Three-Dimensional Pharmacophoric Pattern Searching

Peter Gund

I. Introduction

The medicinal chemist has long known that modifying a drug's chemical structure may cause modified biologic response. If we accept the view that biologic reactions are subject to normal chemical forces, then the usual physical chemical concepts—e.g., thermodynamic and kinetic theory—should apply. However, early attempts to correlate biologic activity quantitatively with chemical structure, using the Hammett equation or Hückel molecular orbital calculations, met with only limited success (GOODFORD, 1973; PURCELL et al., 1973). It is now clear that this failure was due to the greater complexity of bioreactions compared to "normal" chemical reactions in homogeneous solution.

A. Differences between Bioreactions and Solution Reactions

It has proved difficult to find a consistent and valid measure of biologic action, due to the variable response of individual test subjects, and the problem of devising tests in which the measured response is simply and unequivocally related to drug dose (GOURLEY, 1970; GOODFORD, 1973). Drug may be transported to site of action by passive diffusion and partitioning through various membranes and fluids, by active (energy-coupled) transport, or by facilitated transport or selective diffusion through a channel (GOURLEY, 1970; WILBRANDT and ROSENBERG, 1961). Drug may bind to plasma or "false receptors"; it may be rapidly excreted or inactivated; or it may encounter such mysterious obstacles as the blood-brain barrier (OLDENDORF, 1974). Metabolic transformation (McMAHON, 1970) may increase as well as decrease drug potency (SINKULA and YALKOVSKY, 1975).

Biologic processes generally are enzyme-mediated and are subject to greater stereochemical constraints on reagent structure than the corresponding solution reactions (CASY, 1970; HANSON and ROSE, 1975). Although solution reactions usually may proceed to completion (thermodynamic equilibrium), total bioequilibrium is reached only after the organism has died. In order to maintain the chemical driving force necessary for life itself, bioreactions exhibit extremely complex kinetics, involving feedback control loops and profoundly interdependent reaction pathways (CITRI, 1973; SEGAL, 1973). Nature appears to go to great lengths to make biochemical processes discontinuous; a cell process (e.g., mitosis, muscle contraction, nerve stimulus

transmission) tends to be in an "off" state until "turned on" by what may be viewed as a "pattern recognition" response to an activating molecule or stimulus.

Despite all these complications—and more—in biologic processes, quantitative structure-activity relationships (QSARs) are often found for a family of drugs.

II. Quantitative Structure-Activity Relationships

QSARs are derived by means of a number of statistical techniques, which may broadly be categorized as multiple regression (MR) or pattern recognition (PR) analysis.

A. Multiple Regression (MR) Analysis

This well-founded empirical procedure applies standard regression techniques to the problem of correlating activity with several possible structural parameters (REDL et al., 1974; HANSCH, 1973; PURCELL et al., 1973; GOODFORD, 1973). Three major subcategories are *de novo* correlations, linear free energy relationships, and substructural analysis.

The FREE-WILSON (1964) *de novo* analysis assumes that substituents on a drug parent system influence activity independently of each other. In this procedure, a substituent (say, *para*-chloro) is assigned a contribution of 1 to the activity when it is present, 0 when absent. The resulting simultaneous equations (equated to observed activity) are solved by least squares techniques, and the derived coefficient indicates whether a given substituent enhances or hinders activity. Where correlations are found, the method offers the possibility of predicting activity of compounds with all combinations and permutations of the given substituents.

The linear free energy relationship (LFER) approach (WELLS, 1968; SHORTER, 1973) assigns known physicochemical parameters to the substituents and again solves simultaneous equations. This approach has been utilized extensively to relate the rate and equilibrium of solution reactions to chemical structure since it was introduced by HAMMETT (1937). TAFT and LEWIS (1959) reformulated LFER's as a multiparameter equation, and HANSCH et al. (1962) was able to extend the method successfully to biologic reactions. The empirical technique is based on the "extrathermodynamic" assumption—i.e., the assumption is not justified by thermodynamic theory (LEFFLER and GRUNWALD, 1963)—that the free energy of a reaction can be partitioned into a linear sum of individual energetic contributions. Although many variations of the method have been published, and different research groups have "favorite" substituent constants that they try to correlate, the normal analysis includes an electronic term, a partitioning term, and perhaps a steric term (HANSCH et al., 1973a; NORRINGTON et al., 1975).

The electronic term is traditionally HAMMETT's sigma substituent constant (JOHNSON, 1973; SHORTER, 1973), although this is often subdivided into field and resonance terms [$\mathcal{F}$ and $\mathcal{R}$ in the SWAIN-LUPTON (1968) approach]. Partition coefficient is usually represented by the HANSCH pi constant (FUJITA et al., 1964), often with a squared term added to reflect the observed hyperbolic relationship of activity and hydrophobicity (extremely lipid soluble or extremely water soluble compounds have difficulty traversing all biophases to reach site of action: HANSCH, 1969).

Expressing the steric effect in an LFER has long been a problem (CAVALLITO, 1973; SHORTER, 1972; LIEN, 1969; NORRINGTON et al., 1975; SIMON, 1974). Solution reactions usually have been correlated with TAFT's steric constant E_s, but van der Waals radius (CHARTON, 1975), molar refractivity and molecular weight (HANSCH et al., 1973a; NORRINGTON et al., 1975), orientation constants (ÖTVÖS et al., 1975), topologic indices (DUBOIS et al., 1973), calculated congestion (WIPKE and GUND, unpublished results), principal molecular axes or moments of gyration (ZANDER and JURS, 1975), three-dimensional Fourier transform (DIERDORF and KOWALSKI, 1974), and model hydrocarbon strain calculations (DETAR, 1974a,b)—to name just a few recent examples—have also been used as measures of steric environment. Nevertheless, correlation of bioactivity with steric effects has been accomplished with only a few of these indices, and in a limited number of cases (e.g., HANSCH, 1973). This is perhaps not surprising, given the strong influence of reagent geometry on enzymic reactions (HANSON and ROSE, 1975).

Recently, substructural analysis has been added to the arsenal of MR techniques. CRAMER et al. (1974) were partially successful in correlating drug activity with molecular substructure as represented by a fragment code. In a similar approach, ADAMSON and BUSH (1974) correlated antibacterial activity of 79 penicillins with substructural fragments. This technique has also been used recently in an attempt to elucidate chemical structure from ^{13}C nmr spectra (BREMSER et al., 1975). In an attempt to correlate activity with three-dimensional steric environment, SIMON and SZABADAI (1973) defined a minimum steric difference (MSD) parameter essentially as the number of main row atoms remaining when a model was constructed to overlap as many atoms of the reference compound as possible. The correlation of MSD with activity, again, was encouraging but not overwhelming.

B. Pattern Recognition Analysis

The rapidly expanding field of artificial intelligence includes pattern recognition (PR) techniques (JURS and ISENHOUR, 1975; KOWALSKI, 1974; KOWALSKI and BENDER, 1975) such as learning machines, k-nearest neighbor calculations, discriminant analysis, and many others. PR techniques have been used to correlate biologic activity with mass spectral fragmentation patterns (TING et al., 1973) and with molecular fragments. Among others, HILLER et al. (1973), STUPER and JURS (1975), CHU (1974), CHU et al. (1975), and KOWALSKI and BENDER (1974) have "recognized" molecular patterns that appear to be associated with

bioactivity. While PR techniques may be used with any numeric information—including physicochemical parameters—it has largely been used to date to correlate bioactivity with molecular fragement information.

C. QSAR Generalizations

MR techniques hinge on the tenuous assumption that variation of chemical structure with bioactivity is describable by a smooth, continuous curve. While this is doubtless not totally true, the method works surprisingly well in many and varied cases for predicting optimally active examples of a homogeneous set of compounds. Nevertheless, an analog containing a substituent with different steric properties, which is metabolized, or which changes site of binding can fail to correlate.

PR techniques are capable of handling activity discontinuities and inactive analogs. In principle, they can correlate activity among highly heterogeneous sets of compounds. In practice, however, the method appears to contain pitfalls, and published analyses have been the subject of controversy. For example, the correlation of bioactivity with mass spectral data (TING et al., 1973) has been called trivial (PERRIN, 1974) and absurd (CLERC et al., 1973); and a correlation of anticancer data with structural fragments (KOWALSKI and BENDER, 1974) has been criticized as valid only for the highly biased set of compounds used in the analysis (MATHEWS, 1975). It has been said that there is a danger with PR techniques of correlating insignificant information (KENT and GAUMANN, 1975). PR results seem to be highly dependent upon patterns chosen for input; and since a very large number of patterns are generally possible, a fair amount of subjective judgment enters into the analysis. Further problems include pattern overlap and redundancy and relative weighting of patterns. Nevertheless, PR has the capability of predicting new classes of drugs with a desired activity, and deserves to be pursued. It is likely that the most satisfactory results will be attained when fairly large substructures are used to provide good discrimination.

Both MR and PR techniques are parametric methods that are subject to dataset bias and to choice of independent parameters. Furthermore, the specification of stereochemical and spatial information in a general way has proved difficult with either technique. The remainder of this chapter develops the concept of a pharmacophoric pattern as a representation of structural information that might be correlated with bioactivity by either MR or PR methods.

III. Description of Pharmacophoric Patterns

We define a *pharmacophoric pattern* as *a collection of atoms spatially disposed in a manner that elicits a biologic response*. The importance of spatial configuration is a consequence of the drug-receptor theory

of biologic action, which is discussed below. Since both topologic and topographic patterns have been discussed in the literature, we will consider them separately. Historically, the term "pharmacophore" was attributed to EHRLICH (1909) by ARIENS (1966), and the distinction between topologic and topographic patterns has been recognized by KAUFMAN and KERMAN (1974), among others.

A. Topologic Molecular Patterns

Topologic patterns are independent of molecular conformation, but dependent upon molecular configuration (positional isomerism). They are specified in terms of atom type and connectivity, as contained in a connection table representation of structure (LYNCH et al., 1971). They correspond to molecular fragments or substructures. They have been defined "graph theoretically" as atom-centered or bond-centered fragments (LYNCH, 1974), or as "environments" (DUBOIS, 1974). They have been defined chemically in terms of functional groups, rings, side chains, and Wiswesser Line Notation symbols (EAKEN and HYDE, 1974; FELDMAN, 1974). These types of patterns have been used extensively in pattern recognition studies of influence of molecular structure on bioactivity.

B. Topographic Molecular Patterns

Topographic patterns are geometry dependent, conformation dependent, and specified by atom type and position in space. Their importance for bioactivity is a consequence of the drug-receptor theory of biologic action (CAVALLITO, 1973; KOROLKOVAS, 1970).

In essence, receptor theory states that a drug must interact chemically with an appropriate site on a biologic macromolecule in order to exert its effect. ("Drug" is used throughout in its larger sense, as any chemical—synthetic or natural—that causes a biologic effect.) As the drug complexes with the receptor site, a biologic response (perhaps due to a macromolecular conformational change) occurs, and the drug is released, allowing the biomolecule to become primed for the next activating episode. Drug action may be categorized as agonist—eliciting a normal biologic response from the receptor—and antagonist—inhibiting normal response by the receptor. Antagonist action may occur by strong or irreversible complexing to the receptor site, or by binding to some other site on the biomolecule, so normal binding to the distorted receptor site is not possible (allosteric inhibition: GRAY, 1971).

C. Pharmacophoric Patterns: Topologic or Topographic?

Bioreceptors are believed to possess a specific geometric structure. Therefore, the receptor site pattern, and the complementary pharmacophoric pattern, are topographic in nature. In recognition of this principle, CROXATTO and HUIDOBRO (1956) molded receptor sites by "surface complementarity" to space-filling models of active drugs in order to predict bioactivity.

Nevertheless, a number of topologic pharmacophoric patterns have been proposed in the literature and have often proved useful in classifying drugs of similar activities. *Topologic patterns will succeed in describing bioactivity insofar as they correspond to the actual (topographic) pharmacophoric pattern.* Thus, patterns involving unsaturated rings or constrained rings possess relatively few degrees of conformational freedom, and the geometric structure is often inherent in the structural diagram. Chiral topologic patterns contain additional spatial information. Furthermore, distance between functional groups (topography) may be approximately represented as a number of single bonds (topology). For example, JANSSEN (1973) proposed that the basic topologic pattern shown in 1, where X may be several different atom types,

$$Ph\text{-}X\text{-}X\text{-}X\text{-}X\text{-}NR_2$$

1

may be associated with antipsychotic activity. This is clearly a simplification, since many compounds containing that substructure are not neuroleptics. JANSSEN (1970) proposed an S-shaped topography for this pattern, and quantum mechanical calculations (KAUFMAN and KERMAN, 1974; KAUFMAN and KOSKI, 1975) and classical calculations (FEINBERG and SNYDER, 1975) have shown that these drugs do tend to assume certain well-defined conformations. Furthermore, in the preferred conformation, at least some of them may be superimposed on the preferred conformation of dopamine, rationalizing their competitive binding to the same receptor (FEINBERG and SNYDER, 1975). The dynamics of conformational change in these drugs has also been related to antipsychotic activity (FENNER, 1974).

The presence of a topographic pharmacophoric pattern is, of course, no guarantee of activity, since transport, metabolism, and steric problems may interfere. The further complication of molecular flexibility will be considered below. Nonetheless, it appears that the most discriminating measure of potential bioactivity will be an appropriate topographic pattern.

D. Functional Definition of Pharmacophoric Pattern

A pharmacophoric pattern may be specified by a collection of atoms and distances between them. Heteroatoms and unsaturated carbon atoms will most often be involved in pattern definition, with hydrogens usually not explicitly considered if heteroatoms are labeled as either hydrogen bond donating or accepting. Of the forces involved in drug-receptor binding—coulombic, dipole, hydrogen bond, charge transfer, hydrophobic, and van der Waals, in order of decreasing energy (KOROLKOVAS, 1970)—only the last two are important for saturated carbon. Computer representation of pharmacophoric patterns is discussed in a later section.

A pharmacophoric pattern may be *chiral*, except for a planar pattern, which may be *oriented*. We define pattern chirality in a manner similar

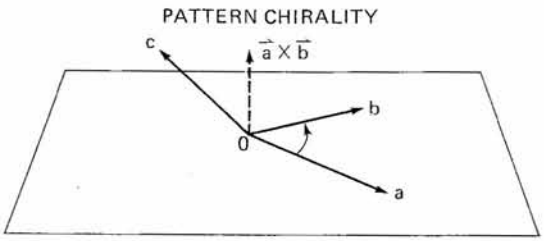


Fig. 1. Definition of pattern chirality

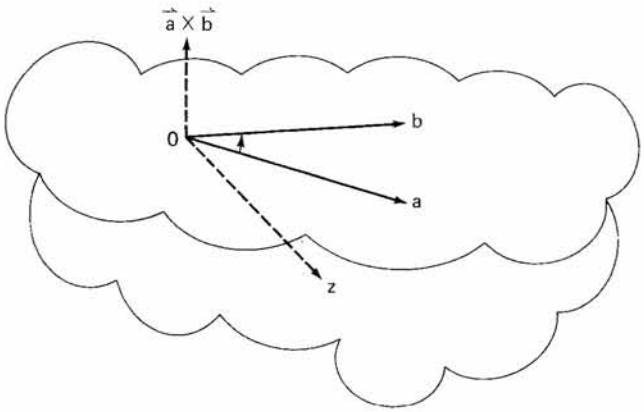


Fig. 2. Definition of oriented facial pattern. z = center of gravity of the molecule containing the oriented pattern

to that used for computer determination of atom chirality (WIPKE and DYOTT, 1974). For a pattern with its origin at atom 1, construct vectors $\vec{a}$ and $\vec{b}$ to atoms 2 and 3 (Fig. 1). The cross product $\vec{a} \times \vec{b}$ defines $\vec{a}$ perpendicular direction. Vector $\vec{c}$ drawn to atom 4 then has the same sign as $\vec{a} \times \vec{b}$ (for parity 0) or opposite sign (for parity 1). If atom 4 is in the plane of atoms 1-3, then it is replaced by the first out-of-plane atom for determining chirality. The orienting of a planar pattern is determined in a similar manner; again, the first three atoms define a direction $\vec{a} \times \vec{b}$, and that direction should point away from the center of gravity of a molecule containing an oriented pattern (Fig. 2). Prochirality of a planar atom is treated similarly in the re/si nomenclature (HANSON, 1966; GOODWIN, 1973).

A pharmacophoric pattern has the property of *accessibility*. Accessibility may be total, as in cations K^+ and NH_4^+. It may be bifacial, as in ethidium bromide and other DNA intercalating drugs (WARING, 1970). It may be facial, or facial oriented, as in tetrodotoxin (MULLINS, 1973). It may be nodal, as in detergents. It may even be inaccessible; for example, an internal amino acid residue in a globular protein. Operational definitions of the various accessibility classes are under development. The related concept of atom congestion has been defined (WIPKE and GUND, 1974). Alternatively, accessibility may be specified as part of the pattern by defining a negative mass atom and requiring that the test molecule have no mass at that pattern position.

Finally, a pattern may be *indefinite*. There may be atom indefiniteness, since often more than one atom type (e.g., NH_2, OH) may be

tolerated at the same position on the receptor surface. This might be specified by indefinite atom types (e.g., D for electron donor atom), or as a fractional electronic charge. An abstraction of this idea defines an interaction pharmacophore as an electrostatic potential field (WEINSTEIN et al., 1973). Also, a functional group (e.g., carboxyl) might be specified as a pattern "superatom," simplifying the number of distance matches to be made. A pattern may also have distance indefiniteness; for example, ZEE-CHENG and CHENG's (1970) antileukemic pattern was reported with distance error limits. The result is a fairly indefinite pattern; as illustrated in Figure 3, the third atom may appear anywhere in the shaded area and still fulfill the defined pattern requirements. Since bonded atoms are relatively close together, they require smaller distance tolerances.

A pattern may also be indefinite because of conformational changes of the drug. For example, the α-adrenergic receptor (KOROLKOVAS, 1970)

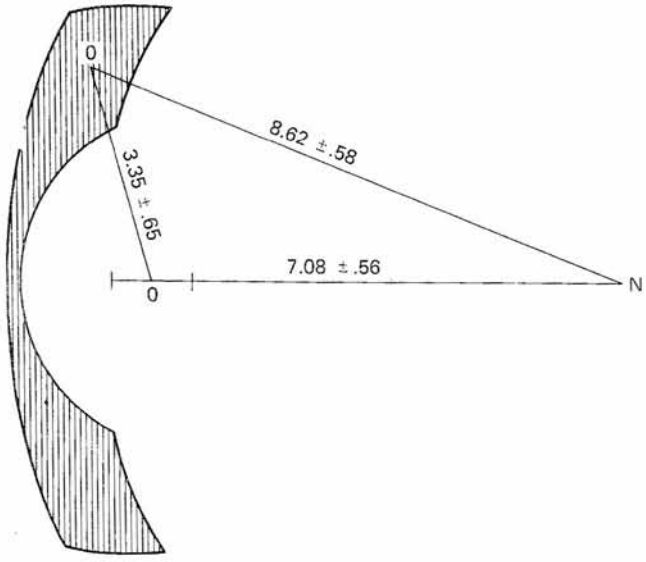


Fig. 3. Error limits for antileukemic triangular pattern (ZEE-CHENG and CHENG, 1970)

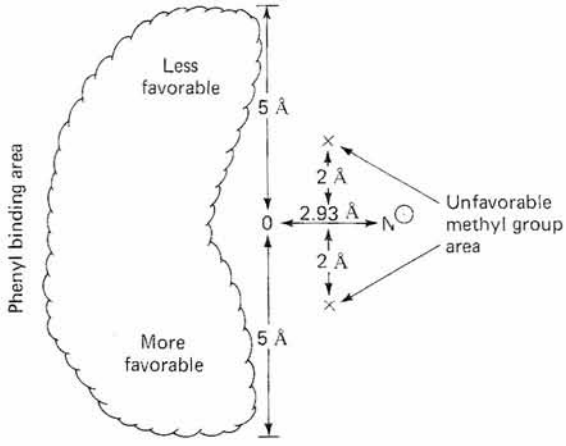


Fig. 4. Proposed α-adrenergic receptor site topology. X = positions where bulk in the drug is not tolerated. From KIER (1968)

has been formulated with a rather large area for binding of a phenyl ring of active drugs (Fig. 4). This may be accommodated in the pattern by assigning large distance error limits, or rather inelegantly by specifying two pattern distances rather than three. Of course the latter will allow more patterns than desired, but at least all proper patterns will be found. Disallowing conformations with mass at positions marked by X in Fig. 4 would further restrict the number of acceptable patterns. These techniques might also prove useful in cases where the active conformation of a flexible drug was not known.

IV. Three-Dimensional Pharmacophoric Pattern Searching

A. Derivation of Pharmacophoric Patterns

Topologic patterns have typically been proposed as a consequence of structure-activity investigations. Topographic patterns have been derived from mapping of a topologic pattern onto a rigid drug analog, or from determining the preferred conformation of a flexible molecule by theoretical or experimental means. Confidence is gained in the reality of such a pattern by finding it in several different drug structures exhibiting similar bioactivity. A long-term objective is to develop an automatic pattern recognizer that finds patterns common to several drugs possessing the same mode of action. Our initial aim, however, was more modest: to search for presupposed patterns to determine their worth. Some patterns already proposed in the literature are collected in Table 1; where available, references are to reviews.

B. Derivation of Drug Molecule Structures

Molecular structures are available from modeling and from experimental studies. About 5000 molecular structures are available from the Cambridge Crystallographic Data File (ALLEN et al., 1973). Transforming unit cell coordinates to cartesian coordinates is trivial (computer program available from the author). The crystallographic method suffers from limitations, however, in that it is a relatively long and expensive procedure performed by experts on specialized equipment; the crystal structure may correspond neither to the solution conformation nor the bioactive form; and crystal structures are not available for many drugs, all liquids and amorphous powders, and nearly all hypothesized new drugs. Structural information is available from a variety of other experimental methods, including electron diffraction, proton and ^{13}C nmr, and lanthanide shift experiments (BARRY et al., 1974).

Molecular coordinates may be derived from measurements on a Dreiding model by using various computer programs that create molecules from standard fragments (such as programs no. 94, 130, 135, 136, 169, 178, 186, and 226 from Quantum Chemistry Program Exchange, Indiana University, Bloomington, Ind.), or by using computer graphics programs that create a structure from an input two-dimensional structural diagram

Table 1. Some proposed pharmacophoric and receptor patterns

Bioactivity	Pattern	References
Local anesthetic	receptor map	KOROLKOVAS, 1970
Muscarinic	receptor map	KOROLKOVAS, 1970;
	pharmacophore	KIER, 1971; KIER, 1973;
Acetylcholinesterase	receptor map	KOROLKOVAS, 1970
Analgesic	receptor map	KOROLKOVAS, 1970; CASY, 1973 PORTOGHESE and WILLIAMS, 1969 TAKEMORI, 1974
Anti-inflammatory	receptor map	KOROLKOVAS, 1970; KIER, 1971
Nicotinic	two-point pattern	KOROLKOVAS, 1970; KIER, 1971; KIER, 1973
Anticholinergic	receptor map	KOROLKOVAS, 1970;
	pharmacophore	MAAYANI et al., 1973; WEINSTEIN et al., 1973; PAULING, 1975
α-,β-Adrenergic	receptor map	KOROLKOVAS, 1970; KIER, 1971; KIER, 1973
	pharmacophore	PULLMAN et al., 1972; COUBEILS et al., 1972; PATIL et al., 1975
Histaminic	receptor map	KOROLKOVAS, 1970
	pharmacophore	KIER, 1971; KIER, 1973
Antihistaminic	receptor surface complement	CROXATTO and HUIDOBRO, 1956
Serotoninic	pharmacophore	KOROLKOVAS, 1970; KIER, 1971; KIER, 1973; KELLY and ADAMSON, 1973
MAO Inhibitory	receptor map	KOROLKOVAS, 1970
Neuroleptic	topological	JANSSEN, 1973
	pharmacophore	KOROLKOVAS, 1970; FEINBERG and SNYDER, 1975
Hallucinogenic	pharmacophore	KANG et al., 1973; GREEN et al., 1973
Convulsant	receptor model	SMYTHIES, 1974
Anticonvulsant	pharmacophore	CAMERMAN and CAMERMAN, 1970
Steroid hormonal	receptor maps	KOROLKOVAS, 1970; KIER, 1971; KIER, 1973; CRENSHAW et al., 1974
Taste	receptor map	BEETS, 1973; GUILD, 1972; KIER, 1973
Antileukemic	triangle	ZEE-CHENG and CHENG, 1970; ZEE-CHENG and CHENG, 1973; ZEE-CHENG et al., 1974
Antipeptic	pharmacophore	BUSTARD and MARTIN, 1972
Hypertensive, hypotensive	receptor surface complements	CROXATTO and HUIDOBRO, 1956
Antimalarial	triangle	CHENG, 1974

(WIPKE, 1974; STILL and LEWIS, 1974; BOLT BERANEK and NEWMAN, 1973; ZANDER and JURS, 1975; DIERDORF and KOWALSKI, 1974). Computerized molecular modeling systems are becoming commonplace in universities (review: MARSHALL et al., 1974), government laboratories (FELDMANN et

al., 1973), and industry [HODGES and NORDBY (Searle), 1975; GUND, ANDOSE and RHODES (Merck, Sharp and Dohme), unpublished].

Several theoretical methods are available for predicting molecular structure (GOLEBIEWSKY and PARCZEWSKI, 1974). In particular, molecular mechanics calculations (WILLIAMS et al., 1968; ENGLER et al., 1973; HOPFINGER, 1973), quantum mechanics methods (POPLE and BEVERIDGE, 1970; KIER, 1970, 1971, 1973; GREEN et al., 1974), and combinations (ALLINGER and SPRAGUE, 1973; WARSHEL and KARPLUS, 1974) have been used. Such calculations, by "conformation mapping," can indicate which structures are most likely to occur (MARSHALL et al., 1974). The magnitude of the current research effort devoted to determining three-dimensional structure of biologically active molecules is apparent from the Proceedings of two recent symposia on the subject (BERGMANN and PULLMAN, 1973, 1974).

The flexibility of many drugs is a worrisome complication. When several conformational states are energetically accessible, then different conformations may prevail in vacuo (the state treated by most theoretical calculations), in solution, in the crystal, and on the receptor surface (GREEN et al., 1974). In fact, it appears that many effector substances (agonists) act by undergoing one or more conformational changes on the receptor surface, while rigid substances are more likely to be antagonists. Indeed, BURGEN et al. (1975) have proposed the "zipper" model for drug-receptor interaction, in which the drug forms a partial "nucleation" complex with the receptor, followed by conformational changes in both drug and receptor to give stronger binding. This model permits fast and selective binding of substrates, and rapid conformational transformation of the biomolecule—which may be required to induce a bioaction.

What, then, is the point of examining molecular models of the non-receptor bound conformation? Basically, because there must be some pattern recognition of the uncomplexed molecule, or there would be no driving force for complexation. BUSTARD and MARTIN (1972) have argued that since according to collision theory the rate of complexation varies with the concentration of the active species, optimal activity will occur for a drug that reacts in its ground state conformation. KIER (1973) has proposed the concept of "remote recognition of preferred conformation," and KIER and HOLTJE (1975) have reported calculations that support the notion that there can be long-range attractive interaction with a drug's ground state conformation.

Our short term solution to the drug flexibility problem is to collect all energetically accessible conformations and pattern search on each one separately. GREEN et al. (1974) recommend considering all conformations within 6 kcal/mole of the global minimum. Alternatively, partially defined patterns (see above) could be used to minimize the importance of a specific conformation. In yet another approach, CHENEY (1974) and SUNDARAM (1975) have used optimization methods to constrain a flexible molecule to superpose upon a more rigid molecule.

C. Finding a Pattern Match

Matching patterns is analogous to the well-known chemical information problem of matching a molecular substructure (fragment). A practical substructure search system generally matches on screens (previously defined fragments or other chemical information) to reduce the problem to a manageable size, then searches for an atom-by-atom match on the reduced set of structures (LYNCH et al., 1971). Similarly, we may save some computer time by counting atom types to assure that the pattern may occur in the present molecule; more sophisticated screening is also possible. We finally need to perform an atom-by-atom match, comparing distances rather than bonds as we go.

Since there are $n(n\text{-}1)$ interatomic distances in a molecule of n atoms, any reduction in the required number of distance calculations would speed the search. We define the minimum number of distances required for a preliminary match as $4(n\text{-}3)+2$ for patterns of 4 atoms or more. An informal derivation is given here.

A pattern of two loci is uniquely defined by the simple distance between them (line). A three loci pattern is specified by three distances (triangle), and a four loci pattern by six distances (sides of a tetrahedron). Every additional atom is then uniquely located in space by its distances to the first four atoms, provided that the first four are not all in the same plane. (If they are, the first out-of-plane atom is taken as the fourth basis locus.)

Once a preliminary match is found, it may be checked for closer match. Initial matching occurs with a rather wide distance error limit (typically 1.5 Å, except for bonded atoms). This may be refined by comparing distances to a closer tolerance. Other criteria—e.g., stereochemistry, accessibility—may also be checked at this point.

If a true match is found, then the "goodness of match" should be determined. We define a *similarity index* straightforwardly as the root mean square (r.m.s.) deviation of *all* interatomic distances ($n(n\text{-}1)$). Similarly, r.m.s. distance differences have been used as a measure of similarity of isomorphous regions of protein crystal structures (DREUTH et al., 1972; RAO and ROSSMANN, 1973; ROSSMAN et al., 1974; NISHIKAWA and OOI, 1974).

Reducing goodness of pattern match to a single number is convenient, and the similarity index may prove to be a useful measure of steric fit in MR and perhaps PR analyses. However, the twin problems of agonist-antagonist activity (too good a fit may turn an agonist drug into an antagonist) and drug flexibility (a pattern will fit some conformations of a drug better than others) require that this new index be applied with caution.

D. Computer Program MOLPAT

The preceding concepts were implemented in a computer program MOLPAT (GUND et al., 1974). This section should be considered a progress

report, since program testing and refinement are incomplete. Patterns were obtained from the literature, or by abstracting the functional atoms from a rigid prototype drug structure. A pattern data file consists of a list of coordinate positions and atom types, and atom-atom combinations for which distances are to be computed. Alternatively, no distances are specified and $4(n-3)+2$ distances are automatically calculated at runtime. Finally, if no coordinate information is given, then distances must appear in the file explicitly. This flexible pattern data format (Fig. 5) allows various descriptions of patterns and partial patterns and has the additional advantage that a pattern may be manipulated and displayed as if it were a molecule.

Drug structures were taken from published crystal determinations or were created from two-dimensional graphically input structural diagrams by a strain minimization program. Molecule file data format (Fig. 6) was similar to pattern data, except bonds and their types are given rather than distances.

In practice, the chemist interactively specifies a pattern and a drug molecule, and the program searches for an occurrence of that pattern in the molecule. If one is found, the occurrence is displayed on a CRT display and match statistics printed. The program is then instructed to search for more occurrences of the pattern in the present molecule, input another molecule, or input another pattern. Flow diagrams outlining the program and pattern search logic are given in Fig. 7 and 8. Although the program is tested and debugged, it cannot be considered to be an effective research tool until a number of

Fig. 5. Pharmacophoric pattern data format

```
ANTILEUKEMIC  PHARMACOPHORE                    title
3   3                                          na  nb  ptype
-   .92    3.22     Ø.     0                   x, y, z coords.,
Ø.         Ø.       Ø.     0                     atom type
7.Ø8       Ø.       Ø.     N
1   2                                          atoms, dist.
1   3
2   3
```

Fig. 6. Molecule data format

```
MORPHINE, X-RAY STRUCTURE
21   25
-2.606  -1.343   1.352   C
-2.352  -2.803   1.423   C
    :
    :
1    2    2
2    3    1
    :
    :
```

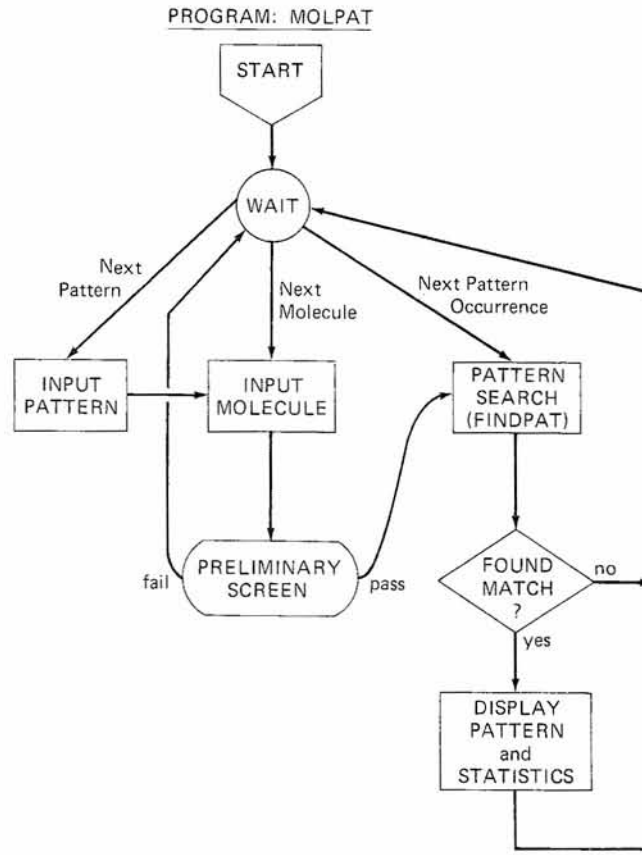


Fig. 7. Flow diagram: Program MOLPAT

program improvements have been implemented—primarily, automatic searching through a file of many drug structures.

The importance of computer graphics for this research should be emphasized (review: MARSHALL et al., 1974). The PDP10/LDS1 facility of the Princeton Computer Graphics Laboratory was utilized for inputting drug structural diagrams and controlling the molecular three-dimensional structure generation, for displaying pharmacophoric patterns in drugs, and for superposing molecules containing similar patterns. While man is the best pattern recognizer, he is best at perceiving two-dimensional patterns. Computer graphics has proven quite helpful in adding a new dimension (the third dimension) to the chemist's perception of structure activity patterns.

V. Pattern Matching Applications

A. Antileukemic Pattern

ZEE-CHENG and CHENG (1970) proposed, from the examination of Dreiding models of drugs, that a triangular arrangement of atoms (Fig. 3) could

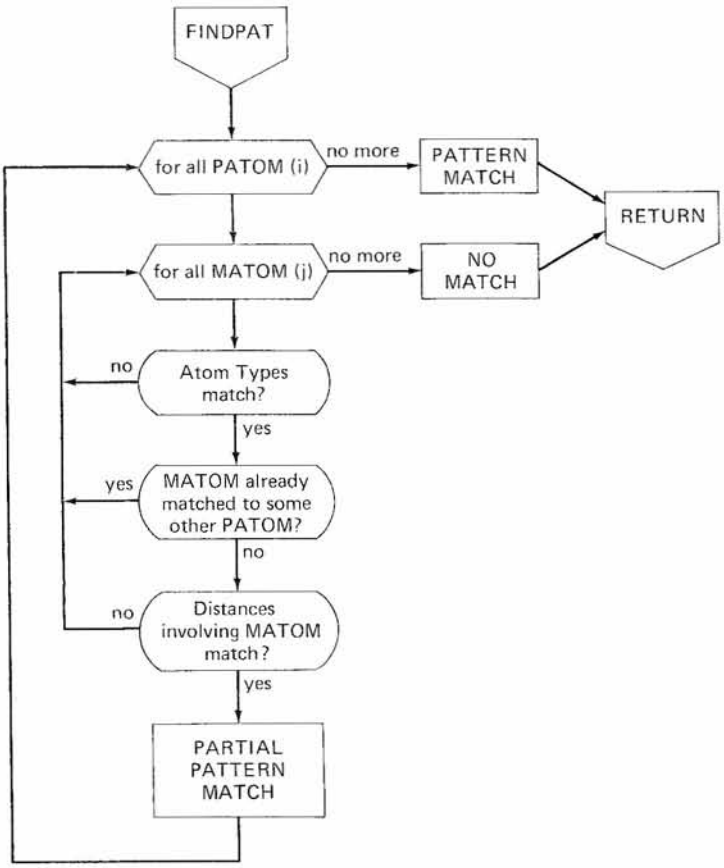
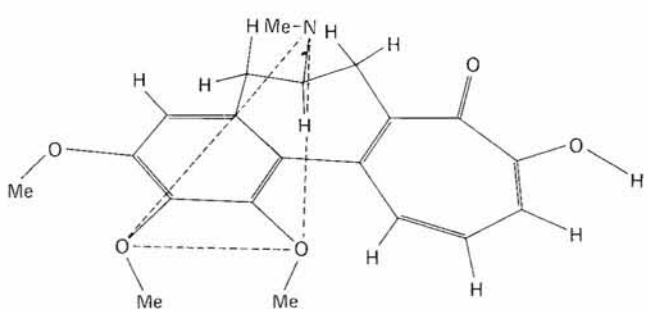


Fig. 8. Flow diagram: Pattern finding subroutine of MOLPAT

be associated with antileukemic activity. They found over 20 antileukemic compounds of various chemical classes that contained the pattern and recently have found others (ZEE-CHENG and CHENG, 1973; ZEE-CHENG et al., 1974). We have modeled many of these structures and have confirmed the presence of the hypothesized pattern in most of them. Figure 9 identifies the pattern in a computer-built model of demecolcine.

Fig. 9. Antileukemic triangular pattern located by MOLPAT in a computer-built model of demecolcine



It is certain that these antileukemic agents do not all have the same mode of action. A subset of these drugs containing a bifacial pattern (tylocrebrine, tylophorine, streptonigrin, demecolcine) may all act by intercalation in DNA (ZEE-CHENG and CHENG, 1973). Another subset consists of known folic acid antagonists (aminopterin, methotrexate). A subset of nucleoside analogs (sangivamycin, 6-mercaptopurine riboside, cytosine arabinoside, 5-azacytidine) may act either by incorporation in DNA or RNA (5-azacytidine: JUROVCIK et al., 1965) or by enzyme inhibition (6-mercaptopurine and its anabolites interact with over 20 enzymes: MONTGOMERY et al., 1970). Other drugs in this group no doubt have yet other modes of action. It would appear that, in order to reduce the number of such redundancies, a proposed pharmacophoric pattern should be as detailed as possible.

B. Analgesic Pattern

An analgesic receptor model has been proposed on the basis of observed drug structure-activity relationships (CASY, 1973; EDDY and MAY, 1973; SIMON, 1973; CLARKE et al., 1974). Its main features include an aromatic ring and a quaternary N^+ atom, separated by an appropriate distance; a hydroxyl group on the aromatic ring enhances activity, but is not essential. I modeled this pharmacophoric pattern by extracting the nitrogen and benzene coordinates from the crystal structure (MACKAY and HODGKIN, 1955) of the prototypical analgesic, morphine.

The resulting pharmacophoric pattern could be located by MOLPAT in models of morphine derivatives, in the crystal structure of the morphine antagonist, 3-hydroxy-N-allylmorphinan (BLOUNT et al., 1973), and, interestingly, in the crystal structure (BAKER et al., 1973) of LSD (Fig. 10). While LSD is reported to give subjects the illusion

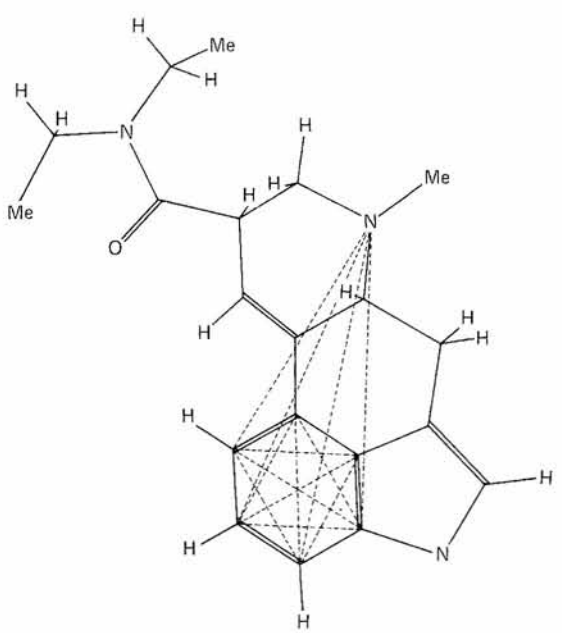


Fig. 10. Analgesic pattern found by MOLPAT in LSD

of increased sensory awareness, it actually reduces tactile sensitivity (EDWARDS and COHEN, 1961); furthermore, LSD has been reported to potentiate response to morphine in a guinea-pig ileum test system (GINTZLER and MUSACCHIO, 1974).

Recently published binding data of analgesics with an isolated rat brain receptor (TERENIUS, 1974) may allow refinement and elaboration of this pharmacophoric pattern, as well as separation of analgesic, narcotic, and narcotic antagonist activity (TAKEMORI, 1974).

C. Prokaryotic Ribosomal Transpeptidase Inhibitors

While not a direct application of MOLPAT, a recent study (HAHN and GUND, 1975) illustrates the power of a three-dimensional model for gaining insight into mechanisms of drug action.

Building on the work of HARRIS and SYMONS (1973), CHENEY (1974) noticed that several antibiotics that inhibit an *in vitro* ribosome transpeptidase system, as well as the system's natural substrates, possess similar *topologic* patterns (Fig. 11). SHIPMAN et al. (1974) used theoretical calculations to find a preferred conformation for the antibiotic, lincomycin. The ground state was confirmed as the active form, since the reaction product of lincomycin and formaldehyde—which has reduced conformational possibilities due to formation of an

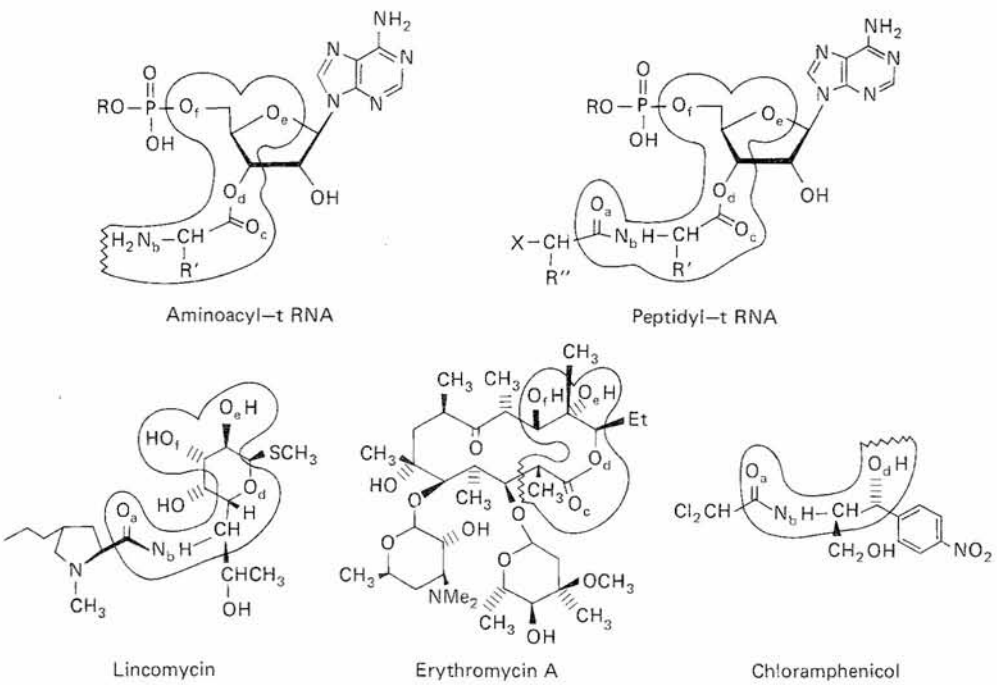


Fig. 11. Topologic similarities in substrates and some inhibitors of peptidyl transferase

additional ring—was also antibacterial (SHIPMAN et al., 1974). By constraining key atoms to their positions in lincomycin and minimizing differences, CHENEY (1974) derived models for aminoacetyl-tRNA and peptidyl-tRNA on the enzyme surface. This effectively defined a topographic pharmacophore for transpeptidase inhibition, if it is accepted that enzymes enhance reaction rates by stabilizing transition state geometry of substrates and that enzyme inhibitors may resemble reactant transition state geometry (LINDQUIST, 1975).

HAHN and GUND (1975) reviewed the large body of information that chemists and biochemists have gathered on chloramphenicol (CM) antibiotic activity in light of CHENEY's (1974) topographic model of transpeptidation inhibition.

HANSCH analyzed CM analog activity in one of his earliest papers on multiple regression analysis (HANSCH et al., 1963); subsequently, at least four other QSAR analyses have been published (CAMMARATA, 1967; HANSCH et al., 1969, 1973b; HÖLTJE and KIER, 1974). Nevertheless, KONO et al. (1969) synthesized some analogs that were as active or more active than CM itself, none of which could have been predicted as being active from the published equations. Another inadequacy of the HANSCH correlations was revealed when comparison of *in vivo* and *in vitro* SAR's (HAHN and GUND, 1975) indicated that steric bulk on the acetyl group gave excellent binding but poor drugs, presumably due to poor facilitated transport. Such a separation of varying steric effects upon binding and upon transport is difficult or impossible to obtain by QSAR techniques. Our three-dimensional structural model also allowed dismissal of HANSCH's hypothesis that CM acts by forming a stable free radical (HANSCH et al., 1969). Superposition of chloramphenicol on CHENEY's model for the arrangement of the natural substrates on the enzyme (Fig. 12) results in a model that accomodates essentially all the SAR data, as well as most of the detailed biochemical data, and that may serve as a framework for the design of novel antibacterial agents.

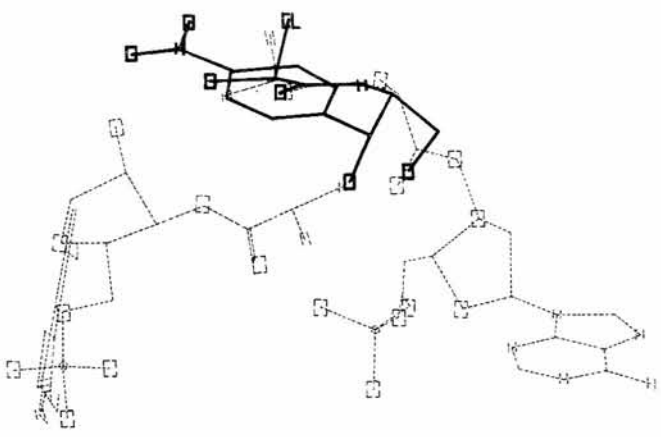


Fig. 12. Model for chloramphenicol inhibition of transpeptidase: Superposition on model of bound substrates (HAHN and GUND, 1975)

VI. Conclusions

The drug-receptor theory of biologic action implies a pattern recognition response of receptor to drug, and that response is to a geometrically well-defined (topographic) pattern. Such pharmacophoric patterns may be defined in terms of atom type and interatomic distances, and, as this work has demonstrated, may be manipulated and matched by computer.

Finding patterns in three-dimensional drug structures is a useful method for testing hypothesized pharmacophores, for predicting activity for new compounds, for designing novel drugs with specific activity, and for proposing and testing mechanisms of drug action. Patterns may be used in multiple regression and pattern recognition analyses of drug activity, either qualitatively as present or absent, or quantitatively via the similarity index. Such patterns appear to provide a satisfactory method of accounting for steric effects in biologic reactions.

Acknowledgments

I am grateful to the National Institutes of Health for a Special Research Fellowship (1971-1973), for the period in which most of this research was accomplished. The Princeton Computer Graphics Laboratory is a Biomolecular Resource supported by the National Institutes of Health and Princeton University. I thank Professor W. T. WIPKE for providing invaluable advice and helpful criticism, and Professor R. LANGRIDGE for his interest and support.

References

ADAMSON, G.W., BUSH, J.A.: Method for relating the structure and properties of chemical compounds. Nature (Lond.) 248, 406 (1974).

ALLEN, F.H., KENNARD, O., MOTHERWELL, W.D.S., TOWN, W.G., WATSON, D.G.: Cambridge crystallographic data centre. II. Structural data file. J. Chem. Doc. 13, 119 (1973).

ALLINGER, N.L., SPRAGUE, J.T.: Calculation of the structures of hydrocarbons containing delocalized electronic systems by the molecular mechanics method. J. Am. Chem. Soc. 95, 3893 (1973).

ARIËNS, E.J.: Molecular pharmacology, a basis for drug design. Progr. Drug Res. 10, 429 (1966).

BAKER, R.W., CHOTHIA, C., PAULING, P., WEBER, H.P.: Molecular structures of hallucinogenic substances: lysergic acid diethylamide, psilocybin and 2,4,5-trimethoxyamphetamine. Mol. Pharmacol. 9, 23 (1973). I thank Prof. Pauling for supplying the molecular coordinates.

BARRY, C.D., GLASEL, J.A., WILLIAMS, R.J.P., XAVIER, A.V.: Quantitative determination of conformations of flexible molecules in solution

using lanthanide ions as nuclear magnetic resonance probes: application to adenosine-5'-monophosphate. J. Mol. Biol. 84, 471 (1974).
BEETS, M.G.J.: Structure-response relationships in chemoreception. In: Structure-Activity Relationships (Ed. C.J. CAVALLITO), vol. I, p. 225. Oxford: Pergamon Press 1973.
BERGMAN, E.D., PULLMAN, B. (Eds.): Conformation of Biological Molecules and Polymers. Proc. Int. Symp., Jerusalem April 1972. Jerusalem: Israel Acad. Sciences and Humanities 1973.
BERGMAN, E.D., PULLMAN, B. (Eds.): Molecular and Quantum Pharmacology. Proc. Int. Symp., Jerusalem April 1974. Dordrecht: Reidel 1974.
BLOUNT, J.F., MOHACSI, E., VANE, F.M., MANNERING, G.J.: Isolation, X-ray analysis and synthesis of a metabolite of (-)-3-hydroxy-N-allylmorphinan. J. Med. Chem. 16, 352 (1973).
BOLT BERANEK and NEWMAN, INC.: PROPHET implementation overview. Report No. 2526, April 1973 to Biotechnology Resource Branch. Div. Res. Resources, Nat. Inst. Health.
BREMSER, W., KLIER, M., MEYER, E.: Mutual assignment of subspectra and substructures—a way to structure elucidation by ^{13}C NMR spectroscopy. Org. Magn. Res. 7, 97 (1975).
BURGEN, A.S.V., ROBERTS, G.C.K., FEENEY, J.: Binding of flexible ligands to macromolecules. Nature (Lond.) 253, 753 (1975).
BUSTARD, T.M., MARTIN, Y.C.: Conformational and structural relationships among antipeptic ulcer compounds. J. Med. Chem. 15, 1101 (1972).
CAMERMAN, A., CAMERMAN, N.: Diphenylhydantoin and diazepam: molecular structure similarities and steric basis of anticonvulsant activity. Science 168, 1457 (1970).
CAMMARATA, A.: An apparent correlation between the *in vitro* activity of chloramphenicol analogs and electronic polarizability. J. Med. Chem. 10, 525 (1967).
CASY, A.F.: Stereochemistry and biological activity. In: Medicinal Chemistry (Ed. A BURGER), 3rd ed., part I, p. 81. New York: Interscience 1970.
CASY, A.F.: Analgesic receptors. In: A Guide to Molecular Pharmacology-Toxicology (Ed. R.M. FEATHERSTONE), part I, p. 217. New York: Dekker 1973.
CAVALLITO, C.J.: Some trends in the development of structure-activity relationships (SAR) and theory of receptors. In: Structure-Activity Relationships (Ed. C.J. CAVALLITO), vol. I, p. 1. Oxford: Pergamon Press 1973.
CHARTON, M.: Steric effects. I. Esterification and acid-catalyzed hydrolysis of esters. J. Am. Chem. Soc. 97, 1552 (1975).
CHENEY, V.: *Ab initio* calculations on large molecules using molecular fragments. Structural correlations between natural substrate moieties and some antibiotic inhibitors of peptidyl transferase. J. Med. Chem. 17, 590 (1974).
CHENG, C.C.: Novel common structural features among several classes of antimalarial agents. J. Pharm. Sci. 63, 307 (1974).
CHU, K.C.: Applications of artificial intelligence to chemistry. Use of pattern recognition and cluster analysis to determine the pharmacological activity of some organic compounds. Anal. Chem. 46, 1181 (1974).
CHU, K.C., FELDMANN, R.J., SHAPIRO, M.B., HAZARD, G.F., GERAN, R.I.: Pattern recognition and structure-activity relationship studies.

Computer-assisted prediction of antitumor activity in structurally diverse drugs in an experimental mouse brain tumor system. J. Med. Chem. 18, 539 (1975).

CITRI, N.: Conformational adaptability in enzymes. Adv. Enzymol. 37, 397 (1973).

CLARKE, R.L., GAMBINO, A.J., DAUM, S.J.: In pursuit of analgetic agents. Hydro-1,3-ethanoindeno-[2,1-c]pyridines and homologs. J. Med. Chem. 17, 1040 (1974).

CLERC. J.T., NAEGELI, P., SEIBL, J.: Artificial intelligence. Chimia 27, 639 (1973).

COUBEILS, J.L., COURRIÈRE, Ph., PULLMAN, B.: Quantum-mechanical study of the conformational properties of sympatholytic compounds. J. Med. Chem. 15, 453 (1972).

CRAMER, R.D., III, REDL, G., BERKOFF, C.E.: Substructural analysis. A novel approach to the problem of drug design. J. Med. Chem. 17, 533 (1974).

CRENSHAW, R.R., LUKE, G.M., JENKS, T.A., BIALY, G.: Potential anti-fertility agents. 7. Synthesis and biological activities of 2-, 3-, and 6-alkyl-substituted 4-aryl-2-methylcyclohexanecarboxylic acids. J. Med. Chem. 17, 1262 (1974).

CROXATTO, R., HUIDOBRO, F.: Fundamental basis of the specificity of pressor and depressor amines in their vascular effects. Arch. Int. Pharmacodyn. 106, 207 (1956).

DETAR, DeL. F.: Calculation of steric effects in reactions. J. Am. Chem. Soc. 96, 1254 (1974a).

DETAR, DeL. F.: Quantitative predictions of steric acceleration. J. Am. Chem. Soc. 96, 1255 (1974b).

DIERDORF, D.S., KOWALSKI, B.R.: Three-dimensional molecular structure-biological activity correlation by pattern recognition. U.S.N.T.I.S., Ad Rep. No. 785863/2GA (1974).

DRENTH, J., HOL, W.G.J., JANSONIUS, J.N., KOEKOEK, R.: A comparison of the three-dimensional structures of subtilisin BPN' and subtilisin novo. In: Structure and Function of Proteins at the Three-Dimensional Level. Cold Spring Harbor Symp. Quant. Biol. 36, p. 107. Cold Spring Harbor Laboratories 1972.

DUBOIS, J-E.: DARC system in chemistry. In: Computer Representation and Manipulation of Chemical Information (Eds. W.T. WIPKE, S.R. HELLER, R.J. FELDMANN, E. HYDE), p. 239. New York: Interscience 1974.

DUBOIS, J-E, LAURENT, D., ARANDA, A.: Systéme DARC. XVII. Théorie de topologie-information. II. Méthode PELCO-procédure d'établissement de corrélation de topologie-information. J. Chim. Physique 70, 1616 (1973).

EAKIN, D.R., HYDE, E.: Evaluation of on-line techniques in a sub-structure search system. In: Computer Representation and Manipulation of Chemical Information (Eds. W.T. WIPKE, S.R. HELLER, R.J. FELDMANN, E. HYDE), p. 1. New York: Interscience 1974.

EDDY, N.B., MAY, E.L.: The search for a better analgesic. Science 181, 407 (1973).

EDWARDS, A.E., COHEN, S.: Visual illusion, tactile sensibility and reaction time under LSD-25. Psychopharmacologia 2, 297 (1961).

EHRLICH, P.: Über den jetzigen Stand der Chemotherapie. Chem. Ber. 42, 17 (1909).

ENGLER, E.M., ANDOSE, J.D., SCHLEYER, P.v.R.: Critical evaluation of molecular mechanics. J. Am. Chem. Soc. 95, 8005 (1973).

FEINBERG, A.P., SNYDER, S.H.: Phenothiazine drugs: structure-activity relationships explained by a conformation that mimics dopamine. Proc. Nat. Acad. Sci. 72, 1899 (1975).
FELDMANN, R.J.: Interactive graphic chemical structure searching. In: Computer Representation and Manipulation of Chemical Information (Eds. W.T. WIPKE, S.R. HELLER, R.J. FELDMANN, E. HYDE), p. 55. New York: Interscience 1974.
FELDMANN, R.J., BACON, C.R.T., COHEN, J.S.: Versatile interactive graphics display system for molecular modelling by computer. Nature (Lond.) 244, 113 (1973).
FENNER, H.: EPR studies on the mechanism of biotransformation of tricyclic neuroleptics and antidepressants. In: The Phenothiazines and Structurally Related Drugs (Eds. I.S. FORREST, C.J. CARR, E. USDIN), p. 5. New York: Raven 1974.
FREE, S.M., Jr., WILSON, J.W.: A mathematical contribution to structure-activity studies. J. Med. Chem. 7, 395 (1964).
FUJITA, T., IWASA, J., HANSCH, C.: A new substituent constant, π, derived from partition coefficients. J. Am. Chem. Soc. 86, 5175 (1964).
GINTZLER, A.R., MUSACCHIO, J.M.: Interaction between serotonin and morphine in the guinea-pig ileum. J. Pharmacol. Exp. Ther. 189, 484 (1974).
GOLEBIEWSKI, A., PARCZEWSKI, A.: Theoretical conformational analysis of organic molecules. Chem. Rev. 74, 519 (1974).
GOODFORD, P.J.: Prediction of pharmacological activity by the method of physicochemical-activity relationships. Adv. Pharmacol. Chemother. 11, 51 (1973).
GOODWIN, T.W.: Prochirality in biochemistry. In: Essays in Biochemistry (Eds. P.N. CAMPBELL, F. DICKENS), vol. IX, p. 103. London: Academic Press 1973.
GOURLEY, D.R.H.: Biological responses to drugs. In: Medicinal Chemistry (Ed. A. BURGER), 3rd ed., part I, p. 25. New York: Interscience 1970.
GRAY, C.J.: Enzyme-Catalyzed Reactions, p. 86. London: van Nostrand 1971.
GREEN, J.P., DRESSLER, K.P., KHAZAN, N.: Mescaline-like activity of 2-amino-7-hydroxytetralin. Life Sci. 12, part I, 475 (1973).
GREEN, J.P., JOHNSON, C.L., KANG, S.: Application of quantum chemistry to drugs and their interactions. Ann. Rev. Pharmacol. 14, 319 (1974).
GUILD, W., Jr.: Theory of sweet taste. J. Chem. Educ. 49, 171 (1972).
GUND, P., WIPKE, W.T., LANGRIDGE, R.: Computer searching of a molecular structure file for pharmacophoric patterns. In: Proc. Int. Conf. Computers in Chem. Res. and Educ., Ljubljana, July 12-17, 1973, vol. 3, p. 5/33. Amsterdam: Elsevier 1974.
HAHN, F.E., GUND, P.: A structural model of the chloramphenicol receptor site. In: Topics in Infectious Diseases, Vol. I Drug Receptor Interactions in Antimicrobial Chemotherapy (Eds. J. DREWS, F.E. HAHN), p. 245. Vienna: Springer 1975.
HAMMETT, L.P.: The effect of structure upon the reactions of organic compounds. Benzene derivatives. J. Am. Chem. Soc. 59, 96 (1937).
HANSCH, C.: A quantitative approach to biochemical structure-activity relationships. Accounts Chem. Res. 2, 232 (1969).
HANSCH, C.: Quantitative approaches to pharmacological structure-activity relationships. In: Structure-Activity Relationships (Ed. C.J. CAVALLITO), vol. I, p. 75. Oxford: Pergamon Press 1973.

HANSCH, C., KUTTER, E., LEO, A.: Homolytic constants in the correlation of chloramphenicol structure with activity. J. Med. Chem. 12, 746 (1969).

HANSCH, C., LEO, A., UNGER, S.H., KIM, K.H., NIKAITANI, D,, LIEN, E.J.: "Aromatic" substituent constants for structure-activity correlations. J. Med. Chem. 16, 1207 (1973a).

HANSCH, C., MALONEY, P.P., FUJITA, T., MUIR, R.M.: Correlation of biological activity of phenoxyacetic acids with Hammett substituent constants and partition coefficients. Nature (Lond.) 194, 178 (1962).

HANSCH, C., MUIR, R.M., FUJITA, T., MALONEY, P.P., GEIGER, F., STREICH, M.: The correlation of biological activity of plant growth regulators and chloromycetin derivatives with Hammett constants and partition coefficients. J. Am. Chem. Soc. 85, 2817 (1963).

HANSCH, C., NAKAMOTO, K., GORIN, M., DENISEVICH, P., GARRETT, E.R., HEMAN-ACKAH, S.M., WON, C.H.: Structure-activity relationships of chloramphenicols. J. Med. Chem. 16, 917 (1973b).

HANSON, K.R.: Applications of the sequence rule. I. Naming the paired ligands g,g at a tetrahedral atom Xggij. II. Naming the two faces of a trigonal atom Yghi. J. Am. Chem. Soc. 88, 2731 (1966).

HANSON, K.R., ROSE, I.A.: Interpretations of enzyme reaction stereospecificity. Accounts Chem. Res. 8, 1 (1975).

HARRIS, R.J., SYMONS, R.H.: On the molecular mechanism of action of certain substrates and inhibitors of ribosomal peptidyl transferase. Bioorganic Chem. 2, 266 (1973).

HILLER, S.A., GOLENDER, V.E., ROSENBLIT, A.B., RASTRIGIN, L.A., GLAZ, A.B.: Cybernetic methods of drug design. I. Statement of the problem—the perceptron approach. Computers Biomed. Res. 6, 411 (1973).

HODGES, D., NORDBY, D.H.: MOLOCH-3: the G. D. Searle molecular modeling system. Abstr. COMP-7, 169th Am. Chem. Soc. Nat. Meeting, Phila., April 1975.

HÖLTJE, H-D., KIER, L.B.: A theoretical approach to structure-activity relationships of chloramphenicol and congeners. J. Med. Chem. 17, 814 (1974).

HOPFINGER, A.J.: Conformational Properties of Macromolecules. New York: Academic Press 1973.

JANSSEN, P.A.J.: Chemical and pharmacological classification of neuroleptics. In: The Neuroleptics (Eds. D.P. BOBON, P.A.J. JANSSEN, J. BOBON), p.33. Basel: Karger 1970.

JANSSEN, P.A.J.: Structure-activity relationships (SAR) and drug design as illustrated with neuroleptic agents. In: Structure-Activity Relationships (Ed. C.J. CAVALLITO), vol. I, p. 37. Oxford: Pergamon Press 1973.

JOHNSON, C.D.: The Hammett Equation. Cambridge: University Press 1973.

JURS, P.C., ISENHOUR, T.L.: Chemical Applications of Pattern Recognition. New York: Interscience 1975.

KANG, S., JOHNSON, C.L., GREEN, J.P.: Theoretical studies on the conformation of psilocin and mescaline. Mol. Pharmacol. 9, 640 (1973).

KAUFMAN, J.J., KERMAN, E.: Quantum chemical and other theoretical techniques for the understanding of the psychoactive action of phenothiazines. In: The Phenothiazines and Structurally Related Drugs (Eds. I.S. FORREST, C.J. CARR, E. USDIN), p. 55. New York: Raven 1974.

KAUFMAN, J.J., KOSKI, W.S.: Physicochemical, quantum chemical and other theoretical techniques for the understanding of the mechanism

of action of CNS agents. Psychoactive drugs, narcotics, and narcotic antagonists and anesthetics. In: Drug Design (Ed. E.J. ARIENS), vol. 5, p. 251. New York: Academic Press 1975.
KELLY, J.M., ADAMSON, R.H.: A comparison of common interatomic distances in serotonin and some hallucinogenic drugs. Pharmacology 10, 28 (1973).
KENT, P., GAÜMANN, T.: Considerations on the interpretation of mass spectra via learning machines. Helv. Chem. Acta 58, 787 (1975).
KIER, L.B.: The preferred conformations of ephedrine isomers and the nature of the *alpha* adrenergic receptor. J. Pharmacol. Exp. Ther. 164, 75 (1968).
KIER, L.B.: Receptor mapping using molecular orbital theory. In: Fundamental Concepts in Drug-Receptor Interactions (Eds. J.F. DANIELLI, J.F. MORAN, D.J. TRIGGLE), p. 15. New York: Academic Press 1970.
KIER, L.B.: Molecular Orbital Theory in Drug Research. New York: Academic Press 1971.
KIER, L.B.: The prediction of molecular conformation as a biologically significant property. Pure Appl. Chem. 35, 509 (1973).
KIER, L.B., HÖLTJE, H-D.: A stochastic model of the remote recognition of preferred conformation in a drug-receptor interaction. J. Theoret. Biol. 49, 401 (1975).
KONO, M., O'HARA, K., HONDA, M., MITSUHASHI, S.: Drug resistance of staphylococci. XI. Induction of chloramphenicol resistance by its derivatives and analogues. J. Antibiot. (Tokyo) 22, 603 (1969).
KOROLKOVAS, A.: Essentials of Molecular Pharmacology, Background for Drug Design. New York: Interscience 1970.
KOWALSKI, B.R.: Pattern recognition in chemical research. In: Computers in Chemical and Biochemical Research (Eds. C.E. KLOPFENSTEIN, C.L. WILKINS), vol. 2, p. 1. New York: Academic Press 1974.
KOWALSKI, B.R., BENDER, C.F.: The application of pattern recognition to screening prospective anticancer drugs. Adenocarcinoma 755 biological activity test. J. Am. Chem. Soc. 96, 916 (1974).
KOWALSKI, B.R., BENDER, C.F.: Solving chemical problems with pattern recognition. Naturwissenschaften 62, 10 (1975).
LEFFLER, J.E., GRUNWALD, E.: Rates and Equilibria of Organic Reactions. New York: Wiley 1963.
LIEN, E.J.: The use of substituent constants and regression analysis in the study of structure-activity relationships. Am. J. Pharm. Educ. 33, 368 (1969).
LINDQUIST, R.N.: The design of enzyme inhibitors: transition state analogs. In: Drug Design (Ed. E.J. ARIENS), vol. 5, p. 24. New York: Academic Press 1975.
LYNCH, M.F.: The microstructure of chemical data-bases and the choice of representation for retrieval. In: Computer Representation and Manipulation of Chemical Information (Eds. W.T. WIPKE, S.R. HELLER, R.J. FELDMANN, E. HYDE), p. 31. New York: Interscience 1974.
LYNCH, M.F., HARRISON, J.M., TOWN, W.G., ASH, J.E.: Computer Handling of Chemical Structural Information. London: MacDonald 1971.
MAAYANI, S., WEINSTEIN, H., COHEN, S., SOKOLOVSKY, M.: Acetylcholine-like molecular arrangement in psychomimetic anticholinergic drugs. Proc. Nat. Acad. Sci. 70, 3103 (1973).
MACKAY, M., HODGKIN, D.C.: A crystallographic examination of the

structure of morphine. J. Chem. Soc. 3261 (1955).
MARSHALL, G.R., BOSSHARD, H.E., ELLIS, R.A.: Computer modeling of chemical structures: applications in crystallography, conformational analysis, and drug design. In: Computer Representation and Manipulation of Chemical Information (Eds. W.T. WIPKE, S.R. HELLER, R.J. FELDMANN, E. HYDE), p. 203. New York: Interscience 1974.
MATHEWS, R.J.: A comment on structure-activity correlations obtained using pattern recognition methods. J. Am. Chem. Soc. 97, 935 (1975).
McMAHON, R.E.: Drug metabolism. In: Medicinal Chemistry (Ed. A. BURGER), 3rd ed., part I, p. 50. New York: Interscience 1970.
MONTGOMERY, J.A., JOHNSTON, T.P., SHEALY, Y.F.: Drugs for neoplastic diseases. In: Medicinal Chemistry (Ed. A. BURGER), 3rd ed., part I, p. 680. New York: Interscience 1970.
MULLINS, L.J.: The use of models of the cell membrane in determining the mechanism of drug action. In: A Guide to Molecular Pharmacology-Toxicology (Ed. R.M. FEATHERSTONE), part I, p. 1. New York: Dekker 1973. See esp. p. 34.
NISHIKAWA, K., OOI, T.: Comparison of homologous tertiary structures of proteins. J. Theoret. Biol. 43, 351 (1974).
NORRINGTON, F.E., HYDE, R.M., WILLIAMS, S.G., WOOTTEN, R.: Physicochemical-activity relationships in practice. I. A rational and self-consistent data bank. J. Med. Chem. 18, 604 (1975).
OLDENDORF, W.H.: Blood-brain barrier permeability to drugs. Ann. Rev. Pharmacol. 14, 239 (1974).
ÖTVÖS, L., MORAVCSIK, E., KRAICSOVITS, F.: Stereochemistry of the reactions of biopolymers. V. Steric effect of chiral substituents in enzyme-catalyzed reactions. Tetrahed. Letters 2485 (1975).
PATIL, P.N., MILLER, D.D., TRENDELENBURG, U.: Molecular geometry and adrenergic drug activity. Pharmacol. Rev. 26, 323 (1975).
PAULING, P.: The shapes of cholinergic molecules. In: Cholinergic Mechanisms (Ed. P.G. WASER), p. 241. New York: Raven 1975.
PERRIN, C.L.: Testing of computer-assisted methods for classification of pharmacological activity. Science 183, 551 (1974).
POPLE, J.A., BEVERIDGE, D.L.: Approximate Molecular Orbital Theory. New York: McGraw-Hill 1970.
PORTOGHESE, P.S., WILLIAMS, D.A.: Stereochemical studies on medicinal agents. VII. Absolute stereochemistry of methadol isomers and the role of the 6-methyl group in analgetic activity. J. Med. Chem. 12, 839 (1969).
PULLMAN, B., COUBEILS, J-L., COURRIÈRE, Ph., GERVOIS, J-P.: Quantum mechanical study of the conformational properties of phenethylamines of biochemical and medicinal interest. J. Med. Chem. 15, 17 (1972).
PURCELL, W.P., BASS, G.E., CLAYTON, J.M.: Strategy of Drug Design: A Molecular Guide to Biological Activity. New York: Interscience 1973.
RAŠKA, K., Jr., JUROVČÍK, M., ŠORMOVÁ, Z., ŠORM, F.: Inhibition by 5-Azacytidine of Ribonucleic Acid Synthesis in Isolated Nuclei of Calf Thymus. Collec. Czech. Chem. Commun. 30, 3215 (1965).
RAO, S.T., ROSSMANN, M.G.: Comparison of super-secondary structures in proteins. J. Mol. Biol. 76, 241 (1973).
REDL, G., CRAMER, R.D., III, BERKOFF, C.E.: Quantitative drug design. Chem. Soc. Rev. 3, 273 (1974).
ROSSMANN, M.G., MORAS, D., OLSEN, K.W.: Chemical and biological

evolution of a nucleotide-binding protein. Nature (Lond.) 250, 194 (1974).
SEGAL, H.L.: Enzymatic interconversion of active and inactive forms of enzymes. Science 180, 25 (1973).
SHIPMAN, L.L., CHRISTOFFERSEN, R.E., CHENEY, B.V.: *Ab initio* calculations on large molecules using molecular fragments. Lincomycin studies. J. Med. Chem. 17, 583 (1974).
SHORTER, J.: The separation of polar, steric, and resonance effects by the use of linear free energy relationships. In: Advances in Linear Free Energy Relationships (Eds. N.B. CHAPMAN, J. SHORTER), p. 71. London: Plenum Press 1972.
SHORTER, J.: Correlation Analysis in Organic Chemistry: an Introduction to Linear Free-Energy Relationships. Oxford: Clarendon Press 1973.
SIMON, E.J.: In search of the opiate receptor. Am. J. Med. Sci. 266, 160 (1973).
SIMON, Z: Specific interactions. Intermolecular forces, steric requirements, and molecular size. Ang. Chem. Int. Ed. 13, 719 (1974).
SIMON, Z., SZABADAI, Z.: Minimal steric difference parameter and the importance of steric fit for structure-biological activity correlations. Studia Biophys. (Berlin) 39, 123 (1973).
SINKULA, A.A., YALKOWSKY, S.H.: Rationale for design of biologically reversible drug derivatives: prodrugs. J. Pharm. Sci. 64, 181 (1975).
SMYTHIES, J.R.: Relationships between the chemical structure and biological activity of convulsants. Ann. Rev. Pharmacol. 14, 9 (1974).
STILL, W.C., LEWIS, A.J.: Computerized molecular modeling. Chemtech. 4, 118 (1974).
STUPER, A.J., JURS, P.C.: Classification of psychotropic drugs as sedatives or tranquilizers using pattern recognition techniques. J. Am. Chem. Soc. 97, 182 (1975).
SUNDARAM, K.: Applications of an optimization technique in submolecular biology. Int. J. Quantum Chem. 9, 393 (1975).
SWAIN, C.G., LUPTON, E.C.: Field and resonance components of substituent effects. J. Am. Chem. Soc. 90, 4328 (1968).
TAFT, R.W., LEWIS, I.C.: Evaluation of resonance effects on reactivity by application of the linear inductive energy relationship. V. Concerning a σ_R scale of resonance effects. J. Am. Chem. Soc. 81, 5343 (1959).
TAKEMORI, A.E.: Biochemistry of drug dependence. Ann. Rev. Biochem. 43, 15 (1974).
TERENIUS, L.: Contribution of "receptor" affinity to analgesic potency. J. Pharm. Pharmacol. 26, 146 (1974).
TING, K-L.H., LEE, R.C.T., MILNE, G.W.A., SHAPIRO, M., GUARINO, A.M.: Applications of artificial intelligence: relationships between mass spectra and pharmacological activity of drugs. Science 180, 417 (1973).
WARING, M.: Binding of drugs to supercoiled circular DNA: evidence for and against intercalation. Progr. Mol. Subcell. Biol. 2, 216 (1970).
WARSHEL, A., KARPLUS, M.: Calculation of $\pi\pi^*$ excited state conformations and vibronic structure of retinal and related molecules. J. Am. Chem. Soc. 96, 5677 (1974).
WEINSTEIN, H., MAAYANI, S., SREBRENIK, S., COHEN, S., SOKOLOVSKY, M.: Psychotomimetic drugs as anticholinergic agents. II. Quantum-mechanical study of molecular interaction potentials of

1-cyclohexylpiperidine derivatives with the cholinergic receptor. Mol. Pharmacol. 9, 820 (1973).
WELLS, P.R.: Linear Free Energy Relationships. London: Academic Press 1968.
WILLBRANDT, W., ROSENBERG, T.: The concept of carrier transport and its corollaries in pharmacology. Pharmacol. Rev. 13, 109 (1961).
WILLIAMS, J.E., STANG, P.J., SCHLEYER, P.v.R.: Physical organic chemistry: quantitative conformational analysis; calculation methods. Ann. Rev. Phys. Chem. 19, 531 (1968).
WIPKE, W.T.: Computer-assisted three-dimensional synthetic analysis. In: Computer Representation and Manipulation of Chemical Information (Eds. W.T. WIPKE, S.R. HELLER, R.J. FELDMANN, E. HYDE), p. 147. New York: Interscience 1974.
WIPKE, W.T., DYOTT, T.M.: Simulation and evaluation of chemical synthesis. Computer representation and manipulation of stereochemistry. J. Am. Chem. Soc. 96, 4825 (1974).
WIPKE, W.T., GUND, P.: Congestion: a conformation-dependent measure of steric environment. Derivation and application in stereoselective addition to unsaturated carbon. J. Am. Chem. Soc. 96, 299 (1974).
ZANDER, G.S., JURS, P.C.: Generation of mass spectra using pattern recognition techniques. Anal. Chem. 47, 1562 (1975).
ZEE-CHENG, K.Y., CHENG, C.C.: Common receptor-complement feature among some antileukemic compounds. J. Pharm. Sci. 59, 1630 (1970).
ZEE-CHENG, K.Y., CHENG, C.C.: Interaction between DNA and coralyne acetosulfate, an antileukemic compound. J. Pharm. Sci. 62, 1572 (1973).
ZEE-CHENG, K.Y., PAULL, K.D., CHENG, C.C.: Experimental antileukemic agents. Coralyne, analogs, and related compounds. J. Med. Chem. 17, 347 (1974).