

JP04 反応評価機能を持つ反応設計

クライアント・サーバシステムの開発

(豊橋技術科学大学) ○太田圭輔

(東京大学) 船津公人*

(三井化学) 塚本幸治

1. はじめに

これまで我々は、有機合成化学者の創造的研究に寄与する3つの有機合成経路設計システムの開発を進めている。有機合成経路設計システム AIPHOS (Artificial Intelligence for Planning and Handling Organic Synthesis) [1]は、情報・経験指向および論理指向戦略の利点を活用し新規合成経路の創出を得意とするシステムであり、KOSP (Knowledgebase-Oriented Synthesis Planning) は、同様の戦略を用いながらも反応データベース (DB) から誘導した知識ベースを用いて戦略部位を認識する機能が加わったことにより、AIPHOS よりも経験指向的な要素の強いシステムである。TOSP (Transform-Oriented Synthesis Planning) は、既存反応から構築した Transform データベースを利用した経験指向型のシステムである。

一方、反応 DB から取り出した部分構造変化情報を活用し、生成物を予測する反応評価システム [2]の開発も行っている。このシステムの開発目的は、前述の合成経路設計システムで提案された前駆体構造に対して反応生成物構造を予測することで、目的化合物以外の反応生成物の確認やその生成条件、さらに安全性などに着目した合成経路の評価を行うことである。そこで現在我々は、合成研究者にとっての一層の利便性を考え、合成経路設計システムと反応評価システムを統合した反応設計システムの開発を進めている。

今回報告する反応設計システムでは、複数反応物間の反応予測を可能にし、反応評価として反応部位や反応条件に着目した反応式のソーティング機能の追加を行った。その結果、実際性の高い反応経路探索が可能になり、反応部位や反応条件に着目した反応評価を容易に行えるようになったので報告する。

2. AIPHOS・KOSP・TOSP

我々の研究室で開発されている3つの合成経路設計システムの1つである AIPHOS は、論理指向により誘導された前駆体に対して、反応知識ベース [3]による経験的な観点からの実際性評価を行うことにより、提案される逆合成ルートの有用性

の保証するもので、経験指向・論理指向の戦略の融合を図ったシステムである。本研究室の AIPHOS が利用している反応知識ベースは Distributed Chemical Graphics 社の反応 DB 「SYNLIB」から自動誘導したもので、反応部位から生成部位への変化を抽象化表現した DB である。AIPHOS は、合成前駆体構造を導く際の標的構造中の着目部位 (戦略部位) を獲得する STRATEGY モジュール、得られた戦略部位から合成前駆体構造を発生する GENERATOR モジュール、反応 DB から自動誘導された知識ベースを元に、上で創出された前駆体構造から標的構造へ反応スキームの評価を行う KNOWLEDGE モジュールで構成されている。AIPHOS は、これら3つのモジュールを利用することで、既知の反応だけでなく新規前駆体構造や反応経路を創出できる。

KOSP も AIPHOS と同様に3つのモジュールから成り立っているが、戦略部位を論理的に獲得する AIPHOS に対して、KOSP は知識ベースから標的構造に適用可能な反応ケースを検索し戦略部位を決定する。そのため AIPHOS よりも経験的な観点から合成前駆体構造を導ける。

TOSP はまず標的構造中に存在する環や官能基などのトポロジカルな構造特徴の認識を行い、あらかじめファイルに格納されている transform の適用の可否を判定する。さらにマッチングした transform 全てを標的構造中の該当個所に適用することで、相当する前駆体候補構造の全てを出力するとともに、適用された transform の形式 (name reaction 名) を表示する。TOSP は実験化学者に素直に馴染む情報を、簡便かつ効率的に提案することができる。このように、TOSP は経験型の合成設計システムである。

以上のように AIPHOS、KOSP、TOSP はそれぞれ合成経路創出の戦略の異なる相互補完的な関係にある。これら3つのシステムを自在に組み合わせながら合成経路設計が行えることで、合成化学研究への寄与が期待できる。

3. 反応評価システム

反応評価システムは、AIPHOS から出力される合成経路の反応評価を目的として開発された。こ

*funatsu@chemsys.t.u-tokyo.ac.jp

のシステムは、反応予測システム SOPHIA (System for Organic Reaction prediction by Heuristic Approach) [3][4]の反応部位認識機能と、反応 DB の部分構造変化情報 (Transform) を組み合わせることで、副反応生成物を予測する。図 1 に反応評価システムの概念図を示す。

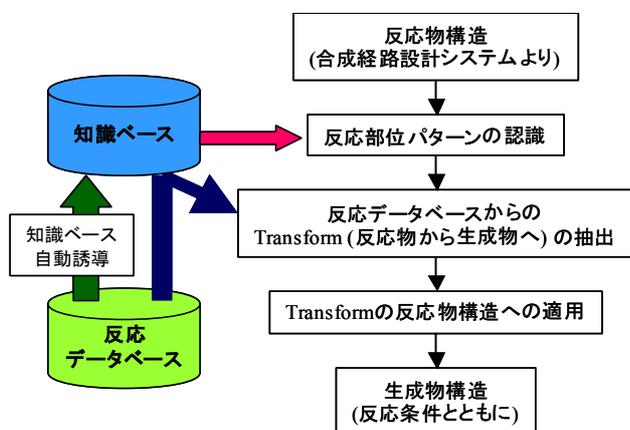


図 1 反応評価システムの概念図

反応物構造の入力により、知識ベースを利用して可能な反応部位パターンの認識が行われる。このとき、認識された反応部位パターンを支持した知識の元となった反応 DB 中の該当する反応データより、反応部位から生成部位への部分構造変化情報が取り出される。この部分構造変化情報を、先に認識された反応部位パターンに対して適用することにより、対応する生成物構造を誘導することができる。また同時に、反応条件や文献情報も得られる。

この反応評価システムを用いて合成経路の評価を行う場合、合成経路設計システムによって提案された前駆体構造を入力構造とする。このとき、出力された可能な生成物候補構造の中に合成経路設計システムの入力構造である合成標的構造以外の生成物構造が提案された場合、その評価対象の合成経路には反応条件によっては標的構造以外に副反応物が生成される可能性があることを示唆する。そして、もし副生成物構造の中にたとえば毒性を有する構造が含まれていれば、いま検討している合成経路の見直しが必要となる。

このように、反応評価システムは副反応生成物に着目した合成経路の評価を容易に行うことができる。

4. 動作環境

3 つの合成経路設計システムおよび反応評価システムは、クライアント・サーバ (C/S) 型を採用している。クライアントはサーバに対する処理の

要求や表示のみを行い、サーバが DB へのアクセスや計算などの負荷の大きい処理を請け負うため、負荷分散が可能である。また、クライアントとサーバ間の通信は TCP/IP を用いており、インターネットを経由した利用などが可能である。なお、サーバは C 言語と FORTRAN 言語、クライアントは C/C++ 言語を用いて開発を行った。サーバは Linux または IRIX、クライアントは Microsoft Windows で動作する。

5. 動作

実行方法

本 C/S システムでは、クライアントを起動後、ユーザ認証を行うと図 2 のようなメイン画面が表示される。

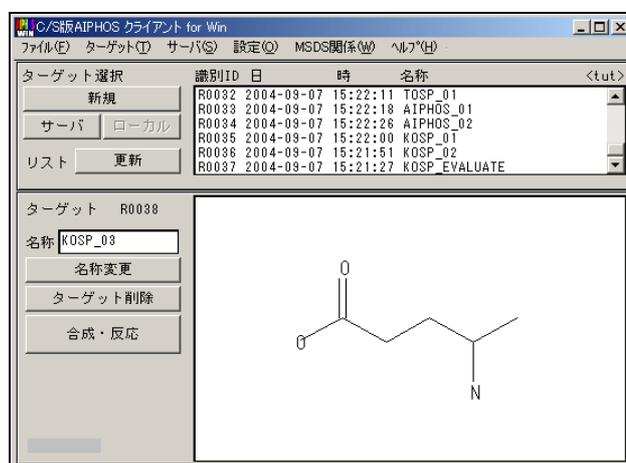


図 2 メイン画面

メイン画面は、上部のターゲット選択部と下部のターゲット表示部で構成される。ターゲット選択部では、その右側のリスト部分に以前の実行結果のログが表示される。ここで閲覧したい実行結果を選択後、「合成・反応」ボタンを押すことで、過去の結果を閲覧できる。

新しく合成経路設計や反応評価を行う場合は、ターゲット選択部の「新規」ボタンを押し、本 C/S システムに実装されている構造式エディタに直接描画するか、または MDL の MOL ファイルとして入力構造を用意する。描画後の入力構造は、ターゲット部右側の入力構造表示部に表示される。次に、メイン画面上の「合成・反応」ボタンを選択すると、図 3 のような実行条件設定画面が表示される。ここでは、各システム (AIPHOS, KOSP, TOSP, 反応評価システム) および Transform や知識ベースの選択を行う。また、逆合成ステップ数などのオプションを指定することができる。

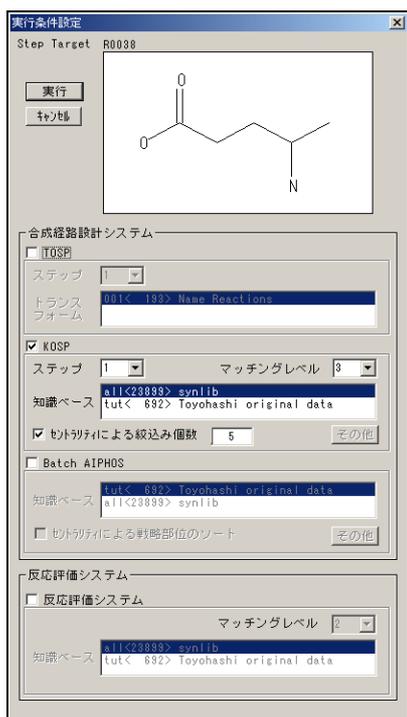


図 3 実行条件設定画面

設定後、「実行」ボタンを押すことで、クライアントはサーバに実行要求を行う。計算終了後、サーバより出力結果がクライアントに送信され、図 4 上のような結果表示画面を表示する。

結果表示画面

図 4 上の結果表示画面は、合成経路設計システム KOSP の結果である。合成経路設計の結果表示画面は、左側に経路樹、右側に逆合成反応が表示される。逆反応表示部では、左上に合成標的構造、矢印の右側に合成前駆体構造が表示される。合成標的構造中の戦略部位および、前駆体構造中の反応部位は赤く表示される。合成経路の下にはその反応を支持する反応 DB 中の反応データ ID 番号が表示される。この番号を右クリックすることで、図 4 下に示すような反応 DB 参照画面が表示され、この逆合成反応を支持した知識の元となった個別反応データを見ることができる。反応データには反応条件や文献情報が含まれている。

また、各前駆体構造を右クリックすることで、その前駆体構造を新たな合成標的構造として設定でき、引き続き 3 つの合成設計システムのどれかを用いて合成経路探索を行える。

反応評価システムの結果画面も同様の出力形式だが、合成経路設計システムの場合、反応式中の矢印は逆合成を表す白で、反応評価システムの場合は順合成を表す黒で示される。

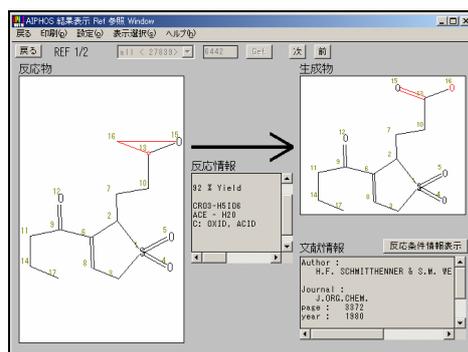
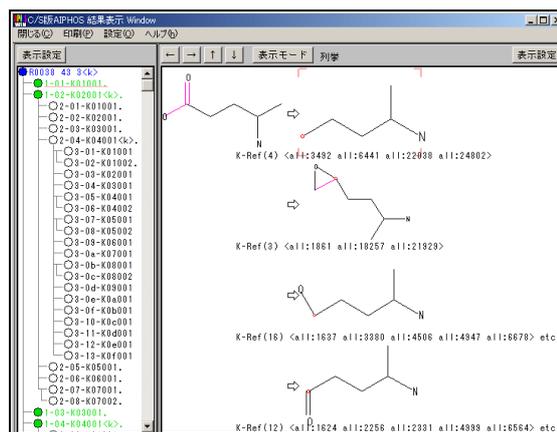


図 4 KOSP の結果表示画面

6. 反応評価システムの新機能

複数反応物間の反応予測機能

反応評価システムはこれまで、入力可能な反応物構造を 1 つに限定してきた。しかし、合成経路設計システム (AIPHOS・KOSP・TOSP) は、分子間反応を意味する、複数の前駆体構造を提案する場合がある。このように、分子間での反応も考慮した反応生成物構造予測の必要が出たことから、これに対応できるように反応評価システムの改良を行った。この結果、たとえば反応物構造 A と B との間の反応や、個々の反応物構造のみの反応の両方の結果を予測することが可能となった。

反応式のソーティング機能

有機化学者が提案された反応経路を閲覧する際に注意する点は、反応物構造中のどの部位が反応しているか、どのような生成物構造が得られるか、またそれはどのような反応条件によってかという点である。このことは、入手可能な試薬やコストの低い試薬による反応経路などを探し出すための基準にもなる。これまで開発してきた反応評価システムでは、提案された反応経路が知識ベースの反応スキーム順に列挙されていた。そのため、提案される反応経路の数が少ない場合は良い

が、その数が多くなると、特定の反応部位が反応してできる生成物構造や特定の反応条件下での反応生成物構造を結果の中から探し出すことは容易ではなかった。この問題を解決するため、反応部位および反応条件に注目することで、利用者が所望する反応経路や生成物構造が結果の中のどこにあるのかを示す機能の実装を行った。

まず、反応部位に注目した結果の閲覧を容易にする機能を実装した。反応経路を、反応部位 AB のグループ 1、反応部位 CD のグループ 2…のように、ソーティングを行い、そのグループ毎に、結果表示画面左部の反応経路樹で使用する文字色の変更を行った。反応評価システムの結果表示画面を図 5 に示す。

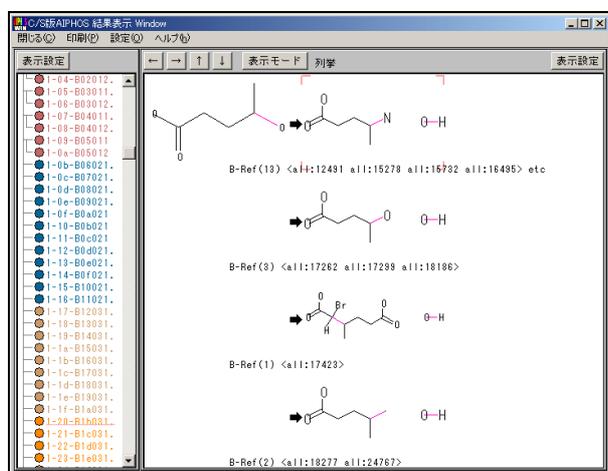


図 5 反応評価システムの結果表示画面

反応評価システムの反応表示部では左上に反応物構造、矢印の右側に予測された生成物構造が列挙される。反応部位および、生成部位は赤く表示される。

この機能の実装により、利用者が同じ反応部位の反応経路を一目で識別できるようになった。さらに、提案された反応経路に複数の反応条件が支持されている場合は、反応条件を一覧表示する機能も実装した。

反応条件検索機能

次に、反応条件に着目した反応条件検索機能の実装を行った。その結果、同じ反応条件で他の反応部位を含む反応経路や他の生成物構造が予測される経路を探ることを容易にした。反応条件検索機能は、利用者が試薬名や溶媒名など、任意の文字列を入力することにより、反応予測結果の中から適合する反応条件下で起こる反応経路を抽出し、図 6 のような検索結果画面上に表示する。

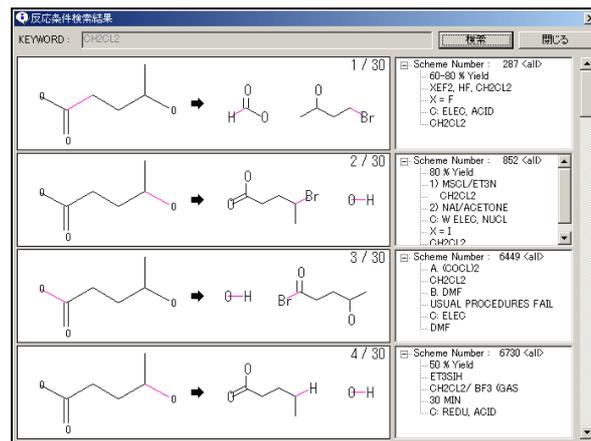


図 6 反応条件検索結果画面

反応条件検索結果画面には、ユーザが指定したキーワードを含む反応条件が画面右部に、その条件下で予測される反応経路が画面左部に表示される。反応経路表示部は、結果表示画面と同様の出力形式をとっている。

このように、利用者は同じあるいは類似の反応条件下での反応経路や反応生成物構造の可能性を容易に閲覧できるようになった。

7. おわりに

今回、複数反応物間の反応予測機能や反応部位に注目した反応式のソーティング機能、反応条件に注目した反応条件検索機能を実装したことにより、実際性の高い反応経路探索や、副反応生成物予測、そして反応条件に注目した反応評価を容易に行えるようになった。

今後は、反応評価システムを用いた反応予測のマルチステップ化を実現することで、より高度な反応設計支援システムの開発を目指したい。

謝辞

反応部位認識に関する機能について有益なコメントを下さった、国立情報学研究所の佐藤寛子博士に感謝致します。

参考文献

- [1] 船津公人, 佐々木慎一, AIPHOS-コンピュータによる有機合成経路探索, 共立出版株式会社, 1994
- [2] 塚本幸治, 佐藤寛子, 船津公人, 第 24 回情報化学討論会講演要旨集, 131-132, 2001
- [3] H. Satoh, K. Funatsu, J. Chem. Inf. Comput Sci., 35, 34, 1995
- [4] H. Satoh, K. Funatsu, J. Chem. Inf. Comput Sci., 36, 173, 1996