# Spartanの化学シフト計算による天然物構造決定

弘前大学 農学生命科学部 分子生命科学科
教授 **橋本 勝**
hmasaru@cc.hirosaki-u.ac.jp

1978年パーソナルコンピュータ（当時はマイコンと呼ばれていました）が発売されて35年が経過しました。その間、CPUクロックだけで0.75KHzから3.5GHzと4万5千倍に計算速度が向上しました。1年かかっていた計算が10分で終了することになります。当時は8ビットCPUでしたが、現在のパソコンの主流は64ビットで、一度に扱えるデータ量も$10^{14}$倍となり、さらに、CPUマルチコア技術や並列計算など計算機数学の発達を合わせると、コンピュータの性能は比較にならないほど向上しました。現在のパソコンは15万円程度のもの（50-100 GFLOPS）ですら、1997年チェス世界チャンピョンを負かせたとして名を馳せたスパコンDeep Blueの計算速度（11.4GFLOPS）を凌駕します。計算化学の分野では大澤映二先生が1980半ばに分子力学ソフトMM2をパソコン用に移植されました。当時のパソコン（16bit, 8MHz）では、筆者の学位論文の題材であったトリテルペン一配座を計算するのに一晩かかっていましたが、同じ分子を現在のパソコン（64bit, 3.5GHz）で、分子力学法MMFFにより計算してみたところ、0.03秒で終了しました。計算速度の向上は、当時は大型計算機が必要であった計算を日常的なパソコン上でも十分可能としました。大型計算機でも厄介であった計算すら、現在のパソコンは実行可能です。

分子モデリングソフトSpartanはSpartan'08 for Windowsからマルチコアに対応し、計算の内容によってはコア数に対応した並列化計算が可能になり、著しい高速化が実現されました（積載可能メモリー量の関係から64bit OSを推奨、最大8スレッド（並列）まで）。これによって密度汎関数法を用いた分子軌道計算も分子量500程度までパソコンの守備範囲となり、多くの天然物の計算が現実的になりました。分子構造を議論するのであればHartree-Fock法でも十分ともいわれますが、スペクトルなど物性に関する議論では密度汎関数法による計算が必要になるそうです。言い換えれば、ソフトを含めた最近のパソコン環境は多くの天然物のスペクトルを、分子軌道法に基づいて計算することを可能にしました。筆者もこれまで天然物の化学シフトについて計算で議論をした経験がありますが、その適用範囲は限定的だと考えていました。[1-3]しかし密度汎関数法を利用することにより多くの天然物に適用できるレベルになったと実感しています。[4]

筆者は分子計算についても、コンピュータについても専門家ではありませんが、Spartanを覚えて分子モデリングに挑戦してきました。すでに多くの人によってその有用性が示されていますが、分子軌道計算に基づく化学シフト予想は想像以上に正確で、天然物構造決定の際、十分に有用であることを実感しました。ここではSpartan10を用いた化学シフト予想を、筆者の経験と共に紹介します。使用しているパソコンは、事務作業と兼用のWindowsマシンです。近くのパソコン屋さんにお願いして15万円程度の予算作ってもらったものです。中程度のゲームマシンといったスペックだそうです。バックグラウンドで計算していても、ワープロなど事務作業に使っていますが、ワープロが遅くなったという感覚はほとんどありません。

分子には複数の配座が存在しますが、配座ごとに化学シフトは異なります。$^{13}$C化学シフトで5-7ppm異なることも珍しくなく、精密な議論にはスペクトルに影響する配座すべてを考慮する必要があります。以下に化学シフトを計算する場合の一般的なスキームを示しました。

**化学シフト計算の流れ**

① 配座解析　② 構造精密最適化
③ 化学シフト計算
④ ボルツマン分布に基づく化学シフトの補正

### ① まず、安定配座を求めます。

Spartanシリーズは環構造を含めた配座解析が可能な数少ないパソコン用の分子モデリングソフトです。すべてゴーシュ配座と仮定した場合でも、回転可能な結合数Nとした場合、可能な配座数は$3^N$個と膨大です。アルゴリズムにもよりますが、配座解析では、原理的にそれらすべての配座を評価する必要があり、高速計算が

要求されます。小さな分子でも設定次第では数十万の配座を検索することになりますが、途中からはほとんどの配座がそれまでに見出した配座と一致するようになります。Spartanのでは、プログラムが「もう新しい配座は見つからない」と判断すると配座解析は終了します。このアルゴリズムが重要な配座を見逃すことにならないのか、筆者には判りません。Spartanにおける構造最適化では分子力学法から分子軌道法まですべての計算手法の選択が可能ですが、開発者であるHehre氏はMMFFの使用を推奨しています。いくつかの分子を想定してMMFF計算をしてみましたが、簡単な分子では化学シフト計算で用いるEDF2/6-31G*法とエネルギーについて強い相関がみられ、とても優秀といえます（図1、compound 1）。しかし、官能基が複雑な分子では成績が極端に悪くなる場合があります（compound 2）。言い換えれば、EDF2/6-31G*法が正しい配座を示すと仮定して、MMFF計算では最安定配座を逃す可能性が無いとは言えないわけです。HF/STO-3GやHF/3-21Gは、官能基の有無にかかわらず比較的良い成績を示しました。これらはMMFF法に比べ約500～1000倍計算時間を必要としますが、密度汎関数法に比べれば10倍程度高速です。分子量500、100配座程度であれば、一晩程度で構造最適化が完了します。したがって、取りこぼしが不安な時は、保存する配座数を増やしてMMFで配座解析を行った後（Spartan配座解析のデフォルトでは10 kcal/mol以内の最大100配座を保存します）、HF/STO-3GやHF/3-21G で構造最適化しなおすことによって、これらを用いて配座解析したと同程度の結果が得られるはずです。また、MMFF配座解析では別の配座とされていたものでも、HF/STO-3GやHF/3-21Gで最適化し直した時。一つの配座に収束する場合もあります。重複した配座を除くことも、後の計算時間短縮には重要です。実際には重要な配座でもMMFF配座解析で漏れてしまう場合も、時々あります（理由は、精度が原因で安定配座のウィンドウに入らない、本来なら安定配座であるもののMMFFエネルギーカーブでは極小値ではない、などが考えられます。）。このような場合、手動で配座を足す必要があります。計算機実験とはいえ、化学の目でチェックすることは大切です。分子によっては直接HF/3-21Gで配座解析することも可能です。

尚、Spartan08以前のバージョンでは、二重結合の異性体も配座異性体として安定配座リストに出てきてしまうバグがありましたが、Spartan10では修正されています。環内に二重結合を有する分子を配座解析する場合、現在のバージョンでの配座解析を推奨いたします。

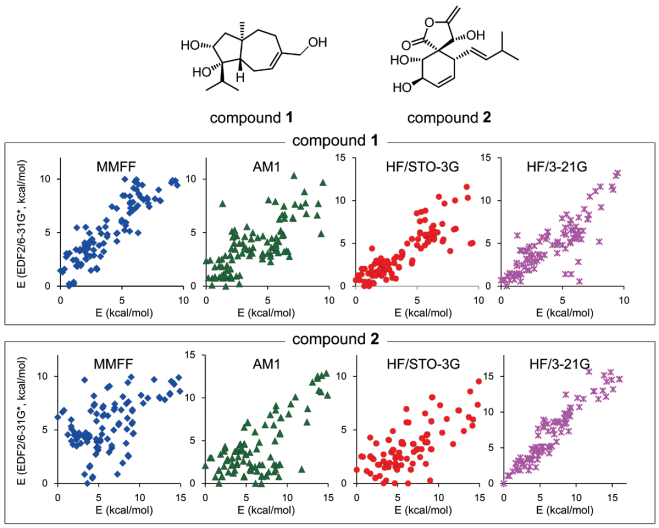



**図1. 計算精度の比較**
MMFF法で配座解析し、安定100配座を得た後、それぞれの方法で、構造を最適化し直し、得られた立体エネルギーをEDF2/6-31G*法のものと比較した。同一配座削除の操作は行っていない。

## ②、③ 得られた安定配座について精密に構造を最適化し、化学シフトを求めます。

Spartanでは計算準備のパレットでNMRのボックスをクリックするだけです（図2）。

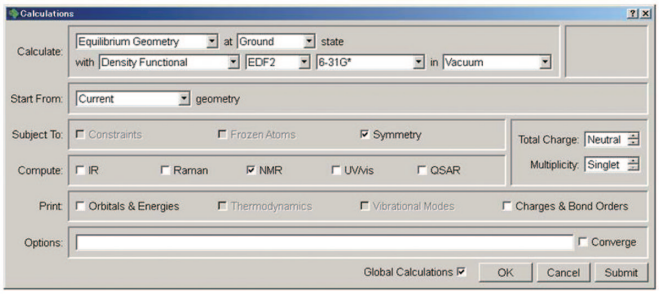



**図2. Calculationsパレット**

以前のバージョンではパソコンの性能が不十分であったのが原因なのか、密度汎関数法で構造最適化することはできず、Hartree-Fock法で構造最適化し、その構造を固定して立体エネルギーと化学シフトを求めていました。現在のバージョンでは、密度汎関数法で構造最適化と化学シフト計算を同時に計算することができるようになり、作業が単純化されただけでなく、計算精度も向上しました。一般に密度汎関数を用いた化学シフト計算ではB3LYP/6-31G*を用いた例が多く文献に紹介されていますが、開発者のHehre氏はEDF2/6-31G*の使用を強く勧めています。EDF2法は赤外線吸収スペクトル再現に優れた計算手法だそうですが[5]、開発者にはNMR化学シフトについても、EDF2法にアドバンテージを示すデータがあるものと想像します。筆者もB3LYP/6-31G*とEDF2/6-31G*を比較してみましたが、これまでのところ後者の方がいくらか成績が良い感触を持っています。Hehre氏によれば基底関数は6-31G*より精度を高くしても意味がないそうです。計算が速くなったとはいえ、分子量500ともなると、一配座を最適化・化学シフト計算すると、その初期構造にもより

ますが、12時間程度要する場合も少なくありません。Spartanでは配座解析が見出したすべての配座が一つのファイルにまとめられて保存されおり、一回の設定でそのすべてを計算することが可能です。しかし、配座のセットすべてを計算しようとすると、その所要時間は容易に数週間に達します。途中、落雷停電があれば水の泡です。OSの自動更新をOFFにしていても、夜中に事務作業用アプリケーションの更新・再起動してしまったといった事故も何回か経験しました。現在は、被害を最小にするため、配座解析の結果を数配座ごとに分割して計算するようにしています。化学シフトに影響しない高エネルギーの配座を計算しても意味がありません。また通常の計算では溶媒効果を考慮しません。（Spartanでは溶媒効果を考慮して計算することは可能ですが、これによって精度が高くなる保証はありません。計算時間は確実に増大します。）したがって、計算の精度は$^{13}$Cで1.5ppm, $^{1}$Hで0.2ppm程度と想定されますことから、最安定配座から2kcal/mol以内の配座を計算すれば十分であると思います。各配座の計算した化学シフトと画面上に示すことも可能です（図3. 化学シフトの表示の一部は重なっていますが、パソコン画面上では分子を回転させることにより、正しく読むことができます）。またSpreadsheetを利用することによって、すべての配座の化学シフトの一覧を作成することも容易です。（図4.）

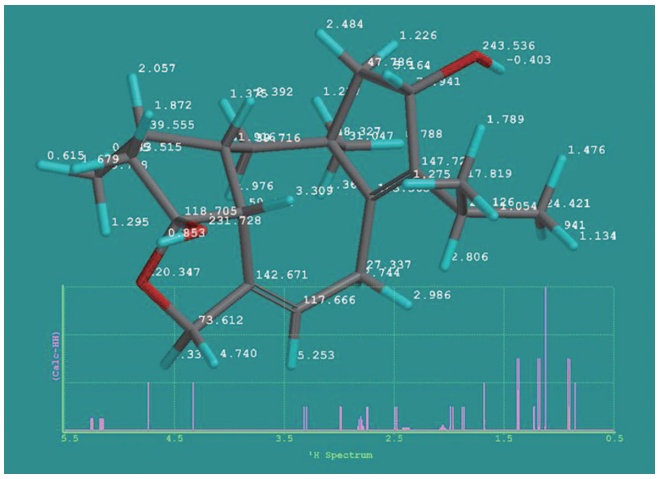
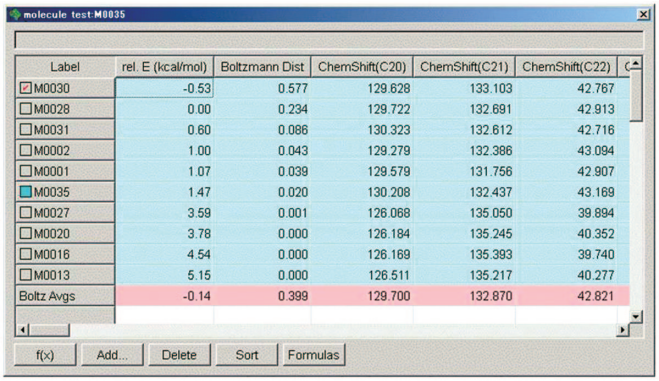


**図3. 計算結果の化学シフトとスペクトルを表示させたパソコン画面**

### ④ 得られた各配座の化学シフトをボルツマン分布に基づく加重平均を算出することで、計算化学シフトを得ます。

Spartanでは、Spreadsheet上で計算することができます（図4、最終行）。また、Spartan上に計算したスペクトルを再現させることもできます。結合定数も考慮して再現させることもできます。マニュアルによれば、結合定数は経験値に基づくもので、理論計算によるものではないそうです。理論的に計算することも可能ですので、この機能の追加を期待してしまいます。なお、配座のBoltzmann分布考慮したスペクトルを表示させることも可能ですが、この場合、結合定数を考慮した表示はできません。



molecule test-M0035

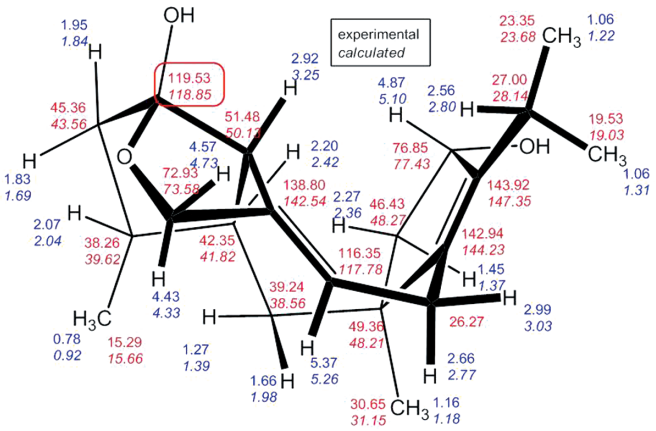
| Label | rel. E (kcal/mol) | Boltzmann Dist | ChemShift(C20) | ChemShift(C21) | ChemShift(C22) |
|---|---|---|---|---|---|
| M0030 | -0.53 | 0.577 | 129.628 | 133.103 | 42.767 |
| M0028 | 0.00 | 0.234 | 129.722 | 132.691 | 42.913 |
| M0031 | 0.60 | 0.086 | 130.323 | 132.612 | 42.716 |
| M0002 | 1.00 | 0.043 | 129.279 | 132.386 | 43.094 |
| M0001 | 1.07 | 0.039 | 129.579 | 131.756 | 42.907 |
| M0035 | 1.47 | 0.020 | 130.208 | 132.437 | 43.169 |
| M0027 | 3.59 | 0.001 | 126.068 | 135.050 | 39.894 |
| M0020 | 3.78 | 0.000 | 126.184 | 135.245 | 40.352 |
| M0016 | 4.54 | 0.000 | 126.169 | 135.393 | 39.740 |
| M0013 | 5.15 | 0.000 | 126.511 | 135.217 | 40.277 |
| Boltz Avgs | -0.14 | 0.399 | 129.700 | 132.870 | 42.821 |

f(x) Add... Delete Sort Formulas

**図4. 計算結果から必要な情報を取り出したSpreadsheet**

計算結果と実験結果を比較して構造解析を評価しますが、筆者はデータをエクセルに移して、相関係数、平均誤差など統計解析しています。（計算結果はSpreadsheetからコピーアンドペーストすることができますが、実験値は手入力しか方法がありません。）

以下実際の計算結果とその解析例を示します。図5には、最近子嚢菌から単離したroussoellol Aの計算結果を示しました。[4] 驚くほど実験結果を再現していることがわかります。この構造決定では、赤枠で括ったヘミケタール炭素の化学シフトが119.53ppmと通常予想される化学シフトより高周波数に現れたため、帰属の裏付けをするため、計算を行ったものですが、計算値も118.85ppmと良い一致をしています。




| | $^{13}$C | $^{1}$H |
|---|---|---|
| 重決定係数 ($R^2$) | 0.99928 | 0.9886 |
| 標準誤差 | 1.22 ppm | 0.15 ppm |

**図5. 計算および実験値の化学シフトとその重決定係数 ($R^2$) および標準誤差**

計算により立体化学の検証も可能です。図6はchaetoglobosin Aと、マクロ環部の水酸基のジアステレオマーについて計算値と実験値を比較したものです。この分子サイズになると、密度汎関数法での構造最適

化に半日程度要してしまいます。この分子サイズでは、多数の配座を密度半関数法計算で計算することは現実的でありません。そこでMMFF配座解析後、HF/STO3Gで構造を最適化し直し、安定な配座から一つずつ計算しました。10配座程度計算したところで、スペクトルはほぼ収束した（新しい配座の化学シフト情報を追加しても変化が見られなくなった）ので、そこで計算を終了しました。両異性体と計算するのに4日要しました。解析では、異性体間でインドール部分の化学シフトはほぼ同一なため、分子全体の化学シフトを比較すると、議論する部位の化学シフトの違いが統計処理で希釈されてしまうためか、その違いは$^{13}$C, $^{1}$Hとも優位とは言えませんでした。そこで構造中、赤で示した大環部分のみで比較したところ、$^{1}$H化学シフトにおいて、天然物構造の標準誤差が顕著に小さくなることがわかりました。この分子に限らず、カルボニル炭素の場合7ppm程度ずれてしまうことも少なくありません。したがって、Δδ値値での単純な議論は危険な場合もあると思います。統計処理を行って、相関係数や標準誤差といった値で比較した法がよいかもしれません。$^{13}$C化学シフトでは、相関係数（R）が0.999を超えるので、重相関係数（$R^2$）を用いた方が違いの議論がしやすいと思います。判り易い、判りにくいといった違いはありますが、相関係数の議論と標準誤差の議論と間で異なった結果を示したことはありません。効果的かつ確実なデータ解析法を議論する適切な処理方法について模索中な状態です。

計算精度を考慮すると、複数の構造が実験値との平均誤差で$^{13}$C化学シフトで3ppm、$^{1}$H化学シフトで0.3ppmを下回った場合、計算結果だけで判断することはできないと感じます。異性体間の比較では$^{1}$H化学シフトの方が、$^{13}$C化学シフトに比べ顕著な差が出る傾向

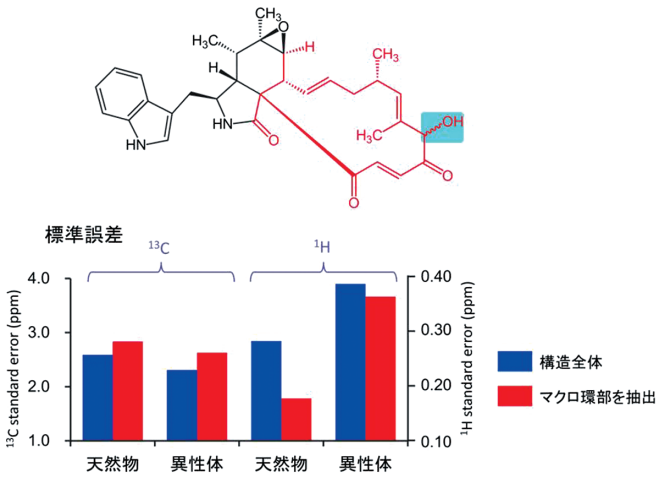



**図6. chaetoglobosin Aと、マクロ環部水酸基のジアステレオマーについて実験値との標準誤差**

にある感触を得ていますが、結論するにはデータの蓄積が不十分です。$^{1}$H化学シフトでは溶媒効果が顕著なことはよく知られています。これまでのところ、計算値は重クロロホルムより重ベンゼンの方が良い成績を与えているように感じています。

モデリング実験では、正しく入力・実行さえすれば放っておいても計算機が結果を出してくれます。計算には試薬もクロマト操作もいりません。学生時代に恩師白濱晴久先生が『計算のアルバイトはタダだ。』とおっしゃっていたこと、真実だと思います。筆者は、決してコンピュータに詳しいわけではなくMS-DOSもUNIXも使えません。挑戦してすぐに挫折しました。Spartanは、ほぼすべての操作についてWindowsのようにGraphic User Interfaceで実行できる秀逸な分子モデリングソフトといえます。オペレーションはとてもシンプルですが、かなり高度な計算も対応できます。マニュアルも他の分子モデリングソフトと比較して比較にならないほど薄く、とても優しい表現で書かれています。筆者のようなモデリング初心者でも短時間である程度の機能を習得することができました。細かな計算条件設定にはキーワードを入力しますが、日本語での説明がWEBで公開されています。また日本支社の内田氏には、様々な技術的な質問にも対応していただけるなどサポートも充実しています。

残念ながら、計算機科学者と天然物研究者の間には、互いに（言葉には現しにくい）理解しえない溝があること、否定できないと思います。計算に必要なコンピュータ環境は既に整いました。分子モデリングをより現実的なものにするためには、実際の化合物を扱っている天然物研究者が、様々な分子を計算・化学的視点から結果を解析して、計算法に対するscoop and limitationをレポート、そして計算機科学者に戻すといったフィードバックを繰り返すことが重要であると思います。

## 参考文献


1. Murakami, T.; Takada, N.; Hehre, W., J.; Hashimoto, M. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4547.
2. Hashimoto, M.; Murakami, T.; Funahashi, K.; Tokunaga, T.; Nihei, K.; Okuno, T.; Kimura, T.; Naoki, H.; Himeno, H. *Bioorg. Med. Chem.* **2006**, *14*, 8259.
3. Honma, M.; Kudo, S.; Takada, N.; Tanaka, K.; Miura, T.; Hashimoto, M. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 709.
4. Takekawa, H.; Tanaka, K.; Fukushi, E.; Nehira, T.; Hashimoto, M. *J. Nat. Prod.* **2013**, in press.
5. Lin, C. Y.; George, M. W.; Gill, P. M. W. *Aust. J. Chem.* **2004**, *57*, 365.



米国法人 WAVEFUNCTION, INC. 日本支店
〒102-0083 東京都千代田区麹町3-5-2 BUREX麹町
Tel 03-3239-8339 Fax 03-3239-8340
E-mail japan@wavefun.com