二次構造情報を生かしたRNA三次構造予測手法の開発

山崎　智
分子機能計算チーム
産総研特別研究員

近年になって、mRNAの翻訳調節・分解といった重要な機能に関わる多種多様の小分子RNAの存在が明らかにされるようになりました。その機能がどのような機構で発揮されるのかを明らかにするためには、そのRNAの三次構造を知るという過程が極めて重要になります。しかしながら実験により構造が決定されていないRNA分子も数多くあります。そのため、RNAの塩基配列からその三次構造をコンピュータで予測する、RNA三次構造予測手法の必要性が高まっています。

三次構造予測手法の中で最も有力なものの一つは、フラグメント・アセンブリ法です。フラグメント・アセンブリ法では、既知三次構造群から得た数塩基長の断片構造をパズルのように巧く組み合わせ、いくつかの構造評価指標に基づいて作られた簡易なエネルギー関数を最小化するような組み合わせをメトロポリス・モンテカルロ法によって繰り返し探索する事で、未知の三次構造の予測を行います。タンパク質の三次構造予測において大きな成果を残したこの手法は、一般に単鎖のフラグメント構造を利用しています。しかしながら、単鎖のフラグメント構造という枠組みでは、RNAの三次構造の要素として最も重要な、塩基対からなる構造、つまり二次構造を直接的に内包する事は出来ません。そのためこの方法では構造探索過程において正解の二次構造を持たない構造に何度も寄り道しながら、正解構造を目指す事になります。『計算対象となるRNAの二次構造が既知であるとは限らない』、という前提の下であるならば、その『寄り道』は多種多様な二次構造パターンを試行錯誤する過程であり、決して無駄なものではありません。しかしながら、RNAの二次構造の情報は、塩基置換法などの実験手法や、様々なRNA二次構造予測プログラムによって、比較的容易に知る事が出来るため、それを最大限有効に利用し、『寄り道』を減らす事が出来れば、三次構造予測の精度・速度の向上に繋がります。

そこで、本研究では、単鎖のフラグメント構造ではなく、二次構造を基本としたフラグメント構造を利用した、新しいフラグメント・アセンブリ法に基づく三次構造予測手法を開発しました。開発手法での構造探索に利用するフラグメント構造群は、既知構造をその二次構造に基づいて分解し、helix、hairpin loop、single strand loopの3種類のカテゴリに分けて用意しました。三次構造予測は、入力された塩基配列と二次構造情報に基づいて、これらのフラグメント構造を組み合わせ、最適化していく事で実現されます。この二次構造を基本としたフラグメント構造を利用する手法によって、従来の単鎖のフラグメント構造を利用する手法よりも、特にステム領域の構造探索の効率を大幅に向上させる事に成功しました。従来法と同等の予測精度を維持しながらも、より少ない計算時間での予測が行えます。これにより生じた余剰の計算時間を使い、更に多くの構造を予測する、あるいは高精度な力場を用いた構造評価を行う、などの事が可能になりますので、更なる予測精度の向上も実現されます。今後の課題としては、従来法においても予測が困難とされる、複数のステム領域から成る複雑な三次構造の予測が挙げられますが、これに関しては、それらのステム同士の高次の相互作用の情報を利用する事で、解決が出来ると考えています。

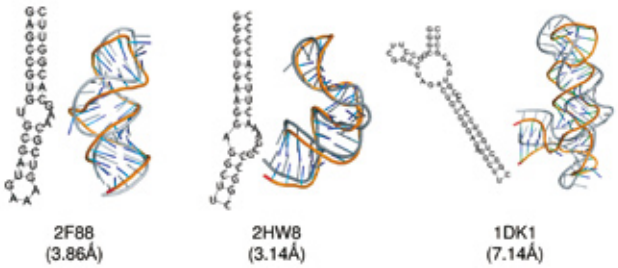

図　入力配列・二次構造と三次構造予測結果の例