

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年 6月 8日現在

機関番号：82626

研究種目：若手研究（B）

研究期間：2010～2011

課題番号：22700317

研究課題名（和文） コンピュータシミュレーションを利用した小分子機能性RNAの構造及び作用機構の解明

研究課題名（英文） Structural and functional analysis of small non-coding RNAs using computational simulations.

研究代表者

山崎 智（YAMASAKI SATOSHI）

独立行政法人産業技術総合研究所・生命情報工学研究センター・産総研特別研究員

研究者番号：80463876

研究成果の概要（和文）：数十～百塩基長程度の小さなRNA分子が、細胞内でタンパク質合成抑制等の機能を担っている事が、様々な生物種において知られている。これら小分子機能性RNAの作用機構を明らかにするためには、まずその立体構造を得る事が必要不可欠である。本研究では、数塩基からなる小断片をコンピュータ上でパズルのように組み合わせて全体の立体構造を得る手法に対して、小断片の扱い方や構造評価関数へ改良を加える事で、高速かつ高精度な小分子機能性RNAの立体構造予測を可能にした。

研究成果の概要（英文）：Recent studies have revealed that some small RNA molecules play important roles such as repression of gene translation in many kinds of living cells. To clarify the mechanism of these small functional RNAs, the tertiary structures of those RNAs are required. In this study we developed a new method of tertiary structure prediction of RNA based on the fragment assembly method. We achieved fast and high-precision prediction by improving fragment libraries and potential energy functions.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	700,000	210,000	910,000
2011年度	600,000	180,000	780,000
総計	1,300,000	390,000	1,690,000

研究分野：総合領域

科研費の分科・細目：情報学、生体生命情報学

キーワード：バイオインフォマティクス、生体機能利用、核酸、立体構造予測、分子シミュレーション

1. 研究開始当初の背景

近年、それ自身はタンパク質の配列情報をコードしないが、他のタンパク質や他のRNA分子などと相互作用する事によって、タンパク質の翻訳反応を制御する機能を持つRNA小分子に注目が集まっている。小分子機能性

RNAと呼ばれるこれらのRNAは、細菌類からヒトに至るまで様々な種で確認されており、遺伝子の発現調節機構の一部、あるいはウイルスなどに対する防御機構の一部として機能していると考えられている。このような小分子機能性RNAは、ノーザンブロッティング

を始めとした生化学的手法によって数多く見つかっている。また、ゲノム配列の中から機能性 RNA となる部分を予測する研究も進められている。大量の小分子機能性 RNA 分子の存在とその塩基配列が知られていく中で、その立体構造を高速・高精度に求める手法の確立が求められている。立体構造を知る事で、小分子機能性 RNA が関わる反応の詳細まで理解する事ができ、その知見を応用することで我々の生活や健康の改善に繋がる事が期待されるからである。X 線結晶構造解析法や核磁気共鳴法などの実験による立体構造決定は、「高精度」という目的は果たすものの、「高速」に行う事はまだ困難である。そのような背景から、コンピュータを使って RNA の塩基配列からその立体構造予測する手法に期待が集まっている。

2. 研究の目的

コンピュータを使った RNA の立体構造予測手法を新規に開発する事を目的として研究を行った。先にも挙げた「高速・高精度」という要件を満たすためには、「効率的な構造空間の探索手法」、そして「優れた構造評価ポテンシャル」の開発が必要であった。これらに重点をおいて研究を進めた。

3. 研究の方法

研究開始当時の数少ない先行研究において利用されていた「フラグメント・アセンブリ (FA) 法」が有力であると考え、この手法をベースに新規手法を開発した。FA 法は、数塩基から成る小断片構造をパズルのように組み合わせる事で、全体構造を得る手法である。この際の最適組み合わせの探索にはモンテカルロシミュレーションを利用する。従来法では、ここで利用する小断片は「単鎖」の構造に限られていた。その実装下において、RNA の立体構造を構成する要素の中で重要と言える「塩基対」を形成するためには、多くの探索を行わなければならなかった。仮に塩基対を形成する位置が予め分かっていたとしても、その塩基対を形成するような断片構造の組み合わせの探索が必要となる。その点を改善するために、開発手法では単鎖の小断片だけでなく、塩基対の形をとった小断片をも利用する事で、構造空間の探索の効率化を図った。この戦略には、RNA の二次構造情報が予め必要になるが、今日においては高速・高精度な RNA 二次構造予測手法を併せて利用する事で解決出来る。この手法をとることによって、立体構造予測に必要な過程の一つ、構造空間の探索に費やす計算時間を大きく減らす事が出来ると期待された。これにより浮いた時間を、より多くの問題を解くことに

費やす事も出来るが、それだけではなく、予測結果の精度を高める事にも利用出来ると考えた。予測結果の精度は、予測結果の尤もらしさを評価するポテンシャル関数に大きく左右され、精度と計算時間は概ねトレードオフの関係にあると言えるからである。実際に全原子ポテンシャル関数を利用した構造評価を取り入れる事で、予測結果の改善も試みた。また、ポテンシャル関数についても、既知のものに加え、塩基対間の相対配置に基づく値を導入する可能性を検討した。

4. 研究成果

研究目的を満たすプログラム、RNA Assembler using Secondary Structure Information Effectively (RASSIE)を開発した。単鎖の断片だけでなく、二次構造を基本とした断片をも利用出来るように実装した結果、従来法と比して特に計算速度の面で改善が得られた。開発手法の性能をここに簡易に示すために、ある配列をもつ RNA について、その立体構造を「既存手法」及び「開発手法」によって 5000 構造ずつ予測し、その予測構造と正解構造の間の RMSD の分布をそれぞれ示した。この性能評価実験は、計算条件・計算時間を 2 手法で可能な限り揃えて行った。その結果得られたこの図において、開発手法の方が正解構造に近い構造をより多く得ている事が分かり、これは既存手法と比して、開発手法が高効率に構造探索を行えている結果であると言える。

次いで、このように高速に得られた 5000 構造の中から最尤の構造を選び出す事が、高精度の予測、という目的を果たす事に必要である。構造クラスタリングに基づく方法、簡易なポテンシャル関数を使った評価に基づく手法なども試したが、最も良い結果を導いたのは、AMBER parmbsc0 力場を用いた全原子モデルでのエネルギー評価法を用いた場合である事が示された。全原子モデルでのエネルギー評価には計算時間が多く掛かるもの

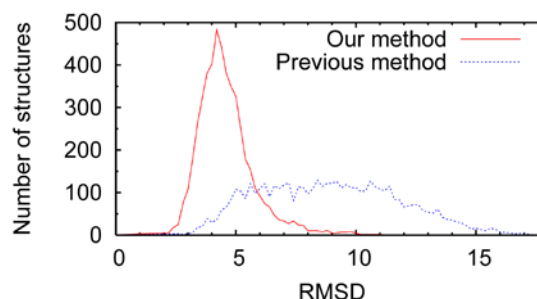


図 1: 予測構造の正解構造からの RMSD 分布
従来法 (青) では、正解構造から大きく離れた構造が得られる事も多くあるが、開発手法 (赤) では、正解構造に近いものがより多く得られる。

の、開発手法の高速・高効率な構造探索と併せて使う事で、実用的なものになると結論付ける事が出来た。図2に、これらの予測手法の結果として得られた構造を例示する。

さらに、開発手法の汎用性を示すために、50塩基長までのRNAについて開発手法を用いて構造予測を行い、正解構造からのRMSD値を塩基長に対して図示したものが図3である。ここで予測した全172ターゲットのうち131ターゲット(約76%)において、正解構造に類似していると言えるRMSD値5 Å以下の予測構造が得られた。これは十分に高い汎用性を持つものとして評価されている。

開発手法の更なる精度向上のためには、構造評価に利用するポテンシャル関数の改善が必要である事が、研究開発途上にて示唆された。そこで、核酸構造に関する新たなポテンシャル関数として、塩基対間の相対配置に基づくポテンシャルの適用可能性を調査した。このポテンシャルは、塩基対間の相対配置の分布傾向が、塩基対を構成する塩基の種類によって異なる事を利用するものである。本研究においては、RNAと類似の塩基対構造を作り、RNAよりも構造上安定であるDNAを利用した。様々な塩基配列を持つDNA分子を分子動力学法によるシミュレーションを使って構造サンプリングを行い、そのトラジェクトリから塩基対間の相対配置の分布を得てポテンシャル関数とした。得られたポテンシャル関数の評価を、既知のDNA-タンパク質相互作用の系を用いて行ったところ、既存の知見とよく合致する事が示された。

本研究の主目的としたRNAの立体構造予測研究は、2010年度の時点では世界的にも限られた報告しかなかったが、ここ2年の間にその注目度は著しく上昇し、特に昨年はRNAの立体構造予測手法に直接関連する論文だけ

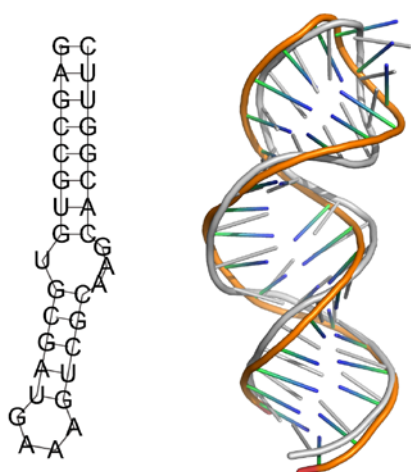


図2：(右) 開発手法によって得られた予測構造(カラー)と、正解構造(灰)の重ね合わせ。(左) ターゲット(2F88)の配列と二次構造。

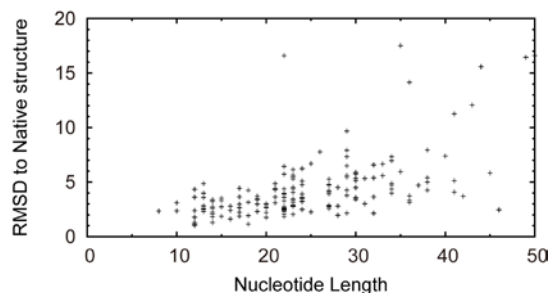


図3：最尤予測構造のRMSD値
様々な塩基配列・配列長を持つターゲットの立体構造の予測に成功した。

でも、本研究によるものを含めて10報に近いものが報告されている。国内においては、同種の研究報告は本研究以外には確認出来ていない。世界的に高い注目度を集めたこの分野に、本研究が幾分かでも食い込めた事は、今後の研究の進展に大きな意味を持つと期待している。今後は、予測性能の更なる向上を目指し、従来法と融合した新しい構造探索手法の開発、ポテンシャル関数の改善を始めた構造評価法の新規開発などを目指す。それと共に、予測結果として得られた構造を利用した次の研究、RNA-RNA相互作用、RNA-タンパク質相互作用研究を進める事が出来ると期待している。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計2件)

- ① Satoshi Yamasaki、Tohru Terada、Hidetoshi Kono、Kentaro Shimizu、Akinori Sarai、A new method for evaluating the specificity of indirect readout in protein-DNA recognition、Nucleic Acids Research、査読有、Vol. 40、2012、DOI: 10.1093/nar/gks462
- ② Satoshi Yamasaki、Shugo Nakamura、Kazuhiko Fukui、Prospects for tertiary structure prediction of RNA based on secondary structure information、Journal of Chemical Information and Modeling、査読有、Vol. 52、2012、557-567、DOI: 10.1021/ci2003413

[学会発表] (計7件)

- ① 山崎 智、RNA tertiary structure prediction based on secondary structure and fragment structure

- library、第34回日本分子生物学会年会、2011年12月14日、神奈川
- ② 山崎 智、二次構造情報とフラグメント構造群に基づくRNA三次構造予測、第10回新しいRNA/RNPを見つける会、2011年09月15日、沖縄
 - ③ 山崎 智、RNA三次構造予測の現状と展望、次世代バイオインフォマティクス研究会、2011年08月03日、北海道
 - ④ 山崎 智、二次構造予測結果に基づくRNA三次構造予測システムの開発、平成22年度LS-BT合同研究発表会、2011年02月01日、茨城
 - ⑤ 山崎 智、Attempts to predict RNA tertiary structures using fragment structural library based on secondary structures、第48回日本生物物理学会年会、2010年09月21日、宮城
 - ⑥ 山崎 智、Attempts to predict RNA tertiary structures using fragment structural library based on secondary structures、CBRC2010、2010年07月28日、東京
 - ⑦ 山崎 智、Information entropy as a measure of the contribution of indirect readout in protein-DNA recognition、第10回日本蛋白質科学会年会、2010年06月18日、北海道

6. 研究組織

(1) 研究代表者

山崎 智 (YAMASAKI SATOSHI)

独立行政法人産業技術総合研究所・生命情報工学研究センター・産総研特別研究員

研究者番号：80463876