

課題名 (タイトル) :

NMR メタボロミクス用理論化学シフトデータベースの構築

利用者氏名 : 近山 英輔

所属 : 横浜研究所 植物科学研究センター メタボローム研究推進部門 先端 NMR メタボミクスチーム

1. 本課題の研究の背景、目的、関係するプロジェクトとの関係

植物科学研究センター先端 NMR メタボミクスチームでは NMR では世界最高度のメタボロミクス解析プラットフォームを構築してきた。そこでは代謝物の標準化合物による化学シフトデータベースが重要な要素となっており、現在 270 代謝物以上の実験的計測データが蓄積されている。しかし、代謝物総体は例えば植物で 20 万種以上あると言われており、その全ての化合物を入手することはできない。また、例え入手できたとしても、その全てを NMR で計測することは不可能である。従って、計算機を用いて大量の代謝化合物についてその化学シフト予測値を計算し、データベース化することは、予測値の精度を犠牲にしても、未知試料の候補代謝物同定時に非常に有用となる。RICC には Gaussian が利用可能となっており、Gaussian では NMR 化学シフトの理論値が複数の量子化学計算アルゴリズムで計算可能となっていることから、本課題の最終目的は、量子化学計算による大量の化合物の化学シフトデータベースを構築することである。

2. 具体的な利用内容、計算方法

昨年度までに用いた手法を応用して、セルロースの構造解析に重要な、セロビオース分子の分子動力学と Gaussian による遮蔽定数計算、回帰直線による化学シフトの概算を行い、構造と化学シフトの関係を調べた。その手法とは、具体的には、本継続課題では、29 化合物に対し、RICC の計算リソースを用いて、各々の化合物で 10 ns の古典定温定積分子動力学(MD)計算により生成された 100 構造に対し、HF/B3LYP(DFT)/MP2 のメソッドと 6-31G/aug-cc-pVDZ の基底セットによる理論レベルを用いた第一原理量子化学計算により遮蔽定数を求めている。主成分分析による遮蔽定数・分子

構造相関解析を行った。各化合物は、化合物ファイルで与えられている初期座標から、エネルギー最小化を真空中で行い、1 ns の平衡化の後、10 ns の MD シミュレーションにかけられた。MD に用いた力場は CACTUS (<http://cactus.nci.nih.gov/>) でエタノール分子の Mol2 フォーマット(Tripes)ファイルを取得するか、又は PubChem(<http://pubchem.ncbi.nlm.nih.gov/>)にて取得した sdf ファイルを Open Babel (http://openbabel.org/wiki/Main_Page) を用いて Mol2 に変換し、acpype(<http://code.google.com/p/acpype/>) を用いて GROMACS 用 GAFF 力場を生成した。MD は倍精度の GROMACS 4.0.5 で修正 Wolf 法を静電力計算に用い、真空中、298.15 K の定温シミュレーションを Nose-Hoover 法で行った。100 ps 毎に各化合物の分子構造をサンプルし、各化合物につき計 100 構造を得ている。MD で得られた各構造の座標についての遮蔽定数の第一原理計算は、RICC 上の Gaussian09 で HF、B3LYP、MP2 の 3 種の理論レベルと 6-31G、aug-cc-pVDZ の 2 種の基底セットについて計算している。故に 1 理論レベル、1 基底セットにつき 100 構造の各化合物の分子内原子の遮蔽定数を得ている。計 2900 構造である。29 化合物は、in-house で作成している化学シフトデータベースの中から、分子量 200 以下の化合物を全て選択している。

3. 結果

セロビオースの計算結果により、構造と化学シフトの関係を見積もった。その結果、NMR 実験で観測された化学シフトを用いて、3 種類の推定安定構造 CS1 (赤) ,CS2 (青) ,CS3 (紫) を分類した。

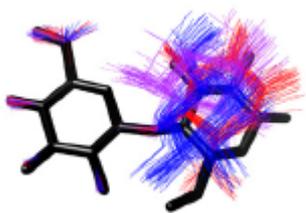


図 1 化学シフトと分子動力学によって分類されたセロビオースの推定安定構造

29 個中、化学シフトの実験値が存在する 26 個の化合物の $^1\text{H}/^{13}\text{C}$ 化学シフトについての、理論遮蔽定数と実験値との回帰直線を求めた(図 2)。このような回帰直線を理論レベル毎に評価し、最も精度と費用対効果の高い、アルゴリズムを選定して、理論化学シフトデータベースを構築してゆく予定である。

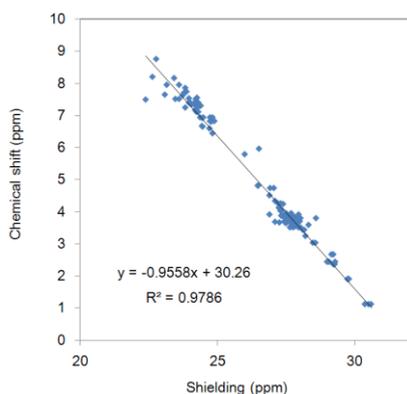


図 2 遮蔽定数と ^1H 化学シフトの回帰直線

4. まとめ

今年度は図 2 のような回帰直線を、全ての 29 化合物の ^1H 原子、 ^{13}C 原子で決定し、理論レベル毎に評価し、最も精度と費用対効果の高い、アルゴリズムを選定する予定であったが、NMR 実験を進める共同研究者のスケジュール等の関係で NMR 実験を進めることができず、データベースのための網羅的遮蔽定数の計算の前段階で中断している状態であった。

5. 今後の計画・展望

全ての理論レベルにおける回帰直線を決定するために、NMR 実験の終了と化学シフトの帰属解析を終了させる必要があるため、まずそれを行う。その後、費用対効果の高い理論レベルを選定し、

化合物数を 10 倍程度にスケールアップして NMR メタボロミクス用理論化学シフトデータベース理論データベースの試作データの構築を行いたい。

6. 利用がなかった場合の理由

データベース用の網羅的計算のためのアルゴリズム選択が終了しなかった。その理由は、アルゴリズム決定のための 29 個の標準化学物質の化学シフト帰属のための NMR 計測実験とその解析が、共同研究者の異動等の関係で中断されていたことにある。次年度はこの実験の遂行者、あるいは代替方法を考慮し、これを終了させ、利用を再開する予定である。

平成 24 年度 RICC 利用研究成果リスト

【論文、学会報告・雑誌などの論文発表】

T. Mori, E. Chikayama, Y. Tsuboi, N. Ishida, N. Shisa, Y. Noritake, S. Moriya, J. Kikuchi, "Exploring the conformational space of amorphous cellulose using NMR chemical shifts", *Carbohydrate Polymers* **90**, 1197–1203(2012)