

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 6 日現在

機関番号：82110

研究種目：若手研究(B)

研究期間：2014～2015

課題番号：26840054

研究課題名(和文)高選択性人工転写因子の開発

研究課題名(英文)Designing a highly specific transcription factor

研究代表者

角南 智子(Sunami, Tomoko)

国立研究開発法人日本原子力研究開発機構・原子力科学研究部門 量子ビーム応用研究センター・研究副主幹

研究者番号：50554648

交付決定額(研究期間全体)：(直接経費) 2,100,000円

研究成果の概要(和文)：ゲノム中の任意の塩基配列に高選択的に結合する分子を作製する事ができれば、ゲノムDNAの編集や修飾、イメージング等多くの応用が考えられる。我々は、既存のゲノム編集タンパク質とは異なる性質を有する新たな分子を設計するために、Engrailed homeodomainタンパク質に注目し、B1H法を適用してその認識配列長を倍加しうるリンカーを同定することに成功した。一方、タンパク質とDNA間の相互作用に重要なDNA構造の柔らかさをより正確に調べるための手法として、X線結晶構造の電子密度マップからDNAの主鎖のゆらぎの情報を抽出する手法の開発を行った。

研究成果の概要(英文)：Our purpose is to develop proteins with properties different from widely used genome-editing proteins such as TALEN and CRISPR. In this study, we rationally designed novel proteins based on engrailed homeodomain and confirmed the binding specificity using bacterial one-hybrid assay. We successfully designed molecules with a strong affinity and high specificity for the target sequence which has two engrailed homeodomain binding sites. Also, we developed a method to evaluate molecular flexibility of DNA using electron density map of X-ray crystal structures.

研究分野：分子生物学、構造生物学

キーワード：転写因子 タンパク質・DNA相互作用 Bacterial One-Hybrid法

1. 研究開始当初の背景

転写因子をはじめとする DNA 結合タンパク質と DNA との相互作用は生命活動の最も基本的なステップである。ゲノム中の任意の塩基配列に高選択的に結合する分子を人工的に作製する事ができれば、ゲノムの編集や修飾、イメージング等多くの応用が考えられる。このような分子として、現在、TALEN、CRISPR といった分子が特にゲノムを編集する用途で広く使われ始めている[1]。しかし、既存の転写因子には依然として選択性が不十分であったり、活性があまり高くなかったりという問題点がある。また、分子量が大きすぎるため、活性・選択性を向上させるためにリピート領域を多く繋ぐことも現実的ではない。

もしも、分子量が小さく、活性・選択性の面で既存の転写因子を上回る分子を創成することができれば、特に現在困難であるとされているイメージングの領域で大きな発展が期待できる。このような分子を用いて、ある特定の DNA 領域が細胞内でどのような動態を有しているのかに関して、詳細な情報を得ることが期待できる。

2. 研究の目的

本研究では、最終的に既存の転写因子よりより活性や選択性が向上した分子を創成するために、2つの着眼で目的を設定した。

(1) 一つは、既存のゲノム編集酵素よりも分子量が小さいタンパク質を骨格に用い、実験的なアプローチによって新たなゲノム編集酵素を開発することである。本研究では骨格として Engrailed homeodomain タンパク質(以下 En と略す)に注目した。このタンパク質は約 80 残基と小型ながら、6bp の認識配列に単量体で非常に強く結合する性質がある。更に、このタンパク質は変異体やホモログの網羅解析によって、様々な配列を認識可能な変異体を作成可能であることが既にわかっている。我々は、このタンパク質を利用してリピート構造を作成することで、ゲノム中の単一の箇所を認識できるだけの活性・選択性が高い分子が創成できるのかどうかを調べることが重要であると考えた。

(2) もう一つは、DNA を計算機上でより正確に取り扱う方法を開発することである。本研究の範囲では、DNA の選択性に重要な役割をはたすと言われている DNA の構造の柔らかさを既存のデータから抽出する方法論の開発を行うことに注力した。このような方法は、将来的には、DNA とタンパク質との相互作用を予測する方法論の開発に繋がると考えられる。

3. 研究の方法

我々は一番目の着眼では、Meng らによって報告された Bacterial One-Hybrid 法 (B1H 法) [2] を適用することで DNA への選択性を調べることにした。人工転写因子の研究では、設計したタンパク質が期待通りの塩基配列特異性を有することを実験的に確認する必要がある。塩基配列特異性の測定には SELEX と呼ばれる手法が広く使われているが、手間・コストが大きく、サイクルを回すことに多くの時間がかかってしまう問題点がある。そこで本研究では、B1H 法を塩基配列特異性検討のために採用した。この方法は、大腸菌を用いて結合配列を同定する方法であるが、手間・コストがそれほど小さくなく、その上、生体内で配列特異性を検討するため、*in vivo* での配列特異性を同定できる点に大きな利点がある。なお、*in vitro* での活性選択性の評価は EMSA 法を用いて行った。

二番目の方法論の開発では、PDB データベース中の既知の DNA 構造を統計的に処理することで、DNA がどの程度の構造ゆらぎを有しているのかを調べた。電子密度図の計算等には、Phenix や CCP4 プログラムパッケージを用い、統計計算には Python の Pandas ライブラリ等を用いた。

4. 研究成果

Engrailed homeodomain タンパク質のリピート構造作成について

本研究では、2つの En を様々な長さのリンカーで繋いで、野生型の倍長の 12bp の配列を特異的に認識可能な分子を作製できるかどうかを検討した。

(1) モデリング

まず、ホモロジーモデリングによってどのようなリンカーを用いれば2つのドメインが DNA にタンデムに結合しうるのかを計算機上で推定した。同時に、ドメイン間の立体障害の可能性について検討した。得られたモデルを図 1 に示す。我々はこのモデルから、リンカーとして数残基程度のフレキシブルリンカーを用いれば、ドメイン間に立体障害を生じることなく目的とする 12bp 長のターゲット配列との結合が可能であると考えた。

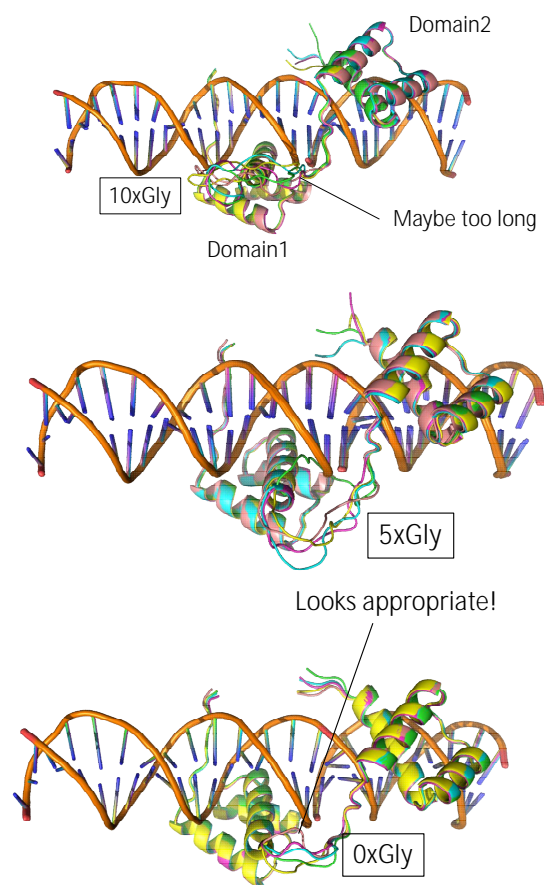


図1. タンデム型 Engrailed homeodomain の配向

(2) 塩基配列特異性の評価

コントロール実験として単量体の En を用いて B1H 法の構築を行った。種々の大腸菌株を検討して、ライブラリ作製および結合配列の選択に適した菌株を同定した。これらの菌株を用いて、28bp 長のランダム配列から成るライブラリを作成し、B1H 法で塩基配列特異性を検討した。その結果、野生型および既知の変異体(Q50K)の双方で、今までの報告同様の配列特異性が得られることを確認した。

この系を用いて、我々は、Q50K 変異体同士を Gly₁₀, Gly₅ リンカーで繋いだもの、および、直接繋いだもの(Gly₀)、更に、2 番めのドメインの最初の 3 残基を削って繋いだもの(ΔDEK)の 4 種類のタンパク質を作成し、これらのタンパク質がどのような配列に結合するのかを調べた。実験の概要と結果を図 2 に示す。この図は、作成したタンパク質が、ランダム配列に比べて 12bp 長のターゲット配列とどの程度特異的に結合しやすくなったのか(P/P_{rand})を示している。モデリングから期待したとおり、Gly₀, Gly₅ 配列がリンカーとして適切であることを実験的にも裏付けることが出来た。

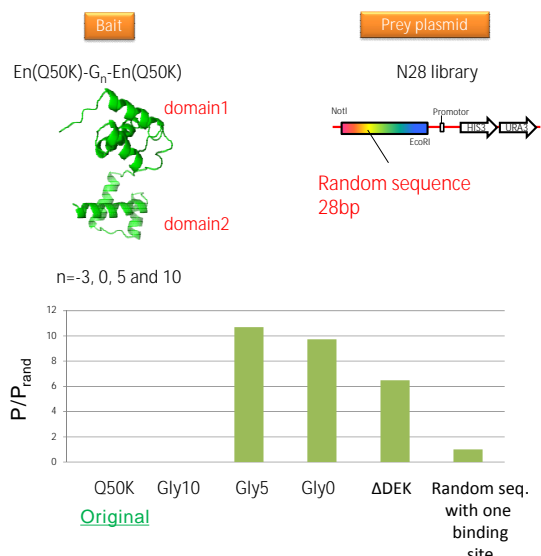


図 2. B1H 法と検討したリンカーの長さの効果

(3) *in vitro* での結合定数の評価

B1H 法では結合活性についての定量的な情報が得られないため、我々は(2)で塩基配列特異性を調べたタンパク質を大腸菌にて発現・精製し、EMSA 法を用いて解離定数を測定した。結果の一部を Table1 に示す。リンカーで繋いだタンパク質はいずれも目的配列に対して元のタンパク質の 10 倍と極めて高い結合活性を持っており、特に Gly₅ リンカーの場合には、その結合活性は非常に良好であった。更に、このアッセイの条件下でも、リンカーで繋いだタンパク質は目的配列に対して高い選択性が確認できた。

Table1. タンパク質・DNA 間の解離定数

D ⁺ N A	Protein	Protein		
		Original En(Q50K)	Gly ₅	Gly ₁₀
TAATCC		0.49 ± 0.10 nM	-	-
(TAATCC) ₂		-	0.05nM	0.12 ± 0.05 nM
Specificity**		3 * 10 ³	2 * 10 ⁵	9 * 10 ⁴

* TAATCC: d(CGCAGTG-TAATCC-CCTCGAC)
(TAATCC)₂: d(CGCAGTG-TAATCC-TAATCC-CCTCGAC)

** Calculated by apparent K_d with or without Calf thymus DNA.

本研究で得られタンパク質は依然として、目的以外の配列にも結合する活性があり、改良の余地は大きい。しかし、本研究によって、Engrailed homeodomain を骨格に用いて分子設計を行うことで、現在使われているTALENやCRISPRとは異なる選択肢を与える可能性を示すことが出来た。

DNA 構造ゆらぎの統計解析

タンパク質と DNA との塩基配列特異的な相互作用には、DNA とタンパク質間の直接的な相互作用だけではなく、DNA それ自体がどの程度構造変化を起こしやすいのかが重要な役割を果たしている[3]。今までに数多

くの DNA の結晶構造が得られているものの、DNA のどのような配列・どのような構造がどの程度柔らかいのか（どの程度ゆらぎやすいのか）を網羅的に検証した研究はほとんどない。

本来、結晶は多数の分子の集まりであるため、その X 線結晶構造で得られる回折データには DNA が持つゆらぎの情報が含まれていることが期待される。しかし、今までの解析では結晶構造は静的であると考えられてきたため、その中の動的な情報が見落とされてきた可能性が高いのではないかと考えた。

我々は結晶中の DNA のゆらぎの情報を取り出すために、FoFc 差フーリエマップに注目した。このマップは、測定された回折データと、モデルから予測される回折データとの差であるため、今までに見逃されてきた DNA 構造のゆらぎを直接反映していると考えられる(図 3)。この電子密度マップを網羅的に再解析することで、今までに見落とされてきた分子のゆらぎに関する情報を集めてくることが可能になると考えた。

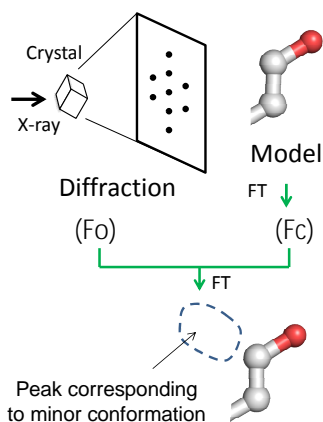


図 3. FoFc 差フーリエマップと分子のゆらぎ

まずは、差フーリエマップ上でどの程度の強度のピークを同定すれば、ノイズではなくシグナルを同定可能であるかを調べるために、DNA の代表的な 3 つの構造を選んで、構造因子に適切なガウシアンノイズを加えて差フーリエマップを計算することで、電子密度図上でノイズの強度と存在確率を推定した。分解能は 1.5Å とした。結果を図 3 に示す。この図から、3.5σ をクライテリアに用いた場合には、ノイズ由来のピークがある原子の周りにたまたま生じる確率は 5%程度と十分に低いことが分かった。また、その確率は原子種によらなかった。

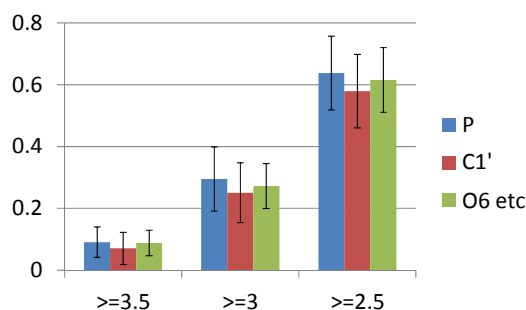
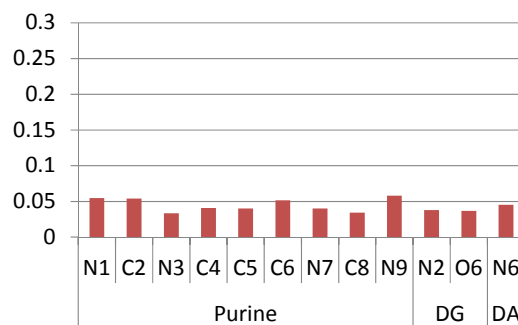
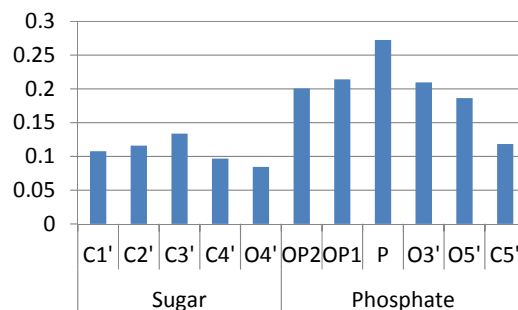


図 4. シミュレーションで得られた FoFc 差フーリエマップ上のノイズの確率

次に、我々は、PDB データベースから 1.5Å 以上の分解能を有する DNA の結晶データを入手した。そして、1.5Å 分解能での FoFc 差フーリエマップを計算した。

そして、DNA の各原子の周りの 3.5σ 以上のピークの存在確率を調べたところ、リン酸や糖の周辺にシミュレーションで推定したよりも高い頻度でピークが観察されることが分かった(図 5)。この方法を用いれば、DNA の主鎖のゆらぎの情報が抽出できることを示している。また、特に、約 25%ものリン原子周辺でピークが観察されたことは興味深い。この事実は、結晶中であっても、DNA の主鎖は今まで考えられていた以上に揺らぎやすいことを示している。



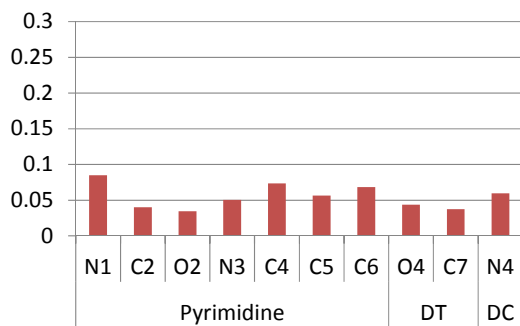


図5. DNA の各原子周辺での FoFc 差フーリエマップ上のピークの割合

そこで、我々はリン原子周辺のピークの位置と頻度を詳細に分析した。図6に結果を示す。この図の ABCD はピークがリン酸基のどの3つの酸素原子が近いのかに対応している。見つかるピークの位置や頻度は DNA の構造によって全く異なっていることがわかった。特に、他の構造と比較して、Z-DNA の RpY ステップではピークの存在確率が高く、更に、ピークの位置が ZI では A (major groove 側), ZII では C(minor groove 側)に局在していることが明らかになった。このピークの位置が、ZI 型と ZII 型との間の構造ゆらぎに相当していることを確認した(図7)。

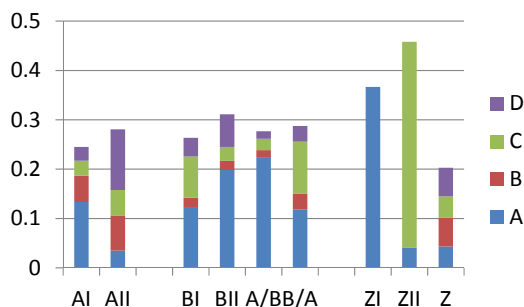


図6 DNA の各原子周辺での FoFc 差フーリエマップ上のピークの割合

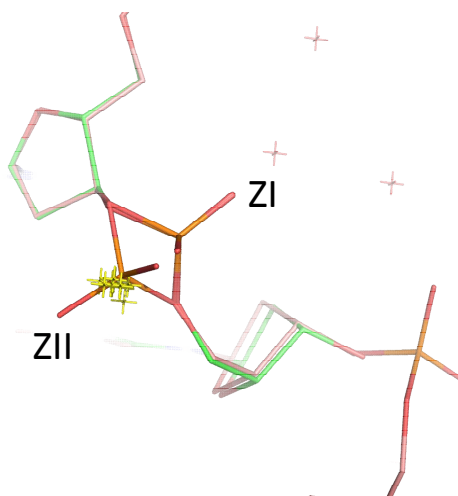


図7 既存の ZI-ZII の multi conformation と観察されたピークの重ねあわせ

また、この解析を通じて、ポリアミンや塩の局所的な環境が、Z-DNA の主鎖の構造ゆらぎに重要な役割を果たしていることが明らかになった。一例を図8に示す。この図は Mg の局在環境では、ゆらぎが見られるが、Ca と共在する際にはゆらぎが抑制されることを表している。同様のゆらぎの抑制がスペルミンなどポリアミンとの結合の際にも観察された。興味深いことに、Z-DNA 結合タンパク質との相互作用においては、このようなゆらぎは抑制されており、タンパク質による B-Z 構造転移の際に、主鎖の構造ゆらぎの抑制が重要な役割を果たしていることが示唆された。

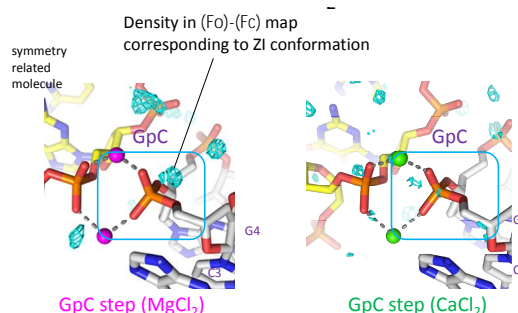


図8. FoFc 差フーリエマップで見られる ZI-ZII 構造ゆらぎの差異

今後は、この方法を適用して、DNA の配列による違いや、Z-DNA 以外のタンパク質・DNA 相互作用を詳しく調べることで、DNA のゆらぎやすさとタンパク質との相互作用の関係を解析し、高選択的な転写因子の創製につなげていきたいと考えている。

現在、追加的な実験・統計解析を行い、論文発表を予定している。

[1] Gaj T, *et al.* (2013) Trends Biotechnol. **31**, 397-405.
 [2] X. Meng, *et al.* (2005) Nat. Biotechnol. **23**, 988-994.
 [3] Michael Gromiha M., *et al.* (2004) J. Mol. Biol. **337**, 285-294.

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)

[学会発表](計4件)
角南 智子, 河野 秀俊, Designing a new artificial transcription factor based on engrailed homeodomain, 第38回日本分子生物学会, 神戸ポートアイランド(兵庫県神戸市), 2015年12月1-4日

角南 智子, 河野 秀俊, Reconsidered DNA conformations in crystal structures, 第 52 回生物物理学会年会, 金沢大学 (石川県金沢市), 2015 年 9 月 13-15 日

角南 智子, 河野 秀俊, Designing a new artificial transcription factor based on engrailed homeodomain, 第 15 回日本蛋白質科学会年会, あわぎんホール (徳島県徳島市), 2015 年 6 月 24-26 日

角南 智子, 河野 秀俊, Designing a new artificial transcription factor based on engrailed homeodomain, 第 51 回生物物理学会年会, 札幌コンベンションセンター (北海道札幌市), 2014 年 9 月 25-27 日

6. 研究組織

(1) 研究代表者

角南 智子 (SUNAMI, Tomoko)

国立研究開発法人日本原子力研究開発機構
原子力科学研究部門 量子ビーム応用
研究センター・研究副主幹

研究者番号: 50554648