

[研究成果]

密度汎関数理論によるナノ・バイオ計算

押山 淳

東京大学大学院工学系研究科物理工学専攻

1. はじめに

物理学においては、実験物理学と理論物理学ふたつのアプローチが古くから認識されていたが、ここ四半世紀における計算機の発達と計算方法論の開拓により、計算物理学という第三のアプローチが重要になってきている。この状況は、他の自然科学および社会科学の諸分野においても同様であろう。特にナノサイエンス、バイオサイエンスの領域においては計算科学的アプローチが欠かせない。それは、ナノ・バイオ構造体では、空間的にはナノメートル(10^{-9} m)のスケール、時間的にはフェムト(10^{-15})秒からピコ(10^{-12})秒で起こる現象の解明が重要であり、実験的にはなかなかアクセスできない状況に突入しているからである。精度の高い信頼できる計算の果たす役割は大きい。

物質は原子核と電子で構成されている。それらは量子力学で記述される粒子あるいは波動であり、互いに相互作用し合っている。その意味で物質は量子多体系ということができ、従って、物質科学あるいはナノ・バイオサイエンスにおける精度の高い計算科学的手法は、量子論に立脚していることが必要である。しかも経験的なパラメータを含まない物理学の第一原理に基づいた計算科学的手法により、初めてその精度が裏打ちされる。しかしそれは、云うは易く行うは難いことであって、相互作用している量子多体系ほど難しい系はこの世の中にあまり存在しない。

密度汎関数理論はこうした量子多体系を扱うひとつの手法である。多体系を有効一粒子系の問題に還元するひとつの手法であり、比較的に簡便なわりには、実際の物質の諸性質を驚くほど精度良く説明、さらには予測することができる。本稿では、密度汎関数理論、および我々が次世代スーパーコンピュータを視野に入れて開発している実空間密度汎関数法の簡単な説明に統いて、ナノ・バイオ系におけるいくつかの計算例を紹介し、現時点での、密度汎関数理論に立脚した計算科学的アプローチの現状を報告したい。

2. 密度汎関数理論

2.1 基礎

密度汎関数理論 (Density Functional Theory: DFT) によれば、相互作用している多電子系の全エネルギー E は電子密度 $n(\vec{r})$ の汎関数として、

$$E[n] = \int v_{\text{ext}}(\vec{r})n(\vec{r})d\vec{r} + \iint \frac{n(\vec{r})n(\vec{r}')}{|\vec{r}-\vec{r}'|}d\vec{r}d\vec{r}' + F[n] \quad (1)$$

と書けることが厳密に証明できる[1]。ここで $v_{\text{ext}}(\vec{r})$ は外場ポテンシャルであり、原子核からのポテンシャルエネルギーあるいは電磁場のポテンシャルエネルギーを表している。右辺第二項は古典的な意味での電子間相互作用エネルギーである。 $F[n]$ は電子系の運動エネルギーおよび量子論的な相互作用エネルギーである。どのような物質であれ、その物質の個性に依らない電子系の相互作用の性質だけで決まるようなユニバーサルな汎関数 F が存在するというのが密度汎関数理論の骨子である。電子密度は、一般に多電子系の波動関数を用いて表すことができるが、ここで準粒子（相互作用の衣を着た電子）の軌道関数 $\varphi_i(\vec{r})$ の絶対値の2乗で表すことにする[2]。

$$n(r) = \sum_i |\varphi_i(\vec{r})|^2 \quad (2)$$

すると(1) は、

$$E[n] = \int v_{\text{ext}}(\vec{r}) n(\vec{r}) d\vec{r} + \iint \frac{n(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' + \sum_i \int \varphi_i^*(\vec{r}) \left(-\frac{\hbar^2 \nabla^2}{2m} \right) \varphi_i(\vec{r}) d\vec{r} + E_{\text{XC}}[n] \quad (3)$$

となる。 $E_{\text{XC}}[n]$ のことを交換・相関 (exchange-correlation) エネルギーとよび、電子系における量子効果を表している。交換・相関エネルギーを2体の密度ゆらぎ相関関数で表すことはできるが、実際は近似を導入する。局所密度近似 (Local Density Approximation: LDA)、一般化密度勾配近似 (Generalized Gradient Approximation: GGA)、ハイブリッド近似 (Hybrid Approximation) などがよく用いられる。

さて次のステップは、(3)で表される系の全エネルギーを、波動関数が規格化されているという拘束条件のもとで、最小化することである。それは軌道関数についての変分をとった変分方程式にはかならず、

$$\left\{ -\frac{\hbar^2 \nabla^2}{2m} + v_{\text{ext}}(\vec{r}) + \int \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + \frac{\delta E_{\text{XC}}[n]}{\delta n} \right\} \varphi_i(\vec{r}) = \varepsilon_i \varphi_i(\vec{r}) \quad (4)$$

を得る。ここで ε_i は変分法における Lagrange の未定乗数であり、今の場合、状態 i の固有エネルギーに対応している。(4)は Kohn-Sham 方程式とよばれ、1電子に対するシュレディンガー方程式と形式的に同じ形をしているが、量子多体効果である E_{XC} を含んでいる。また左辺の演算子は電子密度の汎関数であり、電子密度は(4)を解いて得られる軌道関数を用いて(2)のように表されているので、(2)と(4)は consistent に解く必要がある。これは SCF (Self Consistent Field) 計算とよばれ、再帰的手法により行われる。

DFT 計算において、最も時間が費やされるのは Kohn-Sham 方程式の解を得るプロセスである。標準的には $\varphi_i(\vec{r})$ を基底関数系によって展開し、その展開係数を求めることが行われる。一昔前までは、左辺の演算子の行列を構成し、その行列を対角化するのが普通であったが、現在では計算の大規模化に伴い、(3)そのものを展開係数について再帰的に最小化することにより、(4)の解を得ることが行われている。再帰的最小化の方法としては、共役勾配法、残差最小化法などが使われる。

電子と原子核から構成される系の全エネルギーは、(3)で表される電子系のエネルギー、電子一原子核の相互作用エネルギーに加えて、原子核系のエネルギーをえたものである。その全エネルギーを、原子核の位置について微分し符号を反転させたものが、その原子核にかかる力である。その力に沿って原子核の位置を変化させ、全エネルギーの極小値を得る事ができれば、それが自然界で出現する物質の安定あるいは準安定構造に相当する。また原子核に働く力を用いて、分子動力学計算を行えば、物質での原子のダイナミクスを追えることになる。

2.2 実空間密度汎関数理論コード

Kohn-Sham 軌道関数 $\varphi_i(\vec{r})$ を展開する基底関数系としては平面波基底がよく用いられる。それは平面波の数を系統的に増やして全エネルギーの収束値をみることにより、実際に、完全基底関数系での計算値を得ることができるためである。しかし平面波基底にはふたつの困難がある。第一は軌道関数の満たすべき境界条件に関連したものである。平面波で展開した場合、軌道関数は必然的に周期的境界条件を満たすことになる。結晶の Bloch 波動関数は、実際周期的境界条件を満たしているが、分子・クラスター、結晶中の欠陥などに関連した波動関数はそうではない。平面波で展開するためには、

スーパーセルというターゲットを周期的に並べた系を扱うしかない。第二の困難は、コンピュータのアーキテクチャに関連したものである。ペタフロップスを超える演算能力が期待される次世代スーパーコンピュータは、各プロセッサがいわゆるベクトル型であるか否かに関わらず、必然的に 10000–100000 ノードの超並列アーキテクチャとなる。平面波基底を用いた計算においては、フーリエ変換が頻繁に行われるが、Fast Fourier Transform (FFT) の演算は、全ノード間の通信を伴い、超並列アーキテクチャでは大きな重荷となる。

こうした平面波展開による計算の問題点は、実空間に格子を導入し、波動関数、電子密度等をこの格子上の各点で表現し、計算をその格子上で実行する実空間密度汎関数法によって回避できる。この手法では、実空間のセルを各ノードに割り当てるにより、自然な形の並列コードが作成可能であり、FFTを必要としないために、通信は本質的に局所的なものとなる¹。また、計算結果の完全性は、格子点間の間隔を系統的に狭めていき、収束値を検討することによって達成される。軌道関数が満たすべき境界条件に対しては、実空間処理ではいかなる条件も対応可能であり、その点においても既存の基底関数系による展開に比して魅力的である。我々は、並列アーキテクチャ上でこうした実空間処理コード [RSDFT (Real-Space Density Functional Theory) コードとよんでいる] の開発・高精度化を計算機科学者、工学者との共同で行っている[3,4]。

計算機アーキテクチャを意識したコードの最適化には、計算機科学分野の研究者との連携が重要である。RSDFTは、ベクトル機、並列機のいずれでも高いパフォーマンスを示しているが、並列機での最適化は、筑波大学PACS-CSシステム²をプラットフォームとして、計算機科学分野との共同で行われ、1000–2000 ノード使用時の全体の実効性能はおよそ 20%程度である。また最も負担の重いグラム・シュミット正規直交化計算では、理論ピーク性能の 7、8割程度の実効性能を示している。一例としてそこで行ったアルゴリズム変更を紹介しよう。通常のグラム・シュミット直交化計算は、 ϕ_i をターゲットベクトル、 ψ_i を正規直交化されたベクトルとすると、

$$\begin{aligned}\psi_1 &= \phi_1 \\ \psi_2 &= \phi_2 - \psi_1(\psi_1, \phi_2) \\ \psi_3 &= \phi_3 - \psi_1(\psi_1, \phi_3) - \psi_2(\psi_2, \phi_3) \\ \psi_4 &= \phi_4 - \psi_1(\psi_1, \phi_4) - \psi_2(\psi_2, \phi_4) - \psi_3(\psi_3, \phi_4) \\ \psi_5 &= \phi_5 - \psi_1(\psi_1, \phi_5) - \psi_2(\psi_2, \phi_5) - \psi_3(\psi_3, \phi_5) - \psi_4(\psi_4, \phi_5) \\ \psi_6 &= \phi_6 - \psi_1(\psi_1, \phi_6) - \psi_2(\psi_2, \phi_6) - \psi_3(\psi_3, \phi_6) - \psi_4(\psi_4, \phi_6) - \psi_5(\psi_5, \phi_6) \\ &\dots\end{aligned}\tag{5}$$

という手順で行われる。これは典型的なベクトル積計算であり、BLAS (Basic Linear Algebra Subprograms) のレベル 1 のルチーンで実行される。ところが、この右辺の 4 行目以降の演算の最初の 4 項目までは、

$$(\phi_4, \phi_5, \phi_6) - (\psi_1, \psi_2, \psi_3) \begin{pmatrix} {}^t \psi_1 \\ {}^t \psi_2 \\ {}^t \psi_3 \end{pmatrix} (\phi_4, \phi_5, \phi_6)\tag{6}$$

¹ 実空間での並列化に加えて、各軌道関数の処理に関する並列化も可能であり、計算規模と計算を実行するプラットフォームの能力の 2 つの要素に依存した最適な方針の検討が重要である。

² PACS-CS (Parallel Array Computer Systems for Computational Sciences)は、筑波大学計算科学研究センターで開発された並列計算機であり、2.8 GHz の Xeon を用いた独自開発の 2560 ノードが 3 次元 hyper cross bar network で接続された、理論ピーク性能 14.3 TFLOPS のシステムである。

という行列積の形に書き換えることができ、これにより BLAS3 のルチーンを用いることで、格段に高速化される。そこで図 1 のように、ブロック化処理とそこで最適データ量をチューンすることにより、理論ピーク性能とほぼ同等の実効性能が実現されている。

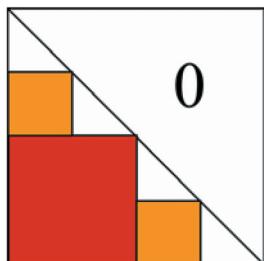


図 1：グラム・シュミット正規直交化を行列積計算で行う概念図。下三角行列を計算する際に、正方形領域を図の赤、橙のように再帰的に構成し、その中の演算を行列積計算として実行する。

3. ナノ系への応用

さて本節では、前節で説明した DFT による物質計算の例をいくつか紹介しよう。

3.1 シリコンナノドットのイオン化エネルギー、電子親和力、励起エネルギー

物質の基本的な量としてイオン化エネルギー I と電子親和力 A があげられる。この二つの量から物質の第一励起エネルギー Δ は、 $\Delta = I - A$ と与えられる。 I と A は、 N 個から成る電子系の全エネルギーを E_N とすると、 $I = E_{N-1} - E_N$ 、 $A = E_N - E_{N+1}$ 、と書ける。無限結晶の第一励起エネルギーを、電子数の異なる系での全エネルギー差である I と A とから求めることは原理的に不可能であり、密度汎関数理論においては、(4)の Kohn-Sham 準位の最高占有状態と最低非占有状態との差として求めることが多い (HOMO-LUMO ギャップ)。しかしクラスター計算ではイオン化エネルギー、親和力ともに直接計算可能であり、さらに系のサイズを大きくした収束値を得ることにより、無限結晶での値をシミュレートできる。図 2 は Si ナノドット (直径 5.6 nm ~ 7.6 nm、Si 原子数 4006 ~ 10701、端端水素原子数 1000 ~ 2000) に対する、そうした全エネルギー差による第一励起エネルギーと、従来からの HOMO-LUMO ギャップによる値のサイズ依存性である。これは世界最大規模の DFT 電子状態計算である。図 2 からわかるように、両者は系統的に異なる。詳細な解析により、その差は、ナノドットに電子を 1 個加えた時のエネルギー上昇に相当していることが明らかとなった。

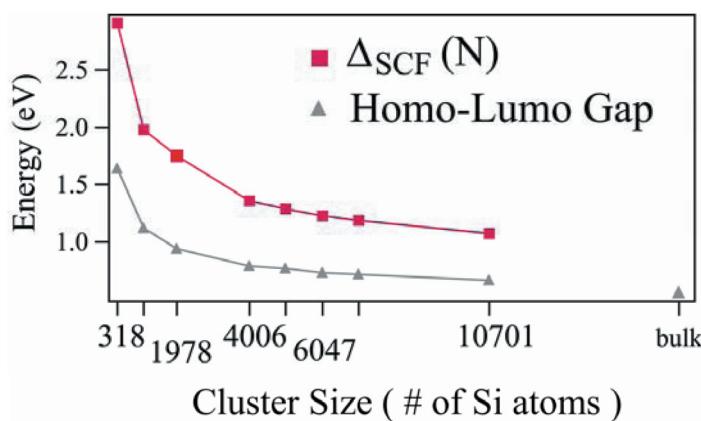


図 2：Si ナノドットの第一励起エネルギーのドットサイズ依存性。赤はイオン化エネルギーと電子親和力の差から求めた第一励起エネルギーで、青は Kohn-Sham 準位から求めた HOMO-LUMO ギャップ。

3.2 新しい体心正方シリコンとゲルマニウムの電子物性

DFT は物質構造を正確に予測できる利点があるが、交換相関エネルギーに対する現時点での近似では、電子励起エネルギーの定量的計算は難しい。現在もっとも実際的な処方箋は、DFT によって構造を定め、GW 近似 (無限次摂動理論における自己エネルギーをグリーン関数 G と繰り込まれた相互作用 W の積で表す近似) によって励起エネルギーを計算するアプローチである。このアプローチによ

り、Si および Ge の新しい結晶相を見出した[5]。

ポストスケーリングのシリコンテクノロジーにおいては、Si および関連物質に加えて高誘電体物質などの新たな材料系をテクノロジーに取り込み、その物性を活用することが重要に成ってきている。それに加えて、前節で紹介したような、ナノ構造における新奇物性の発現も重要である。さらにはダイヤモンド構造以外の新たな構造の Si あるいは Ge が発見されれば、そしてそれらが新奇な物性を示すならば、そのインパクトは計り知れない。

図 3 は我々の DFT+GW 計算によって発見された体心正方 (body center tetragonal: bct) 格子をもつ結晶相の原子配置である。最安定のダイヤモンド構造に比べて原子 1 個あたり 0.1 eV 程度エネルギーは高いが準安定な構造であることが見出された。注目すべきは、全ての原子は、典型的な半導体と同様に最近接原子と 4 配位を形成しているにも関わらず、bct-Si ではギャップは顕著に狭くなり、bct-Ge ではそれが消失し半金属となることである (図 4)。すなわち 4 配位では、 sp^3 の軌道混成により、tetrahedral な配置の最近接原子は結合軌道と反結合軌道を形成し、それらの間にはギャップが生じるというのが、固体物理の教科書に書かれていることである。今回の我々の結果は、たとえ 4 配位であって、ボンドの結合角度の変調により、金属が出現することを示している。

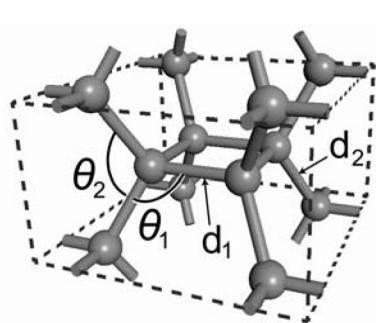


図 3：DFT 計算で見出された体心正方格子を形成する Si と Ge の新しい結晶相。

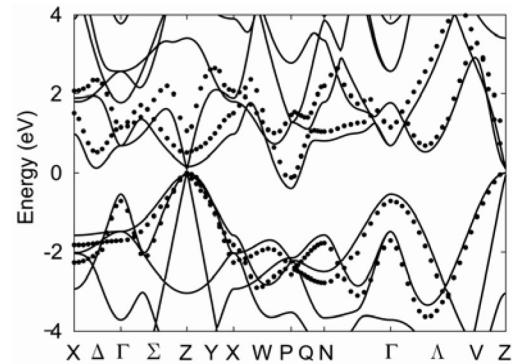


図 4：体心正方 Ge のバンド構造。フェルミ準位の位置をエネルギーの原点にとっている。ブリルアンゾーンの Z 点に正孔、P 点付近に電子が存在する反金属となっている。実線が LDA 近似、ドットが GW 近似による計算値。

3.3 塗化物半導体中カチオン空孔の強磁性

GaN、InN、AlN などの塗化物半導体は直接ギャップ半導体として赤外から紫外までの波長領域をカバーする光デバイスの基幹材料である。他の化合物半導体に比べて、有害な元素を含んでいない点からも将来有望な物質群と云えよう。さらにこれら物質群は、近年スピントロニクス材料としても有望視されている。すなわち、磁性不純物原子をドープすることによって強磁性が室温において発現していることが報告されている。しかし、磁性不純物原子 1 個当たりの磁気モーメントが異常に大きいなど、従来の常識では理解できない現象が報告されている。我々は、電子相関の効果を積極的に取り入れた DFT+U (Gd の局在した 4f 電子状態での電子相関効果をオンサイト・エネルギー U で取り入れる方法) の手法により Gd ドープ GaN の電子状態を計算した。すると、図 5 のように Gd 周囲のカチオン空孔が強磁性的にスピン偏極することが見出された。さらにカチオン空孔のみが存在する系でも強磁性状態が出現することが見出された。これは塗化物半導体においては、カチオン原子と塗素原子の原子半径が顕著に異なることにより、カチオン空孔周囲の塗素原子は、半導体に特徴的なヤーン・テラー歪を起こすことができず、代わりに電子同士の量子論的相互作用により、スピン偏極により安定化するためである。半導体における全く新奇な強磁性相の出現が予測される[6]。

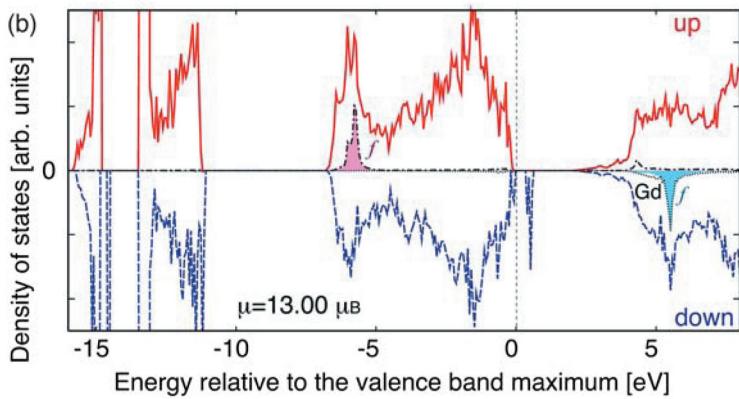


図 5 : DFT+U 計算による Gd+Ga 原子空孔での電子状態密度。上段が多数スピン、下段が少数スピンに対する状態密度であり、13.0 ボア磁子のスピン偏極が見出された。

3.4 ナノキャパシターでの量子効果

電子デバイスにおいて、キャパシタンスはコンダクタンスと並んで基本的な特性量である。微細化と高機能化の進む現状では、デバイス構造はナノメートルスケールであり、それは波動関数の広がりに匹敵するサイズとなっている。キャパシタンスにおける量子効果の解明が急務である。円筒状のチャネル、ゲート構造（サラウンディング・ゲート構造）は、ポストスケーリング時代のシリコンテクノロジーにおいて、電界効果トランジスターの究極構造と目されている。そこでは炭素ナノチューブの特性的優位性を用いたサラウンディング・ゲート電界効果トランジスターの開発が行われている。

我々は、多層炭素ナノチューブからなる円筒構造に対するキャパシタンスを、DFT によって調べ、ふたつの独立な量子効果が見出された[7]。ひとつは波動関数の浸みだしによる有効電極間の減少に伴う、キャパシタンスの増大である。もうひとつは、1 次元構造に特有な、電子状態密度における Van-Hove 特異性による、キャパシタンスの顕著なバイアス電圧依存性である（図 6）。さらに図 7 のような三重炭素ナノチューブの最外側と最内側に電極をつないで電荷を注入すると、中に挟まれた炭素ナノチューブは誘電分極し、キャパシタンスの増大が見られた。この誘電分極から見積もった、ナノチューブ 1 本の誘電率は 1.88 となった。これは SiO_2 などの典型的ゲート絶縁膜の値と同程度である[8]。

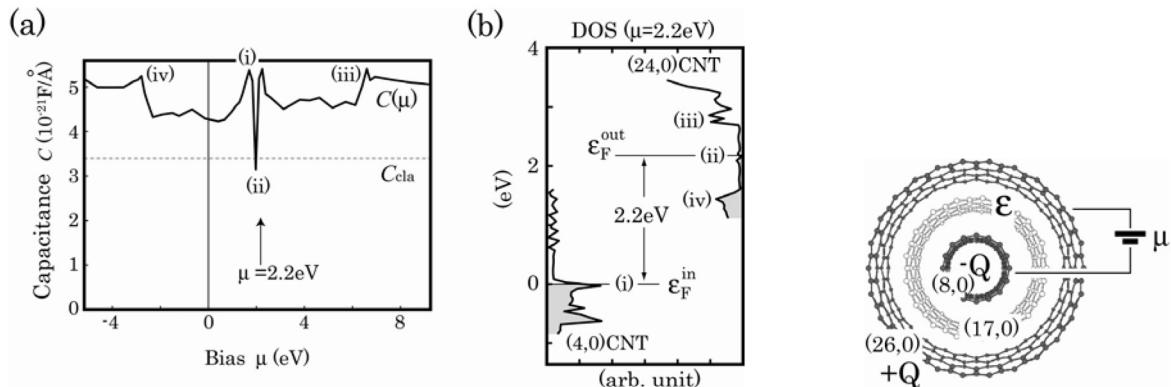


図 6: (a) 二重壁炭素ナノチューブ(4,0)@(24,0)におけるキャパシタンスのバイアス電圧依存性。スパイク状のバイアス電圧依存性は、(b)に示す状態密度の特異性と 1 対 1 に対応している。

図 7: 三重炭素ナノチューブの模式図。バイアス電圧 μ により、外側チューブに $+Q$ 、内側チューブに $-Q$ の電荷が蓄えられている状況。

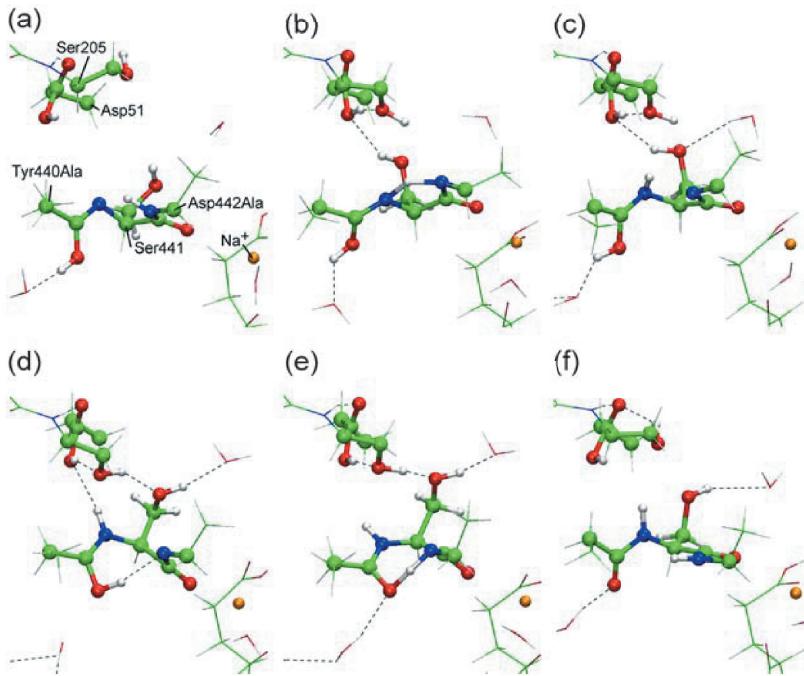


図 8: ペプチド結合を横切るプロトン輸送反応の原子プロセス。緑、青、赤、白丸が、炭素、窒素、酸素、水素原子を表している。シトクローム酸化酵素の内膜から移動してきたプロトンは、Tyr440 の酸素原子に配位する(a)。一方ペプチド鎖の上方に位置するプロトンは Asp51 に引き抜かれ、目的地に到達する。ペプチド鎖バリケードの下方に位置するプロトンは隣のアミノ酸の助けを借りて上方に移動する [(b) – (f)]。

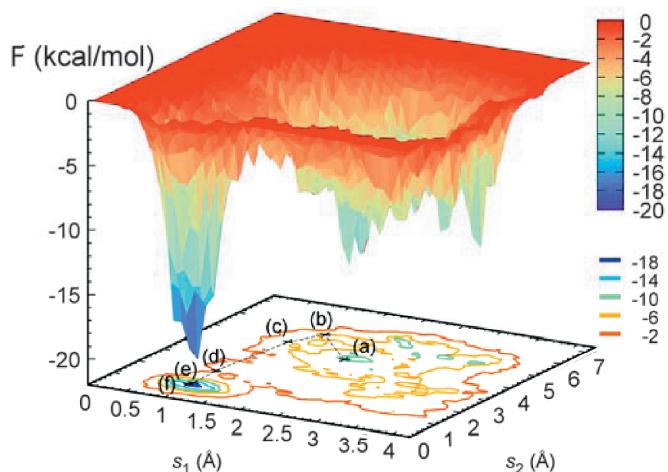


図 9: 図 8 の化学反応[(a)-(f)]に沿った自由エネルギーのプロフィール。(a) から (f)への反応障壁は、その逆のプロセスのそれより小さく、この反応の一方向性を示している。

4. バイオ系への応用

密度汎関数理論はバイオ系に対しても有効である。蛋白質等のバイオ物質での反応中心では、量子力学で記述される化学反応が起きている。計算生命科学には量子論が必要なのである。一方、生命科学においてはダイナミクスが重要である。量子論で記述される電子のダイナミクスはフェムト秒のオーダーである。一方、原子構造のダイナミクスはピコ秒、ナノ秒さらにはミリ秒のスケールである。そこには 10^{12} ものスケールの違いがある。そのマルチスケール現象を扱うのに、我々は Car-Parrinello 分子動力学法とメタダイナミクス法を結合した手法を採用している。詳細は原著論文を参照していただくことにして、計算例をひとつだけ紹介しよう。

4.1 シトクローム酸化酵素でのプロトン移動機構

シトクローム酸化酵素はミトコンドリア内膜に位置し、生物の呼吸作用の最終段階で、プロトンを膜内から膜間に輸送し、それによってATP合成反応を促す極めて重要な蛋白質である。牛由来のシト

クローム *c* 酸化酵素の結晶構造を元に、X線回折構造をもとにして見出されたプロトン移動経路 (H-path) では、Tyr440-Ser441間のペプチドグループが連續的な水素結合のネットワークを妨げており、従来のD-またはK-pathにおけるプロトン輸送とは異なる機構が必要となっていた。我々は、Car-Parrinello 分子動力学法とメタダイナミクス法を結合した手法により、Tyr440-Ser441間およびSer441-Asp442間の空間的に近接したふたつのペプチドグループを介した多段階反応により、共有結合を介したプロトン輸送が可能であることを示した[9,10]。これは、水分子などの他の因子を必要としない新規のプロトン輸送機構である (図8、図9)。

5. おわりに

以上駆け足で、我々のグループでの密度汎関数法計算の紹介を行った。密度汎関数理論は有用なものではあるが、完全ではない。実際、上述した体心正方格子 Si, Ge の研究では、準粒子スペクトルに対する、より精度の良い近似である GW 近似が必要であった。その意味で、計算手法の開拓は、計算物質科学、生命科学の推進には欠かせない。逆に応用計算と実験との比較により、新たな手法開拓の必要性が認識されてくる。理論、計算、実験が三位一体となって科学の進展を促すという構図である。さらに、我々の実空間密度汎関理論コードの開発に見られるように、計算手法を計算機上で具体化、高速化するためには、計算機科学 (computer science) と計算科学 (computational science) の密な共同が重要である。まさに異分野間の共同が、豊かな可能性を生み出していくひとつの例と言えよう。

謝辞

本稿で紹介した研究は、内田和之、合田義弘 (以上東京大学)、岩田潤一、高橋大介、白石賢二 (以上筑波大学)、藤本義隆 (東京工業大学)、Boero Mauro (Louis Pasteur 大学)、神谷克政 (兵庫県立大学) との共同によるものです。また、本稿で紹介した計算は東北大学サイバーサイエンスセンターの SX7-C および SX9 システムを主に利用して行ったものです。

参考文献

- [1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, (1964) B864; M. Levy: Proc. Natl. Acad. Sci. (USA) **76**, 6062 (1979).
- [2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
- [3] J.-I. Iwata, K. Shiraishi and A. Oshiyama, Phys. Rev. B **77**, 115208 (2008).
- [4] J.-I. Iwata, D. Takahashi, A. Oshiyama, T. Boku, K. Shiraishi, S. Okada, adn K. Yabana, J. Comp. Phys, to be published (2009).
- [5] Y. Fujimoto, T. Koretsune, S. Saito, T. Miyake, and A. Oshiyama, New Journal of Physics **10**, 083001 (2008).
- [6] Y. Gohda and A. Oshiyama, Phys. Rev. B **78**, 161201 (R) (2008).
- [7] K. Uchida, S. Okada, K. Shiraishi and A. Oshiyama, Phys. Rev. B **76**, 155436 (2007).
- [8] K. Uchida and A. Oshiyama, Phys. Rev. B (2009) in press.
- [9] K. Kamiya, M. Boero, M. Tateno, K. Shiraishi and A. Oshiyama, Journal of the American Chemical Society **129**, 9663 (2007).
- [10] K. Kamiya, Y. Shigeta, A. Oshiyama, AIP conference proceedings "Theory and Application of Computational Chemistry", AIP Conference Proceedings **1102** 257 (2009).