

次世代生体巨大分子の量子化学

関野 秀男^{a, b}

^a 豊橋技術科学大学・知識情報工学系, 〒 441-8580 豊橋市天伯町雲雀ヶ丘 1 - 1

^b 科学技術振興機構 (JST) CREST, 〒 332-0012 埼玉県川口市本町 4-1-8

e-mail: sekino@tutkie.tut.ac.jp

(Received: March 23, 2007; Accepted for publication: June 7, 2007; Published on Web: July 20, 2007)

巨大生体高分子への量子論の適用は新しい多体論を必要とするかもしれない。既存の量子化学的手法のままの形での拡張適用は理論的にも技術的にも破綻する可能性がある。生命科学に必要な精度を保障するにはシミュレーションにおける空間表現の問題が重要となる。次世代量子化学では半世紀以上にわたって使われているガウシアン基底を越える表現法が必要であろう。多重解像度多重ウェーブレット基底による分極率のような分子物性予測により、新しい効率の良い空間表現の試みを示した。また現在大規模量子化学計算の主流である密度汎関数法での物性算定の問題にも空間表現が関連している。

キーワード: ナノスケール, 空間表現, 多重解像度, ウェーブレット基底

20世紀初頭に出発した量子力学は一世紀を経た現在でもなおミクロ世界の自然原理を説明する最も信頼ある理論として多くの者に認められている。20世紀における量子力学の目覚ましい成功はサブナノスケール領域での物理を定量的に説明予想したことである。生命現象が生命の構成要素としての生体分子の物理によって解明されるはずであるという期待は量子力学の誕生とほぼ時を同じくして構築された。残念ながら量子力学の基本理念である不確定性や確率解釈にかかわらず、そこから得られるものは古典力学と同様な決定論的、可逆的考え方であり、生命現象特有の非可逆的特性を抽出するのは困難のように思えた。そこには生命現象を担う単位である生体巨大分子の非平衡な化学についての膨大な研究が解明される必要があるように思える。量子力学が成功をおさめるサブナノ領域のいわゆる“few-body” problem とは質的に異なる物理的問題、あるいはより技術的な問題が横たわっている可能性は大きい。本稿では現在実測がより大きな分子へと広がっている、生体巨大分子の物性算定についての量子化学計算について現状と展望を述べる。

上にもものべたとおり、生命現象の非平衡的な物理化

学にたいする、より信頼性のあるシミュレーションを提供するのが理論化学、計算化学の使命であるが、過去30年にわたって蓄積された量子化学計算のノウハウは全エネルギーという定常状態の言語に深く関わっている。実際 few-body シミュレーション方程式のように算定可能である場合、物理は全エネルギーにより判断されるより exact な解にしたがっているように思える。しかし重原子化学などシュレーディンガー方程式自身の精度が問われる分野に踏み込むまでもなく多電子問題という生体分子でごく一般的な状況でも多電子シュレーディンガー方程式にたいする一般的近似理論が確立されているわけではない。実際全エネルギーを指標とした巨大多電子系でのシミュレーション(例えば Ab-initio dynamics) が無数の局在最小値のために few-body 問題に比べ格段に難しいことは多くの研究者が経験している。もちろん位相空間でのアンサンブル操作により、より有効なシミュレーションを行おうという努力は進行中であるが、こうした困難が全エネルギー計算という few-body problem 解法の負の遺産であるという事実は否定しがたい。

巨大分子の全エネルギー計算にはより技術的な困難

も潜んでいる。現在のデジタルコンピューターでのフローティングポイント計算における1ワードの情報量の限界のため巨大分子の全エネルギーの有効桁数が確保できないからである。計算アルゴリズムの改善により有効桁数を改善することは可能であるが、現実には巨大原子の全エネルギーの絶対値にどれだけ有意義な情報があるかという問いを発するほうがより健全であると思われる。生体巨大分子量子化学計算の動機のひとつは geometry optimization であるかもしれない。然し上述のダイナミクスやモンテカルロシミュレーションの例が示すように、より現実的な、より大きくて、複雑な分子系での量子問題が全エネルギーを指標にした解析の不自然さを物語っている。

上にのべたような悲観的観測にもかかわらず、量子化学シミュレーションの有用性はもちろん存在している。とくに多次元NMRなど現在急速に発展している、より大きな分子量をもつ生体巨大分子のNMR実測情報の解析や電気分極特性の算定などである。そうした従来の基礎化学では既存の確立された技術がバイオ・ナノ領域の分子科学として注目されているが、そういったフロンティア開発へのシミュレーションからのサポートといった立場である。そうした分野での分子物性値算定に絶対値3桁以上の精度が要求されることはほぼない。多くの場合その物性は相対値としての重要性のみを有しているが、相対値といえども信頼性が保障されなければ無意味である。系が大きくなることで物性を決めている(ここでは電子)実体の物理が質的に変化してしまう可能性に対する細心の注意が必要となってくる。例えば一般中規模分子での電子状態記述にとって平均場近似(Hartree-Fock (HF)理論)が良いゼロ次近似であっても、金属的な電子状態にたいしては全く現実性を欠く理論であるため、そうした理論、あるいはHFを出発点(ゼロ次近似)とした摂動理論などはシミュレーションの道具として不向きであろう。一方凝縮系の電子状態理論に伝統的につかわれている密度汎関数法 Density Function Theory (DFT) は最近量子化学の分野でも主流になりつつあるが、長距離相互作用などの弱い相互作用の記述には現在のままの形では適用できない。生体巨大分子はその分子内相互作用としてこうした長距離力を介する結合を有し、その結合の解離・結合がまさに生体機能のオリジンとなっている場合も多い。ファンデルワールス力のような力もHF理論を出発点とすれば分子内電気分極の相互作用という形式で記述されることになり、その取り込みには高次の電子相関補正が必要であることになる。

電子相関という言葉自体の定義がHFからの補正であり、HF理論が出発点として良い近似でない物理が進行している領域ではその技術的取り込みが天文学的努力となることは想像に難くない。量子化学の創世期とその発展期に多大な貢献をのこした故 Pople 博士は量子化学による化学の定量的予測のためのダイヤグラムを提示して議論することが多かった。博士によれば量子化学発展は縦軸を空間表現の厳密性に対応させ、横軸を電子相関の取り組みの厳密性に対応させ、多くの研究者が横軸に平行な方向に進むか、縦軸に平行な方向で進んでいるが、量子化学のより実りある貢献のためには対角線上に右上がりに進展させるべきであることを説いた(Figure 1)。横軸の電子相関の理論的扱いについては高次摂動論、Coupled Cluster (CC) 理論、Configuration Interaction (CI) 理論などいわゆる厳密電子相関理論が提案応用され、少なくとも中規模以下の分子では大きな成果がもたらされてきた一方その計算時間や資源にたいする莫大なコストのため大規模分子系へのそのままの形での適用は不可能である。

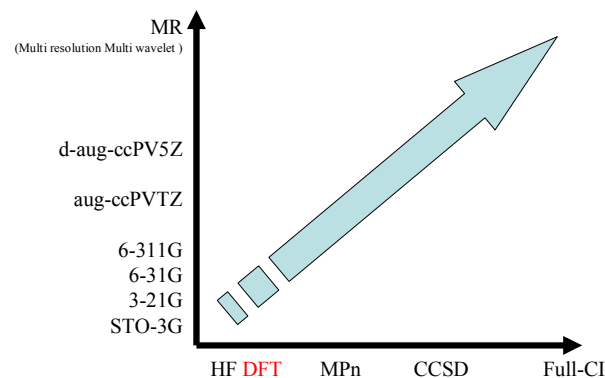


Figure 1. Quantum Chemistry by Dr. Pople

横軸の最右端にある Full-CI は与えられたヒルベルト空間の有限代数近似空間で最も厳密な解を与えるもので理論に要求されるすべての条件をみたしているが、非常に小規模な量子多体問題にしか適用できず現実的ではない。Pople 博士は空間表現の改善に対しては Linear Combination of Atomic Orbitals (LCAO) の考え方の延長路線をとり、2電子積分の扱いが比較的簡単なガウシアン基底関数による空間表現改善を試みた。原子軌道の非常に良い近似であるスレーター関数のガウシアン関数へのマッピングである STO-nG をその最も粗い空間表現とし、分極関数、ディフューズ関

数などを付け加えることによって化学的により精度の高い理論が有効性を発揮できる空間表現を構築した。ガウシアン基底関数に基づいた一般分子の量子化学計算プログラム開発は 20 世紀後半のデジタルコンピューター技術の飛躍的進歩と相まって 21 世紀をまたぐ 10 年間で成熟期といえる状態に成長した。電子相関の取り組みの中で晩年 Pople 博士は計算負荷の比較的少ない DFT 法を選び多くの分子の種々の物性算定を網羅的にを行いその有効性を示し、多くの化学者に量子化学の有効性を認めさせた。こうした量子化学の流れは現在も主流となっており、計算対象をナノスケール領域の巨大分子に広げることを目的とした巨大量子化学プログラム開発が進行している。然し、空間表現としてガウシアン基底関数を使うことに対する是非の議論が積極的にされているとは思われない。ガウシアン基底の示す柔軟性・一般性と 2 電子積分の容易さは量子化学創成期により一般的であったスレーター基底関数を凌ぐものであり、全エネルギーの差分から算定される化学反応の精度を保障するものであることが一般的に認識されたが、巨大分子系での議論に適用できるかどうかの判断には客観性はない。冒頭で述べた如く、こうした 20 世紀に成功をおさめた量子化学技術の巨大分子系への直接応用には落とし穴が潜んでいる可能性もある。空間表現にたいするより数学的、システムティックな改善方法が必要であると思われる。

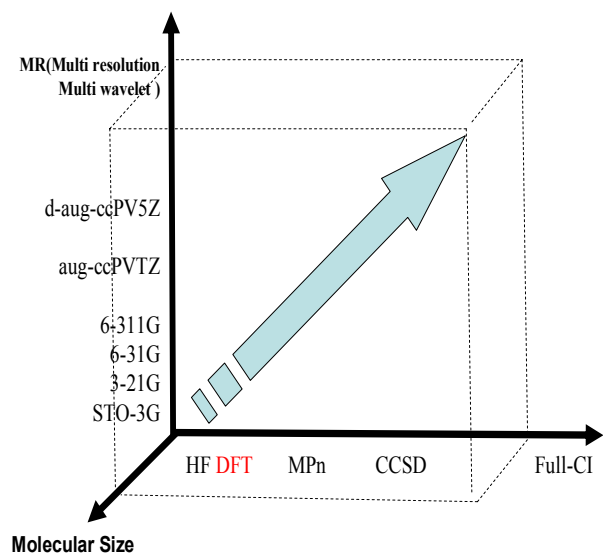


Figure 2. Quantum Chemistry for Large Systems

Figure 2 に Pople 博士のアナロジーとしての巨大系

での量子化学についてのダイアグラムを示す。ここで云いたいのは分子のサイズが化学精度や予見性などに対する理論、空間表現の有効性に質的影響を及ぼすかどうかの検証が必要であるということであり、それは新たな独立の第 3 軸目としてダイアグラムに加えられべきものであるということである。特に従来の空間表現での精度向上が大規模系でそのまま通用するかの判断を全エネルギー算定のみ任せずよりシステムティックに行うことの提案である。

最近有限要素法など局在基底関数により空間表現を行い、波動関数の評価を空間上の各グリッドでの数値評価を行う方法も進展しているが、我々は更にそうした局在関数を基としたスケーリング関数により階層的空間を構築し、多重解像度解析の考え方のもとにウェーブレット基底関数を構築する方法を量子化学の問題に応用する方法を提案している。

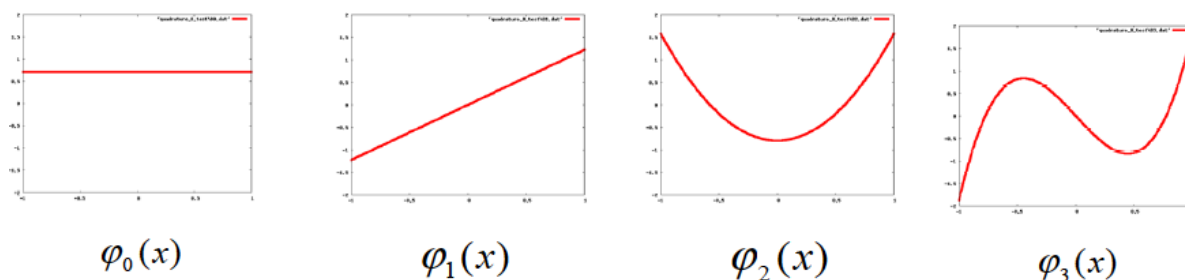
以下に局在基底関数としてルジャンドル多項式を用いたスケーリング関数により構築されるウェーブレット基底の例を示す。

$$\phi_i(x) = \begin{cases} \sqrt{2i+1}P_i(2x-1) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

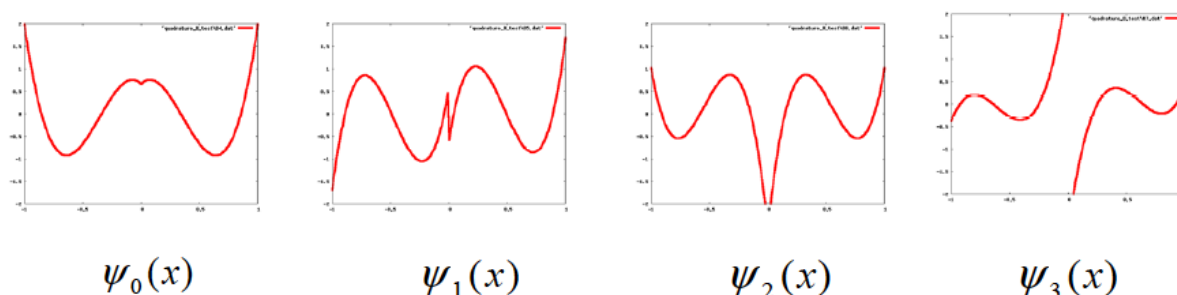
$$\phi_i^n(x) = 2^{n/2}\phi_i(2^n x - l)$$

空間表現の厳密性は解像度の階層 n と各ボックス内での完全性への近似度の指標である多項式の位数 k のみである。上に $k=4$ の場合のマルチウェーブレットの具体的形を示す。ガウシアン関数と異なり各ボックス内、またボックス間での disjoint 性にもかかわらず、システムティックに空間表現の改善が行われ、ガウシアン diffuse 関数を加えたときに起こるいわゆる over-complete problem を回避しつつ核から離れた空間の表現改善を行うことが可能である。こうした特性は基底状態の全エネルギーにほとんど影響をおよぼすことがないに関わらず、系の応答特性算定では重要である空間部分の改善に特に役に立つ。我々は多重解像度多重ウェーブレット基底をもちいて分子の分極率の算定を行い以下のようにシステムティックな改善が得られ非常に大規模なガウシアン基底でも到達できない空間表現にせまれることを見出した [1-6]。今後さらに大規模な計算や量子ダイナミクスシミュレーションにおいてもこうした方法が有効である [7, 8] ことが示されていくことが期待される。

Scaling function (V_0) (Legendre polynomial)



Wavelet (W_0)



These 8 functions construct an orthogonal set

Figure 3. Multi-Wavelet ($K=4$)

Table 1. Molecular Polarizability

以下にマルチウェーブレット基底 $K=9$ $\epsilon = 10^{-6}$ に対する誤差の割合の平均値を示す。

対象分子: C_2H_4 , CO , F_2 , H_2O , HF , NH_3 , SO_2 , CH_4 , CO_2 , H_2 , HCl , N_2 , PH_3

Hartree-Fock		DFT(Svwn5)	
Gaussian function basis			
cc-pvdz	33.539%	cc-pvdz	39.373%
cc-pvtz	19.061%	cc-pvtz	24.224%
cc-pvqz	10.948%	cc-pvqz	15.322%
cc-pv5z	6.723%	cc-pv5z	9.916%
aug-cc-pvdz	3.630%	aug-cc-pvdz	5.128%
aug-cc-pvtz	0.852%	aug-cc-pvtz	1.629%
aug-cc-pvqz	0.496%	aug-cc-pvqz	0.276%
aug-cc-pv5z	0.102%	aug-cc-pv5z	0.334%
d-aug-cc-pvdz	0.295%	d-aug-cc-pvdz	0.330%
d-aug-cc-pv5z	0.013%	d-aug-cc-pv5z	0.014%
MRMW basis		MRMW basis	
k=5 $\epsilon = 10^{-4}$	0.123%	k=5 $\epsilon = 10^{-4}$	0.094%
k=7 $\epsilon = 10^{-5}$	0.001%	k=7 $\epsilon = 10^{-5}$	0.002%

非ガウシアン基底関数に基づいた量子化学計算法は本稿で紹介した以外にも幾つか存在するが、30年以上にわたって蓄積されたガウシアン基底による計算法のノウハウは大きく簡単な計算時間や算定能力の現時点での比較では新手法にとって圧倒的に不利なものとなるであろう。然し、一旦プログラム開発の方向が新しい空間表現の方向に向えば多くの研究者や技術者の参入により状況が一転することを考えるのはそれほど難しくはない。本稿で紹介した局在関数によるウェーブレット基底関数はその局在性故に、超並列アルゴリズム

ムなどの新しいアーキテクチャーに向けた特性をもち、いっきょに巨大分子算定プログラムの中心的役割を担う存在になる可能性を持っている。

現行の多くの巨大分子量子化学計算はその全体を分割することにより、 N - オーダースケーリングを実現しようとしている。北浦等によって提唱されたフラグメント分子軌道法 (FMO) もそのひとつで最近タンパク、DNA、またそれらの複合体のモデル計算もなされるようになってきた。もちろん分割することによる代償は償わなければならない FMO 法の場合、フラグメントのペア補正という摂動論的方法により分割によって損なわれたフラグメント間電子移動効果の補正を行う。こうした分割法においても現在主となる空間表現はガウシアン基底関数によっており、その良悪両面を引き継ぐことになっている。特に分極物性算定に最重要である領域での空間表現が diffuse 関数という巨大分子計算にとって不自然な基底関数によって規定されるため、分割操作自体がむずかしくなっている。多重解像度解析のようなより合理的な分割スキムを導入することによりそうした方法論の更なる発展を期待するところである。

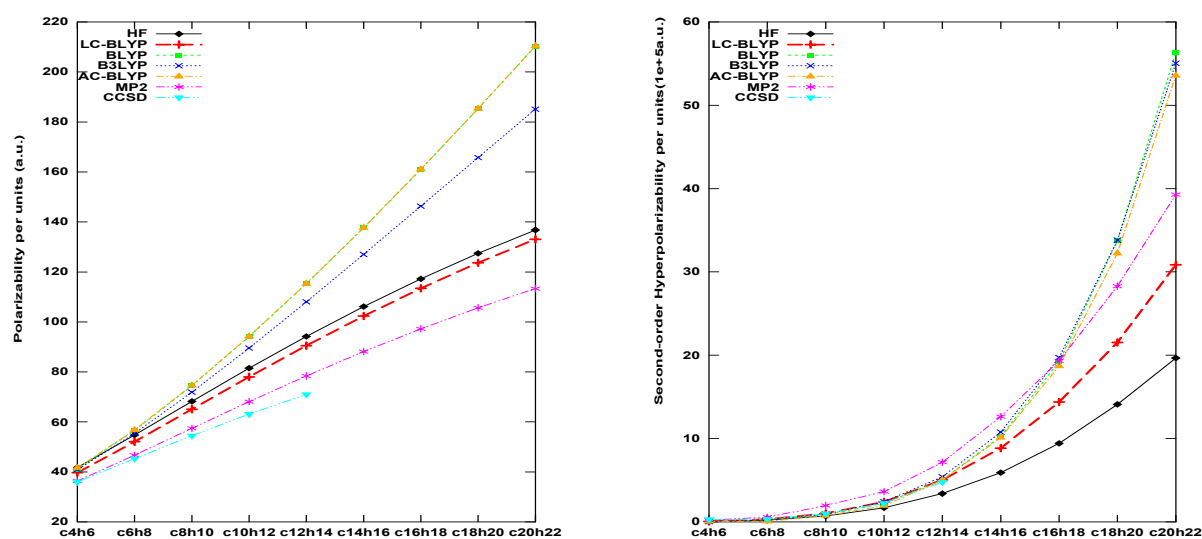


Figure 4. Polarizability α (left) and Second hyperpolarizability γ (right) of Polyene

最後に、現在巨大系量子化学計算では密度汎関数法 (DFT) が主流であるが、本稿で述べた理論及び技術的問題がこの方法論 (DFT) を使った応答物性の算定に顕著に現れており、いわゆる brute force な本手法の大規模系への応用は定性的な予想すら不可能であることを付け加えておく。ポリマーでの分極率、あるいは超分極率の DFT 計算は分子のサイズを大きくすると発散してしまう [9]。応用面からの現行の DFT の改善も試みられている [10, 11] が (Figure 4)、そうした物性算定には、その汎関数が基底状態をうまく記述できるだけでは不十分であり系の応答を的確に記述できなければならない、その構築には正確な空間表現が必要であることも分かってきた [12]。

参考文献

- [1] R. J. Harrison, G. I. Fann, T. Yanai, Z. Gan and G. Beylkin, *J. Chem. Phys.*, **121**, 11587 (2004).
- [2] H. Sekino, Y. Maeda and S. Hamada, *IC-CMSE2005*, 686 (2005).
- [3] Y. Maeda, H. Sekino, T. Yanai and R. Harrison, PACHEM2005 (Hawaii)
- [4] H. Sekino, PACHEM2005 (Hawaii), 分子構造総合討論会 2006 (静岡)
- [5] T. Yanai, R. J. Harrison, Y. Maeda and H. Sekino, ICQC2006 (Kyoto)
- [6] H. Sekino, Y. Maeda, T. Yanai and R. J. Harrison, to be published
- [7] S. Hamada and H. Sekino, ICQC2006 (Sendai), 日本物理学会 2006 秋季年会 (千葉)
- [8] N. Sano, S. Hamada and H. Sekino, 分子構造総合討論会 2006 (静岡)
- [9] B. Champagne, E. A. Perpète, S. J. A. van Gisbergen, E.-J. Baerends, J. G. Snijders, C. Soubra-Ghaoui, K. A. Robins and B. Kirtman, *J. Chem. Phys.*, **109**, 10489 (1998).
B. Champagne, E. A. Perpète, D. Jacquemin, S. J. A. van Gisbergen, E.-J. Baerends, C. Soubra-Ghaoui, K. A. Robins and B. Kirtman, *J. Phys. Chem. A*, **104**, 4755 (2000).
- [10] M. Kamiya, H. Sekino, T. Tsuneda and K. Hirao, *J. Chem. Phys.*, **122**, 234111 (2005).
- [11] H. Sekino, Y. Maeda, M. Kamiya and K. Hirao, *J. Chem. Phys.*, **126**, 012107 1-6 (2007).
- [12] H. Sekino and H. Matsumoto, to be published

Next Generation Quantum Chemistry for Large Bio-Molecules

Hideo SEKINO^{a, b}

^aKnowledge Based Information Engineering, Toyohashi University of Technology

1-1 Hibarigaoka Tenpakucho Toyohashi, Japan

^bJapan Science and Technology Agency, CREST

4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

e-mail: sekino@utkie.tut.ac.jp

The application of quantum theory to the investigation of large bio-molecular systems may require an alternative general theory of many-body problems. Straightforward extension of rather established quantum chemical methodologies may encounter serious problems both theoretically and technically. In order to fulfill the required precision for bio-science, the representation of space in simulation is critical. For next generation quantum chemistry, there should be more refined tools than Gaussian basis functions which have been used over half a century. Preliminary application of multi-resolution multi-wavelet basis functions proved its efficiency in the prediction of molecular properties such as polarizability.

Keywords: Nano scale, Space representation, Multi-resolution, Wavelet basis