

SYNSUP, Synthetic Route Design System

Sumitomo Chemical Co., Ltd.
Organic Synthesis Research Laboratory
Tetsuhiko TAKABATAKE

We have developed SYNSUP and the supporting systems for users to access SYNSUP through both Intranets and the Internet. We herein summarize the history and the current status of the SYNSUP service provided to chemists in our company and also some group companies.

はじめに

化学・製薬企業では日々新規化合物の開発研究が行われている。分子設計・スクリーニングで目的の物性・生物活性を有する候補化合物が決まると、その合成法検討、化学・生物・物性評価、特許出願、安全性試験、工業化検討等を経て、製造、上市されることになる。この一連のプロセス中、化合物のスクリーニング、合成法検討、および工業化検討において、合成経路の検討が行われる。通常、研究者は自身の知識・経験・直感にもとづいて、使用する反応や原料を考え、反応データベース、文献検索により得た反応例を参考に具体的な候補ルートを考え実験を行う。思うように反応しない場合、付随的な問題が発生した場合は、別ルートを考えて実験するというやり方をとっている。最適なルートを得るには、机上検討の段階でできるだけ可能性を網羅した合成ルートマップを検討するのが望ましい。しかし、時間的な制約がある中で机上検討にかけられる時間は限られている。また、優秀な研究者でも人名反応や反応機構をすべて記憶し、化学雑誌で報告される新しい試薬を用いた反応についてキャッチアップするのは至難の業である。

コンピュータを利用すれば、膨大な反応情報を整理して、それをシステムティックに組み合わせて合

成ルートマップを作成することができるのではないだろうか？この思いをもって、我々は合成経路設計システムSYNSUPの開発を進めてきた。その経緯を振り返るとともに、現状のシステムと利用環境の概要について、これまで開発に従事した社内外の関係者を代表して報告する。

合成経路設計システムの歴史¹⁾⁻³⁾

1960年代から欧米の大学でコンピュータによる合成経路設計に関する研究が始まった。Harvard大のE. J. CoreyとTodd Wipkeにより発表されたOCSS⁴⁾が嚆矢である。両者により後にLHASA⁵⁾が開発された。これらは、既知の反応を解析したtransform(逆反応)をもとにretrosynthesis(逆合成)を行ういわゆる情報指向型システムである。Toronto大のMalcolm Bersohnも同時期に合成経路設計システムの研究に着手し、Ashmeed Esackと初の自動型逆合成設計システムを開発した^{6), 7)}。

1970年代にIvar UgiとJohann Gasteigerは、反応を結合と電子対の組み換えで取り扱うCICLOPSを報告した⁸⁾。後に、GasteigerによりEROSに発展した⁹⁾。これらは、論理指向型システムの先駆けである。欧州企業がWipkeのシステムに注目し、企業連合を結成してCASPの共同開発を行った¹⁰⁾。

1980年代に入ると、情報指向型システムの派生技術として、反応データベース検索システムREACCS¹¹⁾が開発された。その後継システムISISが今日普及している。

1990年代になると欧米企業は合成経路設計システムから分子設計に関心を移した。一方、国内では船津公人によるAIPHOS¹²⁾の開発が始まった。

2000年代に入ると、医薬品開発のワークフローの一部として合成経路設計システムの利用を提唱する会社が現れた（ROW2 Technologies, Inc.のChem-Spire¹³⁾, Simulated Biomolecular Systems Inc.のARChem Route Designer¹⁴⁾）。各システムの変遷についてはFig. 1参照。

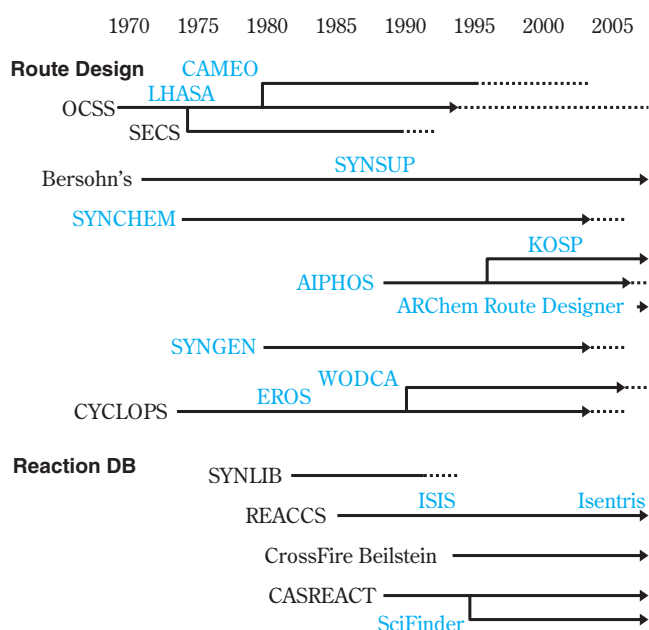


Fig. 1 Transition of major computer aids for organic synthesis

当社での取り組み^{15), 16)}

当社では、1980年代初めに合成経路設計システムの調査を契機として、この分野の研究を開始した。文献調査の他、代表的なシステムの研究者には直接コンタクトを行い、その結果、論理指向型のEROSを1983年に、情報指向型のBersohnプログラムを1984年に導入した。前者は、官能基をもつ化合物の逆合成には不向きであることが分かり、後に、逆合成に特化したWODCA¹⁷⁾を導入した。Bersohnとは、SYNSUP (SYNthesis SUmitomo Program) プログラムとして共同開発を開始した¹⁸⁾。その他、反応予測システムCAMEO¹⁹⁾を導入しpKa予測に利用している。また、不斉化合物の原料提案を行うCHIRON²⁰⁾、

反応機構をモデル化したルールを用いて逆合成を行うSYNGEN²¹⁾の評価を行った。反応データベースとしては、REACCS, ISISなどを導入し合成研究者が利用している。

SYNSUPの評価を目的に、1990年に、LHASAを導入している会社と共同でサンプル化合物のテストランを行い2つのプログラムの結果を比較した。また、1998年には、同様に、SYNCHM²²⁾と比較検討を行った。何れの場合も、SYNSUPがこれら代表的な情報指向型プログラムと遜色がないことを確認した。そこで、ユーザ実行環境を整備して2000年に全社に公開した。その後、順次住友化学グループ会社に利用サービスを拡大してきた。現在、年間700件程度実行されている。Bersohnとの共同研究は2007年末で終了し、それ以降、Bersohnと住友化学(株)の双方で独自にシステム改良を続けている。

SYNSUPの概要

1. 動作概念

SYNSUPにおける処理の流れをFig. 2に示す。目的化合物が与えられると、分子の構造特徴を認識して反応部位認識を行う。ひとつの反応部位を取り上げ関連の反応ルールを適用して前駆体（原料化合物）を生成させる。このとき位置選択性や化学的選択性のチェックを行う。問題がなければひとつの逆合成が完成する。得られた前駆体が入手可能化合物ファイルに見出されるか、ユーザが指定した条件を満たせば、ひとつのルートが完成したことになる。前駆体が条件を満たさない場合は、それを目的化合物として同様の操作を繰り返す。ひとつのルートが完成すると、一段階前の前駆体に戻って別の反応部位を

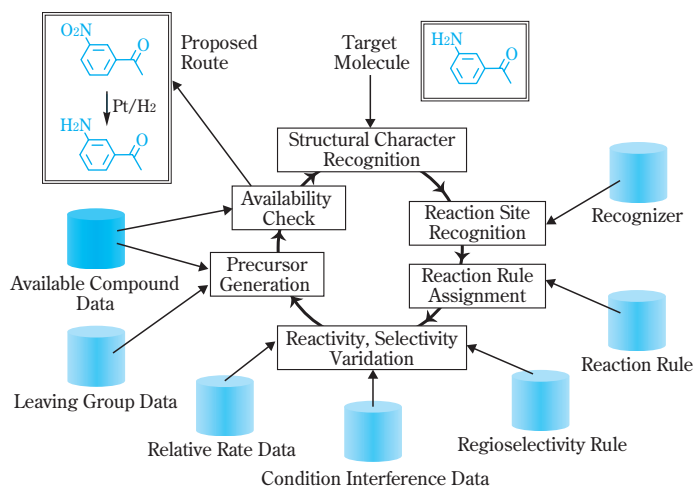


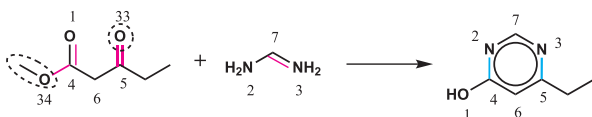
Fig. 2 Illustrative diagram of the search mechanism of SYNSUP

とりあげ同様の操作を行う。最終的に目的化合物、およびすべての前駆体のすべての反応部位が試されたところで探索が終了する。以下、主要な構成要素について概説する。

2. 反応ルール

逆反応を行うために必要なデータを反応ルールと呼ぶ (Fig. 3 参照)。反応の前後で変化する原子およびその反応に必要な部分構造を含めて反応部位と呼び、これを定義したものが反応部位認識コード recognizer である。反応ルールを参照するために反応部位番号を用いる。競争反応のチェックを行うために、原料中で変化する部分構造 (これも recognizer) の番号を列記する。化学的選択性のチェックに用いる反

Reaction Scheme



Reaction Rule

REACTION NUMBER 5632 formamidine cyclization
 SUBNUMBERS 2262¹⁾ 136
 NEXTSAME 0 NO PREVIOUS REACTION
 CONDITIONS 20²⁾ TEMP 25 25
 YIELD 67
 TESTS 169³⁾ 5 5080 169 6 5080
 OPTIONS 1⁴⁾ 9 1 8 10⁵⁾ 2
 MAKEBOND 2 4 3 5
 BREAKBOND 5 33 5 33 4 34 1 4 3 7
 REACTANTFUNG 1 0 4 144 5 136 7 315
 TITLE: A 4-hydroxypyrimidine from the condensation of a beta ketoester and formamidine acetate with NaOMe in MeOH at room temp, J. Heterocyclic Chem., 29,1369(1992),M. Butters, Org. Lett., 7,3801(2005),J.C. Yoburn.
 Two-step procedure: condensation of a beta ketoester and thioure, followed by desulfurization by Ni/H₂, Org. Synth. Coll. Vol. iv 1963,638.
 1) recognizer number
 2) alkoxide
 3) atom 5 must not be fused with aromatic ring
 4) leaving group: 9 (unsaturated oxygen); 8 (methoxy)
 5) aromatize the ring with atom 2

Recognizer

case 65: //phenolic OH, oxygen
 ...
 case 282: //N-C=N, central carbon
 makesb(mol,2262, // 4-hydroxypyrimidine
 1,i,1, //sb[L][1]=i
 1,j,7, //sb[L][7]=j
 33,1,7,4, //sb[L][4]= NEAREST NBR OF ATOM 1 TO ATOM 7
 113,4,6,1,0,0,2,2, //ATOM 4:ARYL C, 2 HETERONBR
 3,4,7,2, //sb[L][2]= COMMON NBR OF 4 AND 7
 113,2,7,1,0,0,0,0, //ATOM 2:UNSAT'D N, NO HETERONBR
 57,2,2582, //RINGINDEX MUST BE 2582
 13,4,1,2,6, //sb[L][6]= A THIRD NBR OF 4 NOT 1 OR 2
 113,6,6,1,0,1,0,1, //ATOM 6:ARYL C, 0-1 H, 0-1 HETERONBR
 7,6,4,5, //sb[L][5]= OTHER RING NBR OF 6 BESIDES 4
 113,5,6,1,0,1,1,1, //ATOM 5:ARYL C, 0-1 H, 1 HETERONBR
 3,5,7,3, //sb[L][3]= COMMON NBR OF 5 AND 7
 113,3,7,1,0,0,0,0, //ATOM 3:UNSAT'D N, NO HETERONBR
 0); //SUBSTRUCTURE END

Fig. 3 An example of a reaction rule and the recognizer for it

応条件 (試薬) 番号、および温度範囲を指定する。文献にある典型的な収率を記載する。反応部位を定義するための情報として recognizer だけで不十分な場合は、別に用意したテスト番号を指定する。より基本的な規定は前者で、類似反応の区別には後者を用いる。脱離基が必要な場合は、脱離基情報を指定する。前駆体の構造を生成させるときに必要な結合の生成 / 切断に関わる原子対を指定する。反応の前後で変化 / 消滅する不斉中心、幾何異性がある場合は指定する (Fig. 3 の例ではこの項目はない)。反応の前後で変化する官能基について指定する。最後に、反応の説明および参考文献の書誌情報を記載する。

収録している反応は、Organic Synthesis, Collective Volume, I ~ X、および教科書記載の反応をほぼ網羅している。過去10年は、医・農業、電子材料で多用されるヘテロ環合成反応に注目して有用な反応ルールを構築してきた。反応ルール合計は5900件余りである。

3. 反応部位認識

反応ルール登録時に、既存の recognizer で適切なものがない場合は新規追加する。反応部位中の2つの構造特徴に着目してその中心原子をアンカーとして、反応部位に含まれる原子の特性を定義する。Fig. 3 に示した recognizer 例では、芳香族ヒドロキシ基の酸素とアミジンの炭素をアンカーとして、ピリミジン環とヒドロキシ酸素の合計7原子からなる反応部位を定義する。

SYNSUP 実行時に、分子の構造が与えられると、分子の結合表の規格化が行われた後、官能基、環、立体的な特徴が認識され、官能基番号、環番号、不斉、幾何異性、環上置換基の cis/trans などの構造特徴データが分子の情報として記録される。反応性に関係する allyl 位、benzyl 位、橋頭位なども記録される。次に、分子中の構造特徴データを用いて、登録されている recognizer 中で該当するものを検索し反応部位番号のリストを作成する。

4. 反応性、選択性チェック

一般的な経験則はルールとしてプログラムしている。例えば、可能な反応点が複数あり、それらが等価でない場合は副生成物が予想されるため反応を却下する。芳香族親電子置換反応のルールでは、環の置換基により原料化合物の反応性を推算し反応ルール適用可否を判断する。

反応条件チェックにおいて、反応部位以外の官能基が反応するとみなされる場合は、副生成物を与えるため却下する。具体的には、2段階でチェックが行われる。まず、反応阻害データ、および競争反応

データにより化学的選択性について詳細なチェックを行う。前者は、反応条件と反応部位番号との組みで、後者は、反応条件と比較する2つの反応部位番号との組みで、それぞれ根拠となる文献書誌情報とともに記載している。もし、該当データがなかった場合は、反応条件の官能基に対する作用について、反応条件180 × 構造特徴150の配列に整理したデータを参照してチェックを行う。

5. 入手可能化合物ファイル

市販試薬データベースから、SYNSUPで取り扱える、原子数72以下で有機塩やポリマーを除く有機化合物を抽出し、化合物名とサプライヤー名、およびデータがある場合は、価格データ、CAS NUMBERを収録している。具体的には、試薬データベースCAP²³⁾を使用して逐次データ更新している。現状、入手可能化合物ファイルには、90万化合物が登録されており、内30%は価格データ有、12%はCAS NUMBERデータ有である。

入手可能化合物ファイルは、ルート探索時に参照してルート完成の判定に用いる。また、提案ルート出力時に、入手可能な出発物質、および副原料に相当する試薬データの出力に用いる。

6. 探索アルゴリズム

目的化合物で認識された反応部位リストに対して反応ルールを無制限に適用すると、膨大な数のルートが生成することになる。これを防止するために種々の仕組みを有している。反応ルールを9種類に分類し骨格形成反応を優先している (Table 1)。官能基変換反応をやみくもに適用してもルート完成に近づくことができないからである。最下位にある分解反応は、出発物質を指定した場合にのみ参照するようにしている。

より望ましいルートをできるだけ網羅するために、合成樹の探索中に枝の剪定 (pruning) を行っている。目的化合物、およびこれから生成される前駆体すべてについて官能基数、不斉中心数などに基づいてcomplexityを計算する。一つルートが完成すると、各ステップの前駆体のcomplexityが判断指標として利用される。すなわち、逆合成を行う際に前駆体のcomplexityが同ステップ数における最小値を超える場合は、非効率なルートの可能性が高いため探索を中断して次の枝の探索に移るようにしている。以上が網羅的探索アルゴリズムの概要である。

複雑な目的化合物の場合、反応部位が多数含まれ、かつ長大な合成ステップ数を要するため、組み合わせ爆発が深刻である。この解決策として、選択的探索アルゴリズムが開発された。目的化合物の構造特

Table 1 Classification of reactions for the reaction rule

priority	categories	descriptions about the reaction categories
1	specialr	enantioselective or makes more than 1 skeletal bond
2	hbuilder	diastereoselective or makes a skeletal C-Hetero or Hetero-Hetero bond
3	ccbuidler 1	makes a skeletal bond by losing at most 1 leaving group
4	ccbuidler 2	makes a skeletal bond by losing more than 1 leaving group
5	ufgi	makes a reactive group by functional group interchange (e.g., from allyl alcohol to allyl halide)
6	removeprotection	removes a protecting group
7	changfun	functional group interchange, usually trivial
8	removefun	removes a functional group by being replaced by H (e.g., deoxygenation of a ketone)
9	degrade	breaks off a carbon fragment (e.g., ozonolysis)

Reactions to introduce a protective group are classified as one of the categories with priority 1 to 7.

徴認識後、分子のcomplexity中心をもとに戦略的に重要な結合 (key-bond) を判定し、これを含む反応部位を形成する反応を優先的に適用することにより、分枝状 (convergent) 合成ルートの探索を行う戦略である²⁴⁾。このとき、反応性・選択性チェックにより骨格形成反応が適用できない場合は、脱保護、もしくは官能基変換反応を適用してkey bondに反応が適用できるように誘導される。網羅的探索と比較して極めて短時間で終了するのが特徴である。

7. 実行オプション

SYNSUPは目的化合物を与えて実行を開始したらユーザの介在なしに合成経路探索を行うため、探索の停止ルールが必要である。実行オプションの例をTable 2に示す。基本的なオプションは、ステップ数の上限を示すSTEPLIMITと出発原料の条件CATALOGである。STEPLIMIT 2、およびCATALOG 2 (出発原料は入手可能化合物のみ) を初期設定してあるので、ユーザがオプションを指定しなくても実行できる。ルート提案がない場合は、自動的にSTEPLIMITを増やして再実行が行われる。ユーザが不必要に大きいSTEPLIMITを指定すると、冗長なルート探索を行い探索時間が過大になる。また、短工程の望ましいルートが提案されない恐れがある。このような問題を回避するために、STEPLIMITの自動設定方式は非常に有効である。

Table 2 Typical Execution Options for SYNSUP

STEP LIMIT	maximum number of steps acceptable in a route
CATALOG	1 : cataloged compound is treated as available, 2 : only cataloged compound is acceptable as the starting material, 3 : the starting material and all coreactants in a route have to both be cataloged compound and satisfy other options specified
RING LIMIT	number of rings acceptable in the starting material
ATOM LIMIT	number of atoms acceptable in the starting material
FUNCTIONAL GROUP LIMIT	number of functional groups acceptable in the starting material
AROSUBST LIMIT	number of aromatic substituents acceptable in the starting material
NO_LG_VARIATIONS	route variations with only the difference of leaving groups to be rejected
STARTING MATERIAL	exact starting material to be used
ASYMMETRIC ONLY	only enantio- or diastereo-selective reactions to be applied
INDUSTRIAL	only industrially applicable reactions to be used
OMIT	omit the use of specified reactions

ユーザ実行環境

1. ユーザインタフェース

1996年に公開した初期のSYNSUPユーザ実行環境は、各ユーザのPCからUNIXサーバにログインしてグラフィックユーザインタフェースを起動する方式で

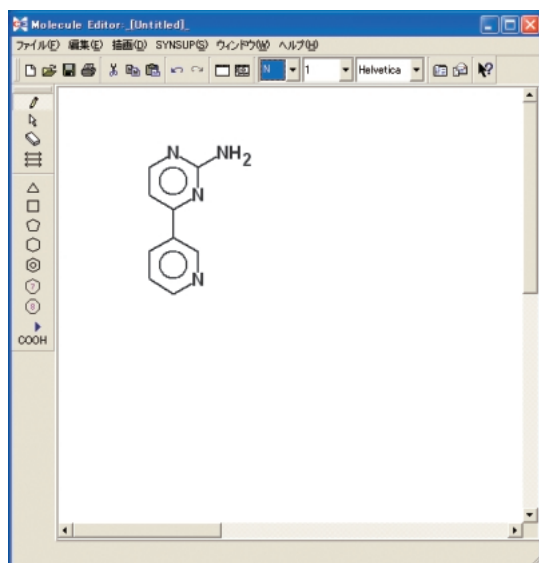


Fig. 4 CMBedit –Molecule editor

あった。しかし、研究所外からのアクセスは通信速度が遅く実用的ではなかった。そこで、PC上で利用できる新インタフェースCMBedit (Chemical Memory Bank edit) を開発した。

CMBeditは、目的化合物の構造を描画するMolecule Editorと提案ルートを表示するSynthesis Viewerから構成されている。前者で、テンプレートと鉛筆ツールを使用して構造式を描画した後 (Fig. 4)、「標準オプション」ウィンドウで、STEPLIMIT、CATALOGなど実行オプションを指定する (Fig. 5)。さらに、「拡張オプション」でAROSUBST LIMIT (芳香環の置換基数の上限) 1を指定して実行した結果をFig. 6に示す。4種類のルート表示方法の一つ「ルートマッ

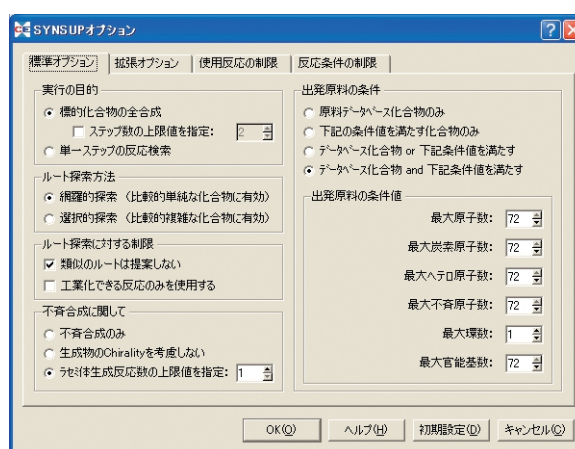


Fig. 5 CMBedit –Molecule editor–Option settings dialog box

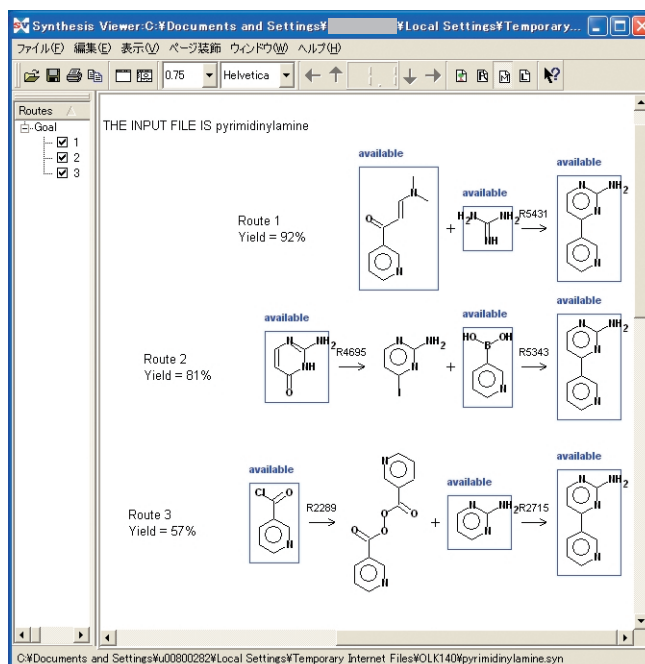


Fig. 6 CMBedit –Synthesis viewer–Map view

「表示」である。右上に目的化合物が描画され、反応式の左端が出発原料である。青ボックスで囲まれた化合物は試薬として入手可能であることを示している。その構造式をダブルクリックするとポップアップウィンドウに試薬データが表示される (Fig. 7)。反応式の矢印をダブルクリックすると同様に参考文献の書誌情報が表示される (Fig. 8)。ルートを開覧しながら不要なルートを削除して好みのルートだけをプリントする機能がある。

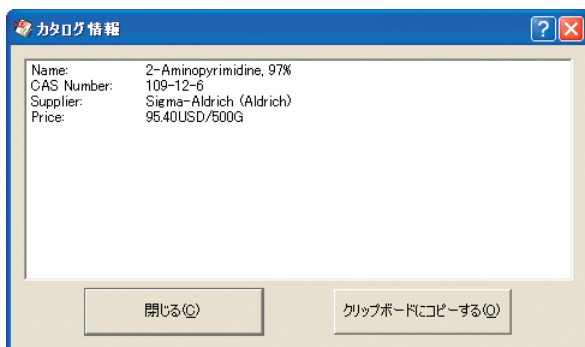


Fig. 7 CMBedit –Synthesis viewer–Catalog info

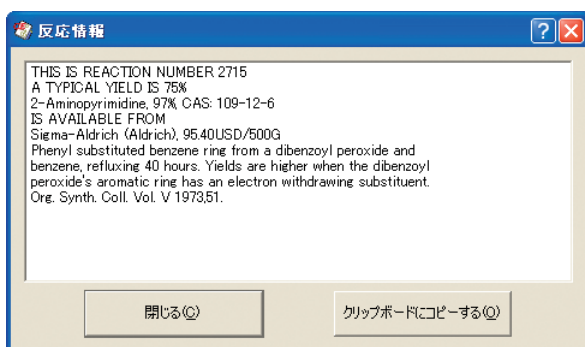


Fig. 8 CMBedit –Synthesis viewer–Reaction info

2. e-mail実行システム

ユーザがPC上で作成したSYNSUP入力ファイルを電子メールでSYNSUPサーバに送信し、サーバ上でSYNSUPを実行した結果を電子メールでユーザに返信するシステムである (Fig. 9)。メールで届いた入力ファイルを処理してジョブ管理システムに渡すプログラムと、実行が終了すると出力ファイルを電子メールでユーザに返信するプログラムを開発し、オープンソースのジョブ管理システムTORQUE²⁵⁾と組み合わせたものである。

3. ユーザ公開

CMBeditとe-mail実行システムを組み合わせることにより実用に耐えるシステムが完成したので、2000年

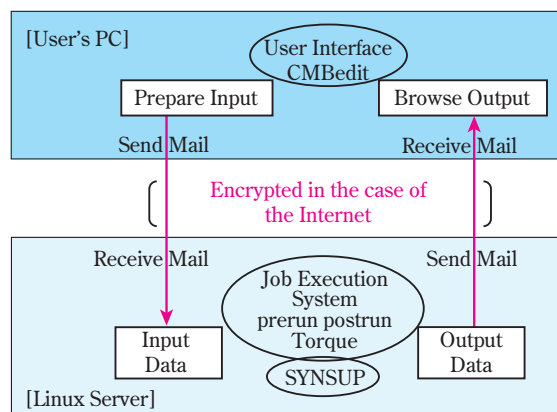


Fig. 9 SYNSUP Job execution system utilizing e-mail for input/output data transfer

からユーザ公開の範囲を拡大した。当社の有機合成研究所、農業化学品研究所を皮切りに、他研究所、および住友化学LANに接続可能なグループ会社に公開した。さらに、SYNSUPホームページを開設し、マニュアル、FAQ等の参照、およびユーザ登録、CMBeditのダウンロード機能を提供してきた。また、ユーザサポート、および統計データ取得を目的にユーザ実行ログ収集システムを運用している。

住友化学LANに接続できないグループ会社からの利用については、インターネット経由でメール送信するため、メールのセキュリティ確保が課題であった。暗号化機能をCMBeditおよびe-mail実行システムに組み込むことにより2008年からグループ会社向けにもe-mail実行システムの運用を開始した。

SYNSUPの実行例

SYNSUPは保有する反応ルールに依存するシステムであるため、必ずしも可能性のあるルートをすべて提案できるわけではないが、合成法のバリエーションを多数示すことで研究者の見落としを補完して最適な合成ルートの選択に寄与することができる。実際に工業化で寄与したものではないが、典型的な実行例を紹介する。

1) アレスロロン²⁶⁾

当社の基幹製品のひとつである合成ピレスロイドの中間体である。STEPLIMIT 3, CATALOG 2の条件で実行した結果、23ルート提案された。公知の代表的な2ルートのみ表示させた画面をFig. 10に示す。アルドール反応を用いた閉環反応Route 4など天然物合成的なルートが多い中で、当社の独自技術であるフランカルピノールのフラグメンテーション反応を用いたRoute 17も含まれていた。システムティックに反応

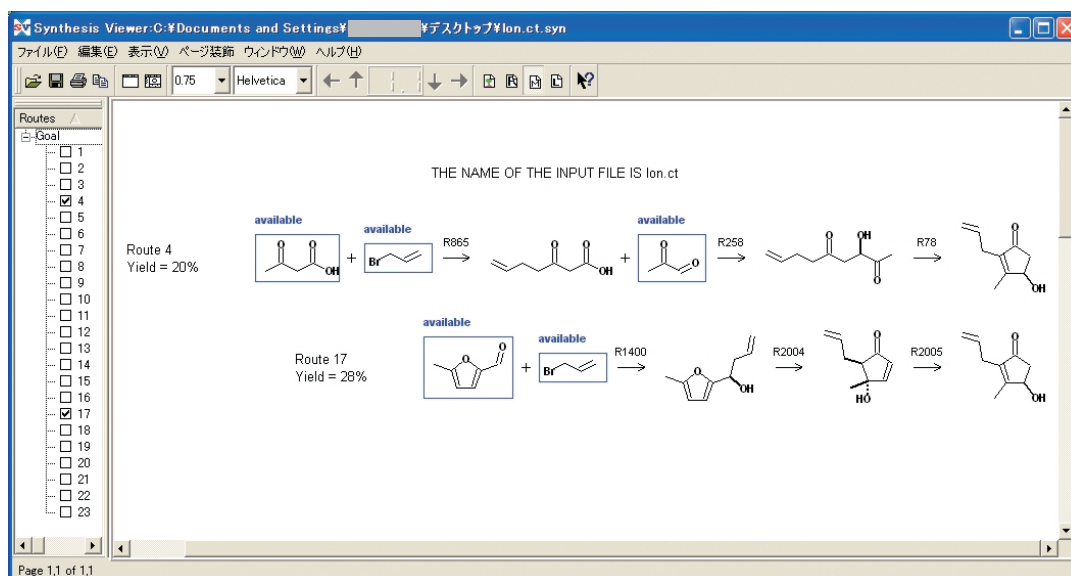


Fig. 10 Execution example 1: allethrolone

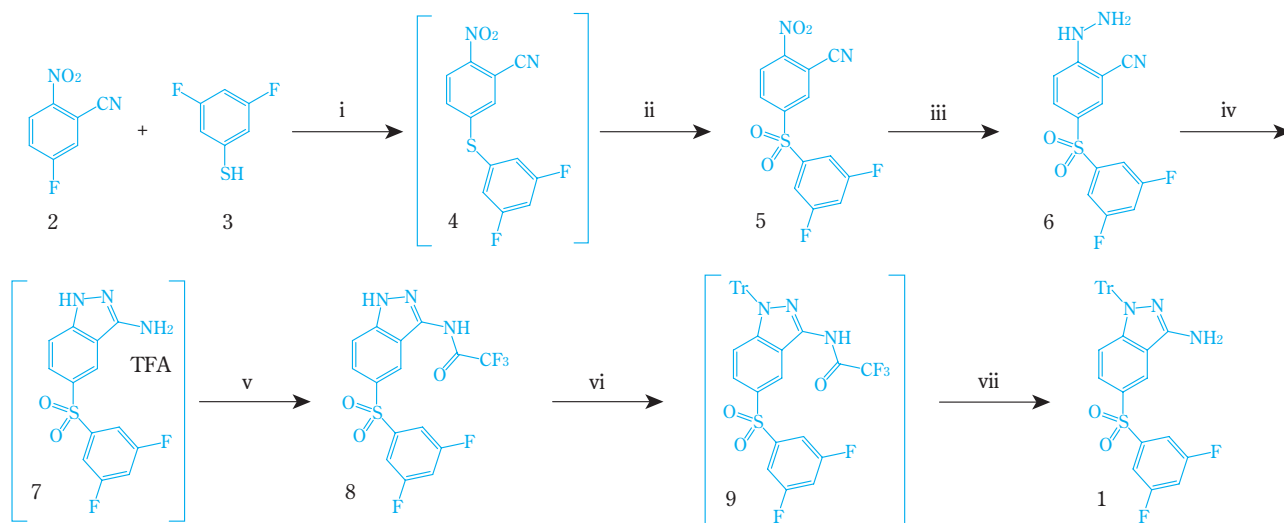
部位を探索することにより、一見隠れている転位反応の可能性を見出すことができるのは、コンピュータシステムを利用するメリットである。

2) a new insulin-like growth factor 1 receptor inhibitor (amine part)

最近文献で報告された化合物についてテストランを行った。Scheme 1²⁷⁾の1をアミド化、脱保護したものが最終目的化合物である。1のトリチル基を除いた7を目的化合物として、STEPLIMIT 3, CATALOG 2の条件で実行した結果、100ルートを超える提案があった。芳香族クロライドのフッ素置換反応を用いるル

トが頻出していたため、実行オプションで該当反応の OMIT (使用禁止の指定) をつけて再度実行したところ、25ルートに絞り込むことができた。類似のルートが含まれているため、代表的な5ルートを残して表示を消したものをFig. 11に示す。

Route 1, 8, 20, 21はアミノインダゾール骨格を有する市販試薬からスタートしている。スクリーニング目的の場合は考慮すべきルートであろう。ただし、何れのルートも保護基の検討が必要である。Route 3は、保護基の有無を除けば、文献ルートと同様のルートである。この例のように、好ましくない反応があった場合は、それらを OMIT して再実行を行うことで、ルー



Reagents and conditions: i) EtOAc, DIPEA (1.05 equiv), 0 C to rt, 0.5 h; ii) MeCN, water, Oxone (2.5 equiv), 40 C to rt, 24; iii) THF, N₂H₄, 35%, rt; iv) THF, TFA, 5 C, 15 min; v) DCM, TFAA; vi) DCM, TrCl, TEA, rt, 4h; vii) MeOH, TEA, refluxed, 5h.

Scheme 1 Synthesis of a new insulin-like growth factor 1 receptor inhibitor (amine part)²⁷⁾

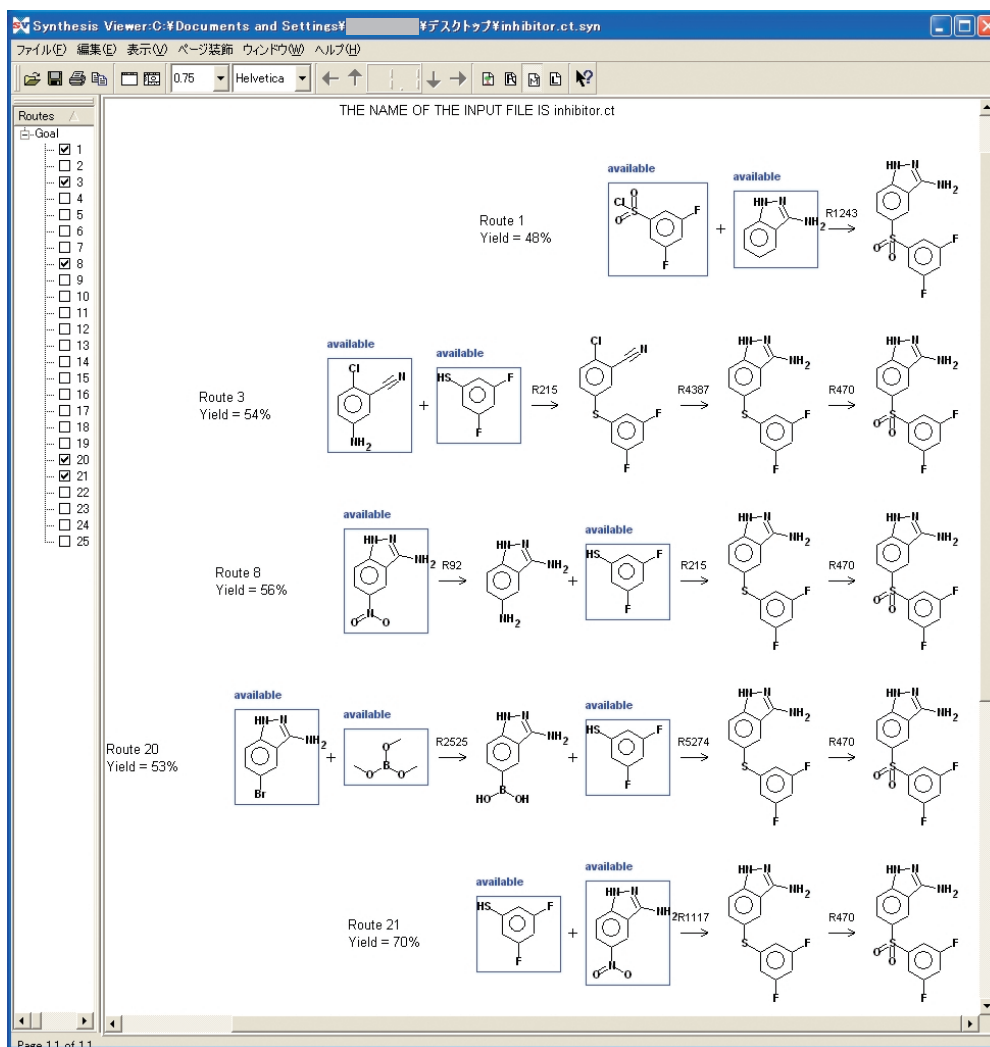


Fig. 11 Execution example 2: an inhibitor

トを絞り込むことができる。また、実行オプションの追加指定により合成樹の剪定が行われた結果、最初の実行では出力されなかったルートを提案することが可能である。

この実行には、Quad-Core Intel Xeon processor X5460 (3.16 GHz) 搭載ワークステーションを使用した。2)の例では、実行時間は1分であった。STEP LIMIT 3の場合、数分以内で終了するのが通例である。しかし、1)の例では1.5時間要した。2)と比較すると一見単純な構造に見えるが、実際は2)よりも反応部位の数、および適用可能な反応ルール数が多いためである。

今後の課題

2006年にユーザアンケートを実施した。回答者の三割は、SYNSUPが「有用な情報ツールである」と回答した。「他のツールを補完するシステムとして有

用」との回答を含めると、ほとんどすべてが肯定的な評価であった。利用回数が多いほど評価が高くなる傾向があった。一方、SYNSUPに対する改善要望として下記の指摘があった。

- ・提案ルートに不適切な反応が使用される場合がある
- ・類似ルートが多数出力される
- ・最近の反応（特に1990年以降）が不足している

SYNSUPに対するユーザ満足度を向上させるためには、可能なルートをできるだけ網羅すると同時に不適切なルートをできるだけ減らすことが重要である。

第1に、反応ルールの拡充が必要である。現状、文献から有用な反応を選択した後、反応ルール作成ツールを用いて対話形式で反応ルールの作成を行っている。また、反応の選択に関しては社外の専門家の協力を得ている。一方、市販の反応データベースから自動的に反応ルールを作成して利用するシステムがいくつか知られている（例、ARChem Route Designer¹⁴、SYNCHEM²⁸、AIPHOS²⁹）。人手をかけずに大量の反

応ルールを作成できるのはたいへん魅力的である。しかし、反応データベース中には類似反応例が多く、また、一部ではあるが誤ったデータが含まれている。有用な反応をどのようにして選別・収集するか、もしくは、実行時に適切な反応ルールをいかに選別するかが課題になっている。さらに、反応データベースのライセンス料を考慮する必要もある。我々は、当面は、有用な文献反応の選択は専門家がを行い、反応ルール作成ツールの自動化を進めるのが実際的であると考えている。

第2に、副反応チェックの精度向上が重要である。反応阻害・競争反応データの拡充が必要である。しかし、文献既知の反応条件を利用した場合でも、原料の構造が異なると反応成績が大きく変わる場合があるため、反応ルール作成時に反応の適用範囲を精確に決めるのは容易ではない。よって、SYNSUP提案ルートで選択性の問題が明らかになった場合に、その都度、recognizerの修正、あるいは反応阻害・競争反応データを登録して適用範囲を制限するようにしている。SYNSUP提案ルートの問題点を知るためにユーザのコメントが望まれる。ユーザに負担をかけずに、不適切なルートについてのフィードバックが得られるシステムを今後開発したいと考えている。

おわりに

合成経路設計は本質的に複雑な思考を要する挑戦である。例えば、未知化合物の構造同定の場合、解は一つであるが、合成ルートの場合多数の解が可能である。もし、合成樹の適切な剪定を行わなければ膨大なルートが出力される。一方、新規なルートを網羅するためには、文献や特許で日々更新される新しい反応の知見を収集する必要がある。さらに、目的によりベストなルートの評価基準が異なる。すなわち、スクリーニング化合物の合成には、信頼性の高い反応を組み合わせた最短ステップのルートが望ましい。一方、工業化検討では、原料コスト、設備コスト、安全性等の点から総合的にルートが選定される。

合成経路設計システムの歴史は40年を超えるが、まだ商用ソフトが広く普及する段階には至っていない。合成経路設計システムが有機合成の現場で真に役立つようになるためには、研究者に評価されるようなルートをしばしば提案できるよう、システムのレベルアップをはかる必要がある。そのためには、文献から有用な反応を収集するとともに、合成の達人・成書から合成戦略の発想を学び、そのアルゴリズムを取り込むことが必要である。

今日、大学、および企業の研究室では天然物の全

合成から、新規で有用な機能をもつ化合物の設計と用途開発に関心が移っている。しかしながら、目的化合物が決まったときに経済的な製法を確立するためには、依然として合成経路設計が必要である。今後、文献検索SciFinderや反応検索CrossFire Beilstein、ISISとあわせて合成経路設計システムSYNSUPが合成研究者の有用なツールとして普及していくことを期待する。

謝辞

本研究のベースとなる合成経路設計システムを開発し、24年間にわたり共同開発をご指導いただいたMalcolm Bersohn名誉教授 (University of Toronto) に感謝する。

引用文献

- 1) W. L. Chen, *J. Chem. Inf. Model.*, **46**, 2230 (2006).
- 2) J. Gasteiger, Ed. "Handbook of Chemoinformatics", Wiley-VCH: Weinheim, Germany (2003).
- 3) 高島 哲彦, 山近 洋, "応用化学講座4「有機合成化学」", 朝倉書店 (1997), p. 209.
- 4) E. J. Corey and W. T. Wipke, *Science* **166**, 178 (1969).
- 5) E. J. Corey, W. T. Wipke, R. D. Cramer, III and W. J. Howe, *J. Am. Chem. Soc.*, **94**, 421 (1972).
- 6) M. Bersohn, *Bull. Jpn. Chem. Soc.*, **45**, 1897 (1972).
- 7) M. Bersohn and A. Esack, *Computers & Chemistry*, **1** (2), 103 (1977).
- 8) J. Blair, J. Gasteiger, C. Gillespie, P. D. Gillespie and I. Ugi, "Computer Representation and Manipulation of Chemical Information", W. T. Wipke, S. R. Heller, R. J. Feldmann and E. Hyde, Eds., Wiley, New York (1974), p. 129.
- 9) J. Gasteiger and C. Jochum, *Topics Curr. Chem.*, **74**, 93 (1978).
- 10) P. Gund, E. J. J. Grabowski, D. R. Hoff, G. M. Smith, J. D. Andose, J. B. Rhodes and W. T. Wipke, *J. Chem. Inf. Comput. Sci.*, **20**, 88 (1980).
- 11) S. E. French, *CHEMTEC*, February, **1987**, 111.
- 12) 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 (1994).
- 13) <http://www.row2technologies.com/>
- 14) J. Law, Z. Zsoldos, A. Simon, D. Reid, Y. Liu, S.Y. Khew, A.P. Johnson, S. Major, R.A. Wade and H. Y. Ando, *J. Chem. Inf. Model.*, **49**, 593 (2009).
- 15) 吉田 元二, 高島 哲彦, 石田 雅也, 住友化学, 1990-II, 39 (1990).

- 16) 高島 哲彦, 竹村 年男, 銅金 巖, 住友化学, 1994-I, 69 (1994).
- 17) W.-D. Ihlenfeldt and J. Gasteiger, *Angew. Chem., Int. Ed. Engl.*, **34**, 2613 (1995).
- 18) M. Takahashi, I. Dogane, M. Yoshida, H. Yamachika, T. Takabatake and M. Bersohn, *J. Chem. Inf. Comput. Sci.*, **30**, 436 (1990).
- 19) T. D. Salatin and W. L. Jorgensen, *J. Org. Chem.*, **45**, 2043 (1980).
- 20) S. Hanessian, J. Franco and B. Larouche, *Pure Appl. Chem.*, **62**, 1887 (1990).
- 21) J. B. Hendrickson, *Angew. Chem., Int. Ed. Engl.*, **29**, 1286 (1990).
- 22) H. L. Gelernter, A. F. Sanders, D. L. Larsen, K. K. Agarwal, R. H. Boivie, G. A. Spritzer and J. E. Searleman, *Science*, **197**, 1041 (1977).
- 23) CAP: Chemicals Available for Purchase, Accelrys Software Inc.
- 24) A. Tanaka, T. Kawai, T. Takabatake, H. Okamoto and M. Bersohn, *J. Comput. Aided Chem.*, **10**, 104 (2009).
- 25) TORQUE: <http://www.clusterresources.com/products/torque-resource-manager.php>
- 26) 銅金 巖, 山近 洋, 南井 正好, 有機合成化学, **41**, 896 (1983).
- 27) I. Candiani, G. D'Arasmo, F. Heidempergher and A. Tomasi, *Org. Process Res. Dev.*, **13**, 456 (2009).
- 28) H. Gelernter, J.R. Rose and C. H. Chen, *J. Chem. Inf. Comput. Sci.*, **39**, 316 (1990).
- 29) K. Satoh and K. Funatsu, *J. Chem. Inf. Comput. Sci.*, **39**, 316 (1999).

PROFILE



高島 哲彦
Tetsuhiko TAKABATAKE
 住友化学株式会社
 有機合成研究所
 主席研究員