薬物間の共通パターンの自動認識

高橋 由雅*・佐々木 慎一*

Automated Recognition of Common Structural Patterns among Drug Molecules.

Yoshimasa TAKAHASHI* and Shin-ichi SASAKI*

It is well known that chemical structure diagram or three-dimensional structure of a compound molecule involves much information related to biological activity, reactivity, or various physicochemical properties on itself. It is a very important thing in a wide range of chemical research to compare the chemical compounds on the basis of such structural information. Especially, in the area of structure-activity studies, to investigate structural feature such as functional groups or substructural fragments, common to a group of chemical compounds exhibiting a particular biological activity often provides us with very useful and objective information which may be resulted in hypothetical pharmacophore.

In this article, a couple of approaches for the automated recognition of common structural patterns among drug molecules will be described. One is related to the maximal common substructure search which can be used to find some topological pharmacophore, and the other one is for the common geometrical pattern search in which three-dimensional geometry of molecules are taken account. The latter approach may be used to identify a certain topographical or three-dimensional pharmacophore. The basic algorithms and related software systems will be summarized with some illustrative examples.

1. はじめに

新規有用物質の合理的な開発を進める上においては、その対象となる化合物群に対する構造活性（物性、反応性……）に関する研究は不可欠のものである。特に、医薬品等の薬物開発においてはその毒性・安全性試験まで含めれば多大な労力と時間・経費を要することから、その候補化合物を如何に合理的かつ高い確率で獲得するかが大きなカギとも言える。今日、このような薬物開発研究の合理化を意図したドラッグデザインに関する種々の方法論的研究、あるいはそのコンピュータ支援システムの開発研究が内外を問わず盛んに行われている。

構造活性に関する研究は、類似の生物活性を有する化合物群においては、その活性発現に寄与するであろう何らかの共通性が存在するにちがいないとの仮定のもとにすべて出発している。そのアプローチには大きく分けて二つがある。一つは化学構造を種々の特性値でパラメータ

化し、回帰分析に代表される統計的手法¹⁾やパターン認識手法²⁾の応用など数式モデルに基礎を置くものであり、もう一つは化合物の構造情報それ自身から直接的にその共通部分や相異点などの構造類似性を比較検討しようとするものである³⁾。本稿では後者の立場から、トポロジカルなレベルさらには分子の三次元幾何を考慮したレベルの構造共通性の探索に対し、類似作用を有する薬物間の共通部分構造（あるいは共通幾何パターン）の自動認識の問題について著者らの研究を中心に述べる。

2. 化学構造式にもとづく共通構造特徴の探索

類似作用を有する化合物間の共通部分構造の探索については、従来の古典的な構造活性相関研究の中で経験的に行われてきたものである。しかしながら、コンピュータによる化学構造式の取り扱いに関する種々の工夫の中で、こうした研究者の直感的な考え方をより論理的に、あるいは網羅的に展開できるようになってきた。Cramer⁴⁾は、特定の薬効をもつ化合物群の中から共通に含まれる官能基や部分構造を探索し、それらの有無や数と生物活性との関係を論じる“部分構造解析”と呼ばれる考え方を示し、それらの応用例を報告している。また、コンピュータ利用によるものではないが我が国において

*豊橋技術科学大学・分析計測センター（ ）

*Research Center for Chemometrics, Toyohashi University of Technology（ ）

も岡ら⁵⁾により、官能基対 (Functional Group Pair, FGP) と呼ばれる部分構造群に注目した同様な考えが報告されている。しかしながら、これらのアプローチにおいては解析の対象とする化合物群に対し、研究者が任意に設定した構造フラグメントについて断片的にその有無や数を調べるものであり、そのために、研究者の主観から生じる活性発現に寄与する他の有用な構造因子の見落としが往々に起こり得る。また、これらの考えは設定した個々のフラグメントそれ自身に注目したものであり、フラグメント相互の結合関係などの情報が反映されない。これらはフラグメントコードに基礎を置いた部分構造解析に対する共通の問題とも言える。従って、こうしたフラグメントコードによる構造情報の記述を避け、特定の部分構造を設定することなく、より柔軟な構造特徴の探索が可能なアプローチが望まれる。

3. 最大共通部分構造の自動認識

化学構造式にもとづく共通パターンには様々なとらえ方がある。例えば、ある類似薬効を有する化合物群に対し、その構造共通性を議論する場合に、研究者によっては通常の部分構造としての比較の他にその単純骨格のみに注目したり、あるいはヘテロ原子の結合関係のみに注目して構造の共通性を見出そうとすることもしばしば行われる。一方、共通部分と言っても構成原子数が非常に少ない共通の構造フラグメント (例えば、構造原子がCやNやOのただ1個であったり、あるいはC-Cなど構造特徴としてあまり価値のないもの) については構造活性相関上積極的な意味を見出し難い。従って、構造活性相関の観点からは、対象とする化合物群に対し、構造中に含まれるより大きな共通部分を見出すことに大きな意義が持たれよう。このことから、著者らは種々の構造表現レベルでの探索が可能な最大共通部分構造の自動認識システム MAXFIT の開発を試みた⁶⁾。

3.1. 化学構造式のグラフ表現と共通性の定義 化学構造式にもとづく共通性の定義には多くの方法が考えられる。MAXFIT では化学構造式を化学グラフ*と見なし、グラフ理論的な表現を用い、構造情報の共通性の定義に対し、種々の表現レベルを設定した (図1)。

レベル 1: 単純グラフ表現 構造中のすべての原子および結合のタイプを等価とみなす。

レベル 2: ノード重み付きグラフ表現 ここでは、種類の異なる原子は非等価とみなし、結合はすべて等価とする。

(i) Simple graph

(ii) Node weighted graph

(iii) Edge weighted graph

(iv) Node and edge weighted graph
(not included hydrogen atoms)

(v) Node and edge weighted graph
(included hydrogen atoms)

Fig. 1 Various graph representations of the chemical structure of acesulfame.

て等価とする。

レベル 3: 辺重み付きグラフ表現 レベル 2 とは逆に原子の種類はすべて等価とみなしタイプの異なる結合はすべて非等価とみなす。

レベル 4: ノードおよび辺重み付きグラフ表現 種類の異なる原子および結合はすべて非等価とみなす。ただし、各原子の隣接水素原子を考慮しない。

レベル 5: ノードおよび辺重み付きグラフ表現 (含水素原子) レベル 4 に加えて各原子に結合している水素原子の数を考慮する。

上記 5 つのレベルからなる共通性の定義により、レベル 1 では単純骨格表現での化合物間の共通部分構造を認識でき、レベル 2 およびレベル 3 ではそれぞれ原子種のみ、あるいは結合タイプだけを区別したものが検出できる。レベル 4 では一般に研究者が通常行っているような原子種および結合の多重度 (2 重結合, 3 重結合, 芳香族結合) を同時に区別した共通フラグメントの認識ができる。また、レベル 5 ではレベル 4 に加え、隣接水素の数により原子種のみではなく 1 級, 2 級, 3 級, 4 級な

* 原子を点 (節あるいはノードとも言う)、結合を辺としてグラフ理論で言うグラフとして表現したもの (通常は水素原子を省略)。

どの同一原子種間での環境も区別した認識が可能である。もちろん、原子の重み付けに関して言えばこの他にも混成等を考慮した重み付けも考えることができるが、著者らは開発の便宜上、構造間の共通性を上記5つの表現レベルに限定して最大共通部分構造の自動認識の問題を検討した。

3.2. アルゴリズムの概要 ここで定義する最大共通部分構造とは“化学構造式をグラフ表現した場合に、対象グラフ間に共通に含まれる辺の数が最大の連結部分グ

ラフ”を言う。従って、非連結グラフ（分離構造となるような部分グラフ）は除外される。

ステップ (i) 複数構造（グラフ）の中から基準構造（グラフ）を選ぶ。

ステップ (ii) 基準グラフをサイズ1のサブグラフに分解する。

ステップ (iii) 部分グラフの各々を規範化*し規範表現をもとに同型グラフの有無を調べる。
(サイズ2以上の場合には、これに先立

Fig. 2 A schematic flow of the algorithm of the largest common substructure search.

って (vi) でのグラフ成長過程における重複生成の排除を行う。)

ステップ (iv) 解析の対象とする構造のすべてに共通な部分グラフを見つける（ここでは部分構造検索の手法を応用）。共通部分グラフがなければ探索を終了。

ステップ (v) この段階までに得られた共通部分グラフを候補者リストにしまう。

*例えば、構造式（あるいは部分構造式）を行列表記で記述しようとする場合、各原子の番号付けを変えるだけで、行と列を入れ替えた何通りもの行列表現が可能となる。このことから構造比較を行うためには個々の構造に対応し、一義的に定まる表現法が必要となる。規範表現はその構造表現の方法、およびその規範化のためのアルゴリズムに依存する。有名なものとしてはCAS/Morganの方法 [H. L. Morgan, *J. Chem. Doc.*, **5**, 107 (1965)] がある。

ステップ (vi) リストにしまわれた部分グラフを成長させる。1つだけサイズの大きい可能な部分グラフをすべて生成させる (ステップ (iii) へ)。これより大きい部分グラフの発生が不可能ならば、この時点で終了。

探索操作の終了時に候補者リストにしまわれている部分グラフに対応する部分構造が、目的の探索レベルでの最大共通部分構造となる。図2にこれらの処理の流れを模式的に示す。このアプローチでは、基準構造を、解析の対象とする化合物集合の中から任意に選ぶことが出来る。しかしながら、一般にはアルゴリズムを考慮すれば、結合の数の最も少ないものを基準構造とする方が探索過程で生成される部分グラフの数が少なくなり、効率が良い。

またステップ (ii) で生成されるサイズ1の部分グラフの数は辺の数に等価である。

ステップ (iii) ~ ステップ (iv) の部分グラフの成長に伴う探索の繰り返し過程においては、最終的な結果が探索下の化合物グラフ表現中のすべての可能な共通部分

グラフの中で“最大”であるとの保証を得るためにはステップ (ii) で生成されたサイズ1の部分グラフをもとに網羅的なグラフ成長が行われなければならない。これに対しては、既に著者らによって開発された GROWSUB⁷⁾ (ある与えられた構造式からすべての可能なユニーク部分構造を列挙するためのコンピュータプログラム) のアルゴリズムを用いている。ステップ (iii) での重複チェックは、部分グラフの成長に伴う重複生成の排除および生成された同型部分グラフ (構成辺即ち、ラベル付き辺集合は異なる) によるステップ (iv) での重複探索を回避するための代表部分グラフの選出を含む。ただし、それぞれの代表部分グラフが対象化合物間に共通に存在すれば、ステップ (v) の候補者リストにはその代表グラフと同型の他のグラフも登録される。これらの処理の概念を図3に示す。図3は、図2の探索例においてサイズ1から2へ進む場合を例に、その処理を模式的に示したものである。図中の (5), (6) は重複生成, (7), (8) は同型の関係にある。ここでの部分グラフの重複生成の排除は辺のラベル付けにより生成された部分グラフの全構成辺ラベルの集合が等価なものについては1つを除き他を生成リストから除くことにより処理される。また、構成辺のラベル集合が異なるグラフ間で生じるいわゆる同型チェックには結合スタック法⁸⁾を Hendrickson らの規範化表現法⁹⁾にもとづいて改訂して用いた。

3.3. MAXFIT の実行例 上記アルゴリズムに従って作成した MAXFIT プログラムの実行例を二つほど以下に述べる。最初の例として、プログラムでの種々の探索レベルによる結果の相異を直感的に理解することが容易な二つの合成甘味剤アセスルファムとサッカリンでの簡単な適用例を示す。図4はこの二つの構造中に存在する最大共通部分構造を各探索レベルで示したものである。レベル1では結合数9からなる2種の最大共通部分構造が得られている。また当然のことながらグラフ表現としての重み付けが詳細になるほど結果として得られる最大共通部分構造のサイズが小さくなっていることが判る。これらの探索処理に伴って生成された部分グラフの数を表1に示す。表1から解るように、重み付けが単純なほど生成される部分グラフの数が多くなっている。これは最終的に得られる共通部分グラフのサイズにも依存するが、そればかりでなく、重み付けが単純なほど探索過程での各サイズレベルで生き残る共通部分グラフの数が増加することによるものである。

ここで示した例について、読者は自らの目と脳コンピュータでどのような最大共通部分構造を見つけること

Fig. 3 A schematic flow of the duplication check with labeled-edge.

Fig. 4 Largest common substructures between acesulfame and saccharin found by MAXFIT under different search levels.

Table 1 Results of the common substructure search.

search level	maximal size	the number of unique common subs.
1	9	32 (254) ^{a)}
2	8	44 (91)
3	7	25 (97)
4	6	24 (31)
5	5	15 (18)

a) The figure in parentheses is the total number of common substructures which involve isomorphic ones. In here, acesulfame was used as the reference structure.

ができるであろうか？ 恐らくレベル 4 あるいはレベル 5 については容易に見つけることが出来よう。しかし、レベル 1 の二つを含めすべての場合に正しく共通部分構造を認識することはこのような簡単な例においてすらそれほど容易ではないことをお判り頂けると思う。さらに、対象とする化合物が増えればそれが困難なことは一層明かであろう。一方、部分構造を中心に構造活性相関の問題を検討する場合においても、必ずしもこのような最大共通部分が最も重要な情報であるとは限らない。そこで、本プログラムではアルゴリズムに示されるようにいずれの探索レベルでの処理においても可能なすべての共通部分構造を得ることができるよう工夫されてい

る。参考のために、前述の例に対し、レベル 4 で検出されたすべての共通部分構造を図 5 に示す。

もう一つの例として、代表的なオピエート類 3 種 (Morphine, Meperidine, Methadone) に対する最大共通部分構造の探索例を示す。Morphine は比較的 rigid な分子構造を持つ。これに対し他の二つは構造化学的に大きく異なる分子であるにもかかわらず Morphine 様作用を示す点で興味深い化合物群である。これら化合物の構造式をもとに MAXFIT を用いたレベル 4 での探索の結果、結合数が 14 からなるグラフ的に異なる二種の最大共通部分構造を得た (図 6)。図 6 において太線で示されるような Morphine のベンゼン核およびペリジン環上の窒素を含んだひとつづきの比較的大きな共通部分構造はこれらの化合物の Morphine 様作用の発現に対する "topological pharmacophore" を暗示するものと考えることが出来る。事実、中枢神経系作動薬あるいは鎮痛薬におけるフェニル環および窒素部位の同時存在の重要性は広く指摘されているところである。一方、これらの分子の X 線結晶構造にもとづき、対応する原子の最小二乗 fitting によりその共通部分の構造重ね合わせを行った。コンピュータグラフィックスの利用による重ね合わせ表示の結果からみて、これらの部分が三次元的にも良好な対応関係を示していることが解る (図 7)。

このように、トポロジカルなレベルながら、化合物間

段の探索が必要になってくる。

4. 三次元共通幾何パターンの自動認識

上で述べたトポロジカルなアプローチは、比較的多数の化合物間の構造類似性を構造式レベルで検討する上で十分に有用なものがある。一方、薬物の薬効発現に寄与する構造因子を考えた場合、分子の三次元幾何を考慮したより詳細な構造類似性を検討することは構造活性相関解析においては極めて重要な問題でもある。近年、コンピュータグラフィックスの普及およびその利用技術の進歩とともにこれらを利用した構造重ね合わせ処理を中心とするいくつかのレセプターマッピングの手法が報告されている¹⁰⁾。単なる構造重ね合わせとは異なるものとして、pharmaco pattern search¹¹⁾と呼ばれる三次元部分構造検索の方法がある。これは通常の部分構造検索と同様に query 部分構造 (ただし、三次元構造として) を指定し、その有無および対応部位を調べる方法である。一方、構造重ね合わせによる分子の形状比較を定量的に行おうとする試みもなされている¹²⁾。しかしながら、query 構造や重ね合わせの基準原子はどのようにして選んだらよいのかという問題が残る、これ自体が重要な研究課題でもある。従って、query 構造を指定したり、重ね合わせの基準原子を指定することなく、比較の対象とする分子の共通幾何を自動的に認識できれば、pharmacophore の候補が自動的に検出されるとともに、前者の query 部分構造や重ね合わせのための基準原子そのものを見つけ出すことができるはずである。このような観点から著者らは三次元分子構造間の共通幾何パターンの自動認識の問題について検討し、そのためのソフトウェアシステム COMPASS (three dimensional COMmon geometrical PAttern Search System)¹³⁾の開発を進めて来た。COMPASS では三次元共通幾何パターンの探索に対し、次のような条件、あるいは要求を前提とした。

- 1) 与えられた分子の三次元座標は rigid なものと見なす。
- 2) 分子を三次元直交座標空間における点の集合として扱う。
- 3) 各点の重み付けにより原子種・荷電情報等を考慮可能とする。
- 4) 対掌体幾何の識別
- 5) 三分子以上の同時探索

4.1. アルゴリズムの概要 互いに異なる二つの分子の場合を例に、その共通幾何パターンの探索のためのアルゴリズムを以下に述べる (図 8 参照)。それぞれの分子の三次元座標は既知であるとする。今、すべての原子

Fig. 8 Docking graph from molecule M_1 and M_2 .

を等価と仮定すれば、与えられた各分子の構造は三次元座標空間における点の集合と見なすことができる。従って、個々の分子の幾何はそれぞれの点集合に対応する距離行列のみによって規定される (ただし、対掌体幾何を扱う場合には座標値も必要となる)。それらの距離行列は与えられた分子の三次元座標データから容易に求めることができる。ここで、得られた距離行列はその構成原子数に等しい数の頂点を有する辺重み付きの完全グラフを表わすものと考えられる。このことから三次元共通幾何パターンの探索においてもその処理手続きはグラフ理論的取り扱いが可能となる。これら二つの完全グラフに対し、辺重み共通基準によるドッキンググラフ¹⁴⁾を生成する。生成されたドッキンググラフは先の二つの完全グラフ表現において等しい重みを有する辺の情報を含んだ通常の連結あるいは非連結グラフである。従って、このドッキンググラフのクリークを探索することは、二つの完全グラフ間の辺重み付き最大共通部分グラフを探索することと等価であり、その過程で得られるクリーク候補となる共通部分グラフを構成する頂点の集合がここで求める共通幾何パターンの構成原子の集合に対応する。また、最終的に得られたクリークを構成する頂点の集合は最大共通幾何パターンの構成原子の集合に対応づけられ

る。これらのアルゴリズムの群細については文献 13 を参照されたい。

4.2. COMPASS システムと実行例 現在, 開発を進めている COMPASS システムでは入力データとして分子の三次元座標データのみを必要とし, 必要に応じて各原子の電荷や結合表なども入力される。後者は各原子種やその環境の相違による原子の重み付け操作に利用される。COMPASS で利用可能な原子の重み付けの種類を表 2 に示す。charge については許容度を付けることも出来る。これらの処理により, 共通幾何パターン探索の基礎となるドッキング・グラフの生成過程においては距離の基準だけではなく, 原子のタイプあるいは電子的環境も反映させることが出来る。従って, このような処理を加えることにより, 化学的により詳細な意味での三次

Table 2 Characterization of atom types using various weight in COMPASS..

Atomic species:	C, H, O, N, S,
Hybridization:	sp ³ , sp ² , sp
Valence type:	e. g. Nitrogen -NH ₂ , >NH, =NH, -N ⁻ , =N≡N, >N ⁺ , -N ⁺ , -N ⁺ =
Charge:	net charge on each atom (with allowance, δ)

元共通幾何パターン (例えば pharmacophore と呼ばれるような) を探索することが出来る。この他, 指定した部分構造あるいは幾何パターンを必ず含む最大共通幾何パターンの探索も可能である。

COMPASS システムの適用例としてドーパミン拮抗作用を有する二種の分子 (Haloperidide と Milenperone) 間の三次元最大共通幾何パターンの探索を行った。各分子の入力三次元座標は X 線結晶解析にもとづくデータを用いた。図 9 にそれぞれの化合物の構造式と COMPASS の実行結果を computer output で示した。この場合探索条件は距離の許容度を 0.5 Å とし, 原子の重み付けは原子種のみを考慮したものである。図 9 の結果に示されるように, 構成原子数が 12 からなる 4 種の最大共通幾何パターンが得られた。結果に示されているコード番号 (F02, F10) の後につづく数字は構造式中の各原子に付けられた番号を表わす。また, No.1 の結果は Haloperidide の 2, 3,, 32 の 14 個の原子で構成される幾何パターンに対し Milenperone には 1 つの構成原子のみが異なる二組の同様な幾何パターンが存在することを示している。これは距離の許容度の設定に起因するものである。

一方, これらの結果はコンピュータグラフィックスを

EXECUTION =====

```

INPUT-FILE :IN
OUTPUT-FILE:OUT
DATASET    :DATASET
-- SAMPLES --
SAMPLE(1)= F02(Haloperidide) (32)
SAMPLE(2)= F10(Milenperone) (29)
-- MATRIX SELECTION --
DISTANCE MATRIX
ALLOWANCE= .50
-- MUTUAL CHECK --
OFF
-- STEREO CHECK --
OFF
-- NODE WEIGHT --
ATOM TYPE=ON
BOND TYPE=OFF
CHARGE TYPE=OFF
-- NODE EXCLUSION --
EXCLUSION = NOT
-- FUNC. GROUP PRESERVATION --
PRESERVATION = NOT
-- FUNC. GROUP COMPRESSION --
COMPRESSION = NOT
-- RESULT --
ROTATION 1/ 0
NO. 1
(F02: 1) 2 3 4 5 6 7 20 21 23 29 31 32
(F10: 1) 3 2 11 10 9 4 12 20 18 26 23 27
2) 3 2 11 10 9 4 12 20 18 26 29 27
NO. 2
(F02: 1) 2 3 4 5 6 7 20 23 29 30 31 32
(F10: 1) 3 2 11 10 9 4 12 18 26 28 29 27
NO. 3
(F02: 1) 2 6 7 8 17 18 19 20 21 22 26 30
(F10: 1) 24 21 23 18 16 15 20 19 14 13 10 2
NO. 4
(F02: 1) 2 6 8 16 17 18 19 20 21 22 26 30
(F10: 1) 24 21 18 17 16 15 20 19 14 13 10 2

```

NUMBER= 4
SIZE =12

Fig. 9 Maximal common geometrical patterns found by COMPASS for haloperidide and milenperone. In weighting atoms, only different atomic species were considered for this search. The distance allowance is 0.5 Å.

利用した構造重ね合わせ表示により容易に視覚化することが出来る。図 10 は, 図 9 に示された No.1 の最初の結果の, 対応する原子間の最小二乗 fitting による構造重ね合わせ図を示す。

次に, 上記の二つの分子に対し, 各分子中の原子の電

Fig.10 Graphic representation for the first result in Fig. 5. The left view shows the superimposition for the whole molecules excepting hydrogen atoms, and the right view is for the relevant portions.

Fig.11 A maximal common geometrical pattern found by COMPASS. In weighting atoms, different atomic species and atomic net charges were considered for this search. The distance of allowance is 0.5 Å.

(a) Three-dimensional structures of haloperidide (left) and milenperone (right).

(b) Display of the superimposition of their whole molecules and common parts.

(c) Display of the common parts with electronic property.

荷の情報も併せて考慮しながら同様な探索を行った。ここでは電荷は各原子の net charge (N) として表わした。計算は CNDO 法に基づき、各原子の中心座標での点電荷として求めた。

一方、本システムでは charge の許容度指定に対し、比較される原子間での相対的な差に注目した指定（距離の許容度指定の方法と同様な考えにもとづく）と、範囲指定による自動重み付け操作の二つのモードが利用できる。ここでは後者の方法に従って、 $N \leq -0.100$, $-0.100 < N < 0.100$, $0.100 \leq N$ の三段階での重み付け操作を行った。距離の許容度その他は前述の場合と同じ条件で探索された。この例において COMPASS は最終的に、構成原子数 11 からなる 8 種の最大共通幾何パターンを見つけ出した。得られた結果の中で興味深いと思われる幾何パターンの一例を図 11 に示した。両者の *p*-置換ベンゾイル上のフッ素原子、もう一つのベンゼン環およびイミド基上酸素が良く対応づけられることが解る。探索条件より、ここで描き出された二つの分子の共通部分はその空間幾何がよく一致すると同時に対応する各原子毎の電子的な環境も類似したものである。

上に示した例は三次元構造活性相関に対する COMPASS システム応用の可能性を示したものである。もちろん詳細な三次元構造活性相関を議論するためには類似作用を有するより多くの構造的に異なる分子について解析する必要があることは言うまでもない。また、ここでは分子の三次元座標として X 線結晶構造解析の結果にもとづいたデータをすべてそのまま使用しているが、より現実的な構造活性相関解析への適用を試みる場合には、COMPASS での実行に先立って、各分子のコンホメーション解析が要求されよう。これは、受容体との相互作用を想定した場合に各分子のとり得るコンホメーションが必ずしも X 線解析で得られたそれと同じであるとは限らないからである。従って、エネルギー的に安定な複数のコンホマーが存在する場合にはそれらを考慮しながら解析を進める必要がある。しかしながら、一たん合理的な分子の三次元座標が得られれば、それらをもとに COMPASS は三次元薬理作用団の候補パターンを自動的に認識することが可能である。なお、本システムは分子構造が 3 種類以上であっても適用可能であることを書き添えておく。

現在、COMPASS システムに対し、著者らは自由回転結合軸回りの flexibility の問題に対する系統的な取り扱いについて検討中であるが、ここで述べた COMPASS のアプローチは、最近のコンピュータグラフィックス利用による構造重ね合わせ解析とは異なった三次元構造活

性相関研究への新しいツールを提供するものであると確信している。

5. おわりに

化学構造式は全ての化学者にとって共通の言語であり、そこには実に多くの情報が濃縮されている。加えて、三次元分子構造にはその分子が示す様々な特性（生物活性、反応性、種々の物理化学的特性など）に関連した多くの既知あるいは未知の情報が含まれている。そして、その未知の部分についても実験・計算の両面から次第に明らかにされつつある。

ここでは構造活性相関の観点から、薬物間の共通構造特徴の抽出に対するコンピュータの利用について述べたが、こうした化学構造に関する共通性の認識あるいは類似性の検討といった問題は化学研究における様々な問題解決の中で常に付いてまわるものであり、それぞれの問題に対し、重要な知見を与えることも多々ある。しかしながら、本文中でも触れたようにたとえ二次元構造図のレベルであっても比較する化合物の数が増えればそれらの中に共通に存在する部分構造をすべて見出すということは決して容易なわざではない。まして、三次元分子構造を対象とした場合にはさらに多くの困難が生じてくる。一方では、その作業は“直感とひらめき”による場合を除けば、煩雑で退屈が多い。従って、このような仕事を忠実なアシスタントであるコンピュータが受け持ってくれば、研究者はその間に次の一手を考えることも出来よう。

化学研究においてコンピュータはあくまで単なる道具であり、手段である。言い換えれば分析装置であり、時にはガラス器具そのものでさえある。人がコンピュータに入力データを読ませ、得られた出力を化学の研究に役立たせていくことは、ガラス器具に試薬を注ぎ、生じた物質を解析しながら研究を推し進めるのと全く同じことなのである。そしてこの両者が相補的に働くとき化学は一層の発展が約束される。コンピュータが必要なことは判っている。そして内心使ってみたいと思っている。それにもかかわらず、使うための最初の努力を惜しんでそれを遠巻きにしながら「あんなもの無くて……」というのであればこれは情けない話である。これまでも化学者自身の手によって多くのコンピュータソフトウェアが作成されている。そしてその殆んどは、個々の研究者が自らの問題解決に必要な道具として作成したものである。従ってそれらは自らの実験法や器具や装置を工夫すると何等変わるものではない。

以上、本稿が「コンピュータはどうも……」とお考え

の読者のコンピュータに対するバリアを少しでも低くするものであれば望外の幸せである。

(昭和62年7月22日受理)

文 献

- 1) 構造活性相関懇話会編, 「薬物の構造活性相関」, 化学の領域増刊, 122号, 南江堂, 東京 (1979); 同「薬物の構造活性相関 (II)」, 化学の領域増刊, 136号, 南江堂, 東京 (1982), に多くの解説と引用文献有り。
- 2a) A. J. Stuper, W. E. Brügger, P. C. Jurs, "Computer Assisted Studies of Chemical Structure and Biological Function", John Wiley & Sons, New York (1979)
- b) 佐々木慎一, 阿部英次, 高橋由雅, 高山千代蔵, 宮下芳勝, 「化学者のためのパターン認識序説」, 東京化学同人, 東京 (1984)
- 3) 高橋由雅, 佐々木慎一, ファインケミカル, **15**, No. 1, p. 48 (1986)
- 4) R. D. Cramer, G. Redl, C. E. Berkhoff, *J. Med. Chem.*, **17**, 533 (1974)
- 5) 森田 桂, 岡 良和, 化学増刊, **79**, 141 (1979); 岡 良和, ファルマシアレビュー, **9**, 123 (1982)
- 6) Y. Takahashi, Y. Satoh, H. Suzuki, S. Sasaki, *Anal. Sci.*, **3**, 23 (1987)
- 7) Y. Takahashi, Y. Satoh, H. Suzuki, H. Abe, S. Sasaki, *ibid.*, **2**, 321 (1986)
- 8) Y. Kudo, S. Sasaki, *J. Chem. Doc.*, **14**, 200 (1974)
- 9) J. B. Hendrickson, A. G. Toczko, *J. Chem. Inf. Comput. Sci.*, **23**, 171 (1983)
- 10) N. C. Cohen, "Computer Assisted Drug Design", E. C. Olson, R. E. Christoffersen, Eds., American Chemical Society, Washington, DC, 1979, p. 377
- 11) P. Gund, "Progress in Molecular and Subcellular Biology", vol. 5, F. E. Hahn, Ed., Springer-Verlag, New York (1977), p. 117
- 12) G. R. Marshall, C. D. Barry, H. E. Bosshard, R. A. Dammkoehler, D. A. Dunn, "Computer Assisted Drug Design", E. C. Olson, R. E. Christoffersen Eds., American Chemical Society, Washington, DC (1979), p. 205
- 13) Y. Takahashi, S. Maeda, S. Sasaki, *Anal. Chim. Acta* (in press)
- 14) F. S. Kuhl, G. M. Crippen, D. K. Freisen, *J. Comput. Chem.*, **5**, 24 (1984)

本 誌 広 告 料 金

表 紙 2	1 頁	80,000 円	後 付 1 頁	50,000 円
表 紙 3	1 頁	60,000 円	〃 1/2 頁	30,000 円
表 紙 4	1 頁 (2 色)	100,000 円	綴 込 (1 枚)	100,000 円現物持込
英文目次下	1/2 頁	35,000 円		(62. 4. 改訂実施)

詳細は下記一手取り扱い社へご照会下さい。

(株) 共 栄 通 信 社

本 社 〒 104 東京都中央区銀座 8-2-1 (新田ビル) Tel. 03-572-3381 (代)
大阪支社 〒 530 大阪市北区西天満 3-6-8 (笹屋ビル) Tel. 06-362-6515 (代)