Quantitative structure–activity relationships (QSARs) within the cytochrome P450 system: QSARs describing substrate binding, inhibition and induction of P450s

DAVID F. V. LEWIS *

School of Biomedical and Life Sciences, University of Surrey, Guildford, Surrey, GU2 7XH, UK

Received 10 May 2002; revised 12 June 2002; accepted 28 June 2002

Abstract—Quantitative structure-activity relationships (QSARs) within substrates, inducers and inhibitors of cytochromes P450 involved in xenobiotic metabolism are reported, together with QSARs associated with induction, inhibition and metabolic rate. The importance of frontier orbitals and shape descriptors, such as planarity (estimated by the area/depth² parameter) and rectangularity (estimated by the length/width parameter) is discussed, particularly in the context of the COMPACT system which discriminates between several P450 families associated with the activation and detoxication of xenobiotics. The use of parameters, particularly those derived from homology modelling of mammalian (especially human) P450s that are involved in exogenous metabolism, in generating QSARs for P450 substrates is discussed in the context of explaining differences in the binding affinities of human P450 substrates which are pharmacologically active.

Key words: Cytochromes P450; quantitative structure–activity relationships (QSAR); xenobiotic metabolism.

1. INTRODUCTION

The explanation of biological activity in terms of molecular structure, such that quantitative relationships (QSARs) can be formulated, clearly represents an important branch of science. In 1993, Hansch estimated that the US expenditure in this area exceeds 20 billion dollars per annum, whereas the total amount spent worldwide is probably about 40 billion dollars (Hansch, 1993). Although the Hansch approach, which relies on the importance of the lipophilicity parameter, log P (where P is the octanol/water partition coefficient), as a measure of a compound’s ability to reach the site of biological activity by undergoing transport across cell membranes, has been (and still is) highly successful in deriving QSARs (Hansch and

*E-mail: d.lewis@surrey.ac.uk
Leo, 1979; Kubinyi, 1997; Sarver et al., 1997; Vaes et al., 1998) for many series of compounds across a huge range of biological activities, it is also apparent that other factors contribute to overall activity especially those involving actual binding to the target macromolecule (e.g. nucleic acid, protein or enzyme). Hansch and coworkers also considered electronic and steric factors in biological activity by the use of substituent parameters derived from the Hammett and Taft equations (reviewed in Lewis, 1990). Modern QSAR techniques utilize a whole range of structural descriptors in formulating correlations with various forms of biological activity (Livingstone, 2000).

However, the log \( P \) parameter provides a means of estimating the desolvation component of the free energy of binding for a receptor ligand or enzyme substrate via direct calculation of the partitioning energy, \( \Delta G_{\text{part}} \), given the expression:

\[
\Delta G_{\text{part}} = -RT \ln P,
\]

where \( R \) is the gas constant, \( T \) is the absolute temperature and \( P \) is the octanol–water partition coefficient (Lewis et al., 1998). It would appear that the use of log \( P \) data in this way gives rise to an acceptable value for the desolvation component of the overall binding free energy, and is frequently found to be the major contribution to enzyme-substrate binding in the P450 system (Lewis, 2002). However, one possible way of distinguishing between the two roles of membrane transport and desolvation for log \( P \) contributions lies in the nature of the QSAR expression itself, because a quadratic relationship in log \( P \) is indicative of transport and a linear equation in log \( P \) tends to suggest purely desolvation of the binding site. Such distinctions can be helpful in analyzing the results of QSAR studies to determine the important aspects of biophysical chemistry in the changes in activity across a series of structurally related compounds under scrutiny.

Although databases of measured log \( P \) values exist (Hansch and Leo, 1979; Leo, 1993; Sangster, 1989) it is sometimes necessary to calculate log \( P \), and computer programs are available for generating calculated log \( P \) values, such as the Pallas System (Csizmadia et al., 1997; Tsantili-Kakoulidou et al., 1997). It is possible to calculate log \( P \) values from QSAR analysis although, strictly, this should be referred to as QSPR (where the acronym stands for quantitative structure-property relationship). For example, in a series of nine endogenous and synthetic steroids it is found that experimental log \( P \) data can be accurately reproduced (with a correlation of 0.99) by a linear combination of the number of hydrogen bond donors and acceptors (\( N_A \), \( N_D \)) and number of hydrocarbon atoms (\( N_{\text{CH}} \)) as follows:

\[
\log P = 0.0658N_{\text{CH}} - 0.459N_A - 0.311N_D + 1.391,
\]

\( n = 9; \) \( s = 0.1497; \) \( R = 0.9896; \) \( F = 78.90, \)

where \( s \) is the standard error, \( R \) is the correlation coefficient and \( F \) is the variance ratio. It can be readily appreciated that the number of hydrocarbon atoms make a positive contribution to log \( P \), whereas the hydrogen bond donor/acceptor atoms negatively influence the overall log \( P \) value, thus leading to the basic principle