

博士論文

**A study on design of highly conductive nucleosides for high precision**

**DNA sequencing based on single molecule electronics**

(単分子伝導度計測に基づく高精度な核酸配列決定の実現に向けた

高導電性核酸塩基の設計に関する研究)

古畑 隆史

# A study on design of highly conductive nucleosides for high precision DNA sequencing based on single molecule electronics

## (単分子伝導度計測に基づく高精度な核酸配列決定の実現に向けた高導電性核酸塩基の設計に関する研究)

山東研究室 37-177132 古畑 隆史

### 第1章. 緒言

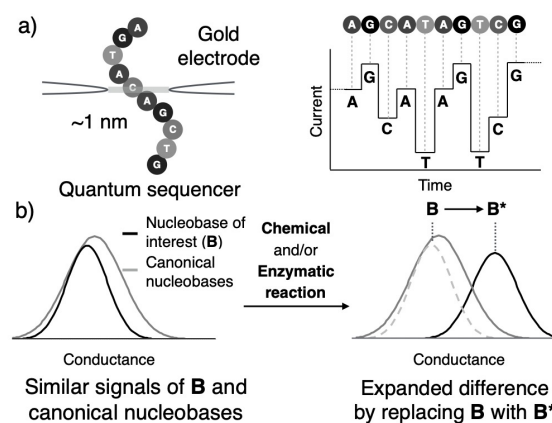
DNA は、生体内において RNA やタンパク質の一次配列を規定する生体高分子である。DNA の配列情報はあらゆるタンパク質の機能と密接に結びついており、DNA 配列を効率的かつ正確に読み取る技術は生命現象の理解に必須である。

近年、DNA の高速な配列解析を実現する手法として、量子シーケンサーと呼ばれる手法が有望視されている。<sup>1</sup> 本手法は、ナノメートルからサブナノメートルスケールに制御された金属 (多くの場合は金) のギャップ電極を用いて電氣的に核酸塩基を検出する単分子伝導度計測の発展的手法である (Figure 1a)。ギャップ間を核酸塩基が通過すると、核酸塩基の導電性に応じて異なる強度のトンネル電流が流れる。その強度差が十分であれば、電流強度の変化から核酸塩基の同定と配列決定を行うことができる。<sup>2,3</sup> 原理上、核酸塩基修飾などの DNA の化学的情報を直接検出できる点は従来法と比べた特色である。<sup>4,5</sup>

しかし、本手法には実応用に向けた大きな課題がある。それは、各核酸塩基に由来するトンネル電流値の分布が大きく重なり、識別が難しいことである。<sup>2,6</sup> この原因として、大きく2つの要因が考えられる。1つ目は、各核酸塩基に由来するトンネル電流値の分散が大きいことである。現在の量子シーケンサーでは、ギャップ電極近傍の DNA の運動がブラウン運動に支配され、電極に対する核酸塩基の配向は制御されていない。そのため、ギャップ電極間で核酸塩基がゆらぎ、電流値のばらつきが大きくなる。この問題は、ナノ空間における DNA の配向を精密に制御可能なナノテクノロジーの発展により将来的に解決できると期待される。<sup>7</sup> 2つ目は、天然に存在する核酸塩基間の導電性の差異が小さいことである。例えば、ゲノムにおいて遺伝暗号を構成する4種の標準核酸塩基 (dA: 2'-deoxyadenosine, dC: 2'-deoxycytidine, dG: 2'-deoxyguanosine, dT: thymidine) のみに着目しても、dA は dC や dG と導電性が類似し、トンネル電流強度の重なりが特に大きい。

<sup>2</sup> この問題は各核酸塩基に固有の物性に由来する原理上の課題であり、有効な打開策が欠如していた。

以上から、本研究では後者の問題 (導電性の類似) を克服し、量子シーケンサーによる正確な核酸塩基の識別を実現する高導電性非天然核酸の創出を目指した。もし、酵素反応や化学反応により特定の核酸塩基を標準核酸塩基に比べて高い導電性をもつ構造に置換することができれば、解析対象となる核酸塩基間の導電性の差異を拡張し、正確な識別が可能となるはずである (Figure 1b)。第2章では、分子設計指針の確立のため、異なる電子物性を有する一連のヌクレオシドを用いて、核酸塩基の電子物性と導電性の相関を検証した。また、得られた知見を元に、高導電性 dA 類似体を用いた4種遺伝暗号 (ATGC) 間の導電性の差異を拡張する戦略の提案を試みた。第3章では、第2章の分子設計戦略を応用し、化学反応を用いた修飾核酸塩基の高導電性化とその効率的検出に向けた検討を行った。第4章では、核酸塩基を介した単分子伝導メカニズムのより詳細な理解と高導電性核酸塩基の設計戦略の拡張を目指し、金電極と標準核酸塩基の相互作用に着目した非天然核酸塩基の設計と評価を行った。



**Figure. 1** Schematic illustration of a) quantum tunneling-based DNA sequencing and b) the general concept of this work to facilitate base-calling in the sequencing.

## 第 2 章. 核酸塩基の物性—導電性相関の系統的評価と 4 種遺伝暗号の識別促進に向けた高導電性核酸の応用

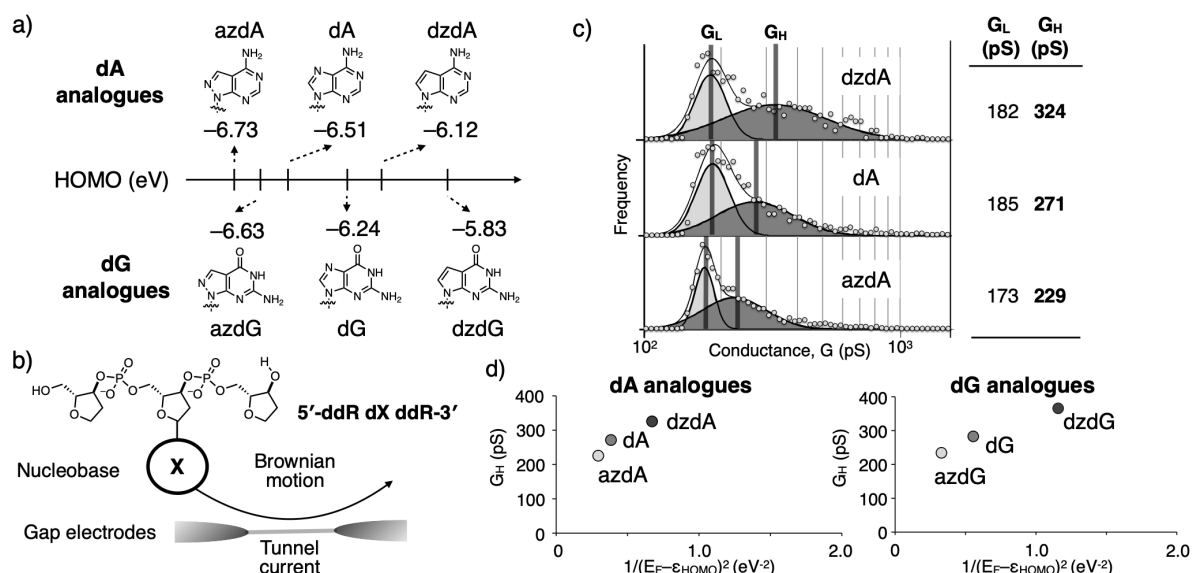
**【HOMO レベルと導電性の相関評価】** 高導電性核酸の分子設計指針として、まず導電性を支配する電子物性の理解が不可欠である。しかしながら、これまで特定の電子物性と導電性の相関を系統的に実証した例は報告されていない。そこで、導電性に重要な核酸塩基の物性として最高被占軌道準位 (HOMO レベル) に着目し、両者の相関を一連の非天然核酸を用いて系統的に評価することとした。一般に、金属のギャップ電極間における分子の導電性 ( $G$ ) は、対象分子の分子軌道準位、および金属電極と分子の電子雲の重なり (電子的カップリング) の 2 つによって決定され、以下の数式で記述される。<sup>8</sup>

$$G = \frac{2e^2}{h} \frac{4\Gamma_L\Gamma_R}{(E_F - \varepsilon_{MO})^2} \quad (1)$$

$e$  は電気素量、 $h$  はプランク定数、 $\Gamma_{L(R)}$  は分子と左右の電極とのカップリング強度である。また、 $\varepsilon_{MO}$  は電極のフェルミ準位  $E_F$  に最も準位の近い分子軌道エネルギーである。解析対象として核酸塩基、電極として金を想定すると、これは HOMO レベル ( $\varepsilon_{HOMO}$ ) に相当する。よって、核酸塩基の導電性は HOMO レベルと強い相関を示すと考えられた。

上記仮説を系統的に検証するため、まず異なる HOMO レベルを有する一連の核酸塩基を設計した。HOMO レベルの寄与を選択的に抽出するには、HOMO 以外の  $\Gamma_{L(R)}$  値に関わるパラメータについては、比較対象の核酸塩基間で差異を最小限に抑えることが望ましい。そこで、このような条件に合う化合物として dA 類縁体、及び dG 類縁体の設計と合成を行った (**Figure 2a**)。これらの核酸塩基は、分子骨格の形状と電子雲の広がり類似しているため、 $\Gamma_{L(R)}$  値の差異を最小限に抑えつつ HOMO レベルの寄与を抽出できると考えられる。

続いて 0.6 nm の金のギャップ電極を用いて、dA、dG 類縁体の単分子伝導度計測を行った (**Figure 2b**)。このとき、ギャップ間のバイアス電圧 ( $V$ ) は 100 mV である。計測後、ベースライン電流の標準偏差  $\sigma$  の 6 倍 ( $6\sigma$ ) を超えた領域をシグナルとして定義し、その最大値をシグナル強度 ( $I$ ) として取得した。また、 $I$  と  $V$  より導電性 (コンダクタンス、 $G = I/V$ ) を算出し、ヒストグラムを作成した (**Figure 2c**)。すると、dA 類縁体、dG 類縁体それぞれで HOMO レベルが金のフェルミ準位に近いものほど、高コンダクタンスなシグナ



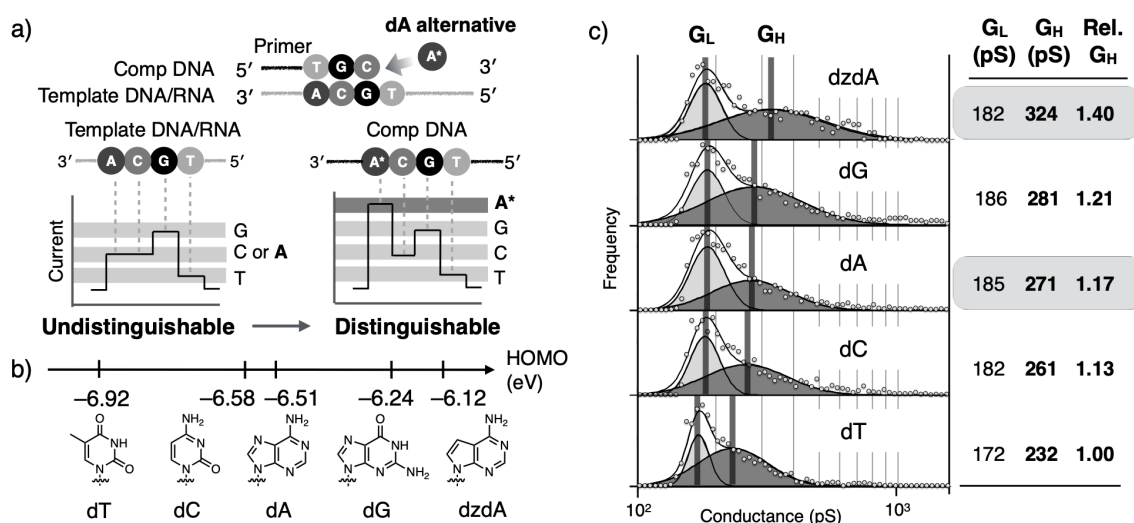
**Figure 2** Systematic evaluation of the correlation between the molecular conductance and the HOMO level. **a)** Chemical structures and HOMO levels of dA and dG analogues. Quantum chemical calculations were conducted by DFT using B3LYP/6-311+G(d,p)/B3LYP/6-31+G(d). **b)** A schematic illustration of the current measurement of the oligonucleotides. **c)** Conductance histograms of the dA analogues and two Gaussian fittings. **d)** Plots of the  $G_H$  values against square inverse of the energy gap between the Fermi level of the gold and the HOMO level of the nucleoside ( $1/(E_F - \varepsilon_{HOMO})^2$ ). Reprinted with permission from *ACS Nano*, 2019, 13, 5028–5035. Copyright 2019 American Chemical Society.

ルの頻度が高くなった。具体的には、350 pS 以上のシグナルの割合を比較した時、dA 類縁体で  $azdA (17.0\%) < dA (23.0\%) < dzdA (37.2\%)$ 、dG 類縁体で  $azdG (13.8\%) < dG (34.6\%) < dzdG (46.6\%)$  であった。

より詳細な検討のため、ヒストグラムの形状に着目した比較を行った。いずれの核酸塩基についても、ヒストグラムは、低コンダクタンス側に頂点があり、高コンダクタンス側に裾の広い構造であった。そこで、ヒストグラムを2つの正規分布の足し合わせ (two Gaussian fittings) によりフィッティングをかけたところ、うまく記述できることが分かった (Figure 2c)。ここで、高コンダクタンス側のピーク ( $G_H$ ) はヌクレオシド間で差が明確であるのに対し、低コンダクタンス側のピーク ( $G_L$ ) ではほぼ差がなかった。このことから、 $G_H$  は各核酸塩基の導電性を反映するのに対し、 $G_L$  は各ヌクレオシドに共通する糖骨格や機械的ノイズに対応すると考えられた。これは、糖骨格の類似構造として 1,2-dideoxyribose を測定した際、 $G_L$  付近に集中する低強度なシグナルが得られたという実験事実からも支持される。そこで、核酸塩基の導電性の指標として高コンダクタンス側のピーク値 ( $G_H$ ) を採用して、 $1/(E_F - \epsilon_{HOMO})^2$  に対してプロットしたところ、dA 類縁体、dG 類縁体のいずれでも正の相関が得られた (Figure 2d)。以上から、導電性と HOMO レベルの相関が実証された。

**【高導電性非天然核酸 dzdA を用いた 4 種遺伝暗号の識別促進戦略】** 続いて、高導電性核酸を用いた 4 種遺伝暗号 (ATGC) の解読促進のための戦略の提案を試みた。前述のように、標準核酸塩基の中でも、特に dA は、dC や dG と導電性が大きく重なり、識別が難しい。そこで、dA とその他核酸塩基の差を拡張するため、dA を酵素的に高導電性の類縁体に置換する戦略を着想した (Figure 3a)。そこで、前節において、標準核酸塩基のうちで最も高導電である dG に比べ、より高い導電性を示す dA 類縁体、7-deaza dA (dzdA) に着目した (Figure 2c,d)。DzdA は DNA 合成酵素による取り込みが dA と同等の精度、効率で行われる dA 類縁体である。dzdA と標準核酸塩基 (Figure 3b) の導電性の比較したところ、dzdA はいずれの標準核酸塩基よりも高い導電性を示す構造であることが示された (Figure 3c)。具体的には、dG と比べても dzdA の  $G_H$  は 42 pS 高い。この値は、標準核酸塩基で最も差の大きい dT と dG の差 (39 pS) に匹敵し、dzdA が標準核酸塩基と効率的に識別可能な遺伝暗号“A”として有望な単分子伝導特性を備えていることを示す。以上から、dA を dzdA へ置換することで、4 種遺伝暗号間の導電性の差異を拡張できることを実証した。

最後に、dzdA を含めた 4 種遺伝暗号からなる DNA 配列の解析可能性の検証のため、遺伝暗号“A”として、dA もしくは dzdA を含むオリゴヌクレオチド (5'-d(TGTXCT)-3', X = dA or dzdA) の単分子伝導度計測を行った。すると、dzdA を含む配列で dG の  $G_H$  値である 291pS を超える高コンダクタンス値の頻度が大きく上昇



**Figure 3 Molecular conductance of dzdA and canonical nucleosides. a)** A schematic representation of enzymatic replacement of canonical dA with conductive dA analogue (dA\*) to prepare decodable oligonucleotide samples. **b)** Chemical structures of nucleobases and their HOMO levels. **c)** Conductance histograms of dzdA and canonical nucleosides and two Gaussian fittings. Reprinted with permission from *ACS Nano*, 2019, 13, 5028–5035. Copyright 2019 American Chemical Society.

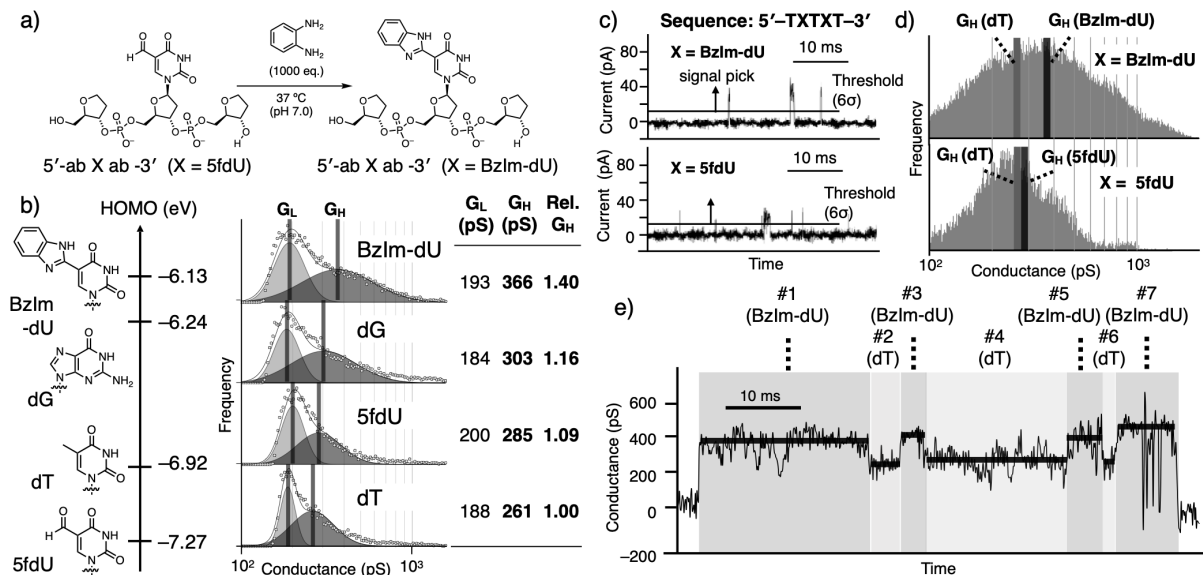
していることが分かった。その割合を比較すると、20%と31%であり、その割合の差(11%)はdzdAがオリゴヌクレオチドに占める割合(6塩基中1塩基, 17%)に匹敵する。このことから、オリゴヌクレオチド中においてもdzdAは高コンダクタンスなシグナルとして検出可能であることが示唆された。

### 第3章. 化学ラベルによる電子物性制御戦略に基づいた核酸塩基修飾の検出

本節では、HOMO レベルに着目した高導電性核酸塩基の設計戦略を応用し、核酸塩基修飾の検出へと展開した。最も有名なDNA修飾であるdCの5位メチル化以外にも、近年新たな核酸塩基修飾として、ピリミジン核酸塩基5位メチル基の酸化産物であるヒドロキシメチル基、ホルミル基、カルボキシル基が発見され、その機能が注目されている。<sup>9</sup>量子シーケンサーは、それら修飾核酸塩基を一塩基分解能で直接検出できる潜在能力を持つ。そこで、修飾核酸塩基の検出促進に向け、化学反応を用いた修飾核酸塩基の導電性の制御を試みた。具体的には、モデル核酸塩基としてdTの酸化産物である5-formyl dU(5fdU)を用い、5fdU選択的な官能基変換反応を用いた高導電性を図った。このような変換反応には、(1) DNAの良溶媒である水中で定量的、かつ5fdU選択的に反応が進むこと、(2) 反応後に5fdUの導電性が標準核酸塩基に比べて高くなること、の2条件が必須である。上記2条件を満たす反応として、1,2-phenylenediamineとホルミル基の酸化的benzimidazole形成反応に着目した。本反応により、5fdUから生じる5位benzimidazole修飾dUをBzIm-dUとする(**Figure 4a**)。 (1)の反応効率について、すでにBzIm-dUの生成反応は生理条件下(pH 7, 37°C)で定量的に進むことが報告されている。<sup>10</sup>実際に、本研究でも2時間でほぼ完全に反応が終了することを確認した。

次に、(2)の導電性について検討した。DFTによる分子軌道計算を行ったところ、BzIm-dUはdGに比べても高いHOMOレベルが算出され、標準核酸塩基に比べて高い導電性が期待された。実際に、単分子伝導度計測を行ったところ、dTやdGに比べてBzIm-dUがより高い導電性( $G_H$ 値)を示す事がわかった(**Figure 4b**)。具体的には、dTに比べて40%、dGに比べて20%程度高い $G_H$ 値を有していた(**Figure 4b**)。一方で、官能基変換前の5fdUは、dT、dGいずれとも導電性の差は10%以下であり、識別が困難であると予想された。

続いて、核酸配列中におけるBzIm-dUと5fdUの検出可能性を検討するため、配列中に5fdU、もしくはBzIm-dUを含むオリゴヌクレオチド(5'-d(TXTXT)-3', X = 5fdU or BzIm-dU)の伝導度計測を行い、その比較



**Figure 4** Chemical labeling-assisted differentiation of the molecular conductance between the modified nucleobase and the canonical nucleobases. **a**) Scheme of labeling reaction to convert 5fdU to BzIm-dU. **b**) HOMO levels and conductance histograms of the nucleotides. **c**) Typical current traces of TXTXT sequences where X is BzIm-dU (upper) and 5fdU (lower). **d**) Conductance histograms of the sequence TXTXT constructed using all data points of signals with maximum intensity in the top 15% and a length over 2 ms. **e**) Partial sequencing of the stepwise signal of TXTXT sequence (X = BzIm-dU). Reproduced with permission from *ChemBioChem*, 2019, doi: 10.1002/cbic.201900422. Copyright 2019 Wiley-VCH.

を行った (Figure 4c)。シグナルのうち、最大強度が上位 15%以内のものを取り出し、そのシグナルに含まれる全データポイントを用いてヒストグラムを作成したところ、 $X = 5fdU$  の場合では、 $5fdU$  もしくは  $dT$  の  $G_H$  付近に一つだけ鋭いピークを持っていたのに対し、 $X = BzIm-dU$  では  $BzIm-dU$  の  $G_H$  付近にもピークが確認された (Figure 4d)。また  $dT$  を  $dG$  に置き換えた配列でも同様の傾向が見られた。以上から、オリゴヌクレオチド中においても  $BzIm-dU$  は高コンダクタンスな値を示すシグナルとして標準核酸塩基と識別可能であることが示唆された。また、シグナルの中には、 $dT$  や  $BzIm-dU$  の  $G_H$  値とおおよそ一致するプラトーを持つ階段状のシグナルも含まれていた (Figure 4e)。これは、各核酸塩基のギャップ電極間の通過に対応する電流変化であると考えられる。以上から、一塩基分解能で核酸塩基修飾の検出が可能であることも示唆された。ただ、このような階段状のシグナルは全体のシグナルのごく一部である。これは、現時点で DNA の挙動がブラウン運動に依存しており、ギャップ間の通過方向や配向が厳密に制御されていないためであると考えられる。現在ナノ空間における DNA の配向や通過方向を制御する手法の開発が盛んに行われており、<sup>7</sup> 関連分野の発展により正確な配列解析が可能になると期待される。

#### 第 4 章. 金-核酸塩基相互作用に着目した高導電性核酸塩基の分子設計とその応用

上記までの検討で、HOMO レベルが核酸塩基導電性の制御において重要な物性であることが示されてきた。一方、HOMO レベルと導電性の順番が必ずしも一致しないケースも見いだされた。前節の検討段階で、 $BzIm-dU$  の他に、HOMO レベルの上昇を期待して benzimidazole 環の末端に Me 基や OMe 基などの電子供与性基を導入した核酸塩基についても導電性の評価を行った。導電性の指標としてコンダクタンス値の 75 パーセント点 (75% (pS)) の比較を行ったところ、置換基を導入した  $BzIm-dU$  類縁体では環外置換基を持たない  $BzIm-dU$  に比べて低い値を示した (Figure 5a)。数式 (1) に立ち返ると、これは環外置換基が分子と電極の電子的なカップリング強度 ( $\Gamma_{L(R)}$  値) に影響し、導電性を制御する要因となり得ることを示唆する。そこで、量子センサーにおける単分子伝導メカニズムのさらなる解明に向け、金-核酸塩基間の相互作用に着目した高導電性核酸塩基の設計と評価を行った。具体的には、2'-deoxyuridine (dU) のピリミジン環 5 位に、電子の供与・逆供与による金との高い親和性が期待される置換基を導入し、それら 5 位修飾 dU の導電性を評価した。

まず、金との親和性が高い置換基として、ethynyl 基、cyano 基、amino 基を導入した 5 位修飾 dU を設計した (それぞれ、EtdU, CNdU, AmdU) (Figure 5b)。HOMO レベルを比較すると  $CNdU < EtdU < AmdU$  の順である。これらについて単分子伝導度計測を行ったところ、EtdU、AmdU、CNdU はそれぞれ  $dT$  よりも高い導電

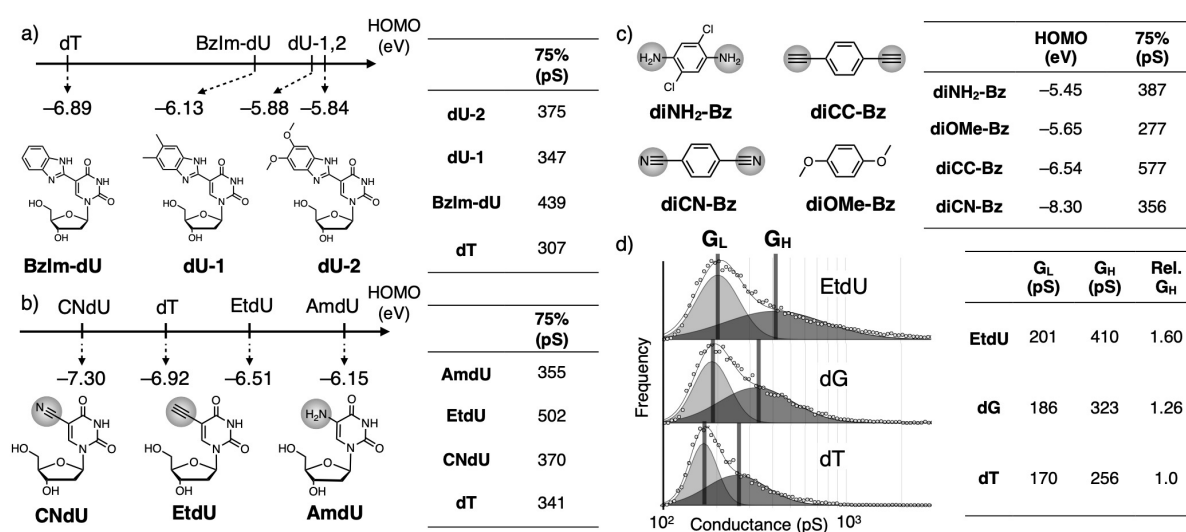


Figure 5 Molecular conductance of 5-modified dU and di-substituted benzene. Chemical structures and 75 percentiles of molecular conductance of a) dU derivatives modified with benzimidazole analogues, b) dU derivatives modified with aurophilic functional groups and c) di-substituted benzenes modified with aurophilic functional groups. d) Conductance histograms of  $dT$ ,  $dG$  and EtdU and two Gaussian fittings to the histograms.

性を示した (Figure 5b)。また、5 位修飾 dU の導電性は CNdU ~ AmdU < EtdU であった。同じ環外置換基を有するパラ二置換ベンゼンを用いて、単分子伝導度計測を行ったところ、この場合も導電性 cyano 基 ~ amino 基 < ethynyl 基の順であり、ヌクレオシドの結果とおおよそ一致する (Figure 5c)。興味深いことにこの順序は HOMO レベルと必ずしも合致しない。例えば、ethynyl 基と cyano 基を持つ分子では比較的 HOMO レベルが低いにもかかわらず、金親和性基を持たない分子に比べて高導電なシグナルが得られていた。一方、amino 基を導入した分子では、導電性は ethynyl 基を持つ分子より低く、cyano 基を持つ分子と同程度であった。以上から、cyano 基や ethynyl 基など平面性の高い  $\pi$  結合を持つ構造は電極とのカップリングが強い、高導電性な金親和性モチーフであると考えられた。これは、amine 基に比べて cyano 基で N 原子周辺の立体障害が小さく、金との結合エネルギーが大きくなることとした先行研究の計算結果から予想される結果とも一致する。<sup>11</sup>

最後に、EtdU と標準核酸塩基の導電性の比較を行った。 $G_H$  値を比較すると、EtdU はいずれの標準核酸塩基に比べても高い値 ( $G_H$ ) を示した (Figure 5d)。これまでに、EtdU は酵素的に DNA 鎖に導入可能なことが知られており、ゲノム DNA の複製領域を標識する分子プローブとしても使われてきた。<sup>12</sup> 従って、将来的に EtdU は量子シーケンサーによる DNA 複製の単分子解析を実現するプローブとしての応用も期待できる。

## 第 5 章. 結言

以上、本研究では量子シーケンサーの応用可能性の拡張に向け、2 つの観点から高導電性核酸塩基の設計指針を探索した。1 つ目は、HOMO レベルに着目した設計戦略である。はじめに、類似した核酸塩基構造を有する一連の dA、dG 類縁体を用いて単分子伝導計測を行い、系統的に HOMO レベルと導電性の正の相関を実証することに成功した。続いて、dzdA を用いた 4 種遺伝暗号における導電性の差異の拡張、選択的化學ラベル反応を用いた修飾核酸塩基の検出促進が可能であることを示した。一方、核酸塩基の芳香族環外に置換基を持つ核酸塩基では、必ずしも HOMO レベルと導電性が相関しない場合があった。そこで、より厳密な導電性の制御に向けた第 2 の着眼点として、金電極と核酸塩基の電子的なカップリングに焦点を当てた設計指針の拡張を目指した。そして、cyano 基や ethynyl 基など平面性の高い  $\pi$  結合を有する置換基を、高導電性核酸塩基の設計に有望なモチーフとして見出した。今後、本戦略を DNA の通過方向や速度、核酸塩基の配向のナノスケールでの制御技術やシグナルの確率統計的な解析技術と組み合わせることで、量子シーケンサーによる正確な DNA 配列解析が実現できると期待される。

## 参考文献

- (1) Di Ventra, M.; Taniguchi, M. *Nat. Nanotechnol.* **2016**, *11*, 117–126. (2) Ohshiro, T. *et al. Sci. Rep.* **2012**, *2*, 501.
- (3) Zwolak, M.; Di Ventra, M. *Nano Lett.* **2005**, *5*, 421–424. (4) Huang, S. *et al. Nat. Nanotechnol.* **2010**, *5*, 868–873.
- (5) Tsutsui, M. *et al. J. Am. Chem. Soc.* **2011**, *133*, 9124–9128. (6) Chang, S. *et al. Nano Lett.* **2010**, *10*, 1070–1075.
- (7) Yokota, K. *et al. RSC Adv.* **2014**, *4*, 15886–15899. (8) Zwolak, M.; Di Ventra, M. *Rev. Mod. Phys.* **2008**, *80*, 141–165. (9) Raiber, E.-A. *et al. Nat. Rev. Chem.* **2017**, *1*, 0069. (10) Liu, C. *et al. Chem. Sci.* **2017**, *8*, 4505–4510. (11) Moreno-Garcia, P. *et al. J. Am. Chem. Soc.* **2013**, *135*, 12228–12240. (12) Yamakoshi, H. *et al. J. Am. Chem. Soc.* **2011**, *133*, 6102–6105.

## 発表論文

### Publication related to the doctoral dissertation

- (1) Furuha, T.; Ohshiro, T.; Akimoto, G.; Ueki, R.; Taniguchi, M.; Sando, S. *ACS Nano*, **2019**, *13*, 5028–5035.
- (2) Furuha, T.; Ohshiro, T.; Izuhara, Y.; Tomoaki, S.; Ueki, R.; Taniguchi, M.; Sando, S. *ChemBioChem*, **2019**, doi: 10.1002/cbic.201900422.

### Publication not related to the doctoral dissertation

- (3) Tsuchiya, A.; Hashim, S. N.; Ise, S.; Furuha, T.; Kawai, K.; Wakabayashi, R.; Goto, M.; Kamiya, N.; Sando, S. *Anal. Sci.*, **2016**, *32*, 543–547.