

組合せ最適化手法による  
シュードノット構造を含んだ RNA の二次構造予測

田中 健夫 若月 光夫 富田 悦次

電気通信大学 電気通信学部 電子情報学科

〒182 調布市調布ヶ丘 1-5-1

e-mail: tanaka@ttl.cas.uec.ac.jp

wakatuki@ttl.cas.uec.ac.jp

tomita@ttl.cas.uec.ac.jp

遺伝子情報解析の問題の一つである RNA の二次構造予測は、生体現象に関与している RNA の平面的な近似構造をその配列から予測する問題である。この問題を動的計画法によって効率良く求める手法が既に提案されているが、この手法では部分問題への分割を困難にするシュードノット構造を扱うことは難しい。これに対して、組合せ的手法はシュードノット構造も扱い得るので、本稿では、これをニューラルネットワークを用いて実現し、塩基間結合を平面的に記述するような制約を導入した予測法を提案する。本手法の特徴は、この制約の強度を温度パラメータで制御することによって、ランダムサーチ的な探索から RNA のフォールディングシミュレーションに近い探索へ徐々に移行できることである。これを用いた計算機実験では、実際にシュードノット構造を持つ RNA 配列に対してその抽出が確認され、本手法の有効性が示された。

**Prediction of RNA Secondary Structure  
Including Pseudoknotting  
Based on Combinational Optimization**

Takeo Tanaka, Mitsuo Wakatsuki, and Etsuji Tomita

Department of Communications & Systems Engineering, University of Electro-Communications  
1-5-1 Chofugaoka, Chofu, Tokyo 182 Japan

e-mail: tanaka@ttl.cas.uec.ac.jp

wakatuki@ttl.cas.uec.ac.jp

tomita@ttl.cas.uec.ac.jp

Prediction of RNA secondary structure which is one of the major genome informatics problems is to obtain approximated RNA structures that cause various biological phenomena. A solution to this problem using dynamic programming is efficient, but it is difficult to deal with pseudoknotted structures. In dealing with such structures, we apply a neural network approach to this prediction, and we introduce a restriction based on an assumption that base pairings are expressed on a plane without crossing. Our approach is able to change the behavior of the simulation little by little from that of a random search to that of a folding simulation by changing a control parameter. The predicted structures by computer simulations which include actual pseudoknotted structures confirmed that our approach is effective.

## 1 はじめに

RNA は遺伝子情報を取めた DNA を補完し、様々な生体機能を担う核酸である。この分子は塩基が重合した一本鎖状の構造として生成されると同時に、離れた塩基同士に生じる水素結合によって折り畳まれる (フォールディングする) ことで、エネルギー的に安定した立体構造を形成する。この構造を知ることは RNA の機能を解明する上で重要な意味を持つが、ごく簡単なもの以外は X 線回折や NMR などの実験的な解析が困難である [1]。そこで、立体構造の予測ではなく、二次構造という RNA を二次元平面上で表現した場合の構造予測が研究されるようになった [2][3][4]。これは、与えられた RNA の塩基配列からエネルギー的に安定した分子状態になる水素結合の位置を求めるエネルギーの最小化問題である。

この問題の解を求める手法で現在広く用いられているものとして、動的計画法を利用した手法 [5] がある。この手法は問題を部分問題に分割して取り扱うことが可能なため、 $O(n^3)$  という多項式時間で実行できる。しかし、分割された部分塩基配列に対する最適な二次構造を求めたとしても、シュードノット構造を対象として考慮しようとする場合には、その配列中の塩基とそれ以外の塩基にまたがる結合を考える必要があるため部分問題への分割が困難であり、動的計画法でこの構造を扱うことは難しい。

水素結合が連続したステム構造を予測構造の単位とし、それらの最適な組合せを求める組合せ的手法 [6][7] では、ホップフィールドネットワークを用いた高速な手法が提案されている [6]。しかし、問題規模が大きくなると、ある特定の局所解に陥り易くなるので妥当な解が得られるとは限らない。また、単なる組合せ抽出の問題形式を採っているのでシュードノット構造も抽出し得るが、このような構造の形成に何の条件も加えないとすると、非結合塩基の数が少ない不自然な予測構造が抽出されてしまい、生物学的に妥当と思われるような構造の抽出は難しい。

ステム構造の候補のエネルギーを比較し、矛盾が起こらないように予測構造を組み立てるフォールディングシミュレーションには、暫定的にシュードノット構造のエネルギーを規定することで比較的良好な結果が得られているものがある [8][9]。しかし、シュードノット構造の形成に伴う熱力学的エネルギーの正確なデータは得られていないので、これを含み、かつ生物学的に妥当な予測を行うことは大変困難である。

本研究では、シュードノット構造も予測対象に含めた、妥当で多様な二次構造予測を行うため、組合せ的手法をボルツマンマシンを用いて実現する。これによって、特定の局所解への落ち込みを避け、ステム構造の候補の様々な組合せを解とすることが期待できる。また、これによって形成される遠距離の塩基間の結合が問題になるので、塩基間結合が平面的に記述できるものと仮定し、それに基づく制約条件を設けている。この制約条件の強度をボルツマンマシンの温度パラメータで制御することによって、徐々に問題例の特徴的な構造を把握していく過程が観察されている。シュードノット構造を含んだ問題例に対する計算機実験では、その構造の抽出が確認され、複数回の試行で実構造とほぼ同等な解も抽出されている。

## 2 RNA の二次構造

RNA には 4 種類の塩基 (A,C,G,U) があり、A-U、G-C (まれに G-U) の組合せで水素結合することによりフォールディングを起こす (図 1 参照)。隣接する塩基結合の形成によって熱エネルギーが外部に放出されるため、連続的な結合領域であるステム構造が形成される [10]。本手法では、このステム構造を予測構造の単位として、この候補を安定化要因とする目的関数の値が最小となる組合せを求める。

一方、塩基結合が形成されない部分では各種 (図 2 (a)-(d)) のループ構造が形成される。また、ステム構造やループ構造以外の特別な構造として、塩基配列上の互い違いの位置に水素結合を形成するシュードノット構造 (図 2 (e)) が存在する。塩基を節点、塩基間の結合を枝として塩基配列を直線状グラフで置き換えると、シュ

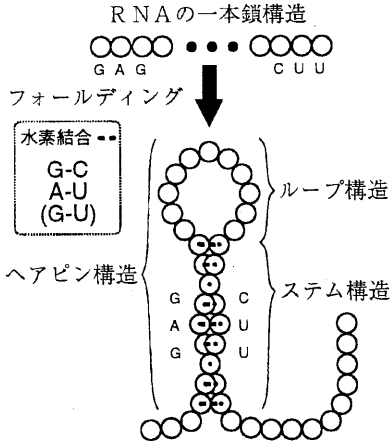


図1: RNAのフォールディング

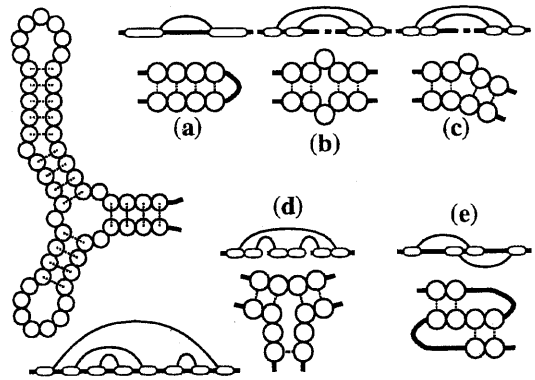


図2: 二次構造の例(左)とループ構造((a)~(d))とシュドノット構造((e))

ドノット構造以外はグラフの上部のみで交差することなく枝を描くことができ(図2(a)-(d))、これらの任意の組合せも同様に表現できる(図2左)。これに対し、シュドノット構造はグラフの上部と下部に分けて枝を描かなければ平面グラフとして描けず、これが複数個存在すると枝を交差させずに平面的に記述することができない場合がある。しかし、シュドノット構造は頻繁には現れないので、ほとんどのRNAの二次構造は平面グラフで表現できるものとみなせる。そこで、本稿で行うRNAの二次構造予測は、RNAが二次元平面内に制限されているものと仮定して、平面内で交差のないフォールディングのうちで最もエネルギーの低い、安定なものを求めるものとする。

### 3 手法

#### 3.1 ボルツマンマシンの導入

解析に用いるボルツマンマシンは、 $n$ 個のユニット同士が対称な結合重みにより全結合したニューラルネットワークである。他のユニットからの入力の総和をユニットの内部活性値とし、その値を出力関数に代入した結果、出力が1になるときを発火状態、0のときを非発火状態と定義する。ユニット $i$ とユニット $j$ との関係を結合重み $\omega_{ij} = \omega_{ji}$ ( $\omega_{ii} = 0$ )、ユニット $i$ の発火基準となる発火しきい値 $\theta_i$ を規定することにより、エネルギー関数 $E$ は次のように定められる。

$$E = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \omega_{ij} x_i x_j - \sum_{i=1}^n \theta_i x_i \quad (1)$$

状態遷移はランダムに選択されたユニット $i$ ( $1 \leq i \leq n$ )に対して実行される。ユニット $i$ が非発火状態から発火状態に遷移するときのエネルギー変化量 $\Delta E_i$ を図3の行番号6の式のように計算し、この値を内部活性値とする。次に、シグモイド関数(図3の行番号7右辺)にこの値を代入して、ユニット $i$ の状態遷移確率 $p_i$ を

```

1: begin
2:   Initialize( $\theta, \omega, n, T$ );
3:   repeat
4:     repeat
5:        $i := \text{random}[1, n]$ ;
6:        $\Delta E_i := \sum_{j=1}^n \omega_{ij} x_j + \theta_i$ ;
7:        $p_i := 1 / \{1 + \exp(-\frac{\Delta E_i}{T})\}$ ;
8:       if  $p_i > \text{random}[0, 1]$ 
9:         then  $x_i := 1$ 
10:        else  $x_i := 0$ 
11:       fi
12:     until stop-criterion
13:      $c := \alpha \cdot T$ ;
14:   until stop-criterion
15: end

```

図3: Boltzmann Machineのアルゴリズム

求める。ここで、この関数中の  $T$  は温度パラメータであり、この値が高ければ  $\Delta E_i = 0$  において  $p_i = 0.5$  となる緩やかな曲線になり、逆に、これが低い程この曲線の形状がステップ関数に近くなる。遷移確率  $p_i$  はユニット  $i$  の次の時点の状態が発火状態になる確率であり、区間  $[0,1)$  の一様乱数値との比較によって出力  $x_i$  を決定する (図3の行番号8,9,10)。このような処理を逐次繰り返すことによってネットワークの状態を更新していく。

ある温度パラメータにおける状態遷移の繰り返しを一定回数だけ行った後で、温度パラメータを  $\alpha$  ( $\alpha < 1, \alpha \approx 1$ ) 倍し微小変化させる。これをランダムな遷移に近い高温の温度パラメータから始め、ネットワークのエネルギーが単調減少するような低温度になるまで繰り返す。これにより、局所解に陥りにくい近似最適解の探索が行われる (最適解とは (1) 式が最小の値をとる解と定義する)。

構造予測の前処理として、ステム構造を形成することが可能な部分塩基配列の対をストリングマッチングによって抽出する。このステム構造の候補をボルツマンマシンの個々のユニットにそれぞれ対応させ、ユニット  $i$  の出力  $x_i$  は対応するステム候補が予測構造中にあるとき発火状態とする。

安定した RNA の二次構造を求めるには、予測構造に含まれるステム候補の自由エネルギーの総和が小さくなるようにし、かつ、それらの候補間に矛盾があってはならない。この要件を満たすエネルギー関数を規定するため、ユニット  $i$  の発火しきい値  $\theta_i$  とユニット  $i, j$  間の結合重み  $\omega_{ij}$  を次のように定める。

$$\theta_i = |(\text{候補 } i \text{ に対応するステム構造のエネルギー})| \quad (2)$$

$$\omega_{ij} = \begin{cases} -\max\{\theta_k \mid 1 \leq k \leq n\} & \text{候補 } i, j \text{ 間の結合塩基が重なる} \\ 0 & \text{候補 } i, j \text{ 間の結合塩基が重ならない} \end{cases} \quad (3)$$

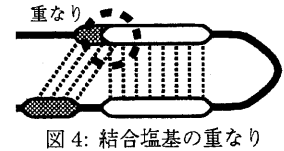


図4: 結合塩基の重なり

ここで、 $\theta_i$  のステム構造のエネルギーには、溶融実験を通じて経験的に求められた値を用いた [11]。また、結合塩基の重なりとは、図4のように両立しない位置関係とする。

さらに、本手法では  $\theta_i$  を2値化することによって、単一のステム構造に挟まれた非結合塩基によって形成されるヘアピンループのエネルギーを考慮している。これは、状態遷移によってヘアピンループが予測構造中に発生 (消滅) するとき、このループに隣接したステム候補  $i$  のしきい値  $\theta_i$  にループのエネルギーを加 (減) 算することにより行う。こうすることで、予測に使用する目的関数を RNA の実際のエネルギーに、より精度良く近似させることができる。

### 3.2 平面的記述による制約

シュードノット構造の形成に関して何の条件も与えずに上述の方法で予測を行うと、実構造とかけ離れた、多数のシュードノット構造が含まれた不自然な解が抽出されてしまう。これを避けるため、ほとんどの二次構造が平面的に記述できると仮定して、これを制約条件に用いる。

予測構造中の候補を、図2に示した直線状グラフの上部、下部に対応した候補集合1(上部)、候補集合2(下部)にそれぞれ割り当てたとき

- 一つの候補集合内の候補同士は、シュードノット構造を形成せず、また結合位置の重なりがない
- 異なる候補集合に含まれる候補同士の位置関係に結合位置の重なりがない

という条件を満たせば、その予測構造は平面的に記述できる。ステム候補の結合塩基の重なりはユニット間の結合重みによるペナルティーによって排除されるので、上記一項目の条件を満たさない遷移を却下しながら状態遷移を行うことによって、平面的記述を保持した予測構造が抽出される。しかし、候補集合1,2に多数の候補が含まれると、状態遷移が却下されやすくなり取り得る解が限られてきてしまうので、これを緩和するため、

補助的に候補集合3を用意する。このとき、上記条件の一項目の候補集合1,2に対する平面的記述の保持とユニット間の結合重みによる結合塩基の重なりの抑制を行いながら、

- 候補集合3には何の制約も課さず、どの候補も取り込むことができる
- 予測構造を構成するステム候補が候補集合に取り入れられる優先順位を候補集合1,2,3の順とする
- 上記の条件のもとで候補集合3に含まれ、予測構造中のどの候補とも同一塩基に結合を持たないとき、その候補を内部活性値を補正する要因(以後は補正要因と記す)とみなし、その発生・消滅時に結合塩基の重なりと同値のペナルティーを内部活性値に加算する(図5参照(xがない場合))

とすることによって、平面的記述を妨げる候補集合3の候補にペナルティーを課しながら、候補の広範囲にわたる組合せを予測構造として取り得ることが可能になる。

ユニット*i*が非発火状態の場合、それに対応する着目中のステム候補を候補集合1,2,3の優先順でどの集合に取り込むことができるかを調べる。このとき、既に集合1に含まれている候補を集合2に移すことによって、そのままでは候補集合3にしか取り込めない着目候補を集合1に取り込める場合がある。その場合は、遷移が受理されたときに候補集合1にある候補を移動することにし、候補集合3に含まれる候補の数を抑える。次に、発火ユニットとの結合重みと着目ユニットのしきい値から、図3の行番号6の式を計算する。この値に、ヘアピループと補正要因の発生・消滅による値を加算し内部活性値を決定する。すなわち、これは着目候補だけではなく、図5に示す集合1中の候補aのようにヘアピループを含まなくなる候補や、集合3中の候補bのように補正要因ではなくなる候補の存在も調べ、これらの状況変化に伴うエネルギー差を先の式の計算値に加算し、着目ユニットの内部活性値を決定する。

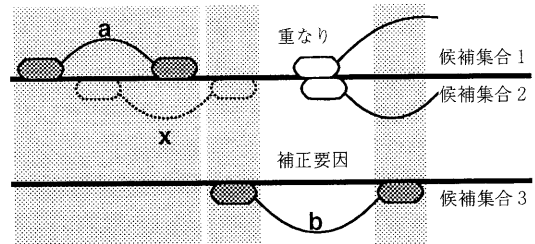


図5: ヘアピループと補正要因の発生

一方、ユニット*i*が発火状態で候補集合1,2に含まれる場合、状態遷移後も候補集合1,2,3の優先順位で候補が候補集合に含まれるように、ステム候補の移動のしかたを決めておく。これは図6のように段階的に行い、それぞれの段階で移動可能な全候補を抽出し、まず、網掛けの○印で示す移動先で平面的記述を崩さず移動候補数が最大となる候補の組合せを求める。移動させる候補を決定した後、続いて

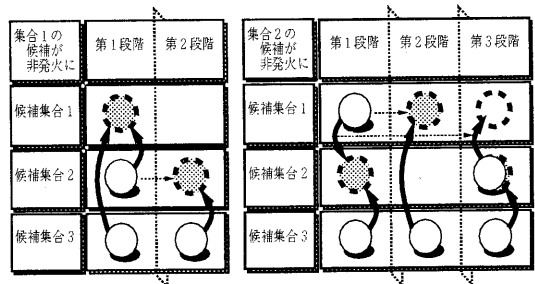


図6: ステム候補の移動

内部活性値の計算を行う。この場合も、図5のように着目候補xが取り除かれることによって、図3の行番号6の式を補正する必要のあるヘアピループを含む候補aや、補正要因となる候補bが予測構造中に発生するので、これらを考慮して内部活性値を決定する。

上記の処理をボルツマンマシンの動作に加えることによって、負の結合重みによるペナルティーと同様に、平面的記述を妨げる候補集合3の補正要因へのペナルティーの効果も温度パラメータを使って制御できる。すなわち、温度が高いときは平面的記述の制約の弱い、集合3に多数の候補を含んだランダムな状態遷移が行われ、一方、温度が低いときは平面的性質を保持し、集合1,2へ逐次的に候補を取り込んでいくフォールディングシミュレーション的な制約の強い探索が行われる。

## 4 計算機実験

実際の問題例に対して本手法による二次構造予測を実験する。本手法の予測結果は候補の着目順に依存するので最適解に収束するとは限らず、また計算上の最適解と実構造とは必ずしも一致しないので、最適解抽出の必要性はそれほど大きくない。従って、温度パラメータの変化率 $\alpha$ はステム候補数500程度の予備実験を行い、問題例の傾向が現れる0.99程度の値に設定した。また、抽出するステム構造の候補は塩基長4以上としたため、2,3塩基長のステム構造は抽出されない。そこで、ここでは実構造と1,2塩基程度のずれがあるものや、抽出されない短いステムを除いたものを実構造に準ずる解(準実構造)と規定した。なお、弱結合のG-U結合が少ない *Bacteriophage T-4* 以外の問題例は、G-U結合を考慮して候補の抽出を行った。

実験には *SUN SPARC Station 10* を使用した。計算機処理時間はいずれの問題例も1試行におよそ1~2時間程度かかった。ここでは問題例1つにつき50回試行した結果について、主に準実構造と最小エネルギー解とのエネルギー関数値の差、準実構造の抽出頻度、および全抽出解の傾向に関して考察することにする。

### 4.1 *Dunaliella parva* 18S rRNA

3000塩基近い塩基長を持つ緑藻の *Dunaliella parva* 18S リボゾーム RNA の部分塩基を解析した。この塩基配列でシュードノット構造が存在するのは塩基番号2145~2565のグループIイントロンであり、実験ではこの構造を含む塩基番号2233~2465を用いた。この部分配列の実構造は図7(左)のとおりであるが、準実構造は図7(右)のように定める。

50回の試行で抽出された最小エネルギー解、準実構造、その他の解に対するエネルギー値と抽出頻度を表1に示す。予測途中のエネルギー値が候補の着目順に依存して違う値を取ることや、ステム候補が852個と多かったことなどが影響して抽出された最小エネルギー解は低い抽出頻度となった。一方、準実構造と一致する解は最も頻繁に抽出された。これは実構造を形成するステム構造の塩基長が比較的長く、そのエネルギーも低いいため、一度予測構造に含まれると定着しやすいことが要因と考えられる。この準実構造と一致した試行は全体の約3割を占め、特にシュードノット構造を形成するステム候補は50回の試行のほとんどの解に含まれていた。従って、この実験ではシュードノット構造を含む塩基配列に対して、妥当な予測結果が得られたものと考えられる。

表1: *Dunaliella parva* 18S rRNA の予測で抽出された解のエネルギーと抽出頻度

解の種類	エネルギー	抽出頻度
最小エネルギー解	-167.1kcal	0.02
準実構造	-163.5kcal	0.32
その他	(平均) -159.6kcal	0.66

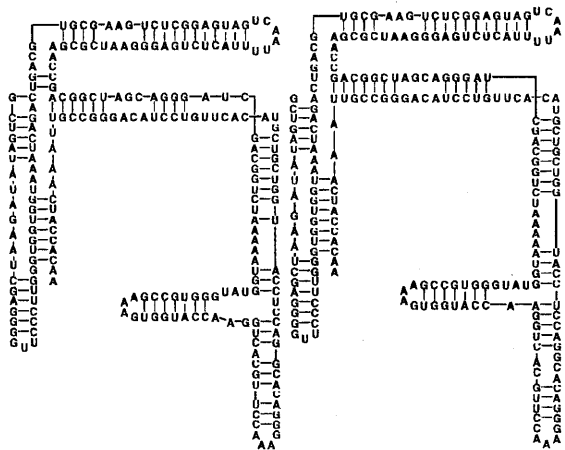


図7: *Dunaliella parva* の実構造(左)と準実構造(右)

### 4.2 *Tetrahymena thermophila* LSU rRNA

次に、同じくシュードノット構造を持つグループIイントロンの問題例として、*Tetrahymena thermophila* LSU rRNA を対象とする。実験で用いた配列の二次構造を図8(左)に、50回の予測結果における最小エネルギー解、準実構造、その他の解に対するエネルギー値と抽出頻度を表2に示す。

抽出された最小エネルギーの二次構造には、実構造に含まれるべき候補を全て含む他に、実構造では塩基長が1および2のステム構造しか存在しないため非結合塩基となる配列位置にステム構造が余計に含まれていた(図8(右)太点線で示す塩基の箇所)。予測構造の中にはこのような余計なステム構造や、そのステムが1塩基ずれたもの(図8(右)四角枠で示す塩基の箇所)など実構造と類似した解が複数種類抽出されており、その準実構造は全試行の3分の1で確認された。

エネルギー値が候補の着目順に依存することに加えて、G-U結合を考慮した結果候補数が多かったため、予測された最小エネルギー解の抽出頻度は低くなった。また、実構造に近い構造が頻繁に抽出されたが、これは実構造に含まれる候補が長い塩基長でG-C結合を多く含むエネルギーの低いものが多く、同じ塩基位置に競合する候補のない特徴的な候補であったためと考えられる。

#### 4.3 Bacteriophage T-4

次に解析する配列は、T4 nrdD の2767 ~ 2968塩基にある202塩基長のグループIイントロンである。シュドノット構造を含むこの部分の二次構造を図9(左)に示す。この実構造から3塩基長以下の候補を除いた構造は、抽出された最小エネルギー解に比べ約8kcal/mol近く高いエネルギー値になるため、50回の試行では抽出が確認されなかった。しかし、抽出された全ての解の随所で実構造に含まれるステム候補が頻繁に確認された。図10は個々のステム候補について、50個の解に含まれていた頻度を示したものである。グラフの縦軸は上が候補の出現頻度、下がそのエネルギーを示し、横軸に沿って出現頻度順に候補を並べた。濃い網掛けの候補は実構造と同じかほぼ同じものであり、それとは若干塩基長が違う候補を薄い網掛けで示した。このグラフからも、実構造に含まれる候補の出現頻度が高いことが分かり、非結合塩基や他の候補の影響によって出現しなかった候補は候補 $\alpha$ 一つだけであった(図9(左))。そこで、後処理として50回の試行にお

表2: *Tetrahymena thermophila* LSU rRNA の予測で抽出された解のエネルギーと抽出頻度

解の種類	エネルギー	抽出頻度
最小エネルギー解	-147.6kcal	0.02
最頻出解	-145.2kcal	0.16
準実構造	(平均) -144.6kcal	0.34
その他	(平均) -133.5kcal	0.66

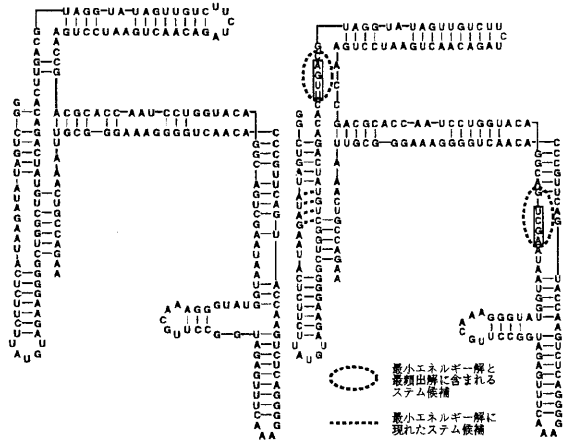


図8: *Tetrahymena thermophila* の実構造(左)と準実構造の一例(右)

表3: Bacteriophage T-4 の予測で抽出された解のエネルギーと抽出頻度

解の種類	エネルギー	抽出頻度
最小エネルギー解	-106.86kcal	0.02
最頻出解	-102.96kcal	0.18
準実構造	-98.81kcal	0.00
抽出された全解	(平均) -99.92kcal	1.00

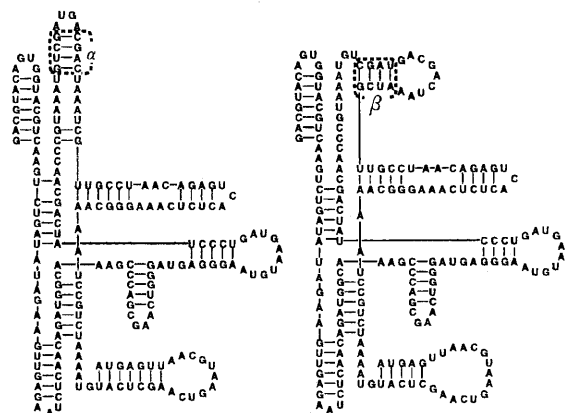


図9: Bacteriophage T-4 の実構造(左)と準実構造(右)

る候補の出現頻度に基づき、深さ優先探索によってその総和順に解の再構成・列挙を試みた。すると、頻度の総和が6番目に高い解として実構造に比較的近い準実構造(図9(右))が抽出された。この構造に含まれる候補 $\beta$ は実構造のステム $\alpha$ と位置が若干ずれている。これはステム $\alpha$ 付近に非結合塩基が多く、他の候補が頻繁に出現したことなどが原因と考えられる。この候補 $\beta$ も低いエネルギーを持ち、問題例の持つ特徴を示している。

このように、ステム候補の出現頻度に基づいて再構成された解は、問題例の特徴をよく表している。実構造に含まれるステム構造と一致する候補が低いエネルギーを持っているなど強い特徴を持っている場合、解を再構成することによって実構造を抽出できることが確認された。

## 5 おわりに

本稿では、塩基間の結合関係に平面的な制約を設け、シュードノット構造も対象としたRNAの二次構造予測を行う組合せ的な手法を提案した。計算機実験の結果から、シュードノット構造を含む妥当な予測結果が得られ、この制約が有効に働いていることが確認できた。今後の課題は、ステム候補の数を極力抑えつつ、長さが3塩基以下のステムまでも候補として考慮することである。実構造中にある長さ3以下の短いステムは、実験で出現頻度が高かった準実構造に含まれる候補と塩基配列上で競合しないので、より実構造に近い予測構造の抽出が期待できる。また、4.3の実験結果から示唆されるように、複数回試行をする必要があると考えられるので、それを効率良く実行する方法も検討する必要がある。これらの課題を解決することによって、大規模な問題にも適用ことができ、RNAの二次構造予測に有用な解析手法を提供できるものと考えられる。

**謝辞** 本稿執筆にあたり、研究に関する御指摘や資料等の提供をしてくださった電気通信大学情報工学科の小林聡助手に深謝いたします。

## 参考文献

- [1] 星田(編): 遺伝子情報への挑戦, 共立出版(1994).
- [2] 中谷, 山本, 米澤: 超並列計算機を用いたRNA2次構造の予測と視覚化, 情報処理, **37**, 10, pp.946-951 (1996).
- [3] 横森, 小林: 計算論的言語理論とDNA計算, 情報処理, **37**, 10, pp.929-934 (1996).
- [4] Y.Takefuji et al.: Discovering RNA interactions (Chapter8) in Neural Computing for Optimization and Combinatorics, World Scientific, pp.137-159 (1996).
- [5] M.Zuker, P.Stiegler: Optimal computer folding of large RNA sequence using thermodynamics and auxiliary information, *Nucl. Acids Res.*, **9**, 1, pp.133-148 (1981).
- [6] 秋山他: 対称相互結合型ニューラルネットを用いた大規模なRNAの二次構造予測, 信学技報NC90-62, pp.57-64 (1991).
- [7] 田中, 若月, 富田: 最大重みクリークの近似抽出解法によるRNAの二次構造予測, 信学大会D-10 (1995).
- [8] J.P.Abrahams et al.: Prediction of RNA secondary structure, including pseudoknotting, by computer simulation, *Nucl. Acids Res.*, **18**, 10, pp.3035-3044 (1990).
- [9] A.P.Gulyaev: The computer simulation of RNA folding involving pseudoknot formation, *Nucl. Acids Res.*, **19**, 9, pp.2489-2494 (1991).
- [10] J.Tinoco et al.: Estimation of secondary structure in ribonucleic acids, *Nature*, **230**, pp.362-367 (1971).
- [11] W.Salsler: Globin mRNA Sequence; Analysis of Base Pairing and Evolutionary Implications, *Cold Spring Harbor Symp. Quant. Biol.*, **42**, pp.985-1002 (1977).

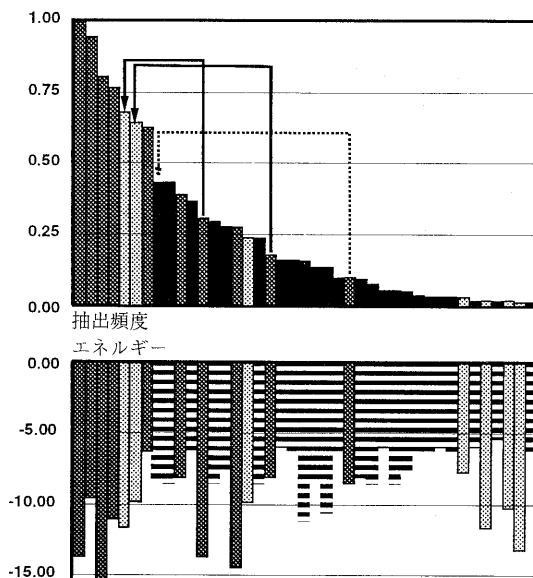


図10: 候補毎の抽出頻度とエネルギー