Can Quantitative Structure Activity Analyses and Molecular Graphics Assist in Designing New Inhibitors of Photosystem II?

Wilfried Draber

Pflanzenschutzzentrum Monheim, Bayer AG, D-5090 Leverkusen, Bundesrepublik Deutschland

Z. Naturforsch. **42c**, 713–717 (1987); received January 5, 1987

Photosynthesis Inhibition, Herbicides, QSAR Methods and Parameters, Molecular Graphics, Metamitron

A calculation, based on various plausible assumptions, shows that under optimal conditions a rate of less than 10 g/ha of a potent photosynthesis inhibitor would be sufficient for good herbidical action. The possibility to achieve this by computer assisted drug design (CADD) is discussed. In order to provide some background, several methods of "classical" QSAR and the most frequently used parameters are discussed. Moreover, some aspects of molecular graphics are described. This powerful new CADD-method gives results that are qualitatively different from QSAR. It is not only intellectually demanding but also expensive with regard to hard- and software. In conclusion, one successful QSAR-based prediction of a herbicide is reviewed.

Introduction

Near the end of their book on biochemical modes of action of pesticides, Corbett, Wright and Baillie [1] present a calculation from which they conclude that one day there might be herbicides which are effective in the field at rates of the order of milligrammes/hectare. This has not yet come true, but certain inhibitors of amino acid biosynthetic pathways are fully active – at least in the greenhouse – at a few grammes/hectare. Among these are the sulfonylureas and the imidazolinones (such as chlorsulfuron and imazapyr) which inhibit acetolactate synthase, and phosphinotricin (an inhibitor of glutamine synthetase).

The calculation on which Corbett and his coauthors base their prediction is disputable in one or the other point. Moreover, when dealing with PS-II-inhibitors it is necessary to take into account the relative high amount of chlorophyll per hectare compared to some enzymes that are required for amino acid biosyntheses. If one assumes a coverage (by crop and weed plants) of 100% as realistic, the number of binding sites can be estimated to approximately 0.03 mol/ha. With the average molecular weight of a good inhibitor (pI_{50} about 7) of 200, it follows that 6.0 g/ha should be sufficient.

This calculation is based on the following assumptions:

1. Post emergence application is assumed.
2. A reasonable estimate is that 1 cm^2 leaf surface contain 0.1 mg chlorophyll. For 100% coverage, 10 kg chlorophyll/ha results from this figure.
3. 10 kg chlorophyll is equivalent to 11.12 mol.
4. According to Tischer and Strotmann [2] 2.7×10^{-3} inhibitor-binding sites correspond to 1.0 mol chlorophyll, giving 0.03 mol binding sites for 11.12 mol chlorophyll.
5. 0.03 mol correspond to 6.0 g of an inhibitor with the mol. w. of 200.

One provision must be mentioned: It is necessary for the inhibitor to reach the site of action – the thylakoids within the chloroplasts – with practically no loss en route.

What is QSAR?

Among various quantitative structure-activity procedures (QSAR = Quantitative Structure Activity Relationship), the Hansch-Fujita approach [3, 4] has been the most widely and effectively used up to now, in medicinal chemistry as well as in the field of agrochemicals. It rests on the assumption that the potency of the specified biological activity of a compound is mathematically related to a (not necessarily linear) combination of different physicochemical properties or descriptors. When the number of compounds is great enough and the biological parameter can be determined with sufficient precision, a correlation equation of the form

$$\log(1/c) = a \cdot \pi + b \cdot \sigma + c \cdot S \ldots . + \text{const.} \quad (1)$$

can be derived by standard statistical methods (least

Verlag der Zeitschrift für Naturforschung, D-7400 Tübingen
0341–0382/87/0600–0713 $ 01.30/0

squares). In this equation, log $(1/c)$ denotes the biological activity which is usually expressed as the logarithm. It was the merit of Hansch and Fujita when they proposed this procedure, to introduce π as the hydrophobic parameter which is derived from the octanol/water partition coefficient as

$$\pi = \log P_S - \log P_H \quad (2)$$

where P_S is the partition coefficient of the compound substituted by S and P_H is the partition coefficient of the unsubstituted parent compound.

Despite many attempts to replace octanol (which is by no means totally inmiscible with water!) by other solvents (e.g. benzene, chloroform, aliphatic hydrocarbons) hitherto no better one in combination with water (or aqueous buffer) has been found, and the importance of the π-values thus defined can hardly be underestimated [5, 6]. Nowadays no chemist who tries to synthetise compounds with any kind of biological activity can avoid thinking in terms of π and to take into account the hydrophilicity/hydrophobicity of the substituents and of the whole molecule. This attitude may go so far that sometimes the second term in equation [1], the Hammett constant σ [7, 8] is neglected. If we agree that σ is just a (statistically more or less rough) measure of the electronic properties of a substituent (though "σ" stands for more than half a dozen constants derived from more specific reactions than acid/base equilibria), it becomes evident that σ may be replaced by coefficients from semi-empirical quantum-mechanical calculations, *e.g.* π- or σ-charge densities on certain atoms. Correlations employing MO-parameters, however, are rather scarce in the literature up to now, perhaps as result of several misinterpretations some time ago. Another piece of information that can be obtained from MO-calculations is the relative stability of tautomeric forms or of various conformations of the molecule. – In equation [1] "S" is a steric parameter which can be the Taft Es-parameter. This has been commented on by Charton [10] and Hancock [11].

Today, Verloop's STERIMOL parameters, which are calculated from Van der Waals radii [12] are much more frequently used. In some cases, squared hydrophobic and steric parameter values are required to obtain a (statistically) good fit. When linear and quadratic terms with alternating signs occur in the same equation, this may point to optima. Sometimes, indicator variables or (less kindly called) "dummy parameters" accounting *e.g.* for hydrogen bonding in the presence or absence of a substituent (or substructure) are needed. In other cases, the theoretical connectivity parameters by Kier and Hall [13] might prove useful. All in all, any measured or calculated molecular property can be included into a QSAR, provided one cares to avoid chance correlations and parameter intercorrelations.

Another method of correlating biological activity and structure was proposed by Free and Wilson [14] almost simultaneously with the Hansch-Fujita approach in 1964. It does not need parameter values with any physico-chemical or theoretical meaning. The equations have the form

$$\log(1/c) = a \cdot X_1 + b \cdot X_2 + c \cdot X_3 + \ldots + \text{const.} \quad (3)$$

In this equation, the coefficients a, b, c, ... which represent contributions of the substituents to the activity, are calculated by least squares multiple regression techniques as in the Hansch approach. The "const." represents the mean potency of the series or that of a reference compound. The X_i indicate the presence or absence of a substituent in a specific position of the molecule.

The apparent strength of this *de-novo* method – prejudices are precluded – is outweighed by two disadvantages: 1. It is necessary to have the same substituent at least twice (if possible more) in the same position; 2. it is implicitly assumed that the contributions made by a specific substituent to activity are additive, which as a rule is not the case. Moreover, the method does not permit extrapolations to substituents that are not in the data base.

The Free-Wilson method is discussed thoroughly by Martin [15], Franke [16], and Seydel [17]. It has been compared with the Hansch approach, and sometimes Free-Wilson parameters were successfully correlated with extrathermodynamic ones [18, 19]. Most authors agree that the predictive value of this and related methods is questionable. – Some other techniques like cluster, principal component and discriminant analysis and also pattern recognition have as yet never been applied to herbicides and are just named here.

In general it can be stated that QSAR studies are relatively inexpensive and easy to carry out. This does not hold true for another methodology – molecular graphics – either in terms of personnel time or hardware plus software costs. It shall be briefly discussed here.

A New Dimension in QSAR: Molecular Graphics

Molecular graphics has brought a new dimension into many aspects of the relationship between chemical structure and biological activity, in more than one respect. By this procedure it becomes possible to tackle the problem of how conformation and size of a molecule affect the biological activity and to visualize this in three dimensions. The input for molecular modelling usually comes either from X-ray crystallography or, if not available, from quantum and molecular mechanics calculations. These do not only need an expert, they are also CPU-time consuming, especially when they have to be applied to larger molecules or peptides. An additional increase in computer time is required if conformational mobility must be accounted for. There are several commercial programmes on the software market that allow not only a multicolored, three-dimensional display of structures but also of surfaces, different charge densities and other structural features. Moreover, it is possible to show in one picture a multitude of favourable conformations of one or more molecules at the same time. The docking of effector molecules to its receptor can be demonstrated, although for such purposes generally an X-ray structure of the receptor is required. Because this technique emphasizes steric relationships between effector molecules and receptors, one cannot expect predictions pointing towards specific substitution as in the Hansch approach. Its predictive power is of another quality. To name one advantage of molecular graphics: quite often it is possible to detect similarities between two or even more chemical classes of effectors that look different on the paper, provided – of course – they act by the same molecular mechanism. From this may result another class of active compounds by a kind of "crossing" of the previous ones.

As a prerequisite for work in this field one needs compounds with known activity in a specific test. The activity should cover a range of several orders of magnitude: this is favourable in all QSAR applications. But in contrast to extrathermodynamic or *de-novo* approaches, no equations are obtained. Instead one can visualize the steric requirements for activity, mark areas which enhance it and others which are unfavourable or even forbidden. It is also possible to display charge densities by different colors signifying their sign and quantity or molecular electrostatic potentials.

In general, a combination of classical QSAR and the advantages that molecular graphics offer will also contribute greatly to the problem of the interaction between inhibitors and the components of the photosynthetic electron transport. A recent X-ray study by Michel, Deisenhofer *et al.* [20–22] of a photosystem has provided a wealth of information and thereby greatly influenced all thinking on the mechanism of photosynthetic inhibition, though the study was performed on the reaction center of a photosynthetic (purple) bacterium, *Rhodopseudomonas viridis,* which differs in certain aspects, *e.g.* sensitivity to inhibitors, from that of green plants. Nevertheless, other similar studies will undoubtedly follow and by this lead to an improvement in knowledge of the plant system.

Are there CADD Success Stories?

The ultimate goal of all attempts at computer assisted drug design (CADD), be it classical QSAR, cluster analysis or the methods that computer graphics is providing, is to extrapolate from a given set of compounds with activity to novel – unknown – ones, whereby the therm "novel" can be a matter for dispute. Certainly everyone will agree that novelty is not to be equated with patentibility, but also not in the sense of the (first) invention of the wheel. A pragmatic definition of "novelty" could be that it has to be something new measured against the highest level of information available at that time. This touches on communication which is restricted not only by certain governments but also by commercial companies, for obvious reasons. Therefore, well-documented CADD-stories are rather scarce, especially if they resulted in marketable compounds. Fujita and Iwamura [23] in 1982 published a compilation of some QSAR-studies in pesticide research which were carried out in Japan. Though they quoted some herbicides, no PS-II-inhibitor was among them. This changed in 1986 when Fujita [24] extended his study by adding more examples, not only from Japan. Among others he mentioned the particularly well-documented metamitron which has now become one of the most important sugar beet herbicides on the European market.

Although this compound – 3-methyl-4-amino-6-phenyl-1,2,4-triazin-5-one – belongs to a class of compounds that were known as PS-II-inhibitors, 3-alkyl substituted triazinones were virtually unknown beforehand. Therefore it appears legitimate to con-

sider these compounds (3-alkyl triazinones) as a novel class, the degree of "novelty" however, as in most cases, being open to dispute. The prediction was based on the pI_{50}-values of 13 4-amino-1,2,4-triazin-5-ones substituted in 3-position by methylthio, methoxy and methylamino groups. As result of a Hansch correlation an equation with log P and (log $P)^2$ was obtained. Since larger 3-substituents did not fit and electronic factors could obviously be neglected, it was concluded that the chemical nature of the 3-substituent was unimportant, but it had to be small [25]. From this followed the logical consequence that 3-alkyl triazinones had to be synthesized [26].

Molecular graphics can demonstrate the overlapping of the volumes of active triazinones such as metamitron and metribuzin (3-methylthio-4-amino-6-*tert*-butyl-1,2,4-triazin-5-one). Electronically very similar compounds, *e.g.* triazinones with a bridge between the 3- and 4-positions, were found to be almost inactive. Since the binding site of triazinones and other photosynthesis inhibitors is located on the D-1-polypeptide (which is the necessary assumption), a superposition of active and inactive triazinones (and other groups of potential PS-II-inhibitors) yields a negative imprint of the binding niche. Thereby much information is collected on the conformation of the binding niche itself and also the interrelationships between different inhibitor classes.

Some Personal Views

The author is convinced that s skilful combination of CADD methods will lead to new PS-II-inhibitors that need not fear a comparison with amino acid pathway blocking agents when it comes to field performance. Use of "classical" QSAR is apparently trivial and comparatively inexpensive though caution has to be observed in its interpretation. One often encountered problem and source of worry is the poor correlation between pI_{50}-values and greenhouse (not to mention field) data. There are some papers that try to overcome this difficulty, but the results are not too encouraging so far.

It is the author's conviction that molecular graphics and its diverse applications in CADD belong in the hands of an expert. Only an expert will – at the present state of the art – have the appropriate judgement to make effective use of all the possibilities which are offered by this expensive and intellectually demanding instrument.

Acknowledgements

I am greatly indebted to my colleagues Dr. Carl Fedtke and Dr. Andreas Zywietz for valuable advice and numerous discussions.

[1] J. R. Corbett, K. Wright, and A. C. Baillie, The Biochemical Mode of Action of Pesticides, pp. 343–344, Academic Press, London, New York 1984.
[2] W. Tischer and H. Strotmann, Biochim. Biophys. Acta **460,** 113–125 (1977).
[3] C. Hansch and T. Fujita, J. Am. Chem. Soc. **86,** 1616 (1964).
[4] T. Fujita, J. Iwasa, and C. Hansch, J. Am. Chem. Soc. **86,** 5175–5180 (1964).
[5] J. Iwasa, T. Fujita, and C. Hansch, J. Med. Chem. **8,** 150–153 (1965).
[6] C. Hansch, Accounts of Chemical Research **2,** 232–241 (1969).
[7] L. P. Hammett, Physical Organic Chemistry, McGraw-Hill, New York 1940, 1970.
[8] J. Shorter, Correlation Analysis in Organic Chemistry, Clarendon Press, Oxford 1973.
[9] R. W. Taft, Separation of Polar, Steric and Resonance Effects in Reactivity, in: Steric Effects in Organic Chemistry (M. S. Newman, ed.), pp. 556–675, Wiley, New York 1956.
[10] M. Charton, J. Am. Chem. Soc. **91,** 615–618 (1969).
[11] C. K. Hancock, E. A. Meyers, and B. J. Yager, J. Am. Chem. Soc. **83,** 4211–4213 (1961).
[12] A. Verloop, W. Hogenstraten, and J. Tipker, Development and Application of New Steric Substituent Parameters in Drug Design, in: Drug Design, **Vol. VII** (E. J. Ariens, ed.), Academic Press, London, New York 1976. For a comparison of steric parameters see also: T. Fujita and H. Iwamura, Applications of Various Steric Constants to QSAR, in: Topics in Current Chemistry **114** (F. L. Boschke, ed.), pp. 119–157, Springer, Berlin, Heidelberg 1983.
[13] L. B. Kier and L. H. Hall, Molecular Connectivity in Chemistry and Drug Research, Academic Press, London, New York 1976.
[14] S. M. Free and J. W. Wilson, J. Med. Chem. **7,** 395–399 (1964).
[15] Y. C. Martin, Quantitative Drug Design, Marcel Dekker, New York, Basel 1978.
[16] R. Franke, Optimierungsmethoden in der Wirkstoff-Forschung Quantitative Struktur-Wirkungs-Analyse, Akademie-Verlag, Berlin 1980.
[17] J. K. Seydel and K. J. Schaper, Chemische Struktur und biologische Aktivität von Wirkstoffen, Verlag Chemie, Weinheim, New York 1979.
[18] A. Cammarata, J. Med. Chem. **15,** 573–577 (1972).
[19] H. Kubinyi and O. H. Kerhahn, J. Med. Chem. **19,** 578–586 (1976).

[20] J. Deisenhofer, O. Epp, K. Miki, R. Huber, and H. Michel, J. Mol. Biol. **180,** 385–398 (1984).
[21] J. Deisenhofer, O. Epp, K. Miki, R. Huber, and H. Michel, Nature **318,** 618–624 (1985).
[22] H. Michel, O. Epp, and J. Deisenhofer, EMBO J. **5,** 2445–2451 (1986).
[23] H. Iwamura and T. Fujita, J. Pesticide Sci. **7,** 289–299 (1982).
[24] T. Fujita, Proc. Intern. Congr. Pesticide Chem. 1986, Ottawa, in press.
[25] W. Draber, K. Dickorè, K. H. Büchel, A. Trebst, and E. Pistorius, Naturwissenschaften **55,** 446 (1986); W. Draber, K. Dickorè, K. H. Büchel, A. Trebst, and E. Pistorius, in: Progress in Photosynthesis Research, **Vol. III** (H. Metzner, ed.), pp. 1789–1795, Tübingen 1969.
[26] W. Draber, H. Timmler, K. Dickorè, and W. Donner, Liebigs Ann. Chem. **1976,** 2206–2221.