

## 研究課題別研究評価

1. 研究課題名：化学反応のニューラルモデル化による定量的反応予測の実現

2. 研究者名：佐藤寛子

3. 研究のねらい

本研究の目的は化学合成研究を効率的かつ合理的に進めるためにコンピュータ上で化学反応を定量的に予測する汎用的なコンピュータプログラムシステムを開発することである。本目的の達成のためには任意の化学反応の生成物と生成比とを予測できるだけの理論が必須である。しかし、化学反応の複雑さと多種多様さゆえにこのような一般的な理論はいまだに確立されていない。そこで、化学の長い歴史の中で蓄積された膨大な反応情報をもとに、反応の法則性や規則性を見出すときの化学者の思考過程を模倣したニューラルネットワークモデルを創製する。ここでは、コンピュータの特色である前提を維持した演算の高速性と網羅性、正確な数値処理、大容量の記憶容量を十分に活かすことで、系統的かつ定量的な予測能力を有し、かつ化学者の定性的な直観や思考と相補的に発展できるモデル化を目指す。これを主軸として活用し、定量的反応予測プログラムシステムを構築する。

4. 研究結果及び自己評価

### 研究結果

化学者の研究形態と、化学反応情報の多様性・複雑性・非線形性に着想を得た、「化学者の記憶と思考の過程を形式的に模倣した化学反応のニューラルモデル化による定量的反応予測」の研究構想を提案した。本研究構想のポイントは、1) 反応に関する種々の支配因子をコンピュータでどのように表現するか、2) そして、それらをどのようなネットワークで有効に関連付け、予測モデルへと繋ぐか、にある。さきかけ3年間において、これらのポイントに重点を置いて得た種々の研究結果を以下に示す。

1) 反応の支配因子の表現法の開発

・ 反応表現法 Reaction Similarity (ReacSim vector) を開発し、従来にはなかった汎用性の高い化学反応の数値表現を可能にした。ReacSim vector を識別子として任意の反応の分類に適用、その有効性を確認した。

・ 立体化学表現法 CAST (CAnonical-rearrangement of STereochemistry)を開発し、従来不可能であった、立体化学の規範的表現を可能にし、立体構造データベースからの正確な立体化学構造検索を可能とした。また、CASTを立体化学表現の基本として利用し、正確な立体化学表現と認識を必要とするデータベース主導型 NMR (核磁気共鳴) 化学シフト予測に応用し、<sup>13</sup>C-NMR 化学シフト予測システム CAST/CNMR を開発した。CAST により、従来不可能であった、立体化学を考慮した精密な<sup>13</sup>C-NMR 化学シフト予測を実現し、CAST の有効性を示した。

・ 仮想反応相手との相互作用に基づいて分子を3次元的に特性値化する FRAU (Field-characterization for Reaction Analysis and Understanding) からの変数選択により、多様な反応試薬を一般的に表現する方法を考案した。本表現を識別子として反応試薬を分類した結果、反応試薬の機能の類似性とよい相関のある分類結果が得られた。

2) 予測モデル化

・ FRAUを識別子とした反応試薬分類結果をもとに、反応試薬の機能予測モデルを構築した。モデル化の手法としては、カウンタプロパゲーションタイプの Kohonen ニューラルネットワークに対して、反応

データを取扱う上で重要である「反応しなかったデータ(ネガティブデータ)がある」ことを考慮できるように拡張・改良を行い、これを利用した。本モデルを用いて、実際に報告されていない試薬も含めて予測を行い、文献ならびに化学実験によって予測結果の妥当性を示した。

#### 自己評価

反応をコンピュータで取扱う上で重要な表現方法 ReacSim vector、CAST、FRAU 等の開発により、研究構想の重要なフェーズである反応分類の研究を大きく進展させることができたと思う。また、各表現方法の汎用的反応分類、NMR 化学シフト予測、反応試薬機能予測等への適用により、それぞれの表現方法の有効性を示すことができた。これらの結果は、今後の研究構想の実現に向けた研究の発展に大きく結び付くことを示すとともに、一方では、アプリケーションそれぞれの実用化の可能性も示した点で、有効な成果であったと思う。

このように、反応を支配する重要な因子の表現については、当初計画していた以上に研究を進めることができた。また、反応試薬の機能予測モデルにより、特性値化 分類 モデル化 予測の流れを示した。しかしながら一方で、さきがけスタート地点での目標であった、「対象とする反応を絞った上で3次元的要素も考慮した反応予測の流れを実現する」ところまでは、3年間で至ることができなかった。今後、さきがけ研究で得た成果を踏まえ、当初の目標に向けて研究を進めたいと考えている。

また、反応を支配する因子としては、これまでに開発したものだけではなく、今後、分子軌道に関する因子、溶媒の因子、触媒の因子等についても考慮できるように表現方法を開発する必要がある。モデル化法についても、従来の情報科学で開発されてきた手法だけではなく、化学反応という複雑かつ多次元のデータを取扱うために必要な手法を、拡張・開発していく必要があると考えている。

反応の系統的理解と予測は、化学者の一つの夢であり、壮大なテーマである。本研究構想は、従来とは異なるアプローチでこのテーマを実現しようとするものである。さきがけ研究の3年間で、本研究構想の実現に向けた、重要な種々の研究成果を得ることができたと思う。実現には今後も多くのブレイクスルーが必要であるが、一つ一つの難題を地道に解決し、最終目標に近づいていきたいと考えている。一方で、それぞれ一つ一つの研究から派生してくる、NMR 化学シフト予測システムや反応試薬機能予測モデルなどの、比較的短い時間で実用化の可能なシステムも成果として得られる。これらの実用化の方向への研究も併せて進めていきたい。さきがけ研究の3年間では、本研究の基幹部分において大きく前進することができた。多謝するとともに、これらを糧に今後も研究を進めていきたいと思う次第である。

#### 5. 領域総括の見解

わが国では、まだ研究分野として歴史の浅い環境の中で、コンピュータ上で化学反応を定量的に予測する汎用的コンピュータプログラムシステムを開発することを目的とし、不断の努力により大きな成果をあげたことは高く評価できる。また、外部に対する成果の発表にも積極的であった。今後の目標もしっかり捉えており、将来の研究推進が大いに期待できる。

## 6 . 主な論文等

### 〔論文〕

- (1) Satoh, H., Itono, S., Funatsu, K., Takano, K., Nakata, T.: "A Novel Method for Characterization of Three-dimensional Reaction Fields Based on Electrostatic and Steric Interactions toward the Goal of Quantitative Analysis and Understanding of Organic Reactions", J. Chem. Inf. Comput. Sci., 39, 671-678 (1999)
- (2) \* Satoh, H., Koshino, H., Funatsu, K., Nakata, T.: "A Novel Canonical Coding Method for Representation of Three-dimensional Structures", J. Chem. Inf. Comput. Sci., 40, 622-630 (2000)
- (3) \* Satoh, H., Funatsu, K., Takano, K., Nakata, T.: "Classification and Prediction of Reagents' Roles by FRAU System with Self-organizing Neural Network Model", Bull. Chem. Soc. Jpn., 73, 1955-1965 (2000) (Headline Article)
- (4) Satoh, H., Koshino, H., Funatsu, K., Nakata, T.: "Representation of Configurations by CAST Coding Method", J. Chem. Inf. Comput. Sci., J. Chem. Inf. Comput. Sci., 41, 1106-1112 (2001)
- (5) 佐藤寛子: "コンピュータによる定量的有機反応予測への挑戦 反応の支配因子による分類から反応の本質を探る "化学と工業" 化学のフロンティア'99 はばたけ若き研究者たち ", vol. 52, 2月号, 146-150 (1999)
- (6) 佐藤寛子: " 化学反応の地図を目指して ", 科学技術ジャーナル, vol. 10, No. 2, 50 (2001)
- (7) 佐藤寛子: " 化学反応の地図と反応予測", 有機合成化学協会誌, 「特集号 今、若手研究者は何を目指しているか」, vol. 59, No. 5, 470 (2001)

### 〔特許〕

- (1) 佐藤寛子、船津公人、中田 忠、鷹野景子「分子の反応特性予測方法」、特許・出願番号 H10-258483.
- (2) 佐藤寛子、越野広雪、中田 忠「分子の立体化学コード化方法」、特許・出願番号 H11-288691.

### 〔受賞〕

- (1) 有機合成化学協会 第一製薬研究企画賞 (1999年2月)
- (2) 科学技術振興事業団 第35回科学技術情報振興賞・論文賞 (2000年10月)
- (3) 第23回情報化学討論会ポスター賞 (2000年11月)
- (4) 日本化学会 BCSJ 賞 (2001年3月)
- (5) 日本薬学会第27回反応と合成の進歩シンポジウムポスター賞 (2001年11月)