[Chem. Pharm. Bull. 30(10)3657—3661(1982)]

Quantitative Drug Design Studies. V.[1] Approach to Lead Generation by Pharmacophoric Pattern Searching

Toshiyuki Esaki

Department of Applied Chemistry, Faculty of Engineering, Nagoya University, Furo-cho, Chikusa-ku, 464, Japan

(Received March 4, 1982)

A graphic program module, which can extract the pharmacophoric patterns from candidate molecules, was developed to rationalize lead generation and was added to the quantitative drug design (QDD) program. This performs molecular mechanical conformational analysis and quantum chemical calculation for a test molecule, followed by the prediction of its pharmacological features by means of retrieval of the pharmacophoric pattern data-base, which contains information on 261 pharmacophores selected according to original criteria using the atomic coordinates derived from crystallographic data. By pattern matching, as a rule, both geometry and quantum chemical indices such as electron densities (total and frontier) and frontier orbital energy are checked and molecular modelling for the test molecule is also available as an option. The example of epinephrine is shown to illustrate the use of the module. It may be useful in the evaluation process of drug activities as one of the prescreening procedures before animal tests.

Keywords——lead generation; pharmacophore; pharmacophoric pattern searching; geometry; electronic state; molecular modelling; conformational analysis; crystallographic data; graphics; data-base

In quantitative drug design, lead generation plays an important part together with lead optimization. Computer-assisted pharmacophoric pattern searching[2] is an important approach in attempts to rationalize the former step. Under the pharmacophore hypothesis, an attempt is made to predict the pharmacological features of the candidate molecule by finding matching patterns in the data-base involving various pharmacophoric patterns. In this paper, the author describes an attempt to add a graphic program module for lead generation to the quantitative drug design (QDD) program.[1,3]

Experimental

Computer——The new module was implemented on the FACOM M-200 computer system at the Nagoya University Computation Center. The graphic display units were a FACOM 9532A (refresh type), a Tektronix 4014 (storage type), and a Tektronix 4027 (color refresh type).

Programming——The module was written in FORTRAN. The graphic subroutine packages used in the module were GSP-3D[4] for the FACOM 9532A, GSP[5] for the Tektronix 4014, and IGL[6] for the Tektronix 4027.

Results and Discussion

Outline of the Module

Fig. 1 depicts an outline of the pharmacophoric pattern searching procedure used in the present module. The module utilizes several existing programs and a data-base. Namely, the test molecule is visualized using the substituent data-base[3] at the molecular modelling option. For conformational analysis, the subroutines in the MMI program[7] are used to estimate the conformation. The quantum chemical calculations are carried out using the extended Hückel method.[8] For the pharmacophoric pattern searching, the subroutines in the MOLPAT program[2] are utilized in an extended form.

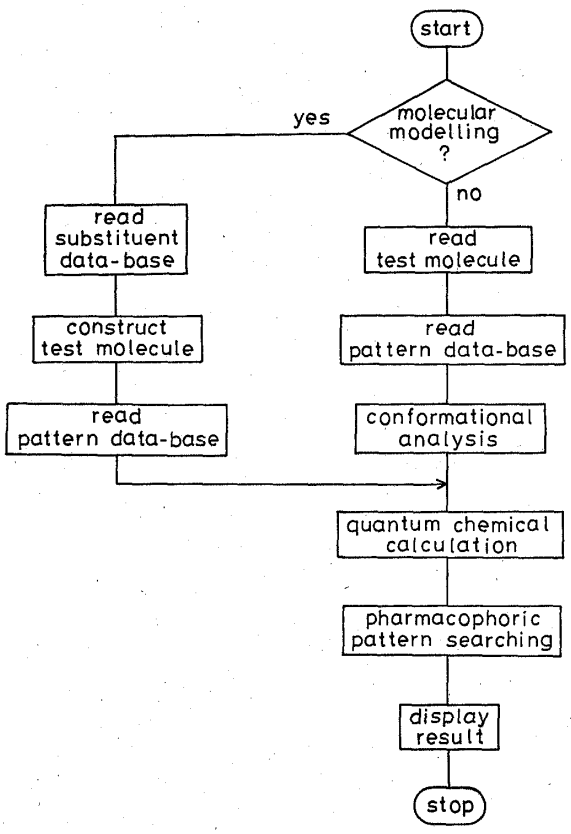


Fig. 1. Flow Diagram of the Module

Pattern matching is done, as a rule, based on a survey not only of the geometry but also of the electron densities (both total and frontier) and the frontier orbital energy level, except for cases in which quantum chemical indices for the pattern in question are not available in the data-base; this type of case will be mentioned later. In order to save computation time, geometrical matching is checked for the minimum number of interatomic distances (P) defined as in Eq. 1 or 2.[9]

$$P = n(n-1)/2 \qquad (n \leqq 5) \qquad (1)$$

$$= 4(n-3)+2 \qquad (n \geqq 6) \qquad (2)$$

where n is the number of atoms in the specified pattern. The frontier electron densities of HOMO and LUMO on the atom A (F_A) are calculated using Eqs. 3 and 4, respectively.

$$F_A^{HO} = \sum_r^{on\ A} \sum_s c_r^{HO}\, c_s^{HO}\, S_{rs} \qquad (3)$$

$$F_A^{LU} = \sum_r^{on\ A} \sum_s c_r^{UL}\, c_s^{LU}\, S_{rs} \qquad (4)$$

where c_r or c_s represents the frontier orbital coefficient of the atomic orbital r or s, and S_{rs} is the overlap integral between the atomic orbitals r and s. These quantities express to what extent an electron may distribute over the corresponding frontier orbital. If the test molecule shows patterns similar to those in the data-base within the specified tolerance (*e.g.*: $\pm 10\%$ for the interatomic distance, $\pm 5\%$ for the total electron density and frontier orbital energy, *etc.*), a list of the matched patterns is displayed on the cathode ray tube (CRT) after retrieval. It is also possible to superimpose a required pattern upon the perspective view of the test molecule. However, differences in pharmacological behavior between enantiomers cannot be distinguished.

The author has designated the present expanded program as GPQDD (Graphic Program for Quantitative Drug Design), as a development of QDD.

Pharmacophoric Pattern Data-Base

Many pharmacophoric patterns have already been proposed for various pharmacological properties.[2] However, in this study, the author tried to prepare a unique pharmacophoric pattern data-based on the following criteria.

(a) Patterns are retrieved from the drugs registered in the crystallographic data-base.[1]

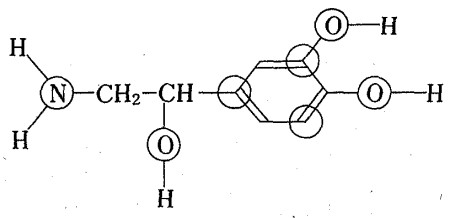


Fig. 2. Assumed Pattern for Noradrenaline

Pattern atoms are circled.

(b) Atomic coordinates obtained from crystallographic data are used as a geometrical source, and one pattern is established per crystallographic data unit.

(c) All atoms having pi or lone pair electrons are regarded as required ones for pattern, except for planar moieties such as phenyl, *etc.*, in which three atoms, the minimum number of atoms to determine the plane, are arbitrarily chosen in the moiety. This is because outer mobile electrons of a molecule such as pi or lone

Table I. Staple Pharmacophoric Patterns in the Data-Base

No.	Pattern	No.	Pattern	No.	Pattern
1	Muscarinic	10	Antihistaminic	19	Anticonvulsant
2	Nicotinic	11	Narcotic	20	Hypnotic
3	Anticholinergic	12	Antinarcotic	21	Diuretic
4	Adrenergic	13	Analgesic	22	Anti-inflammatory
5	Antiadrenergic	14	Antipyretic	23	Anthelmintic
6	Sedative	15	Antidiarrheal	24	Antimicrobial
7	Neuroleptic	16	Antihypertensive	25	Antineoplastic
8	Antidepressant	17	Hallucinogenic		
9	MAO inhibitor	18	Local anesthetic		

pair electrons are considered to participate most in the biophasal reaction. The hyperconjugation effect is ignored. As an example, Fig. 2 illustrates the pattern atoms established for noradrenaline, the α-adrenergic agonist, based on this criterion.

The current data-base stores information on 261 kinds of pharmacophoric patterns collected in the above-mentioned way. Table I summarizes 25 categories of pharmacophoric patterns. Each data unit consists of pharmacological features,[10] atomic symbols and coordinates of pattern atoms, and, when hydrogen atom coordinates were available (152 kinds), electron densities (both total and frontier) and frontier orbital energies were estimated using the extend-

A

No.	Pharmacological features
1	CNS stimulant
2	Sympathomimetic agent
3	Antihypotensive agent
4	α-Adrenergic agonist
5	Dopaminergic agonist
6	Dopaminergic agonist
7	Sympathomimetic agent
8	α-Adrenergic agonist, β-Adrenergic agonist
9	Sympathomimetic agent, Anorectic agent
10	β-Adrenergic agonist

B

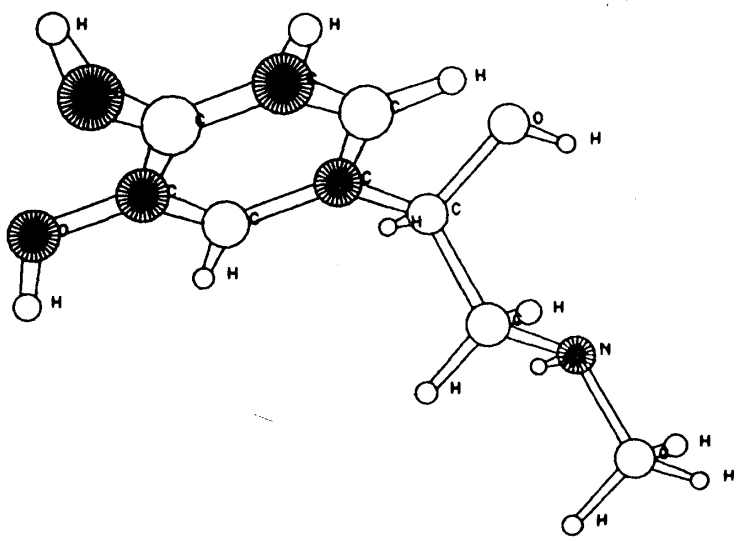
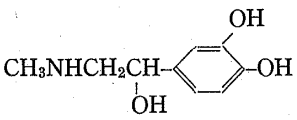


Fig. 3. Result for Epinephrine

A: List of the pharmacological features extracted from the matched patterns.
B: α-Adrenergic pharmacophore [No. 4 in A].
Pattern atoms are marked by radiating lines.

ed Huckel method. As regards molecules for which only crystallographic data of ionic species were available, the molecular orbital (MO) calculation was, as a rule, carried out after modification to the corresponding neutral species (*e.g.*: $-NH_3^+ \rightarrow -NH_2$, *etc.*).



I

Example

Fig. 3 gives an example of the CRT frames when the present module was executed for epinephrine (I), one of the α-, β-adrenergic agonists, the pattern of which had not been stored in the data-base because of lack of valid crystallographic data.

Fig. 3A shows the list of matched patterns obtained after searching; ten features, associated with the patterns, are listed. Judging from this result, it is considered that the pharmacological features of I were correctly predicted. In Fig. 3B, the pattern atoms of No. 4 in Fig. 3A are superimposed on the molecular structure of I.

Conclusions

In the past, lead compounds were usually discovered by chance. The Hansch–Fujita and Free–Wilson approaches are ineffectual for lead generation.[11] Thus, various kinds of pattern recognition techniques have so far been applied to this problem.[12] However, any results obtained from such studies seem to give insight into only limited series of pharmacological properties and chemical structures. On the other hand, the pharmacophoric pattern searching technique may provide a clearer and more rational answer to one of the main practical themes in lead generation, that is, theoretical prediction of pharmacological features of any candidate molecule (actually synthesized or not) on the basis of the molecular structure, though this may not necessarily find lead compounds which have new types of pharmacophores. It is certain that many problems still remain in the present module, such as refinement of patterns, expansion of the data-base, and application of a more sophisticated MO method, *etc.*. However, there is no doubt that such an approach coupled with graphic facilities can provide a useful prescreening tool, before animal tests, during the process of evaluation of drug activities.

Acknowledgement The author is grateful to Dr. P. Gund, Merck Sharp and Dohme Research Laboratories, for the listing of MOLPAT and related materials. He is also indebted to Prof. T. Sasaki, Faculty of Engineering, Nagoya University, for access to his library, and to Dr. Y. Terada, Faculty of Pharmacy, Meijo University, for MMI.

References

1) Part IV: T. Esaki, Proceedings of the Fourth Informational Chemistry Symposium, Tokyo, Oct. 1981, pp. 58—62.
2) P. Gund, "Progress in Molecular and Subcellular Biology," Vol. 5, ed. by F.E. Hahn, Springer-Verlag, Inc., New York, 1977, pp. 117—143; P. Gund, *Annu. Repts. Med. Chem.*, **14**, 299 (1979).
3) T. Esaki, *Nippon Kagaku Kaishi*, **1981**, 868; T. Esaki, *Anal. Chim. Acta*, **133**, 657 (1981).
4) "FACOM OSIV/F4 GSP-3D Manual (64SP-6051-1)," Fujitsu, Ltd., 1981.
5) "FACOM OSIV GSP Grammar Book (64SP-6020-3)," Fijitsu, Ltd., 1980.
6) "Plot 10 4010C01 Interactive Graphic Library User's Manual," Tektronix, Inc., 1979.
7) N.L. Allinger, MMI, Program No. 318, Quantum Chemistry Program Exchange, Indiana University, Bloomington, Ind. (1975).
8) T. Esaki, *J. Pharm. Dyn.*, **3**, 562 (1980).
9) P. Gund, private communication.
10) J.P. Tollenaere, H. Moereels, and L.A. Raymaekers, "Atlas of the Three-Dimensional Structure of Drugs," Elsevier/North-Holland Biomedical Press, Amsterdam, 1979.
11) R.F. Gould (ed.), "Biological Correlations—The Hansch Approach," Advances in Chemistry Series 114, American Chemical Society, Washington, D.C., 1972.

12) K.L.H. Ting, R.C.T. Lee, G.W.A. Milne, M. Shapiro, and A.M. Guarino, *Science*, **180**, 417 (1973); B.R. Kowalski and C.F. Bender, *J. Am. Chem. Soc.*, **96**, 916 (1974); R.D. Cramer, III, G. Redl, and C.E. Berkoff, *J. Med. Chem.*, **17**, 533 (1974); K.C. Chu, *Anal. Chem.*, **46**, 1181 (1974); A.J. Stuper and P.C. Jurs, *J. Am. Chem. Soc.*, **97**, 182 (1975); K.C. Chu, R.J. Feldmann, M.B. Shapiro, G.F. Hazard, Jr., and R.I. Geran, *J. Med. Chem.*, **18**, 539 (1975); A. Cammarata and G.K. Menon, *ibid.*, **19**, 739 (1976); W.J. Dunn, III, S. Wold, and Y.C. Martin, *ibid.*, **21**, 922 (1978); W.J. Dunn, III and S. Wold, *ibid.*, **21**, 1001 (1978); D.R. Henry and J.H. Block, *ibid.*, **22**, 465 (1979); J.T. Chou and P.C. Jurs, *ibid.*, **22**, 792 (1979); A.J. Stuper, W.E. Brügger, and P.C. Jurs, "Computer Assisted Studies of Chemical Structure and Biological Function," John Wiley and Sons, Inc., New York, 1979.