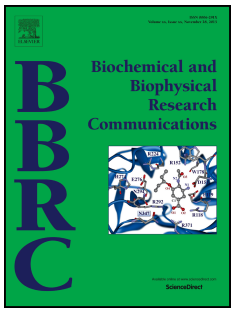
Accepted Manuscript



A method to improve prediction of secondary structure for large single RNA sequences

Ali El-lakkani, Eman M. Ibrahim

PII: S0006-291X(18)30098-6

DOI: 10.1016/j.bbrc.2018.01.086

Reference: YBBRC 39259

To appear in: *Biochemical and Biophysical Research Communications*

Received Date: 10 January 2018

Accepted Date: 11 January 2018

Please cite this article as: A. El-lakkani, E.M. Ibrahim, A method to improve prediction of secondary structure for large single RNA sequences, *Biochemical and Biophysical Research Communications* (2018), doi: 10.1016/j.bbrc.2018.01.086.

A method to improve prediction of secondary structure for large single RNA sequences

Ali El-lakkani[1] and Eman M. Ibrahim[2*]

[1,2]Biophysics Department, Faculty of Science, Cairo University, Giza, Egypt, 12613

Email: Ali El-lakkani- ellakkani@netscape.net; Eman M. Ibrahim*- dremanmibrahim@gmail.com

* Corresponding author

Abstract

The function of a particular RNA molecule within an organic system is principally determined by its structure. The current physical methods available for structure determination are time consuming and expensive. Hence, computational methods for structure prediction are often used.

The prediction of the structure of a large single sequence of RNA needs a lot of research work. In the present work, a method is introduced to improve the prediction of large single sequence RNA secondary structure obtained by Mfold program using the concept of minimum free energy. The Mfold program contains a constraint option that allows forcing some helices in the predicted structure. In our method, some of the firstly formed hairpins that are expected, by a statistical study, to be present in the real structure are forced in the Mfold predicted structure. The results show improvement, toward the real structure, in the Mfold predicted structure and this gives evidence to the RNA kinetic folding.

Introduction

Ribonucleic acid molecules, RNAs, are vital biological molecules, which play important roles in many processes. They consist of nucleotides that are composed of a nitrogen base, five carbon sugars and a phosphate group. There are four types of nucleotides; adenosine (A), guanine (G), cytosine (C) and uracil (U), where pairs of bases form hydrogen bonds. The RNA molecules have numerous well known functions as translation (mRNA), transcription (tRNA), RNA splicing, and gene regulation, it can work also as a catalyst and as a genome in a lot of viruses. An unknown