

Title	有機化合物の化学構造と <sup>1</sup> H-NMRおよび <sup>13</sup> C-NMR スペクトルデータの統合データベースシステム構築に関する研究
Author(s)	奥山, 徹
Citation	大阪大学, 1989, 博士論文
Version Type	
URL	<a href="https://hdl.handle.net/11094/36696">https://hdl.handle.net/11094/36696</a>
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、 <a href="https://www.library.osaka-u.ac.jp/thesis/#closed">〈a href="https://www.library.osaka-u.ac.jp/thesis/#closed"〉</a> 大阪大学の博士論文について <a href="https://www.library.osaka-u.ac.jp/thesis/#closed">〈/a〉</a> をご参照ください。

***Osaka University Knowledge Archive : OUKA***

<https://ir.library.osaka-u.ac.jp/>

Osaka University

氏名・(本籍)	おく 奥	やま 山	とおる 徹
学位の種類	理	学	博 士
学位記番号	第	8 5 1 3	号
学位授与の日付	平成元年3月15日		
学位授与の要件	学位規則第5条第2項該当		
学位論文題目	有機化合物の化学構造と $^1\text{H-NMR}$ および $^{13}\text{C-NMR}$ スペクトルデータの統合データベースシステム構築に関する研究		
論文審査委員	(主査)		
	教授	千原 秀昭	
	(副査)		
	教授	木下 達彦	教授 小田 雅司
	教授	藤原 鎮男	教授 桑田 敬治 (神奈川大学知識情報研究所)

### 論文内容の要旨

現在までに知られている有機化合物の数は数百万と言われている。このような膨大な数の化合物の情報を収集・管理するためにはコンピュータの支援なしでは不可能である。そこで、本研究では有機化合物の化学構造 (実際には2次元構造式) と $^1\text{H}$ -および $^{13}\text{C}$ -NMRスペクトルデータの統合データベースシステムの構築を試みた。

また、データベース構築の際問題となる構造式のコンピュータ上での表記法や構造式やスペクトルデータのコンピュータへの入力方法等について検討した。さらに、実際のデータを登録・管理し、種々の方法で検索するためのデータベース管理システムを構築した。

今回構築されたデータベースは構造式をデータ管理の中心としている。ここで使用されている構造表記法はCANOSTと呼ばれているものである。CANOST表記法自身は以前から用いられていたが、今回これをさらに改良した。特に、CANOSTの規範化の部分は新しい方法を考案した。これにより、実用的な速度で対称性の高い構造式をも規範化できるようになった。

次に、データの入力方法を改善し、これまで手作業であった入力作業をコンピュータ支援で行なうようにした。これは、入力データの品質管理のためには不可欠であり、内外のデータベース構築者が等しく知恵を絞っている部分である。

最後に、実際にデータベース管理システムを構築し、データの登録・検索のための種々の手法を検討した。その結果、データ登録の中心に構造式を据え、すべてのデータ登録作業は構造式を介して行なわれることとした。すなわち、これは構造式のデータベースをキーデータベースとし、種々の物性値を統合する統合データベース管理システムへの発展の可能性を示唆している。ここでは、その手始めとして、 $^1\text{H}$  と

$^{13}\text{C}$ の2つのNMRデータの統合管理を試みている。また、検索のための手段として、3つのモードの構造検索(全構造、部分構造、骨格構造)や化学シフトを中心としたスペクトル検索( $^1\text{H}$ ,  $^{13}\text{C}$ ), 実際のスペクトル中に現われるピーク位置での検索( $^1\text{H}$ )等を提供している。さらには、これらの手法を複合的に使うことや、構造検索に部分的ではあるがGenericな構造表記を用いること等を許している。このような、検索手法の多様性は、今後ますます重要性を増すものと思われる。

以上、本研究の成果が今後のファクトデータベース構築のなんらかの指針となれば幸いである。

## 論文の審査結果の要旨

奥山君の論文は、有機化合物の構造をコンピュータに認識させ、構造とその表記との1:1対応を保証する規範化の方法および構造検索の方法と、これらの構造に対応させたプロトンおよび炭素核高分解能NMRスペクトルデータベースの作成法に関するものである。化学構造のコンピュータ内部表現に関しては種々の方法が実用になっているが、これを一義性をもった表現に変換する、いわゆる規範化については、これまでのところ効率の劣る方法で行われているのが現状である。奥山君は、従来の方法の欠点がどこにあるかを検討しなおした上で、構造の骨格の表現にはじめから対称性を系統的にとり入れる方法を考案し、これによって規範化の速度を格段に速くすることに成功した。一般に分子の対称が高いほど原子-原子マッチングによるだけの従来法では処理が長時間を要する宿命をもっているが、これは対称性を全く利用しないためであった。奥山君の方法による対称性認識はグラフの隣接行列のベキ乗の行列を用いることによってある原子と第二の原子(自分を含む)との間に存在するn個の結合が現実の構造において認められるか否かをベクトル表示(ビット列)に表わしうることに着目した。これによって従来の方法では非現実的な長い時間を要した規範化が実行可能となったことが実証(計算機上で)された。

さらにスクリーンファイルをべつに用意することによって部分構造検索が可能な検索法を実現した。従来行われていない“骨格検索”の概念を導入しグラフのマッチングをとれるスクリーン生成を行った。

これらの構造表記法を実際の有機化合物約2万についてコンピュータに蓄積しNMRスペクトルデータ(化学シフト)と対応したデータベースを作った。このデータの検索の面において、現在一般に行われている完全一致のほか類似検索を許すシステムとした。以上説明したように、奥山君の論文は有機化合物の一義的なコンピュータ表現を導出する一般性のある方法を発案し、その実用システムを作ったデータベース処理システムの研究であって、情報化学の進歩に大きな貢献をしており、理学博士論文として十分な価値をもつものである。