Quantitative Structure-Chromatographic Retention Relationships

Roman Kaliszan

Volume #93 in Chemical Analysis: A Series of Monographs on Analytical Chemistry and Its Applications—J.D. Winefordner, Editor; I.M. Kolthoff, Editor Emeritus

Quantitative Structure–Chromatographic Retention Relationships

ROMAN KALISZAN

Medical Academy
Gdańsk, Poland

A WILEY-INTERSCIENCE PUBLICATION

JOHN WILEY & SONS

New York / Chichester / Brisbane / Toronto / Singapore

Library of Congress Cataloging in Publication Data:

Kaliszan, Roman.
Quantitative structure–chromatographic retention relationships.

(Chemical analysis; v. 93)
"A Wiley-Interscience publication."
Includes bibliographies and index.
1. Chromatographic analysis. 2. Solution (Chemistry) I. Title. II. Series.

QD117.C5K35 1987 543'.089 87–14802
ISBN 0–471–85983–4

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

CHEMICAL ANALYSIS

A SERIES OF MONOGRAPHS ON
ANALYTICAL CHEMISTRY AND ITS APPLICATIONS

VOLUME 93

A WILEY-INTERSCIENCE PUBLICATION

JOHN WILEY & SONS

New York / Chichester / Brisbane / Toronto / Singapore

Quantitative Structure–Chromatographic Retention Relationships

In memory of my father

PREFACE

The quantitative relationships between the structure of solutes and their chromatographic retention data have been extensively studied recently for three main reasons: explanation of the mechanism of chromatographic separations; prediction of retention coefficients; and characterization of the solute physicochemical properties of importance for reactivity and especially for bioactivity. Thus, this research area is truly an interdisciplinary one, and publications may be found in chromatographic, analytical, physicochemical, biochemical, pharmacological, and pharmaceutical journals.

At constant temperature, three main variables determine the distribution of a solute between mobile and stationary chromatographic phases: the chemical structure of the solute, the physicochemical properties of the mobile phase, and the physicochemical properties of the stationary phase. The solute distribution in various chromatographic techniques and modes is easily quantified by means of several free-energy-related retention parameters.

For any given solute the relationships between retention data and the composition of mobile and/or stationary phase have early been observed and met a more or less rigorous theoretical treatment based on chemical thermodynamics. Except for simple homologous series chromatographed under identical conditions, however, attempts to relate quantitatively the structure of an individual solute to its retention parameter were unsuccessful.

There has been an unquestionable trend in chemistry for the past few decades toward quantitation of chemical, physicochemical, and biological activities of various compounds. When computers became commonly available in the 1960s, the studies flourished because of pioneering works by Hansch and others on quantitative structure–biological activity relationships (QSARs). The QSAR methodology, that is, the means of characterizing solute molecular structure numerically, and the statistical procedures applied or developed for QSAR purposes have been successfully employed for quantitative structure–retention relationship (QSRR) studies. Since the late 1970s hundreds of papers have been published that may be categorized under the term QSRR. Certainly now, after about 10 years of intensive development, QSRR studies deserve a thorough review and critical discussion. Such an attempt is undertaken here.

Among the numerous QSRR equations published, there are many with little or no information value. Some may even be misleading if statistical requirements were not fulfilled when the appropriate equations were derived.

Nonetheless, the vast majority of the QSRR studies yielded results of importance for physical, analytical, and medicinal chemistry. From these points of view the collected data are discussed here.

Chromatography is a unique system for studying structure–activity relationships involving intermolecular interactions. In a chromatographic process all the conditions may be kept constant or controlled, and thus the solute structure is the single independent variable in the system. Contrary to biological determinations, chromatography is able to yield readily a great amount of unequivocal, precise, and reproducible data. It may be anticipated that through QSRR studies the still more precise methods of solute structure parameterization will be worked out, which will next be applied to derive reliable QSAR equations allowing rational design of new drugs.

A knowledge of the physicochemical principles of solute–stationary/mobile phase interactions is required for proper understanding of the basis of chromatographic separations. Therefore, a description of the intermolecular interactions known in chemistry is presented first. After this, a chapter is devoted to the most commonly known general theories of the chromatographic distribution process. Then follows a brief characterization of the factors influencing the retention data for an individual solute. Next, the mathematical models employed in QSRR studies are reviewed and their formal requirements are discussed. In the following chapters the reported QSRR are reviewed successively in terms of individual molecular structure descriptors of the solutes. In another chapter the relationships between liquid chromatographic retention and partition coefficients are critically discussed. Finally, application of QSRRs in medicinal chemistry is reviewed in detail.

To help the reader use the literature cited, complete titles of the articles quoted are given. The literature up until March 1986 has been considered.

I am very grateful to Professor Richard A. Hartwick, Rutgers University, Piscataway, New Jersey, who encouraged the writing of this monograph. Thanks are also due to the authors and publishers of copyrighted materials. Without the patience of my wife, Anna, this book would not have been written.

ROMAN KALISZAN

Gdańsk, Poland
1987

Quantitative Structure–Chromatographic Retention Relationships

CONTENTS

CHAPTER

1

INTRODUCTION

The beginning studies of quantitative relationships between the structure of solutes and their chromatographic retention may be dated back as early as 1949. At that time, Martin [1], in his fundamental paper, suggested that a substituent changes the partition coefficient of a solute by a factor that depends on the nature of the substituent and both the mobile and stationary phases employed but not on the remaining part of the molecule. Since practically the birth of chromatography, regularities have been observed of retention behavior among the more or less closely related solutes. Evident have been simple relationships between chromatographic parameters and, for example, carbon number for a series of homologues.

Following Green et al. [2], who found that substituent increments to the thin-layer chromatographic parameter $R_M = \log(1/R_f - 1)$ are additive for a number of benzenoid compounds, Iwasa et al. [3] suggested in 1965 that chromatographic data for studies of quantitative structure–biological activity relationships (QSAR) may be useful. Since that time chromatography has been extensively employed for quantitation of hydrophobicity of bioactive agents [4–7].

In 1977 publications [8–10] appeared in which the QSAR methodology was directly applied for analysis of chromatographic retention data of a series of solutes. Soon the number of structure–retention correlations in the literature started to increase exponentially with time. This has been the result of common accessibility of personal computers as well as of the appropriate programs of statistical calculations. By analogy to QSAR, the term quantitative structure–retention relationships (QSRR) has been proposed [11] to comprise the new expanding area of chromatographic science.

The chromatographic conditions can be modified in several ways. One can change the stationary and/or mobile phase, and the changes in the relative retentions of investigated compounds caused by these modifications can be treated as a source of information about the ability of the compounds to undergo different kinds of intermolecular interactions on a thermodynamic basis [12]. Another approach consists in selecting a suitable group of test compounds for which chromatographic data would be determined at constant conditions or the numerical retention parameters obtained would be normalized to some standard conditions. In that second approach the differences in

chromatographic data reflect the differences in the solute structure. Relationships between chromatographic retention data and the quantities related to the solute structure cannot be solved in strict thermodynamic terms. Such relationships are of the so-called *extrathermodynamic* type.

As Prausnitz has remarked [13], "Classical thermodynamics is revered, honoured and admired, but in practice is inadequate." Prausnitz has suggested the use of molecular thermodynamics for solving real problems—with molecular thermodynamics being seen as a synthesis of classical approaches, statistical thermodynamics, molecular physics, and physical thermodynamics. One way in which such results can be achieved is to employ *extrathermodynamics.*

The term extrathermodynamics means that the science lies outside the formal structure of thermodynamics, although the approach resembles that of thermodynamics in that detailed microscopic mechanisms do not need to be explicitly identified during use [14]. Extrathermodynamic approaches are combinations of detailed models with the concepts of thermodynamics. Since it involves model building, this kind of approach lacks the rigor of thermodynamics, but it can provide information not otherwise accessible. The manifestations of extrathermodynamic relationships are the linear free-energy relationships (LFER). Although LFERs are not a necessary consequence of thermodynamics, their occurrence suggests the presence of a real connection between the correlated quantities, and the nature of this connection can be explored [15].

Generally, LFERs may be regarded as linear relationships between the logarithms of the rate or equilibria constants for one reaction series and those for a second reaction series subjected to the same variation in reactant structure or reaction conditions [16, 17]. Thus, plotting the logarithms of rate or equilibrium constant for one reaction series against the corresponding constants for a second related series frequently gives a straight line, which can be expressed by

$$\log k_i^{\mathrm{B}} = m \log k_i^{\mathrm{A}} + c \tag{1.1}$$

where k_i^{A} and k_i^{B} are rate or equilibrium constants of two reaction series A and B that are subject to the same changes in the structure of reactants or the surrounding medium. It is often convenient to express LFER in terms of ratios of constants by referring all members of a reaction series to a reference member of the series; thus, the correlation in Eq. (1.1) can also be expressed by

$$\log (k_i^{\mathrm{B}}/k_0^{\mathrm{B}}) = m \log (k_i^{\mathrm{A}}/k_0^{\mathrm{A}}) \tag{1.2}$$

where k_0^{A} and k_0^{B} are the constants for the reference substituent or the reference solvent.

The chromatographic retention parameters used in correlation studies are normally assumed to be proportional to the free-energy change associated with the chromatographic distribution process. Not all chromatographic data, however, are suitable for QSRR studies. As is well known, free-energy changes, ΔG, are related to enthalpy, ΔH, and entropy, ΔS, changes by the Gibbs equation

$$\Delta G = \Delta H - T\,\Delta S \tag{1.3}$$

where T is temperature. For LFERs to be found between real and model systems, changes either in entropy or enthalpy must be constant, or the enthalpy changes must be linearly related to entropy changes [18]:

$$\Delta H = \beta\,\Delta S + \Delta G_\beta \quad (\text{at } T = \beta) \tag{1.4}$$

When enthalpy–entropy compensation is observed with a family of compounds in a particular chemical transformation, the values of β and ΔG are invariant and β is called the compensation temperature.

Using the Gibbs relationships [Eq. (1.3)], one can rewrite Eq. (1.4) in order to express the free-energy change ΔG_T measured at a fixed temperature T for isoequilibrium process as

$$\Delta G_T = \Delta H\left(1 - \frac{T}{\beta}\right) + \frac{T\,\Delta G_\beta}{\beta} \tag{1.5}$$

In liquid chromatography the retention parameter, the so-called capacity factor k', is related to the thermodynamic equilibrium constant K for solute binding by $k' = \phi K$, where ϕ is the phase ratio of the column. The free-energy change for the chromatographic process is expressed by

$$\Delta G = -RT \ln K = -RT \ln(k'/\phi) \tag{1.6}$$

where R is a gas constant. As shown by Melander et al. [18], the substitution of Eq. (1.6) into Eq. (1.3) yields, for the capacity factor:

$$\ln k' = -\frac{\Delta H}{RT} + \frac{\Delta S}{R} + \ln \phi \tag{1.7}$$

[5] If the mechanism of the process is invariant over the temperature range studied and the enthalpy is constant, a van't Hoff plot of $\ln k'$ against $1/T$ yields a straight line. From the slope of the line, the enthalpy change ΔH for a given solute can be assessed.

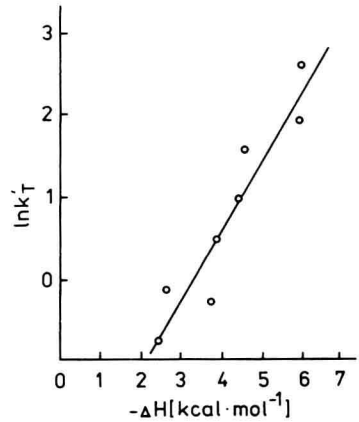

Figure 1.1. Enthalpy–entropy compensation plot for a group of aromatic acids chromatographed in different reversed-phase high-performance liquid chromatographic systems. (After W. Melander, D. E. Campbell, and Cs. Horváth, *J. Chromatogr.*, **158**, 213, 1978. With permission.)

Equations (1.5) and (1.6) can be combined to give

$$\ln k'_T = -\frac{\Delta H}{R}\left(\frac{1}{T}-\frac{1}{\beta}\right)-\frac{\Delta G_\beta}{R\beta}+\ln\phi \tag{1.8}$$

where k'_T is the capacity factor at temperature T. According to Eq. (1.8), plots of $\ln k'_T$ of various solutes measured at a given temperature T under different conditions against the corresponding enthalpy change are linear when the enthalpy–entropy compensation occurs (Fig. 1.1). In such a situation, the reversible binding of the solutes by a stationary phase involves essentially the same mechanism. To avoid statistical artifacts, it is recommended that the reference temperature T in Eq. (1.8) be near the harmonic mean of the experimental temperatures used for the evaluation of the enthalpies by Eq. (1.7) [18–20].

From the slope of compensation plot of $\ln k'_T$ versus ΔH [Eq. (1.8)], the compensation temperature β may be obtained. If values of β for different chromatographic systems approach their 95% confidence limits, the retention mechanism is assumed to be the same [18, 21, 22].

It should be observed here that the prevailing majority of QSRRs reported concern retention parameters as obtained in routine chromatographic measurements. Enthalpy–entropy compensation is checked only occasionally. Reservations are often justified as far as the chromatographic retention parameters are defined and/or determined for individual solutes at specified

conditions. Unfortunately, for many authors the only criterion of value of the QSRR equations derived is the more or less exact conformability of the observed and calculated data. Even that conformability is sometimes treated quite mechanically; that is, the statistical significance of the statistically derived relationships is not analyzed. Nonetheless, certainly enough material has been collected to make a critical discussion worthwhile. It may be anticipated that studies on QSRR will further expand, and the aim of the following chapters is to help the interested reader clarify the present status of QSRR science and to encourage deeper, critical, and scientificaly productive analysis of chromatographic data.

References

1. A. J. P. Martin, Some theoretical aspects of partition chromatography, *Biochem. Soc. Symp.*, **3**, 4, 1950.
2. J. Green, S. Marcinkiewicz, and D. McHale, Paper chromatography and chemical structure. III. The correlation of complex and simple molecules. The calculation of R_M values for tocopherols, vitamins K, ubiquinones and ubichromenols from R_M (phenol). Effects of unsaturation and chain branching, *J. Chromatogr.*, **10**, 158, 1963.
3. J. Iwasa, T. Fujita, and C. Hansch, Substituent constants for aliphatic functions obtained from partition coefficients, *J. Med. Chem.*, **8**, 150, 1965.
4. E. Tomlinson, Chromatographic hydrophobic parameters in correlation analysis of structure–activity relationships, *J. Chromatogr.*, **113**, 1, 1975.
5. R. Kaliszan, Chromatography in studies of quantitative structure–activity relationships, *J. Chromatogr.*, **220**, 71, 1981.
6. R. Kaliszan, High performance liquid chromatography as a source of structural information for medicinal chemistry, *J. Chromatogr. Sci.*, **22**, 362, 1984.
7. R. Kaliszan, Quantitative relationships between molecular structure and chromatographic retention. Implications in physical, analytical, and medicinal chemistry, *Crit. Rev. Anal. Chem.*, **16**, 323, 1986.
8. R. Kaliszan and H. Foks, The relationship between the R_M values and the connectivity indices for pyrazine carbothioamide derivatives, *Chromatographia*, **10**, 346, 1977.
9. R. Kaliszan, Correlation between the retention indices and the connectivity indices of alcohols and methyl esters with complex cyclic structure, *Chromatographia*, **10**, 529, 1977.
10. I. Michotte and D. L. Massart, Molecular connectivity and retention indexes, *J. Pharm. Sci.*, **66**, 1630, 1977.
11. B.-K. Chen and Cs. Horváth, Evaluation of substituent contributions to chromatographic retention: Quantitative structure–retention relationships, *J. Chromatogr.*, **171**, 15, 1979.