

令和 6 年 6 月 17 日現在

機関番号：12701

研究種目：基盤研究(C)（一般）

研究期間：2020～2023

課題番号：20K06504

研究課題名（和文）データサイエンスを導入した原子間力顕微鏡による四重鎖DNA検出法の開発

研究課題名（英文）Development of AFM-based quadruplex DNA detection method using data science technology

研究代表者

田中 陽一郎（Tanaka, Yoichiro）

横浜国立大学・研究推進機構・技術専門職員

研究者番号：70450426

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：本研究では、四重鎖構造を直接検出可能な測定法として原子間力顕微鏡（AFM）を使用し、基板上に固定した多数のDNA分子の粒子形状を測定して各分子の表面形状パラメータを取得した。機械学習によって四重鎖DNAに特徴的なパラメータを同定することで、AFM測定によって高速に四重鎖を判別する新たな測定法を確立した。

確立した四重鎖判別法によって、大腸菌ゲノムに存在する未知の四重鎖DNA分子の評価を行った結果、本方法で四重鎖と判別されたDNAは、NMRや電気泳動実験で実際に四重鎖を形成することが確認できた。これらの成果から、AFM測定によって高速に四重鎖を検出する方法を構築した。

研究成果の学術的意義や社会的意義

近年、DNAの特殊な構造の一つであるグアニン四重鎖構造がヒトゲノム上に多数存在し、疾病の治療ターゲットになり得る事が分かってきた。どのような配列が実際に四重鎖を形成し、疾病と関連しているかを解明することが喫緊の課題である。そのためには、四重鎖形成が推定される多数のDNA配列を高速に解析し、実験的に四重鎖構造の形成を判別する必要がある。本研究では、比較的安価に原子レベルの形状情報を得られる原子間力顕微鏡（AFM）を使用し、DNAの四重鎖形成を高速に同定可能な方法を開発した。得られた結果は実験的なデータに基づいており、疾病の原因となる四重鎖の同定等に高い能力を持つことが期待される。

研究成果の概要（英文）：In this study, we used atomic force microscopy (AFM) as a measurement method that can directly detect quadruplex structure, and measured the particle forms of many DNA molecules fixed on a mica surface to obtain the surface shape parameters of each molecule. By identifying the unique parameters of quadruplex DNA by machine learning, we established a new measurement method to discriminate quadruplexes at high speed by AFM measurement. As a result of the evaluation of unknown quadruplex DNA in the E. coli genome using the established quadruplex discrimination method, it was confirmed that the DNA detected as a quadruplex by our method formed a quadruplex by NMR and electrophoresis experiments. Based on these results, we constructed a method to detect quadruplexes rapidly by AFM measurement.

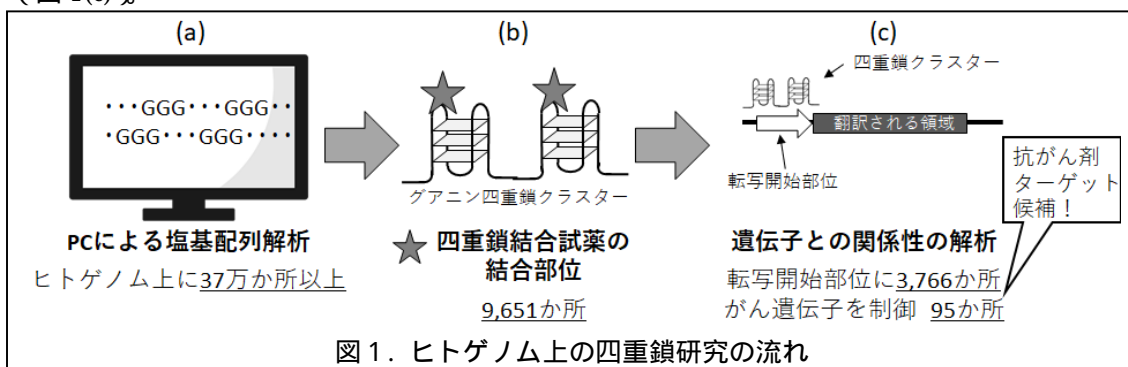
研究分野：構造生物化学

キーワード：グアニン四重鎖 AFM データサイエンス 機械学習 NMR

1. 研究開始当初の背景

生物において、遺伝情報を記録する DNA は、通常は二本鎖による二重らせん構造を形成している。しかし、DNA の配列によっては異なった構造を形成し、例えば、連続したグアノシン残基 (G) の繰り返し配列は四重鎖構造となる場合がある。代表的な四重鎖 DNA としては、染色体末端のテロメア領域の (TTAGGG) n のような繰り返し配列が挙げられる¹⁾。また、がん化に関連が深い「がん遺伝子」である c-myc 等では、遺伝子発現を制御する転写開始部位に四重鎖が存在することが報告され、四重鎖 DNA を標的とした抗がん剤の開発が期待されている²⁾。さらに、難病の ALS (筋萎縮性側索硬化症) でも、四重鎖が発症の鍵となることが分かってきた³⁾。このように、四重鎖は生物学的に重要な役割を持ち、疾病の原因となることを示唆する例が次々に明らかになっている。

従来、ヒトゲノム上には大量の四重鎖候補配列の存在が推定されていたが (図 1(a))、2018 年に、四重鎖に結合する試薬「7OTD」の結合部位がゲノム上の約 1 万か所に存在する事が示唆された⁴⁾ (図 1(b))。さらに、その多くは転写開始部位でがん遺伝子を制御していると推定された (図 1(c))。



しかし、ゲノム上の四重鎖候補配列はそれぞれ異なり、四重鎖構造は複数の種類が存在する。四重鎖結合試薬がどの程度正確に四重鎖を認識しているかは分かっておらず、その信頼性は高くない。このため、四重鎖が生物学的に、又は病気の治療標的として重要な役割を持っている例が明らかになり、しかもゲノム上に大量にあることが示唆されているものの、大量の四重鎖候補配列から、実際に四重鎖構造を形成している配列を決定し、真に重要な四重鎖をスクリーニングする必要がある。

現在の主な四重鎖の解析法は、ゲノムから推定する方法では網羅的に行えるが実験的な裏付けが必要になり、NMR や X 線結晶構造解析による分子構造を元にした方法は確実だが手間と時間がかかるなど、いずれの方法も一長一短があり、高い確実性で、しかも高速に四重鎖の形成と機能を解析する方法は存在しない。大量の四重鎖候補配列から、生物学的に重要な四重鎖を同定するのは困難である。これを踏まえ、本研究課題では、ゲノム上に多数存在する四重鎖候補配列を、高速で、しかも分子構造を元に確実に解析する方法の開発を行った。

2. 研究の目的

四重鎖構造を高い確実性をもって解析するには、実際に分子構造を基にした情報を測定することが重要である。本研究では、構造情報を基に四重鎖 DNA の特徴を抽出し、タンパク質との相互作用を高速に解析する新たな方法の開発を目的とした。そのための方法として、原子間力顕微鏡 (AFM) に着目した。AFM は、微小なカンチレバーによって物質の表面形状を測定する装置で、次のような特徴がある。

1. 分子の立体構造情報を持つ 1 分子ごとの大きさや高さの情報を得られる。
2. 溶液中で容易に測定可能で、DNA とタンパク質の相互作用解析も可能である。
3. 一回の測定は数分程度で、一度に数十～数百分子の情報を得られ、高速である。

しかし、AFM で詳細な DNA の構造を認識する解像度の実現は未だ困難で、単純な形状情報 (高さ、面積等) からの四重鎖形成の判定は誤差が大きい。

この問題の解決のために、研究代表者らは、複雑なデータを統計的に解析して有用な規則や判断基準を明らかにするデータサイエンスの技術に着目した。研究分担者 / 連携研究者の荻野らが開発した、AFM データの解析に機械学習を取り入れて生体粒子「エクソソーム」の形態を解析する方法⁵⁾を応用し、DNA を AFM で 1 分子ずつ測定して機械学習を行い、四重鎖と二本鎖を高精度に判別する方法を確立することを研究の目的とした。

3. 研究の方法

一本鎖、二本鎖または四重鎖を形成する 15 残基前後の DNA 分子を buffer (50 mM KCl、10 mM MgCl₂、20mM HEPES(pH7.0)) で 50 μM に溶解し、95 °C で 5 分間加熱後、12 時間で 15 °C まで徐冷して DNA の構造を安定化した。各 DNA を測定直前に buffer で 10 μM に希釈

し、NiCl₂を10mMになるように加えて剥離したばかりのマイカ基板に固定した。基板表面をH₂Oで洗浄後、乾燥してAFMで吸着粒子の形状像を測定した。各DNA粒子の形状データから、AFMデータ解析ソフトウェア「Gwyddion」を用いて粒子ごとに41種類の形状変数を抽出した。

各DNA粒子から得られた形状変数を特徴量として機械学習によって解析し、一本鎖、二本鎖または四重鎖DNAを識別する分類器を作成した。さらに、各粒子の識別のために重要度が高い数種類の形状変数を抽出した。

次に、大腸菌ゲノム上で遺伝子発現調節部位近傍に存在する、四重鎖を形成すると推定される配列を持つDNAを合成し、四重鎖構造の判別に寄与が高い数種類の形状変数を使用して評価を行った。同時に、同じDNAについてNMRによってイミノプロトンスペクトルを測定し、四重鎖を形成しているかを確認した。また、非変性アクリルアミドゲル電気泳動(PAGE)によってバンド位置に変化があるかを確認し、同時に四重鎖を特異的に染色するチオフラビンT染色によって、四重鎖の形成を確認した。

さらに、四重鎖の形成によってGFPタンパク質の発現を抑制可能なレポータープラスミドを導入した大腸菌を調製し、AFMによって四重鎖を形成すると判定した大腸菌由来のDNA配列が、大腸菌中で実際に四重鎖を形成するかどうかを判定した。

4. 研究成果

(1) AFMによる各DNAの測定と四重鎖を判定するパラメータの同定

2種の本鎖DNA、1種の二本鎖DNA、4種の四重鎖DNAをそれぞれ基板上に固定後、AFMによって高解像度に測定し、表面形状データから各DNA分子の粒子としての形状変数を抽出した。得られた形状変数を機械学習によって解析し、各DNA粒子を識別するRFによる分類器を作成した結果、各粒子を87.1%の精度で識別可能という結果が得られた。この分類器で重要度の高い形状変数を検討した結果、DNA構造を強く反映している数種類を得ることができた。これらの形状変数は、四重鎖構造を形成するDNA粒子と一本鎖、二本鎖となる粒子とでは顕著に値が異なる事が分かった。AFMによる四重鎖構造の測定では、分子の高さ方向の値がある程度異なる事が知られていたが、今回判明した形状変数によって評価を行うことで、より高い精度で四重鎖を識別できる事が分かった。

(2) 大腸菌ゲノム上の四重鎖と推定されるDNAの解析

大腸菌ゲノム上に存在する、四重鎖構造を形成する可能性が高い15~17残基の6種類のDNA配列についてAFMで測定を行った。前項で得られた形状変数によって各DNA粒子を評価したところ、そのうち2種類で四重鎖を形成することを示す結果が得られた。

非変性PAGEによるチオフラビンT染色では、上記の方法で四重鎖であることが示された2種のDNAのみが染色され、四重鎖を形成する結果が得られた。また、NMRによってイミノプロトンスペクトルを測定した結果、上記2種のDNAは四重鎖を示すシグナルが得られたが、AFMによる方法で四重鎖ではないとされた2種のDNAもシグナルが得られた。NMR測定は高濃度で安定した条件で測定を行うため、安定性が比較的低いDNAも四重鎖を形成すると考えられる。このため、本研究で開発したAFMによる四重鎖判別法は、比較的高い安定性が高い四重鎖の形成を判別することがわかった。

(3) 四重鎖を形成するDNAの大腸菌内での四重鎖形成の判定

次に、AFMによる判別法で四重鎖と判定された2種のDNAを、プロモーター部位に導入したレポータープラスミドを大腸菌に導入し、12時間培養後の大腸菌濃度あたりのGFPの蛍光を測定した。四重鎖を導入したプラスミドでは、導入しないプラスミドと比較して1/2または3/4程にGFP発現量が抑制されることがわかった。さらに、GFPの安定性に大きく寄与するカリウムイオンをトラップするクラウンエーテルや、カリウムイオンと競合して四重鎖を不安定化する塩化リチウムを加えて培養したところ、四重鎖による発現抑制が解除され、GFP発現量が増加した。これらの結果から、AFMによって四重鎖であると判定された配列では、大腸菌中でも四重鎖を形成することがわかった。

これらの結果から、本研究で開発したAFMによる四重鎖判別法は、大腸菌ゲノム中の四重鎖候補配列から、実際に四重鎖を形成するDNAを判別する方法として使用可能である事がわかった。今後は、本研究で開発した四重鎖判別法が様々な生物種に適用が可能かを検証し、より汎用的な実験法として確立する予定である。

<引用文献>

1. Cech, T.R., *Cell*, **116**, 273 (2004)
2. Todd, A. K. et al., *Nucleic Acid Res.* **33**, 2901 (2005)
3. Haeusler, A.R., *Nature*, **507**, 195 (2014)
4. Yoshida, W. et al., *Sci. Rep.* **8**, 3116 (2018)
5. Ogino, T. et al., *J. Phys. Chem. B.*, **122**, 6224 (2018)

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計9件（うち招待講演 3件 / うち国際学会 6件）

1. 発表者名 Tanaka Y., Nagata T., Sugimoto, C., Ogino T.
2. 発表標題 Identification of quadruplexes that can regulate gene expression.
3. 学会等名 The 13th International Symposium of Advanced Energy Science (国際学会)
4. 発表年 2022年

1. 発表者名 Tanaka Y., Nagata T., Sugimoto, C., Ogino T.
2. 発表標題 Development of AFM-based G-quadruplex DNA detection method using data science technology.
3. 学会等名 15th International Symposium on Nanomedicine (ISNM2022) (招待講演) (国際学会)
4. 発表年 2022年

1. 発表者名 Tanaka Y., Nagata T., Ogino T.
2. 発表標題 Identification of quadruplexes that can regulate gene expression.
3. 学会等名 The 12th International Symposium of Advanced Energy Science (国際学会)
4. 発表年 2021年

1. 発表者名 田中 陽一郎、永田 崇、杉本 千佳、荻野 俊郎
2. 発表標題 原子間力顕微鏡による機械学習を使用した簡便な四重鎖DNA判別法の開発
3. 学会等名 第44回日本分子生物学会年会
4. 発表年 2021年

1. 発表者名 Y. Tanaka, T. Nagata, and T. Ogino
2. 発表標題 Elucidation of correlation between G-quadruplex-based gene expression regulation and their structure.
3. 学会等名 The 11th International Symposium of Advanced Energy Science (国際学会)
4. 発表年 2020年

1. 発表者名 荻野俊郎, 田中陽一郎, 杉本千佳, 伊藤和希
2. 発表標題 データサイエンスを活用した生体物質の特徴抽出
3. 学会等名 応用物理学会東海支部基礎セミナー (招待講演)
4. 発表年 2020年

1. 発表者名 Tanaka Y., Nagata T., Sugimoto, C.
2. 発表標題 Identification of quadruplexes that can regulate gene expression
3. 学会等名 The 14th International Symposium of Advanced Energy Science (国際学会)
4. 発表年 2023年

1. 発表者名 Chika Sugimoto, Yoichiro Tanaka, Toshio Ogino
2. 発表標題 AFM-based G-quadruplex Detection Using Machine Learning
3. 学会等名 45th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC2023) (国際学会)
4. 発表年 2023年

1. 発表者名 荻野 俊郎、田中 陽一郎、杉本 千佳
2. 発表標題 応用電子物性と周辺分野、工学、そして異分野との融合
3. 学会等名 第71回応用物理学会春期学術講演会（招待講演）
4. 発表年 2024年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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
研究分担者	永田 崇 (Nagata Takashi) (10415250)	京都大学・エネルギー理工学研究所・准教授 (14301)	
研究分担者	杉本 千佳 (Sugimoto Chika) (40447347)	横浜国立大学・大学院工学研究院・准教授 (12701)	
研究分担者	荻野 俊郎 (Ogino Toshio) (70361871)	横浜国立大学・研究推進機構・非常勤講師 (12701)	2020 - 2022年度

	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
研究協力者	荻野 俊郎 (Ogino Toshio)		2023年度

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
---------	---------