

## Gaussian 03を用いた分子軌道計算の実際

神崎, 亮  
九州大学大学院理学研究院化学部門

<https://doi.org/10.15017/1472539>

---

出版情報：九州大学情報基盤センター広報：全国共同利用版. 5 (2), pp. 65-68, 2005-10. 九州大学情報基盤センター  
バージョン：  
権利関係：



# Gaussian 03 を用いた分子軌道計算の実際

神崎 亮<sup>1</sup>

**概要** 現在の化学において分子軌道計算は重要な地位を占めるようになった。分子軌道計算を化学の“ツール”として使うユーザーとしての視点から、まずはその背景を概説し、そして九州大学の情報基盤センターに今回新しく導入された高性能演算サーバーで、どれぐらいの系がどれぐらいの時間で計算できるのか、限られた例ではあるが紹介したい。

## 1. 非経験的分子軌道計算への期待

近年、化学の世界では分子軌道計算は必須のツールとなっている。我々が分子の構造を知りたいと思っても、X 線や中性子線を用いたかなり特殊な方法で“推定”する事しかできないが、分子軌道計算を行うと最安定な構造がたやすく可視化できる。化学者はなんとか分子の世界を垣間見たいと日々励んでいるから、これは魅力的である。

分子軌道計算がこれほどまでに一般的になったのは、計算機の高速化、優秀なパッケージングソフトの出現に依るところが大きい。九州大学箱崎キャンパスでは、最も広く使われているであろう分子軌道計算ソフト Gaussian 03 のサイトライセンスを受けている\*。Windows 版もあるので、小規模な系であれば、普段デスクトップで使っているようなパーソナルコンピュータでさえ十分な計算ができる。九州大学の職員・学生であればすぐにでもその恩恵に与ることができるので、試みたことのない方にはぜひお勧めしたい。一昔前までは限られた研究者によって限られた系を対象としてきたわけだから、便利な世の中になったものである。

一方で、理論や計算アルゴリズムの発展も多くの化学者に受け入れられた理由として挙げられる。分子軌道計算は、多数の原子からなる分子の分子軌道を解く過程で、どうしても近似や仮定が含まれる。ほんの 5 年も前までは、得られたエネルギーと実測値があまりにも乖離しており、議論に耐えなかった。近年ようやく、十分な理論と基底関数を用いれば有用な情報が得られることが、計算を専門にしない化学者にも認識され、議論に組み入れられるようになった。計算化学者のたゆまぬ研究の賜物である。

計算時間と精度の折衷案として用いられてきた半経験的分子軌道計算との対比も俚上に乗せよう。これは多くの計算時間を費やす数値積分を、既知の実験値を再現するような経験的パラメータで置き換える方法である。高速でかつ多数の原子を扱うことができるため、現在においても系によっては威力を発揮する。しかし非経験的に分子軌道を解くことには、

- ・配位結合や水素結合など、非共有結合を記述できる
- ・遷移金属元素を含めた全ての元素を計算できる
- ・活性化状態や反応素過程の中間体など、実験的には作り得ない状態を扱える

<sup>1</sup> 九州大学大学院 理学研究院 化学部門 [violascc@mbbox.nc.kyushu-u.ac.jp](mailto:violascc@mbbox.nc.kyushu-u.ac.jp)

\* [http://www.cc.kyushu-u.ac.jp/scp/system/library/campus\\_licence/Gaussian03.html](http://www.cc.kyushu-u.ac.jp/scp/system/library/campus_licence/Gaussian03.html)

といったメリットがある。これらは現代の化学のトピックスであるから、半経験的方法是既に力不足となりつつある。新規の分子が持つかもしれない新たな機能や、分子集合体の織りなす未知の作用に焦点を当てようとしたとき、半経験的な方法では結論を大きく見誤る危険性を伴う。

## 2. 高性能演算サーバー

さて、より大きな系や基底関数を用いたり、励起状態や振動数解析などの複雑な計算を行ったり、また反応のポテンシャルエネルギー曲面を描くために大量の計算をしたいといった場合は、より大型の計算機資源が必要となるし、また並列化によって計算に要する実時間が少なくなれば、限られた時間でより多くの結果を得ることができる。九大の情報基盤センターでは、これまでもスーパーコンピュータやスカラー並列サーバーによる分子軌道計算環境が提供されてきたが、2005年、新たに IBM eServer p5 (POWER5、1.9GHz) 高性能演算サーバーが導入された。今回は導入と同時に Gaussian 03 がインストールされ、専用のチューニングが施されたとのことである。我々のような分子軌道計算をツールとして“使うだけ”のユーザーに便宜が図られたことについて、関係各位に感謝したい。今後ますます化学の分野で計算の占める割合が高くなるであろうことを考えると歓迎すべきことである。

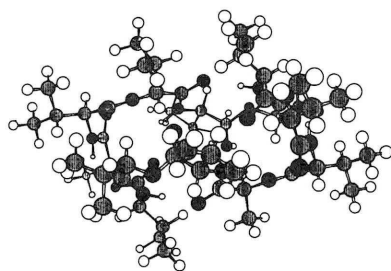
新サーバー移行に伴い、テスト期間が設けられた。そこで期間中に、2つのテストジョブ、および実際に研究で用いている系についていくつかのパターンの並列処理を実行し、どの程度の計算時間が必要か、またどの程度計算速度向上が得られるのかを試みたので、報告したい。テストジョブとは、Gaussian のリリース元であるガウシアン社が提供する数百個の典型的なジョブであり、Gaussian の移植やコンパイルが正しく行われたかを判断するための計算ファイルである。そのうちいくつかは“定番”となってベンチマーク代わりに使われている。

## 3. ジョブ実行時間の比較

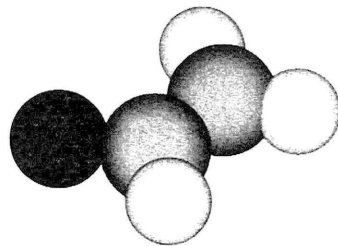
ジョブの完了に要した CPU 時間、および実際にかかった時間を、下に表で示した。単位は全て秒である。テストジョブでは、筆者の研究室で稼動中の Intel Itanium 2 (1.4GHz) ×2 CPU マシンによる結果も示した。この構成は、研究室単位でコストパフォーマンスの良い計算システムを構築しようとした時のモデルの一つであると言ってよいだろう。なお、ここでは得られた結果の妥当性については議論しない。最新の化学でも、気相中に置かれた単分子のエネルギーや構造を正確に決定するのは極めて難しい。

### 3-1. テストジョブの実行

下の表で、kyu-cc と示しているのが、新しく導入された高速演算サーバーによる結果である。スカラー並列サーバー kyu-ss (COMPAQ AlphaServer GS320)、前述の Itanium 2 機 (IA64 と表記) による結果も合わせて示してある。



バリノマイシン、 $C_{54}H_{90}O_{18}N_6$



フルオロエチレン、 $C_2H_3Cl$

今回用いたテストジョブは、以下の 2 つである。

test397 (バリノマイシンの力定数計算) : 抗生物質として有名なバリノマイシンは、6 つのペプチドが環状に連なった分子である。この分子における各共有結合の力の定数を RB3LYP/3-21G のレベル (基底関数 882) で計算する。test397 はかなりの CPU 時間を消費する一方で、ディスクアクセスが少なく、また並列化処理されているルーチンを多用するため、CPU のベンチマークやコンパイルによる最適化の評価指標として用いられている。

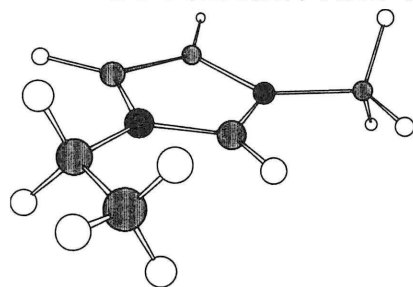
test415 (フルオロエチレンの MP2 計算) : このジョブで取り上げられているフルオロエチレンは原子数わずか 6 個の分子であり、MP2/6-311G(df,pd)レベル (基底関数 108) で、中性分子、および陽イオンラジカルの計算を行っている。巨大な一時ファイルを作成するためディスクアクセスが頻発し、実行時間がディスク性能に依存する。特に後半のイオンラジカルの計算 (Job-2) では、この一時ファイルサイズが 2GB (31 ビット) を超えるため、実行不可能な OS もある。

| Platform | #CPU | Test397      |              |       | Test415, Job-2 |              |       |
|----------|------|--------------|--------------|-------|----------------|--------------|-------|
|          |      | CPU Job Time | in Real Time | Ratio | CPU Job Time   | in Real Time | Ratio |
| kyu-cc   | 2    | 5,999        | 3,076        | 1.95  | 847            | 895          | 0.95  |
|          | 4    | 6,108        | 1,629        | 3.75  | 955            | 905          | 1.06  |
|          | 8    | 6,513        | 1,123        | 5.80  | 1,130          | 896          | 1.26  |
|          | 16   | 6,597        | 1,114        | 5.92  | 1,479          | 914          | 1.62  |
| kyu-ss   | 1    | 19,017       | 19,023       | 1.00  | 2,230          | 2,240        | 1.00  |
|          | 2    | 20,204       | 10,433       | 1.94  | 2,814          | 1,915        | 1.47  |
|          | 4    | 21,934       | 5,841        | 3.76  | 3,929          | 1,961        | 2.00  |
|          | 8    | 30,062       | 4,792        | 6.27  | 13,062         | 3,498        | 3.73  |
| IA64     | 2    | 3,405        | 3,415        | 1.00  | 471            | 469          | 1.00  |

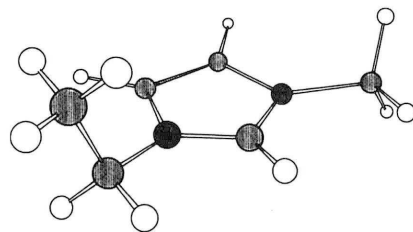
結果は表に示したとおりである。まず test397 では、同じ数の CPU を用いても kyu-ss 機に比べ約 4 倍の速度向上が図られている。また、計算を化学のツールとして考えた場合、実実行時間が問題となるが、8 CPU 程度までは実時間も CPU 数に反比例しており、並列化がうまく機能していることがうかがえる。ただし 16 CPU では実時間の短縮化は見られず、このあたりに限界がありそうである。Itanium 2 機と比較すると、速度は CPU クロック比までは早くなっていない。しかし CPU の数を増やすことで単位時間あたりに多くの計算結果を得られるのはやはり強みである。一方、ディスクアクセスの多い test415 の 2 番目のジョブであるが、予想通り並列化はあまり機能していない。実時間も CPU 数を増やすと逆に増える傾向にある。ただしそれでも kyu-ss に比べ倍程度には高速化されており、CPU 周辺部の能力の底上げも図られていることが窺える。

### 3-2. 実際の研究に用いる系における計算

次に、実際の系について少し触れたい。emim<sup>+</sup> (エチルメチルイミダゾリウム) は、イオン液体と呼ばれる常温で溶液状態となる塩の陽イオンとしての有力な候補である。



emim<sup>+</sup> (0°)



emim<sup>+</sup> (105°)

イミダゾリウム環やメチル基はほとんど固定されているが、エチル基は自由に回転できる。事前の計算で、エチル基の末端炭素がイミダゾリウム環平面上にある場合と、2 位との炭素とのねじれ角が  $105^\circ$  になったときの 2 つの安定構造が存在することが見出されていた。そこで、この 2 つの構造を、MP2/aug-cc-pVDZ レベルで 2 CPU、および 16 CPU で計算した。結果は CPU 数が 8 倍になったにも関わらず実計算時間は半分にも短縮されず、CPU の並列化が汎用的に効果を発揮するとは残念ながら言えなさそうだ。

Gaussian は、Link と呼ばれる PE (Program Executable; 単独で実行可能なファイル) として切り離されたサブルーチンの集合からなる。今回は、計算時間が比較的大きなルーチンにかかった時間も比較してみた。Link 502 は SCF (自己無撞着場) 到達のための繰り返し計算ルーチン、Link 1002 は Coupled Perturbed HF (摂動を含む Hartree Fock) を解く繰り返し計算ルーチン

| <i>emin</i> , $0^\circ$ | 2cpu     |           | 16cpu    |           |
|-------------------------|----------|-----------|----------|-----------|
|                         | CPU-Time | Real-Time | CPU-Time | Real-Time |
| Total                   | 4,087    | 2,065     | 20,167   | 1,430     |
| Link 502                | 1,081    | 544       | 317      | 83        |
| Link 906                | 1,292    | 653       | 19,156   | 1,213     |
| Link 1002               | 1,369    | 692       | 327      | 94        |

| <i>emin</i> , $105^\circ$ | 2cpu     |           | 16cpu    |           |
|---------------------------|----------|-----------|----------|-----------|
|                           | CPU-Time | Real-Time | CPU-Time | Real-Time |
| Total                     | 4,061    | 2,051     | 28,364   | 1,808     |
| Link 502                  | 1,077    | 542       | 1,128    | 77        |
| Link 906                  | 1,289    | 651       | 25,469   | 1,608     |
| Link 1002                 | 1,351    | 682       | 1,403    | 95        |

であり、これらの計算は並列化の恩恵を受けているが、MP2 (電子相関摂動) 計算のルーチンである Link 906 が足枷となっているようだった。しかし、同じ MP2 計算を行っている test415 では並列化が効率的に機能しているようなので、同じ理論やルーチンでも分子や基底関数によって並列化の効果は大きく異なるようである。

#### 4. まとめ

結論としては、今回導入された計算機サーバーは満足の行くパフォーマンスを発揮しているように思う。CPU 単体の能力、およびディスクアクセス速度の向上もさることながら、並列化による高速化も kyu-ss 機より効率的であり、より計算実行時間の短縮化が期待できるであろう。ただし当然のことではあるが、並列化がうまくいくかどうかは系によって異なり、分子の大きさや用いる理論、基底関数によってかなり差があることも明らかとなった。今回いくつかの例を示し、ある程度の目安にはなったと思われるが、やはり最初は自分が計算したい分子、理論で実際に計算を仕掛け、どの程度の並列化が最も効率が良いか、当たりを付ける必要がある。