

令和 4 年 6 月 2 日現在

機関番号：11101

研究種目：基盤研究(C)（一般）

研究期間：2018～2021

課題番号：18K11519

研究課題名（和文）シュードノット構造を考慮したRNA二次構造分布ダイナミクスのシミュレーション

研究課題名（英文）RNA secondary structure dynamics simulation including pseudoknots

研究代表者

種田 晃人 (Taneda, Akito)

弘前大学・理工学研究科・准教授

研究者番号：70332492

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：RNAの機能には二次構造が深く関わっている。転写中に二次構造が順次形成される転写中フォールディングのシミュレーションを行う際に、単一配列の構造変化をモンテカルロ法で追うのではなく、二次構造分布の時間発展（RNA二次構造分布のダイナミクス）をマスター方程式により追うことで転写中フォールディングの全体像を俯瞰することができる。シュードノット構造はRNA二次構造の重要な機能モチーフの一つである。RNA二次構造エネルギー地平の解析に利用できる構造サンプリング手法のシュードノット対応版を開発し、それを応用した二次構造分布ダイナミクスシミュレーションを行った。

研究成果の学術的意義や社会的意義

ゲノム情報が普及しつつある現代社会において、生体配列情報の重要性は高まっている。その一方でRNAの機能はまだ不明なことが多い。二次構造などの高次構造を解析するための手法の発展はRNA機能解析に資することから、RNA構造予測の成果は社会へ還元されることが期待できる。RNA構造ダイナミクスをシミュレートするための方法論はまだ発展途上であり、新手法の開発はRNA構造予測分野における日本の国際的なプレゼンスを上げることにつながる。

研究成果の概要（英文）：Secondary structures play important roles in the biological functions of RNAs. Co-transcriptional folding is a mechanism that can strongly affect functional RNA structures. Co-transcriptional folding simulation based on the master equation is a promising approach for predicting time-dependent behavior of secondary structure distribution during transcription. Pseudoknots are important functional motifs and are dealt with in secondary structure level. In the present study, we developed a computational method for sampling the secondary structures that compose an RNA secondary structure energy landscape; the developed sampling method was applied to co-transcriptional folding simulation that utilizing RNA secondary structure dynamics simulation including pseudoknots.

研究分野：バイオインフォマティクス

キーワード：RNA二次構造 エネルギー地平 転写中フォールディング RNAスイッチ

1. 研究開始当初の背景

(1) 近年、miRNA や lncRNA などの研究の進展により、細胞内での生命現象における RNA の重要性が高まっている。RNA の機能には二次構造が大きく関わっていることが知られており、二次構造予測は RNA 機能解析のための重要なツールとなっている。これまで多数の優れた RNA 二次構造予測アルゴリズム・ソフトウェアが報告されているが、主流は自由エネルギー最少化法や期待精度最大化法など、平衡状態の構造予測を目的とした手法である。これに対し、転写中フォールディングなどの動的な構造変化を伴う、非平衡と考えられる状態を解析するためには平衡状態の構造予測のものとは異なるアプローチが必要となる。

(2) 二次構造エネルギー地平を俯瞰するために、マスター方程式に基づき構造分布の時間変化をシミュレーションする手法が提案されている。マスター方程式により RNA 二次構造分布の時間変化を表現する方法では、局所最適解に着目してエネルギー地平を粗視化することで、効率的に二次構造分布の時間発展を追うことが可能である。

(3) マスター方程式の方法では、ダイナミクスシミュレーションに先立ちバリアー木などの形でエネルギー地平の情報を計算しておく必要がある。しかし、多数の局所最適構造の列挙を行い、さらにその後局所構造間のエネルギーバリアーの値を精度よく求めることはコストの高い計算となる。シュードノット構造を含めた場合にさらに計算コストは跳ね上がることとなる。このエネルギー地平解析の難しさがシュードノット構造を考慮した RNA 二次構造ダイナミクスシミュレーションのボトルネックの一つとなっている。

2. 研究の目的

(1) 本研究の目的は、シュードノットを考慮した二次構造ダイナミクスをマスター方程式に基づき解析するための並列化シミュレーション手法を開発・実行することである。

3. 研究の方法

(1) 本研究課題では RNA 二次構造分布の時間発展を効率的にシミュレートするために、二次構造 x の自由エネルギー関数 $E(x)$ に基づき構成されるエネルギー地平を近似的に表現する粗視化モデルを利用した。塩基対の変化としては塩基対の形成ならびにシフトを考慮した。エネルギー地平を探索することにより、周囲の構造に対し低いエネルギーを持つ構造(局所安定構造)を列挙した。原理的には、与えられた RNA 配列において取りうる全ての二次構造を調べ上げることでエネルギー地平の形状を求めることが可能であるが、短い配列の場合を除き、一般に転写中フォールディングのシミュレーションにおいて要求される二次構造サンプル数は膨大なものとなることから、すべての可能な二次構造を利用することは効率的ではない。全ての可能な二次構造を考慮する代わりに、より少数の低エネルギー構造をそのエネルギー地平を代表する二次構造と捉えて二次構造分布の時間発展を連続時間マルコフモデルでモデル化してシミュレートする手法が Wolfinger らにより提唱されており、本研究課題でも同様なアプローチにより RNA 二次構造分布の時間発展シミュレーションを行った。このアプローチでは得られた局所安定構造群に対する遷移確率行列を計算しマスター方程式を数値的に解くことで、構造分布の時間発展をシミュレートした。本研究課題では、局所安定構造を代表構造としてエネルギー地平を近似的に扱うだけでなく、より高次の粗視化を行うことで、複雑なエネルギー地平の持つマクロな構造の利用を図った。

(2) 転写中フォールディングのシミュレーションでは、時間刻みごとに RNA 配列を 5' 端側から 3' 端側へ伸長し、配列を伸長する度にエネルギー地平の形状を決定した。配列を伸長する際に、二次構造間の対応を考慮してひとつ前の配列長において得られた二次構造分布を継承することで、配列の伸長に伴う二次構造分布の時間発展を連続的に扱った。

(3) エネルギー地平の形状を数値的に決定する際には、二次構造をサンプリングし局所安定構造の集合を得た。この二次構造サンプリングは計算時間を要する処理となり得るため、OpenMPI により並列化を行った。

(4) 2つの構造間のエネルギーバリアー値に基

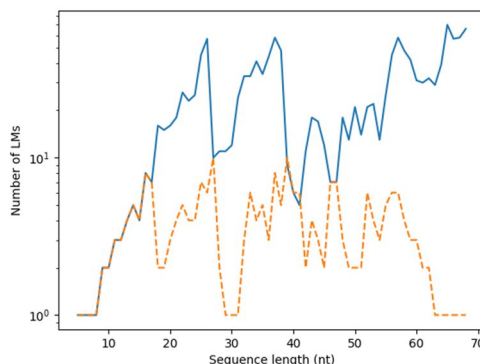


図1 局所安定構造の数

づき状態間の遷移確率を計算することができる。本研究課題では、エネルギーバリアー値の予測をアントコロニー最適化に基づいて行う手法の開発も行った。

(5) シュードノット構造を考慮した RNA 二次構造予測に関連する研究の一環として、指定した分子間シュードノット構造を持つ 2 本の RNA 配列を設計するためのウェブサーバの構築を行った。構造ダイナミクスの影響を考慮した設計を行うために、近似的なエネルギーバリアーを目的関数に取り入れた設計が可能なウェブサーバとした。

4. 研究成果

本研究課題で開発した RNA 二次構造分布の時間発展シミュレーション手法により得られた結果の解析例についての説明を以下に示す。構造変化が機能に深く関わると考えられている人工リボスイッチ配列ならびにシュードノット構造を持つことが知られている RNA 配列を例にとり実行結果を示す。本研究課題で開発した RNA 二次構造分布の時間発展シミュレーション手法の成果については英語論文を投稿する予定である。エネルギーバリアー予測法と配列設計ウェブサーバについての成果も以下に示す。

(1) 局所安定構造の数

図 1 に、配列長 68 塩基の人工リボスイッチ配列に対して実行した転写中フォールディングにおける局所安定構造数の推移を示す。図中実線で示されているのが第一段階の粗視化後の局所安定構造数、点線で示されているのが第二段階の粗視化後の局所安定構造数である。この例は、本研究課題で用いた粗視化がエネルギー地平を 10 個以下の代表的な局所安定構造として近似したことを示している。

(2) 二次構造分布の時間発展

シュードノットを持つ配列長 46 塩基の配列に対する転写中フォールディングシミュレーションで得られた、二次構造分布の時間発展(抜粋)を図 2 に示す(それぞれの配列長において存在確率が最も高かった 2~3 個の構造を表示した)。この粗視化シミュレーションでは、配列長 28 塩基では 3 つの構造(A、B、C)が競合しているが、転写が進み配列長 30 塩基ではネイティブ構造に近い構造(D)が支配的となった(約 9 割がネイティブ構造に近い構造となっている)。配列長 45 塩基(G、H、I)では、ネイティブ構造のシュードノット(H)が現れ、配列全長となる 46 塩基ではネイティブ構造(J)が支配的な構造として得られた。

(3) エネルギー地平の可視化

図 3 に、配列長 68 塩基の人工リボスイッチ配列に対して行われた転写中フォールディングシミュレーションの結果例(配列長 56 塩基でのエネルギー地平を多次元尺度法を利用してプロットしたものを)を示す。○で示されているのが第一段階の粗視化後の局所安定構造である。この可視化により、得られたエネルギー地平全体のマクロな構造が、どのような局所安定構造により捉えられているかを俯瞰できた。

(4) エネルギーバリアー予測法

アントコロニー最適化に基づくエネルギーバリアー予測法 ACOfoldpath を開発した。19 本の RNA 配列に対するベンチマークを行い、従来手法との比較を行った。多くの RNA 種において

A	((.....((((((((.....))))))))).....)	0.23
B	..((...((((((((.....))))))))).....)	0.29
C((.[[[[[[.]].]]]]]).....	0.46
D	(((((.....((((((((.....))))))))).....)))	0.91
E	((.....((((((((.....))))))))).....)	0.07
F((.[[[[[[.]].]]]]]).....	0.01
G	(((((.....((((((((.....))))))))).....))).....	0.55
H	(((((.....((((((((.....[[[[[.]]]]))))).....))).....]]]]]	0.39
I	((.....((((((((.....[[[[[.]]]]))))).....))).....]]]]]	0.05
J	(((((.....((((((((.....[[[[[.]]]]))))).....))).....]]]]]	0.90
K	((.....((((((((.....[[[[[.]]]]))))).....))).....]]]]]	0.09

図 2 二次構造分布の時間発展

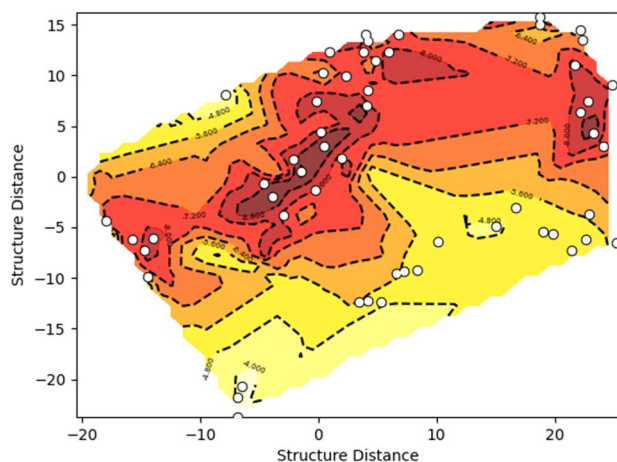


図 3 エネルギー地平の可視化

ACOfoldpath により最良のエネルギーバリアー値を得ることができた。

(5) 配列設計ウェブサーバ

分子間シュードノット構造を安定構造として持つ RNA 配列を二次構造予測に基づき設計するウェブサーバを開発した。指定した二次構造を安定に持つ RNA 配列の設計では、複雑な目的二次構造を指定する必要がある。手入力で目標構造を設定する際の労力を軽減するための新しい二次構造指定方法 (S+記法) を考案し、ウェブサーバに実装した。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Taneda Akito, Sato Kengo	4. 巻 22
2. 論文標題 A Web Server for Designing Molecular Switches Composed of Two Interacting RNAs	5. 発行年 2021年
3. 雑誌名 International Journal of Molecular Sciences	6. 最初と最後の頁 2720-1 ~ -12
掲載論文のDOI（デジタルオブジェクト識別子） 10.3390/ijms22052720	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Takitou Seira, Taneda Akito	4. 巻 83
2. 論文標題 Ant colony optimization for predicting RNA folding pathways	5. 発行年 2019年
3. 雑誌名 Computational Biology and Chemistry	6. 最初と最後の頁 107118-1 ~ -16
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.compbiochem.2019.107118	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計1件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 種田 晃人
2. 発表標題 RNAシュードノット構造を考慮した自由エネルギーランドスケープに関する検討
3. 学会等名 2021年日本バイオインフォマティクス学会年会・第10回生命医薬情報学連合大会（IIBMP2021）
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

分子間シュードノット構造を持つRNA配列設計ウェブサーバ：
<http://rna.eit.hirosaki-u.ac.jp/modena/web/>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------