fraction was allowed to evaporate at room temperature and the residue was dissolved in 2 ml of ethanol for measurement of ultraviolet absorption. For estimation of these steroids with Porter-Silber reagent, the residue was dissolved in 0.5 ml of 60 % aqueous ethanol and then 1.0 ml of Porter-Silber reagent was added and mixed. The mixture was allowed to stand at room temperature for 15 h and the optical density was measured at 410 mμ. The recovery of steroids was 75 to 80 %. The elution pattern is shown in Fig. 1. The column could be used repeatedly.

Since the solutes distributed themselves between the less polar outer liquid phase and the more polar swollen stationary phase which is in equilibrium with the liquid phase, the elution sequence of 17-hydroxycorticosteroids resembled that in straight phase partition chromatography. This type of partition chromatography, which utilizes synthetic cross-linked polymer of medium polarity as the stationary phase and a less polar solvent mixture consisting of smaller amounts of polar components (e.g., ethanol and water) and larger amounts of less polar components (e.g., benzene and n-hexane) as the moving phase, seems to be useful for the separation of neutral organic compounds of moderate and higher polarity.

Department of Genetics, Osaka University Medical School,   TOKUICHIRO SEKI
Osaka (Japan)

1 O. SAMUELSON AND B. SWENSON, Anal. Chim. Acta, 28 (1963) 426.
2 T. SEKI, J. Chromatog., 22 (1966) 498.
3 T. SEKI AND K. MATSUMOTO, Endocrinol. Japon., 13 (1966) 75.
4 T. SEKI AND K. MATSUMOTO, J. Chromatog., 27 (1967) 423.
5 E. NYSTROM AND J. SJÖVALL, Anal. Biochem., 12 (1965) 235.
6 R. VIHKO, Acta Endocrinol., Suppl., 109 (1966) 15.
7 P. DECKER AND H. HOLLER, J. Chromatog., 7 (1962) 392.

The prediction of R_F values

Several theoretical studies have appeared in the chromatography literature. A large number of specific solutions to particular problems are also available. Very few studies have been made using both approaches simultaneously. A plea is made here for more studies applied to the practical problems of the selection of solvents, the determination of the conditions of chromatography, and the presentation of data.

Chromatography is essentially a stochastic process and the statistical knowledge of estimation and prediction could be applied with success here. The aim of this investigation is to show the possibilities of predicting the R_F value of a compound, given a set of conditions.

The physical factors influencing R_F values have been investigated quite extensively on filter paper[1-4]. CONSDEN, GORDON AND MARTIN[4] called attention to the factors paper, temperature, quantity of material, extraneous substances, degree of

saturation with water, supply of solvent, and distance between starting point and source of solvent. CASSIDY[1] reviewed these factors and others at the bulk and molecular levels. Contradictory data are reported in the literature, probably due to the fact that the factors investigated are very limited.

No similar study has been made on thin-layer chromatoplates. The possibility of studying the support by using different adsorbents or binary mixtures of adsorbents has created a new dimension in chromatography[5]. The article analyzes some of the factors affecting and explaining the behavior of compounds on thin-layer chromatoplates.

Material

Thin-layer plates were prepared by mixing MN Cellulose 300 and Silica Gel G in known proportions in water and applying to a thickness of 0.25 mm with the Desaga-Brinkman* applicator. The plates were dried overnight at room temperature to insure reproducibility. The compounds studied were iodine-containing compounds of medical interest: 3-monoiodotyrosine (MIT), 3,5-diiodotyrosine (DIT), 3,5,3'-triiodothyronine (T3), 3,5,3',5'-tetraiodothyronine (T4) and iodide (I⁻). The solvent system used was *tert.*-amyl alcohol–$1 N$ ammonium hydroxide–dioxane ($2:1:2$). Following a 10 cm run in a saturated tank, iodide was visualized by spraying with 0.5 % palladium (II) chloride in $1 N$ hydrochloric acid and the tyrosines and thyronines by spraying with 0.5 % diazosulfanilic acid in 5 % sodium carbonate. The temperature was kept constant ($\pm 1°$) by placing the tank in a constant temperature bath**.

Method

The mathematical technique will be presented in some detail hoping that others will follow this method of analysis.

We are presenting an analysis of the relations among a single criterion measure and two or more predictor measures. The technique of multiple regression is applied to predict the hR_F ($R_F \times 100$) value of a compound from *a priori* knowledge of the conditions of chromatography.

In the general case of an m-dimensional space, the standard regression equation has the form:

$$\hat{z}_{mi} = \beta_1 Z_{1i} + \beta_2 Z_{2i} + \beta_3 Z_{3i} + \ldots + \beta_{m-1} Z_{(m-1)i}$$

The beta weights indicate the relative contribution of the corresponding variables to the prediction of the criterion. The Z's are defined as $Z_{1i} = (X_{1i} - \bar{X}_1/\hat{\sigma}_1)$. It is often desirable to compute the regressed deviation score $\hat{x}_{mi}$ instead of the regressed standard score $\hat{z}_{mi}$; in this case, the b weights are derived from the beta weights by scaling each beta weight by the ratio of the standard deviation of the criterion to the standard deviation of that predictor ($b_j = (\hat{\sigma}_m/\hat{\sigma}_j)\beta_j$). It is possible to compute a multiple correlation coefficient $R_{m.1,2,\ldots,m-1}$ in the range $0 \leq R \leq + 1$. The coefficient R^2 provides an estimate of the proportion of the total variance in the criterion that can be predicted from the known variance in the predictors and is a measure of the overall

---

* Brinkman Instruments, Westbury, New York.
** Forma Scientific Inc., Marietta, Ohio.

effectiveness of the multiple regression. The significance of R^2 is tested by the analysis of variance as follows:

$$F^{m-1}_{N-m-2} = \frac{R^2(N-m-2)}{(1-R^2)(m-1)}$$

where N is the number of subjects in the sample. The standard error of estimate for $\hat{Z}_m$ is given by $\sqrt{1-R^2}$.

The criterion is the hR_F value of a compound. The set of predictors are listed in Table I. They touch upon the support (adsorbent), the solvent (developing time), the environment (temperature) and the compound (molecular weight and number of iodine atoms in the compound).

TABLE I

SET OF PREDICTORS

Z	Description
1	Adsorbent (% Silica Gel G)
2	Temperature (°C)
3	Molecular weight of compound
4	Number of iodine atoms in the molecule
5	Developing time (time for solvent to reach 10 cm)

A stepwise procedure which adds one variable to the prediction equation at a time and thus provides a number of intermediate regression equations is possible. Variables are added or dropped according to the statistical significance of their contribution to the prediction of the criterion. In this case any effort to generalize from sample to population is open to serious danger of capitalization on chance.

All computations were performed at the Harvard Computation Center on the IBM 7094 Digital Computer. The routine for multiple regression was as published by COOLEY AND LOHNES[6]. The stepwise multiple regression program was taken from the share library (SDA 3143).

Results

The results of the correlation analysis on 225 sets of data are summarized in Table II. Apart from the trivial correlation between the number of iodine atoms and

TABLE II

CORRELATION MATRIX

Variables	1	2	3	4	5	6
1	—					
2	0.02	—				
3	0.00	0.00	—--			
4	0.00	0.00	0.96*	—		
5	0.06	—0.95*	0.00	0.00	—	
6	—0.06	—0.26*	0.38*	0.50*	0.25*	—

* $p < 0.001$.

the molecular weight of a compound, we have a high correlation between the temperature and the developing time. Significant correlation coefficients are observed between hR_F value and all other variables except the percent silica gel in the adsorbent, the most important factors in predicting the R_F value of a compound being the number of iodine atoms in the molecule and its molecular weight. This does not mean that there is no relation between the hR_F values of iodine-containing compounds and the modifications of the adsorbent. Very high correlations are observed for a specific compound especially with T_4, T_3 and iodide. The lumping of all the data together is the reason for the vanishing of the high correlation between hR_F and percent silica gel.

The results of the multiple regression analysis are presented in Table III. The importance of each factor (the betas of the regression equation) in the prediction of the hR_F values is given. The coefficient R^2 is 0.4632 ($R = 0.6806$). This coefficient is a measure of the effectiveness of our equation in predicting R_F values from a knowledge of the conditions of chromatography.

Following a stepwise multiple regression analysis, variables 2, 3 and 4 were retained as significant and are quoted in Table IV with their respective b weights and standard error.

TABLE III

SUMMARY OF MULTIPLE REGRESSION ANALYSIS

Variables	$\overline{X}$	S.D.	β weights	b weights
1	53.33	36.11	—0.07	—0.03
2	22.73	11.17	—0.13	—0.22
3	458.99	233.75	—1.43	—0.11
4	2.20	1.17	1.88	29.77
5	103.44	37.06	0.13	0.07
6	36.66	18.50	—	—

In spite of a high coefficient of determination, the solution to the multiple regression equation is not a unique one[7]. As can be seen from the stepwise multiple regression analysis, when variables 2 and 5 are eliminated from the set of contributing factors some of the weights change (compare Tables III and IV). The squared multiple correlation coefficients of these two equations are quite different: it was 0.46 for the complete equation and 0.68 when only the significant variables were included. The efficiency of prediction is different for both equations.

TABLE IV

SUMMARY OF STEPWISE REGRESSION ANALYSIS

Variables	Coefficient(b)	Standard error of the coefficient
2	—0.43	0.08
3	—0.11	0.01
4	29.77	2.97

Discussion

Very few studies have dealt with the problem of predicting R_F values. CONNORS[8] was able to predict the R_F values of a set of compounds in one solvent given this knowledge in a different solvent. To our knowledge, this is the only investigation oriented toward the prediction of R_F values.

Let us not be distracted by the mathematics from the more important chemical observations and let us summarize with this idea in mind.

When we say 68 % of the variance in hR_F is predicted from a knowledge of temperature, molecular weight and number of iodine atoms in the compounds, we mean that this is a measure of the overall effectiveness of our predictive equation. The *b* weights indicate that the number of iodine atoms in the compounds is the most useful predictor. Since in only 68 % of the cases can we predict the R_F value accurately, other factors are involved in determining the R_F values of a given compound; and temperature, molecular weight and number of iodine atoms in the compound are the significant predictors. All other variables did not contribute significantly to the prediction of R_F values. These other factors need to be uncovered before a very accurate prediction can be achieved. Some of these factors might be solubility of the compounds in the solvent, elution strength of the solvent, viscosity, specific gravity and surface tension of the solvent. The most important factors affecting the behavior of chromatographed compounds, such as partition, adsorption and ion exchange, are difficult to measure and express numerically. Only when we can quantify these concepts will it become possible to predict R_F values.

A most interesting aspect of this procedure is the possibility of predicting the R_F value of any member of a chemical series, given the numerical values for the set of conditions of chromatography. Unfortunately the material on which we have accumulated a large body of data comprises a very small series and forbids any attempt at prediction. Amino acids or sugar series would be ideally suited for such a study, and would allow an estimate of the accuracy of prediction. A large body of data accumulated under standardized specific conditions is needed before the usefulness of the technique described can be fully evaluated.

Physics Research Laboratory,   R. P. OUELETTE*
Massachusetts General Hospital,   J. F. BALCIUS
Boston, Mass. 02114 (U.S.A.)

1 H. C. CASSIDY, *Fundamentals of Chromatography*, Interscience, New York, 1957.
2 A. SAIFER AND I. ORESKES, *Anal. Chem.*, 25 (1963) 1539.
3 G. N. KOWKABANY AND H. G. CASSIDY, *Anal. Chem.*, 24 (1952) 643.
4 R. CONSDEN, A. H. GORDON AND A. J. P. MARTIN, *Biochem. J.*, 41 (1947) 590.
5 R. P. OUELETTE AND J. F. BALCIUS, *J. Chromatog.*, 24 (1966) 465.
6 W. W. COOLEY AND P. R. LOHNES, *Multivariate Procedures for the Behavioral Sciences*, Wiley, New York, 1962.
7 M. G. KENDALL, *A Course in Multivariate Analysis*, Hafner, New York, 1961.
8 K. A. CONNORS, *Anal. Chem.*, 37 (1965) 261.

* Present address: Hazleton Laboratories, Inc., P.O. Box 40, Falls Church, Va. 22046, U.S.A.