ELSEVIER

Journal of Chromatography A, 808 (1998) 277–278

JOURNAL OF CHROMATOGRAPHY A

Discussion

Comments on predictive strategies for determining retention indexes

C.T. Peng

Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, CA 94143-0446, USA
Nuclear Science Facility, San Jose State University, San Jose, CA 95192-0163, USA

Received 2 December 1997

Keywords: Group retention factors; Structure–retention relationships; Retention indices; Allylic alcohols

Simpson and Jackson [1] in their study of predictive strategies for determining retention indexes of some allylic alcohols and linalyl and geranyl esters by gas chromatography have relied on the discrepancies, shown in Table 1, between the $(GRF)_{cp}$ and $(GRF)_{p}$ values to support their proposed retention mechanism (GRF=Group retention factor). The $(GRF)_{cp}$ values were either calculated from an equation or taken from our published data and the $(GRF)_{p}$ values were from the intercepts of regression equations of the plots of observed retention index (I_{obs}) versus the number of atoms (Z) of the four allylic alcohols. A brief outline of the concept on which our method [2,3] is based, is necessary for the discussion on the significance of the GRF differences from these sources.

Numerous methods for retention index (I) prediction are given in the literature [4]. The system that we have proposed for I prediction is an algorithm using the Kovats index to account for retention contribution from the atoms and the GRF for contribution from functionalities. The purpose is to reduce a complex organic compound step-by-step structurally to a normal alkane of equal number of carbon atoms. An analyst can then count the number of C, O, and N atoms in an organic molecule multiplied by 100 to give the base value and add to that the GRF values of the functionalities in the molecule to obtain the predicted I. The definition of functionality is broad, including functional groups, such as alcohol, ester groups, etc. as well as structural differences from normal alkane, such as tertiary and quaternary carbons atoms, *cis–trans* configurations, ring formation, etc. to mention just a few. The observed I (I_{obs}) is a function of A and GRF, in addition to Z. All the (GRF) values in our system are obtained by adjusting the regression coefficient (A) to 100 index units (i.u.). The method can accurately predict the I values of monofunctional compounds but is much less accurate for multi-functional compounds. Since the multiple functional groups can interact with each other to affect the retention of the molecule, accurate GRFs for individual functional groups and functionalities, under such situations, are difficult to obtain. In view of the above, the following points in Simpson and Jackson's publication are misleading:

(i) Simpson and Jackson compared the $(GRF)_{cp}$ and $(GRF)_{p}$ values listed in Table 1 to support their hypothesis. The $(GRF)_{cp}$ values were reputedly from our published data which were obtained from regression equations by normalizing the regression coefficient (A) to 100, whereas the $(GRF)_{p}$ values were taken from the regression equations without any adjustment. If one had normalized the A value in these authors' Eqs. 5a and 6a, by changing 98.90 on the polar column and 107.50 on the nonpolar column to 100, the adjusted intercept or the total GRF value would have been 726 for the polar column and 144

0021-9673/98/$19.00
PII: S0021-9673(97)01267-3

for the nonpolar column; these recalculated values would then have been much closer to our published values than those given in Table 1. It is not generally meaningful to compare the intercept or the GRF value of one regression equation with that of another for the same homologous series without comparing at the same time the value of the coefficient (*A*). The residual difference between the recalculated and our published values could be attributed to the purity of the column stationary phases. That means, the polar column used by Simpson and Jackson is more polar and their nonpolar column less polar than ours. Had these authors reported along with their compounds the I_{obs} values of some simple reference compounds, such as benzene, butanol or hexanol, ethyl hexanoate or any monofunctional compounds that are in our published data, we would be in a better position to compare the polarity between their column and ours to determine whether the discrepancy is real.

(ii) Simspon and Jackson showed in Tables 4 and 5 their calculated I_{cp} and I_{p} values for linalool, nerol and geraniol. The authors included in their calculation the tertiary carbon and *cis–trans* configuration in the molecule of linalool, but for no apparent reason left these functionalities out in nerol and geraniol. Had the (GRF) contributions for tertiary and quaternary carbons and *cis–trans* configuration in these molecules been included as they should be, the % difference between their I_{obs}, I_{p} and I_{cp} values would have been considerably smaller than those shown in Tables 4 and 5. The results in Tables 2 and 3 are difficult to evaluate because the details of calculation for the I_{cp} and I_{p} values are not given.

(iii) Simpson and Jackson formulated Eq. (4) in their paper and attributed it to our 1988 publication [2], but I was unable to locate a replica of that equation in the cited publication or in our other publications [3,5]. We have recommended the use of column difference (ΔI) for the identification of unknown compounds [5], but never recommended it for *I* prediction. Column difference (ΔI) is defined as the difference between RIs on a polar and a less polar column [5] and can only be obtained when the *I* values on both columns are known. We stated under the heading *Carboxylic acid esters* that: ''On nonpolar SE-30 column the esters of the fatty acids were found to behave chromatographically in the same way as aliphatic hydrocarbons [1]. The residual polarity and polarizability of the acid ester group may cause additional retention on CW-20M column. The *GRF* on polar column for this molecular moiety is equal to the column difference (ΔI). The methyl, ethyl, propyl and butyl esters have column differences of +294, +270, +259 and +256 units, respectively. This value must be added to the base value to yield the predicted *I* for the ester on polar column.'' [3]. The GRF value for the ester group on nonpolar SE-30 column is essentially zero, thus allowing the value of the column difference (ΔI) to be used as the GRF value for the ester group on polar column; it is a special case. Simpson and Jacksons' generalization of the use of column difference (ΔI) and its formulation into Eq. (4) where GRF should be, is not justified in theory nor valid in practice for *I* prediction of esters on polar columns, when the column difference (ΔI) is between a polar and a less polar column, and the less polar column has a non-zero value of GRF for the ester group.

References

[1] C.I.C. Simpson, Y.A. Jackson, J. Chromatogr. A. 766 (1997) 141–146.
[2] C.T. Peng, S.F. Ding, R.L. Hua, Z.C. Yang, J. Chromatogr. 436 (1988) 137–172.
[3] C.T. Peng, Z.C. Yang, S.F. Ding, J. Chromatogr. 586 (1991) 85–112.
[4] M.V. Budahegyi, E.R. Lombosi, T.S. Lombosi, S.Y. Mészáros, Sz. Nyiredy, G. Tarján, I. Timár, J.M. Takács, J. Chromatogr. 271 (1983) 213–307.
[5] C.T. Peng, J. Chromatogr. A 678 (1994) 189–200.