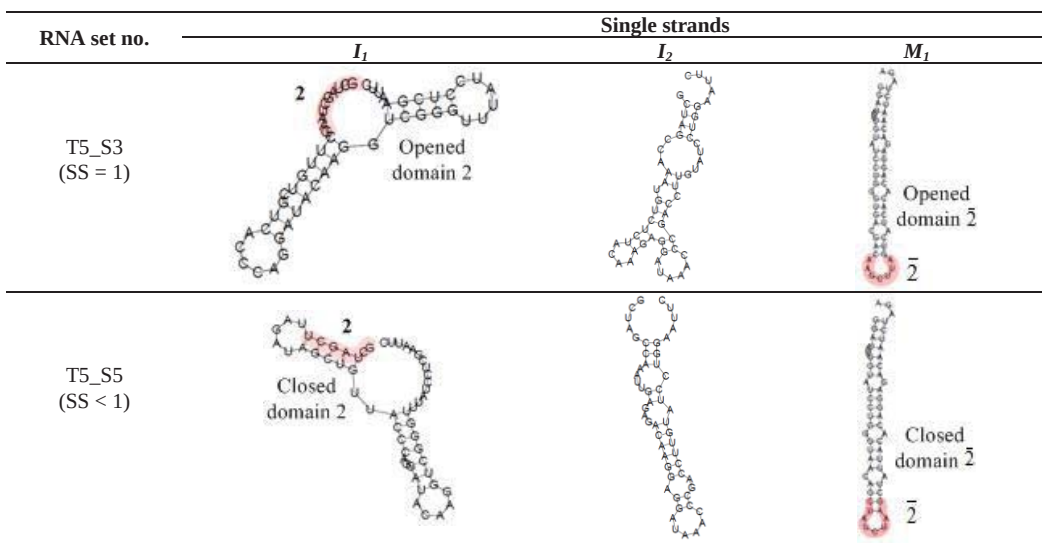


Fig. 6. The predicted secondary structures of two RNA sets: T5_S3 and T5_S5. The lines indicate the location of desired (T5_S3) and undesired (T5_S5) characters.

5. Conclusion

We proposed a strategy for designing an RNA-RNA interaction based system. The conceptual function of comparator was used as a model to demonstrate this strategy. In this strategy, two evaluation scores, SS and AS, relating the secondary structures of RNAs and their thermodynamic property, were determined to select the proper RNA sets. Only four out of 40 RNA sets were selected and predicted to behave as the comparator-like function. This preliminary work offers an alternative strategy for designing a system involving multiple RNAs strands, enabling the creation of novel and complex RNA based devices.

Acknowledgements

The authors thank D. Koolpiruck and B. Keawkammerdpong for their valuable comments and suggestions on this research project. The research was supported by the financial budget of fiscal year 2012 from King Mongkut's University of Technology Thonburi.

References

[1] Anderson, J.C., C.A. Voigt, and A.P. Arkin, Environmental signal integration by a modular AND gate. *Molecular Systems Biology*, 2007. **3**: p. 133.

[2] Win, M.N. and C.D. Smolke, Higher-Order Cellular Information Processing with Synthetic RNA Devices. *Science*, 2008. **322**(5900): p. 456-460.

[3] Tabor, J.J., et al., A synthetic genetic edge detection program. *Cell*, 2009. **137**(7): p. 1272-81.

[4] Elowitz, M.B. and S. Leibler, A synthetic oscillatory network of transcriptional regulators. *Nature*, 2000. **403**(6767): p. 335-8.

[5] Atkinson, M.R., et al., Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli. *Cell*, 2003. **113**(5): p. 597-607.

[6] Tigges, M., et al., A tunable synthetic mammalian oscillator. *Nature*, 2009. **457**(7227): p. 309-12.

[7] Looger, L.L., et al., Computational design of receptor and sensor proteins with novel functions. *Nature*, 2003. **423**(6936): p. 185-90.

[8] Antunes, M.S., et al., A synthetic de-greening gene circuit provides a reporting system that is remotely detectable and has a re-set capacity. *Plant Biotechnology Journal*, 2006. **4**(6): p. 605-22.

[9] Radhika, V., et al., Chemical sensing of DNT by engineered olfactory yeast strain. *Nature Chemical Biology*, 2007. **3**(6): p. 325-30.

[10] Purnick, P.E. and R. Weiss, The second wave of synthetic biology: from modules to systems. *Nature Reviews Molecular Cell Biology*, 2009. **10**(6): p. 410-22.

[11] Saito, H. and T. Inoue, Synthetic biology with RNA motifs. *The International Journal of Biochemistry & Cell Biology*, 2009. **41**(2): p. 398-404.

[12] Dirks, R.M., et al., Paradigms for computational nucleic acid design. *Nucleic Acids Research*, 2004. **32**(4): p. 1392-1403.

[13] SantaLucia, J. and D. Hicks, The thermodynamics of DNA structural motifs *Annual Review of Biophysics and Biomolecular Structure*, 2004. **33**(1): p. 415-440.

[14] Dirks, R.M., et al., Thermodynamic Analysis of Interacting Nucleic Acid Strands. *SIAM Review*, 2007. **49**(1): p. 65-88.

[15] Muckstein, U., et al., Thermodynamics of RNA-RNA binding. *Bioinformatics*, 2006. **22**(10): p. 1177-1182.

[16] Smith, C., et al., Freiburg RNA Tools: a web server integrating IntaRNA, ExpaRNA and LocARNA. *Nucleic Acids Research*, 2010. **38**(suppl 2): p. W373-W377.

[17] Zadeh, J.N., et al., NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry*, 2011. **32**(1): p. 170-173.

[18] Schroeder, R., A. Barta, and K. Semrad, Strategies for RNA folding and assembly. *Nature Reviews Molecular Cell Biology*, 2004. **5**(11): p. 908-919.

[19] Chen, S.-J., RNA folding: conformational Statistics, folding kinetics, and Ion electrostatics. *Annual Review of Biophysics*, 2008. **37**(1): p. 197-214.

[20] Penchovsky, R. and R.R. Breaker, Computational design and experimental validation of oligonucleotide-sensing allosteric ribozymes. *Nature Biotechnology*, 2005. **23**(11): p. 1424-1433.

[21] Tinoco, I., G. Chen, and X. Qu, RNA Reactions One Molecule at a Time. *Cold Spring Harbor Perspectives in Biology*, 2010: p. -.

[22] Isaacs, F.J., et al., Engineered riboregulators enable post-transcriptional control of gene expression. *Nature Biotechnology*, 2004. **22**(7): p. 841-847.

[23] Zhang, D.Y., Towards domain-based sequence design for DNA strand displacement reactions, in *Proceedings of the 16th international conference on DNA computing and molecular programming*. 2011, Springer-Verlag: Hong Kong, China. p. 162-175.

[24] Gruber, A.R., et al., *The Vienna RNA Websuite*. Nucleic Acids Research, 2008. **36**(suppl 2): p. W70-W74.

[25] Dirks, R.M. and N.A. Pierce, Triggered amplification by hybridization chain reaction. *Proceedings of the National Academy of Sciences of the United States of America*, 2004. **101**(43): p. 15275-15278.

[26] Hjalt, T.Å.H., E. Gerhart, and H. Wagner, Bulged-out nucleotides protect an antisense RNA from RNase III cleavage. *Nucleic Acids Research*, 1995. **23**(4): p. 571-579.