

演 題	三次元分子構造-NMR データベースと $^{13}\text{C}$ -NMR 高精度予測システム CAST/CNMR の開発	
発 表 者 ( 所 属 )	○佐藤寛子 <sup>a</sup> , 越野広雪 <sup>b</sup> , 鶴澤 洵 <sup>b</sup> , 中田 忠 <sup>b</sup> ( <sup>a</sup> 国立情報学研究所, <sup>b</sup> 理化学研究所)	
連 絡 先	101-8430 東京都千代田区一ツ橋 2-1-2 国立情報学研究所・知能システム研究系 佐藤 寛子 Fax: 03-3556-1916 Email: <a href="mailto:cheminfo@nii.ac.jp">cheminfo@nii.ac.jp</a>	
キ ー ワ ー ド	NMR 化学シフト, 立体化学, データベース, CAST コード, CAST/CNMR	
開 発 意 図 適 用 分 野 期 待 効 果 特 徴 な ど	立体化学規範的コード化法 CAST を分子構造表記の基盤として用いた, 三次元分子構造-NMR データベースを構築した. 本データベースを利用することで, 立体化学を考慮し, 高精度に $^{13}\text{C}$ -NMR 化学シフトを予測するシステム CAST/CNMR を開発した.	
環 境	適 応 機 種 名	SGI ワークステーション
	O S 名	IRIX
	ソ ー ス 言 語	
	周 辺 機 器	
流 通 形 態 ( 右 の い ず れ か に ○ を つ け て く だ さ い )	<ul style="list-style-type: none"> <li>・日本コンピュータ化学会の無償利用</li> <li>・ソフトとする</li> <li>・独自に頒布する</li> <li>・ソフトハウス、出版社等から市販</li> <li>・ソフトの頒布は行なわない</li> <li>・その他 <span style="border: 1px solid black; border-radius: 50%; padding: 2px;">未定</span></li> </ul>	具 体 的 方 法

## はじめに

有機化合物の構造決定のためにNMRは不可欠な手法である。NMRスペクトルを予測するコンピュータシステムは種々開発され、構造推定システムと合わせて化学系企業における測定・構造解析の自動化に利用されている。NMRスペクトル予測システムの主流は、分子構造と化学シフトの実測値を系統的に蓄積し活用するスペクトル-分子構造データベース主導型のものである。データベースとして蓄積し活用する手法の利点は、膨大な事実データとコンピュータによる網羅的な情報処理に裏付けられた客観的な評価が得られる点にある。しかしながら、これらのシステムで取扱う構造は基本的に平面構造であり、立体異性体は二重結合の幾何異性を除いて的確に評価されない。平面構造が同じ化合物でも立体化学が異なる場合には、その化学シフト値は大きく異なることが少なくない。すなわち、これらのシステムでは、立体化学を考慮できない限り、一般に複雑な立体化学構造を有する天然有機化合物などの生物活性を有する化合物の構造解析に実用的に利用できない。

そこで我々は、立体化学を規範的に表記するコード化手法 CAST(CAnonical-representation for STereochemistry)<sup>1-3</sup>を開発し、これらを分子構造記述の基盤として用いた三次元分子構造-NMR データベースをもとに、立体化学を考慮して高精度に化学シフトを予測するシステム CAST/CNMR を開発した<sup>4</sup>。

## CAST/CNMR による化学シフト予測法

CAST/CNMR のシステム構成を Fig. に示した。システムは大きくデータベース(DB)部分と予測部分からなる。DB 部分は平面構造表記 (CANOST 表記) と立体構造表記 (CAST 表記) で記述された分子構造 DB と実測  $^{13}\text{C}$ -NMR DB で構成されている。立体構造は、立体配座と立体配置の各情報で記述されており、目的によって使い分けができる。本データベースを用いて「同じ部分構造であれば、その部分構造の中心の炭素の化学シフトは近い値を与えるはずである」との考えに基づき、化学シフトの予測を行う。すなわち、予測対象の分子構造について CANOST, CAST コード

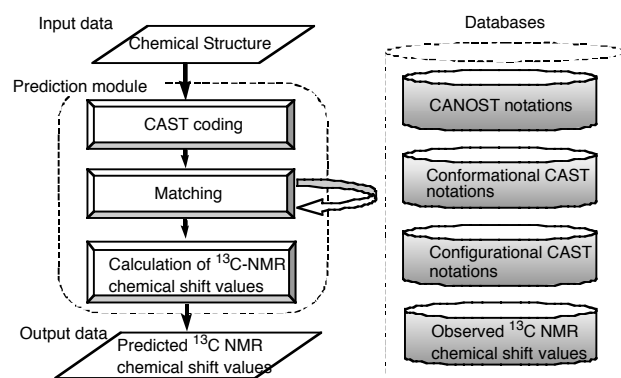


Fig. CAST/CNMR システムの概要

指定された範囲の部分構造の一致するデータの実測  $^{13}\text{C}$ -NMR 化学シフト値の平均値を予測値として出力する。部分構造の範囲は予測炭素原子からの結合数で指定し、一般に  $^{13}\text{C}$ -NMR 化学シフトは  $\gamma$  位までの影響を大きく受けることから最低限 3 結合までの部分構造の一致をデフォルト条件として設定している。

## CAST/CNMR による化学シフト予測結果

天然有機化合物 arisugacin F(1) の CAST/CNMR による予測結果を適用例として示す。テルペノイド, ステロイド, ポリケタイド, ポリエーテル系天然有機化合物とそれらの合成中間体を含む 733 件の分子構造-NMR データベースを用いた。  $\gamma$  位までの平面構造と立体配置の一致を予測条件とした。

予測された  $^{13}\text{C}$ -NMR 化学シフト値と実測値を Table に示した。全炭素原子の 78% は予測誤差が 0.05ppm 以内であり、全炭素の 93% は予測誤差が 0.1ppm 以内であり、全ての炭素原子が 1.5ppm 以内の予測誤差に収まり、高精度の予測結果が得られた。

種々の天然有機化合物にも適用し高精度の予測結果が得られることを確認した。

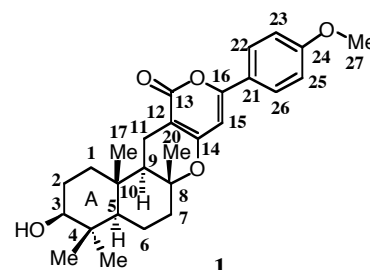


Table Arisugacin F の予測結果と実測値

No. of carbon atom	Predicted $^{13}\text{C}$ chemical shift values /ppm	Observed $^{13}\text{C}$ chemical shift values /ppm	$\Delta$ (predicted - observed values)
1	38.6	37.5	1.1
2	27.3	27.1	0.2
3	78.6	78.5	0.1
4	38.8	38.8	0
5	55.3	55.0	0.3
6	19.4	19.4	0
7	40.1	40.4	-0.3
8	80.6	80.5	0.1
9	51.3	51.6	-0.3
10	37.7	36.9	0.8
11	17.3	17.2	0.1
12	99.2	98.4	0.8
13	164.4	164.7	-0.3
14	163.3	163.5	-0.2
15	97.5	96.7	0.8
16	158.4	158.3	0.1
17	16.1	15.1	1.0
18	27.7	28.1	-0.4
19	15.3	15.5	-0.2
20	20.7	20.7	0
21	124.1	124.0	0.1
22	127.0	127.0	0
23	114.2	114.2	0
24	160.1	161.5	-1.4
25	114.2	114.2	0
26	127.0	127.0	0
27	55.3	55.4	-0.1

## まとめ

立体化学を考慮し、高精度に  $^{13}\text{C}$ -NMR 化学シフトを予測するシステム CAST/CNMR を開発した。複雑な立体化学を有する天然有機化合物に適用し、高精度の予測結果を与えることを実証した。

## 参考文献

1. Satoh, H., Koshino, H., Funatsu, K., Nakata, T. *J. Chem. Inf. Comput. Sci.*, **40**, 622-630(2000)
2. Satoh, H., Koshino, H., Funatsu, K., Nakata, T. *J. Chem. Inf. Comput. Sci.*, **41**, 1106-1112(2001)
3. Satoh, H., Koshino, H., Nakata, T. *J. Comput. Aided Chem.*, **3**, 48-55 (2002)
4. Satoh, H., Koshino, H., Uzawa, J., Nakata, T. *Tetrahedron*, **59**, 4539-4547(2003)