

報告番号	※	第	号
------	---	---	---

主論文の要旨

論文題目 Measurement Informatics Approaches for NMR Signal Deconvolution and Data-Driven Analysis Toward Molecular Complexity (分子複雑系の核磁気共鳴信号分離およびデータ駆動型分析のための計測インフォマティクスアプローチ)

氏名 山田 隼嗣

論文内容の要旨

現代社会の重要課題である水、食料、材料、エネルギーの環境・資源や健康の問題を解決するため、多様な分子で構成された混合物や材料の高次構造などの分子複雑系を対象とした計測技術の発展が求められている。核磁気共鳴 (NMR) 分光法は、様々な取得パラメーターやパルス系列を使用して、原子分解能で非侵襲的試料の化学構造と分子運動性の有益なデータを提供するため、分子複雑系を特徴付けるための強力な計測技術の1つである。近年の情報技術の発展に伴い、蓄積された NMR データを有効活用するための計測インフォマティクスアプローチは、ますます重要になっている。しかし、混合物の NMR スペクトル分析は、ノイズや信号の重複により困難であり、多くの労力を要する。従って、スペクトル分析の前段階にあたるデータ品質管理、ノイズ低減、信号分離などのデータクレンジング、および長年蓄積したデータを活かしたデータ駆動型分析に関する研究が必要である。

本研究では、まず、広幅 NMR スペクトルに対して、ピークを強調および分離する前処理法を援用した信号帰属法を検討した。また、研究室内および他機関で蓄積された NMR データの計測パラメーターと信号対雑音比 (signal-to-noise ratio: SNR) の関係を調査するノイズ要因解析を遂行した。そこで、異なる運動性 (横緩和時間 T_2 が磁場不均一性の影響で短縮した実測値 T_2^*) 成分およびノイズを含む溶液 NMR 信号を対象として短時間フーリエ変換 (STFT) および確立的スパース行列因子分解 (PSMF) を組合せた信号分離法を開発した。さらに、この理論を基盤として、固体状態で複数のドメイン構造、または成分を持つ多成分材料の固体 NMR 信号を対象として信号分離法の適用を検討した。続いて、成分や構造の記述子として NMR 信号および物性を予測するため、生成的地形図回帰 (Generative Topographic Mapping Regression; GTMR) の適用を検討した。以上の手法に基づき、分子複雑系の NMR 分析に有用な計測インフォマティクスアプローチを開発した。

① 分子複雑系の低分解能 NMR 分析のための前処理法および信号帰属法

分子複雑系の NMR 分析では、分子の物理化学特性は多様であるため、様々な試料調製法やパルス系列が使用される。しかし、混合物の NMR 信号帰属は、信号重複の問題、参照スペクトルや信号帰属ツールの不足のために困難であり、多くの労力を要する。そこで本研究では、低分子から高分子や脂質までの分子複雑系の低分解能 NMR 分析における信号帰属を支援するため、広幅スペクトルの前処理、信号帰属およびデータベース (DB) からなるウェブツールとして **InterSpin** を開発した。低磁場の卓上 NMR や固体 NMR から得られる広幅スペクトルの分析を支援する前処理ツールとして、スペクトル積分により感度を改善する **SENSI** および信号を分離する **PKSP** を開発した。**PKSP** では、新規 NMR 信号分離法として非負スパースコーディング (**NNSC**) を実装し、**NMF** などの従来法と比較して、高速かつ正確な信号分離を可能とした。さらに、**SENSI** で得られる各ピークの変動係数 (**CV**) と **PKSP** で得られる分離信号を組み合わせ、魚料理の低磁場 NMR やユーグレナの固体 NMR により得られた広幅スペクトルの信号を分離し、**CV** や標品スペクトルを利用して信号帰属を可能とした。食品、材料、環境、健康などの幅広い研究には、多様な標品スペクトルが必要である。しかし、類似構造を持つ高分子や脂質の固体 ^{13}C CP-MAS や DMSO-d6 溶媒での溶液 NMR に関しては、DB や信号帰属法は確立していない。そこで、固体および溶液状態 (溶媒として、 D_2O 、 MeOD-d_4 、 DMSO-d_6) の低分子から高分子までの多様な標品スペクトル (^1H - ^{13}C 相関、 ^1H - J 分解、 ^{13}C CP-MAS) を蓄積した新規 DB (**SpinLIMS**) を構築し、多様な試料に対して **InterSpin** での信号帰属を可能とした。この DB を基盤として、高分子や脂質の信号を帰属する **SpinMacro** を開発し、固体 ^{13}C ピークや DMSO-d6 溶媒での ^1H - ^{13}C 相関ピークの信号帰属を自動化した。さらに、 ^1H - ^{13}C 相関ピークと ^1H - J 分解ピークの統合により候補分子を絞り込む **InterAnalysis** を開発し、信号帰属を効率化した。分子複雑系の低分解能 NMR 分析のための前処理法および信号帰属法は、信号の重複問題を解決し、DB を活用した NMR 信号帰属を高度化した。

② STFT および PSMF を組合せた NMR 信号分離法およびノイズ要因解析法

データ駆動型分析において、データ品質は結果に影響を与えるため重要である。しかし、NMR 分野で、蓄積された NMR データの品質管理、ノイズ低減、信号分離などのデータクレンジング法は確立していない。そこで本研究では、NMR データの計測パラメーターとノイズに焦点を当て、一次元 NMR の自由誘導減衰 (**FID**) 信号分離およびノイズ要因解析によるデータクレンジング法を開発した。ノイズ低減および信号分離のために、**STFT** を適用することで **FID** を一定時間ごとに分割して各時間でのスペクトルの変化の活用を検討した。**STFT** によって追加された時間軸上の T_2^* 緩和に伴う信号強度減衰の特徴に基づき、行列因子分解により個々の周波数における信号とノイズを識別できた。行列因子分解法として **PSMF** は、他の **NMF** などの手法に比べ、信号とノイズを良好に分離できた。**STFT** および **PSMF** を組み合わせた新規信号分離法は、多変量解析のみを使用する従来のノイズ低減法とは異なり、多くのサンプルま

たは計測パラメーターの FID を必要とせず、ノイズ低減および信号分離を可能とした。本手法により、FID 信号のノイズを分離し、スペクトルの SNR を約 3 倍に改善した。Diffusion-edited NMR スペクトルは、各周波数成分の T_2^* の違いにより、高分子や脂質と低分子の信号に分離できた。蓄積された NMR データを活用するには、データ品質を確認する必要がある。そこで、NMR データの計測パラメーターと SNR の関係を調査するノイズ要因解析を行った結果、溶媒抑制が不十分な場合、スキャン数が SNR を低下させる主な要因であった。STFT を用いた溶液 NMR の信号分離およびノイズ要因解析は、信号の重複問題を解決し、データ駆動型分析を高度化した。

③ 多成分材料の固体 NMR のための信号分離法および信号・物性予測法

近年、海洋プラスチックの海洋汚染、廃棄物処理、地球温暖化などの地球規模の課題から、低炭素社会への研究が重要視されている。石油資源の代替としての微生物製品と植物バイオマスは、プラスチックや原料などの高分子材料の製造に使用できる。ポリ乳酸 (PLA)、ポリ- ϵ -カプロラクトン (PCL)、セルロースなどのポリマーは、複数のドメイン、またはコンポーネントを有する分子複雑系であり、様々な特性を持つ材料として使用される。微生物および植物バイオマスは、複数のドメインを持つ高分子や脂質を含む複数の成分で構成される生化学的システムとして分析する必要がある。従って、微生物製品、植物バイオマス、プラスチックなどの多成分材料の固体 NMR 分析アプローチを開発する必要がある。そこで本研究では、固体 NMR を対象とした信号分離法を開発した。STFT と非負のテンソルおよび行列因子分解 (NTF、NMF) を使用した新規信号分離法を検討した。本手法により、セルロース分解プロセスにおける ^{13}C CP-MAS スペクトルは、STFT と NTF によって T_2^* の違いによりセルロース、タンパク質、脂質の信号に分離できた。また、PCL の異方性スペクトルは、STFT と NMF によって結晶とアモルファスの信号に分離できた。計測時に異方性除去のため適用されるデカップリングに対する代替手段として、異方性の計測データに対して計算科学的手法による信号分離を可能とした。また、STFT データを活用し NMR 信号や物性を可視化および予測するための新規手法として、GTMR を検討した。GTMR は、セルロース分解プロセスにおける生成物である酢酸と CO_2 の NMR 信号強度を予測できた。また、プラスチックの NMR 信号と各物性の統合データに GTMR を適用することで、所望の物性 (ガラス転移点、融点、分解温度) となる NMR 信号を予測できた。STFT を用いた固体 NMR の信号分離および予測法は、信号の重複問題を解決し、多成分材料の特性評価および材料設計を可能とした。

以上の各手法により、NMR 分析において、 T_2^* の異なる成分およびノイズの重複した NMR 信号の分離、続く信号帰属および高次構造・物性予測に関するデータ駆動型分析を高度化した。本研究で開発した計測インフォマティクスアプローチは、健康、食品、材料、環境などの様々な分野における分子複雑系のデータ駆動型の研究、開発、生産、品質管理などへの貢献が期待される。