

H-031

## 自由エネルギーを用いた DNA ナノ構造のための塩基配列設計

## Sequence Design based on Free Energy for DNA tile structure

飯村直記† Naoki Iimura 山本雅人‡ Masahito Yamamoto 田中文昭† Fumiaki Tanaka 亀田充史† Atsushi Kameda 大内東† Azuma Ohuchi

## 1. はじめに

近年, DNA 分子が本来持つ会合能力を利用した自己組織的なナノ構造形成の手法や新たな計算手法として, DNA タイルが注目されている. DNA タイルによるナノ構造形成を行うためには, 個々のタイルが正確に会合する構造がより安定となるような塩基配列が必要であるが, 従来の研究では塩基の相補性のみを考慮した設計手法が提案されているに過ぎない.

本研究では9本の一本鎖 DNA から成り, 4方向の粘着末端(他の一本鎖 DNA と相補結合する一本鎖 DNA) をもった  $4 \times 4$  DNA タイルを対象とし, 自由エネルギーを用いた各一本鎖 DNA 間の会合予測を用いて, 部分的に SA (Simulated Annealing) を用いた最適化を行うことで, より効率的なタイル形成が可能となる塩基配列設計手法を提案する.

## 2. DNA ナノ構造と塩基配列設計

DNA の塩基には AGTC の4種類があり A と T, G と C が結合する性質(ワトソンクリック相補性)がある. DNA はこの性質が並列に作用することにより2本の一本鎖の DNA から螺旋状の1本の二本鎖 DNA を形成する. このことを会合もしくはハイブリダイゼーションという. Winfree らは複数の DNA を組み合わせて会合させることで DNA をタイル状に生成し, さらにそれらを入れ子に会合させることで平面のナノ構造を形成した[1]. 図1はその概念を表したものである. 左図は4本の DNA を会合させて出来た DNA タイルであり, 他のタイルと会合するために一本鎖 DNA が突出した粘着末端を有している. 右図はタイル同士が自律的に会合することで平面構造を形成していることを表している. また Winfree らは DNA タイルによってチューリングマシンと同様な計算が行なえることも示している[2].

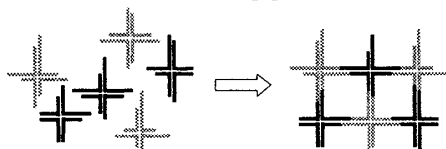


図1. DNA タイルによるナノ平面

Duke大学のHao Yanらは9本の一本鎖DNAから4方向の粘着末端をもった $4 \times 4$  DNA タイル(図2)を組み合わせて, タンパク質を規則的に配置したり, 銀のナノワイヤーを作りあげている[3]. これらは今まで作られて

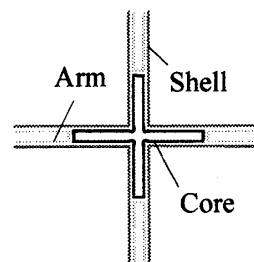
きたものより小さく有益なナノデバイスや回路などに応用が期待されている. このようなDNAナノ構造をより正確に形成するためには, 個々のDNAタイルが正確に会合しなければならない.

DNA の二次構造を利用した DNA タイルの配列設計についての先行研究として SEQUIN[4]があるが, この方法は塩基の相補性を考慮しているのみである. 望みの構造を生成するための DNA の塩基配列設計手法は, DNA コンピューティングの研究分野において多く研究されており, それらでは, 塩基の相補性のみならず, 二本鎖 DNA が一本鎖 DNA に乖離する温度(融解温度:  $T_m$ )や隣接する塩基の並びによって求められる自由エネルギー( $\Delta G$ )による配列の評価が重要であるとされている[5].

本研究では  $4 \times 4$  DNA タイルを対象とし, 自由エネルギーを用いた新たな塩基配列手法を提案する. 本手法によって, 個々の DNA タイルがより正確に会合し熱力学的に安定した構造になるような塩基配列を設計することが可能となる.

## 3. 提案手法

$4 \times 4$  DNA タイルは中心となる1本のCore, 他のタイルと入れ子となって結合する4本のArm, CoreとArmを支える4本のShellから成り立っており(図2), 合計9本の塩基配列を設計しなければならない. 配列設計にあたっては次の3点を考慮することが必要である.

図2.  $4 \times 4$  DNA タイル

1. タイル形成のための相補配列の結合が安定である.
2. 各々の配列が二次構造をとらない.
3. タイルの4本腕の会合の安定性が等しい.

2はDNA分子内の二次構造によって他のDNA間との分子間での会合が阻害されるのを減らすためである. また3は4本の腕がバランス良く会合していくことで, タイルが正確に形成されることを狙っている. 配列は  $\Delta G$  を用いて評価される.  $\Delta G$  が低いものほど安定した構造となるため, 会合してほしくない部分(Arm同士やCore同士, Shell同士, 会合することが望ましくない分子同士あるいは二次構造)の  $\Delta G$  は高くなるよう設計され, 会合してほしい配列同士は  $\Delta G$  が低くなるように設計されなければならない.

本研究では以下の3stepを繰り返し, 配列設計を行う.

†北海道大学大学院 情報科学研究科  
(札幌市北区北13条西8丁目)

‡CREST, Japan Science and Technology Corporation  
(埼玉県川口市本町 4-1-8)

- Step1.  $\Delta G$  を元に Core をある基準を満たすように生成する.
- Step2. Shell において Core と会合する部分配列に相補な塩基を割り当てる.
- Step3. 全ての配列を SA 用いて最適化する.

### 3.1 Step1 Core の設計

Hao Yan らの配列モデル[6]によると Core は図3に示されるように4方向に腕をもっている. 各々の腕の間にはTが4つつながったT4Loopがあり“関節”のような役割を果たすと同時に, 誤った会合による2次構造を減らす役割も果たしている.

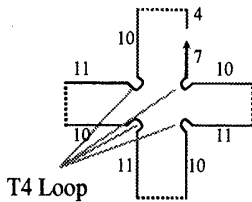


図3 Coreの概形  
(数字は塩基数を表す)

本研究の提案手法では, まずこの腕の11(あるいは10)塩基ずつをランダムに8本生成する. ただし, それぞれの腕とShellとが会合する力のバラつきを最小限にするために, 腕ごとのGとCの含量が50~60%に収まらないものは採用しない. また同じ塩基配列を含まないように重複のチェックも行なう. この作業を8本の塩基配列が完成するまで繰り返す. 次に生成した8本の塩基配列とT4Loopを図3のように繋いで1本のCoreを生成する. こうして生成されたCoreの $\Delta G$ を求め, ある閾値以下の塩基配列は2次構造をとりやすいと判断できるため採用しない. 最後にCoreとCoreを間に無限長に伸びる仮想塩基を挟んで結合させ $\Delta G$ を求める(Pairfold)[7]. Pairfold状態で求める $\Delta G$ は異なるDNA分子間の自由エネルギーとなる. ここでも閾値以下の塩基配列はCore分子同士で会合する可能性が高いので採用しない. こうして篩いに掛けた後に残った塩基配列をCoreとして採用する(図4).

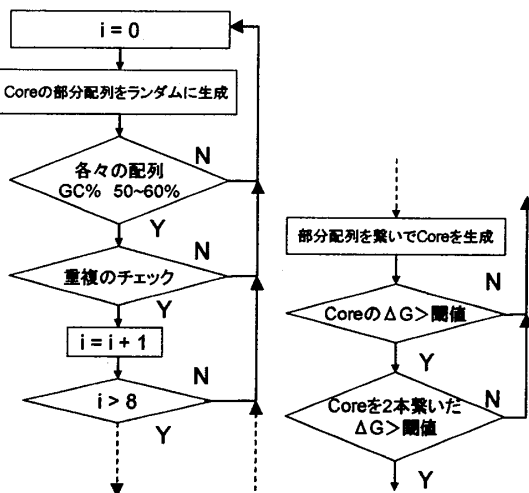


図4. Coreの設計手順

ここで閾値を定めるための予備実験として, この手法によって生成されたCore5000本の $\Delta G$  (25°C, [Na<sup>+</sup>] = 50  $\mu$ M 時)の分布を示すと, 図5のようになる. 図中の破線は $\Delta G$ が高いものから5%を, 実線は $\Delta G$ が高いものから10%を表している. 図5における $\Delta G$ の平均値は-8.15, 分散は9.8である. Hao Yan らがSEQUINを利用して生成したCoreの $\Delta G$ が-6.9であるから $\Delta G$ が高いものから10%以内であればHao Yan らのCoreよりも2次構造を取りにくいものが出来ることが予測される. 次に生成したPairfoldによるCore5000本の $\Delta G$ を同条件下で求め, 分布を示すと図6のようになる. このときの平均は-33.6である. 同様にHao Yan らの場合,  $\Delta G$ は-31.2であるから, 10%以内の配列であればCore同士の会合がおこりにくい配列ができるものと考えられる.

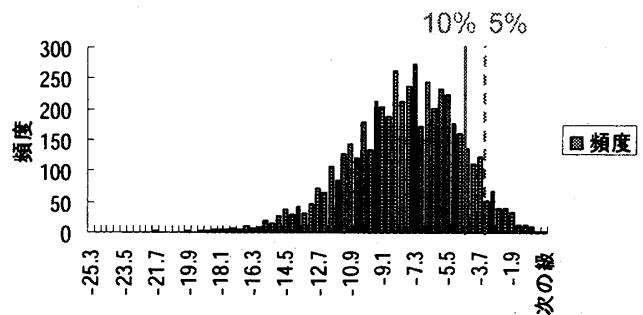


図5. Coreの $\Delta G$ 分布

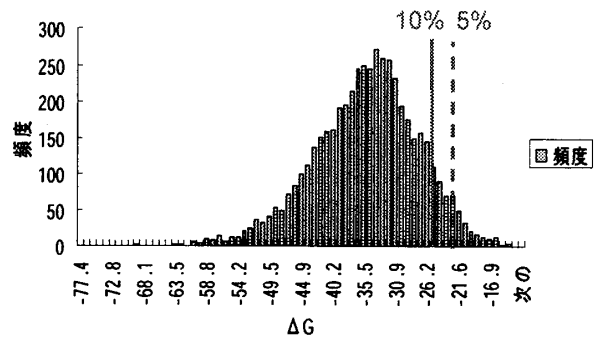


図6. Core2本によるPairfoldの $\Delta G$ の分布

### 3.2 Step2 Shell の部分配列決定

Step1を終えるとShellにおいてCoreと会合する部分(図5.S<sub>1M</sub>, ..., S<sub>4M</sub>)の塩基配列が決定する. ここではCoreと相補な塩基を割り当てる. CoreのGC含量が50~60%の間にあるので, ShellのGC含量も揃うことになる. そのためCoreとShellとが会合するときの安定性が整えられる.

### 3.3 Step3 SAによる最適化

Step3では, まずShellで塩基が割り当たっていない箇所(図5.S<sub>1F</sub>, S<sub>1B</sub>, ..., S<sub>4F</sub>, S<sub>4B</sub>)にランダムに塩基を割り当て, それに対応するArm(図5.A<sub>2L</sub>, A<sub>1R</sub>, ..., A<sub>1L</sub>, A<sub>4R</sub>)に相補塩基を充てる. その後全ての配列について, SAを用いて最適化を行う. 最適化の評価関数には一本鎖DNAが2次構造をとるときの自由エネルギー $\Delta G$ と, Pairfoldのときの自由エネルギー $\Delta G'$ を用いる. また, SAのサイクルによる選

移前のタイルをK, 遷移後のタイルをK'とする。K'は各配列において、ランダムに1塩基を選択し、AならTへ、GならCへ、TならAへ、CならGへとフリップさせた配列から形成されるタイルである。遷移は評価関数Eによって求められる確率Pで起こる。

$$\Delta E = E(K') - E(K)$$

$$P = \begin{cases} 1 & (\Delta E \leq 0) \\ \exp(-\Delta E/T) & (\Delta E > 0) \end{cases}$$

また、E(K)は以下の(1)-(3)式の線形和によって求められる。

K'に対して、

$$E(K) = \alpha E_1 + \beta E_2 + \gamma E_3 \quad (\alpha, \beta, \gamma < 0)$$

$$E_1 = \sum [\Delta G(S_j) + \Delta G(A_j)] + \Delta G(C) \quad (1)$$

$$E_2 = \sum \sum [\Delta G(S_j, \overline{S_j})] + \sum \sum [\Delta G(A_j, \overline{A_j})] + \Delta G(C, C) \quad (2)$$

$$E_3 = \sum [\Delta G(seq, \overline{seq}) - \min \Delta G(seq, \overline{seq})] \quad (3)$$

$\Delta G(x)$ : 配列xが二次構造をとるときの $\Delta G$

$\Delta G(x, y)$ : 配列xと配列yのPairfoldのときの $\Delta G$

$(seq, \overline{seq})$ : 本来会合する配列のペア (seqは $S_j, A_j, C$ のいずれか)

$(seq, \overline{seq})$ : 本来会合しない配列のペア (ただし,  $seq \neq \overline{seq}$ )

$$S_j = \{S_{jB}, S_{jM}, S_{jF}\}$$

$$A_j = \{A_{jR}, A_{jL}\}$$

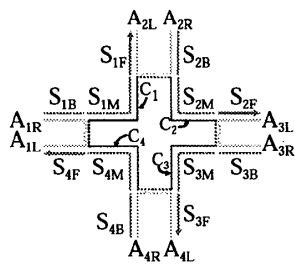


図5. SAによる最適化

(1)式は各々の配列が二次構造をとるときの $\Delta G$ の和であり、値が大きいものほど二次構造をとりにくく良い配列セットであるといえる。

(2)式はShell同士, Arm同士, Core同士によるPairfoldの $\Delta G$ の値の和である。これらの分子同士は会合してはならないので値の大きいものほど良い。

(3)式は本来会合すべき配列同士のPairfoldの $\Delta G$ とそれ以外の(会合し合わない)配列同士のPairfoldの $\Delta G$ の差の和をとったもので、この和が大きいものほど本来会合すべき配列同士が会合しやすくなるので、誤った会合が起こりにくい配列だといえる。

評価関数全体としては、値の小さいものほど、タイルをより正確に形成する配列であると評価できる。

上記の式による評価関数を用いて、DNAタイルの評価を行う予備的実験を行なった。図7はstep1で生成したCoreを用いて、step3でSAによる最適化を行なう前の配列セットを対象とし、評価関数を用いた安定性の評価がどのような分布をとるか示したものである( $\alpha = \beta = \gamma = -1$ )。平均が359.3なのに対し、Hao Yanらの配列の場合は370.1となる。

図7から、先行研究で生成された配列よりも安定なタイル形成を可能とする配列が存在することがわかる。またSAによる最適化を行うことにより、よりこの評価関数が良い配列を設計できると考える。

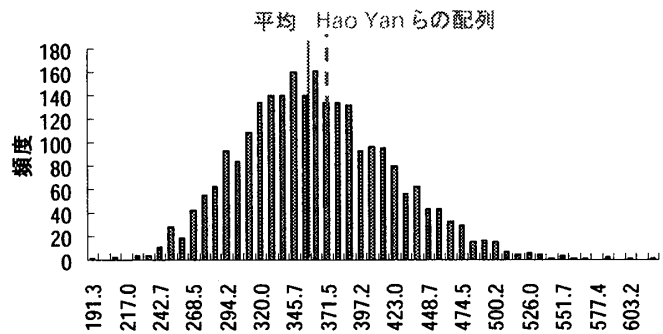


図7. ランダムな配列に対する評価関数の分布

### まとめ

本研究では $4 \times 4$  DNAタイルにおいて、個々のタイルが正確に会合する構造がより安定となる塩基配列の設計手法を提案した。この手法では自由エネルギーを用いた配列の評価によって、従来研究による配列設計よりも熱力学的に安定したタイル形成が可能となると考えられる。また、複数種のタイルを必要とする場合などにより有用となると思われる。

今後は本研究で生成された配列によってできるタイルやナノ構造が効率良く生成されていることを、化学実験やAFMによって検証しなければならない。

### 参考文献

- [1] Erik Winfree, Furong Liu, Lisa A. Wenzler and Nadrian C. Seeman, "Design and self-assembly of two-dimensional DNA crystals", Nature, Vol394, 539-544 (1998)
- [2] Erik Winfree, "DNA computing by Self-assembly", in NAE's The Bridge 33(4):31-38, 2003
- [3] Hao Yan, Sung Ha Park, Gleb Finkelstein, John H. Reif, Thomas H. LaBean, "DNA-Templated Self-Assembly of Protein Arrays and Highly Conductive Nanowires", Science, vol301, 1882-1884 (2003)
- [4] Nadrian C. Seeman, "DeNovo Design of Sequence for Nucleic Acid Structural Engineering" Journal of Biomolecular Structure & Dynamics, ISSN0739-1102, Vol8 (1990)
- [5] Fumiaki Tanaka, Atsushi Kameda, Masahito Yamamoto, Azuma Ohuchi, "Design of nucleic acid sequence for DNA computing based on a thermodynamic approach", Nucleic Acids Research, 2005, Vol33, No3, 903-911
- [6] Supporting Online Material <http://www.sciencemag.org/cgi/data/301/5641/1882/D1>
- [7] Mirela Andronescu, Rosalia Aguirre-Hernandez, Anne Condon, Holger H. Hoos, "RNAsoft: a suite of RNA secondary structure prediction and design software tools", Nucleic Acids Research, 2003, vol.31, No13, 3416-3422 (2003)