Journal of Chromatography, 485 (1989) 349–363
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 21 923

PREDICTION OF RETENTION OF METABOLITES IN HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY BY AN EXPERT SYSTEM APPROACH

K. VALKÓ

Central Research Institute for Chemistry, Hungarian Academy of Sciences, P.O.B. 17, H-1525 Budapest (Hungary)

G. SZABÓ

CompuDrug Ltd., Fürst S. u. 5, H-1136 Budapest (Hungary) and CompuDrug USA, P.O. Box 202078, Austin, TX 78720 (U.S.A.)

J. RÖHRICHT

Chemical Works of Gedeon Richter Ltd., P.O. Box 27, H-1475 Budapest (Hungary)

K. JEMNITZ

Central Research Institute for Chemistry, Hungarian Academy of Sciences, P.O.B. 17, H-1525 Budapest (Hungary)

and

F. DARVAS*

CompuDrug Ltd., Fürst S. u. 5, H-1136 Budapest (Hungary) and CompuDrug USA, P.O. Box 202078, Austin, TX 78720 (U.S.A.)*

SUMMARY

An expert system (HPLC-METABOLEXPERT) has been developed for the simultaneous prediction of the metabolites and the retention data of an organic compound. Prediction of the retention time in reversed-phase high-performance liquid chromatography (RP-HPLC) is provided by the system for an RP-18 column together with advice concerning optimum eluent concentration, pH and expected change in the UV spectra. Prediction is performed using the RP-HPLC data for the parent compound and empirical correction rules based on the structural difference between the parent compound and the metabolites.

The predictive power of the expert system was investigated in this study for seven molecules and their eight main metabolites by measuring their reversed-phase retention data using various reversed-phase columns and conditions. Among the main metabolic transformation routes, N-demethylation and C-hydroxylation were investigated in more detail. The average difference between the measured and predicted retention times was 1.8 min.

INTRODUCTION

Identifying metabolites of drugs, pesticides and their residues has growing importance in medicinal chemistry, agricultural and environmental chemistry[1]. During

0021-9673/89/$03.50

the past few years, high-performance liquid chromatography (HPLC) has acquired a major role in the identification of metabolites. The procedure is normally complicated by the problem that compounds with unknown or incompletely known structures and retention times need to be identified in a chromatogram that contains mainly the peaks of non-metabolites. The final goal of the analysis is to obtain a picture of the metabolic pathways of the substance. Hence the problem is similar to putting together a jig-saw puzzle when the elements of several puzzles are present.

Elements of the puzzle (relevant compounds within the chromatogram) can be identified by predicting first the structures through computer-assisted metabolism prediction (CAMP), then assigning retention times to the structures by using quantitative structure–retention time relationships (QSRR). Putting together the puzzle (building up a tree of metabolites corresponding to the pathway) is also provided by CAMP. We found it of interest to develop an expert system (HPLC-METABOLEXPERT) that solves both problems simultaneously. This paper summarizes the first experiences with this expert system. HPLC-METABOLEXPERT is now commercially available from CompuDrug (Budapest, Hungary) or CompuDrug USA (Austin, TX, U.S.A.).

CAMP is a new method, the purpose of which is to predict the metabolic fate of a substance within a living system[2]. Its underlying ideas are the artificial intelligence-based simulation of the biological fate, and thus metabolite formation, of a compound[3], applying a 'reasoning by analogy' approach[4] or simply using the analogy between the problem of predicting metabolite structures and finding precursors in synthesis design[5].

Technical realization of CAMP are based on computer programs that belong to the family of the expert systems. An expert system[6] is computer software that uses high-level mathematical logics to answer questions and uses a database that stores human expert knowledge (knowledge base). The expert systems for CAMP are members of the METABOLEXPERT family[7].

The prediction of the retention data and chromatographic conditions for compounds with known chemical structures has been an interesting problem in chromatographic science for a long time. Relationships between the chemical structure and chromatographic behaviour have already been investigated[8–10], even in quantitative way[11,12]. Finding quantitative structure–retention relationships (QSRR) is relatively easy in reversed-phase (RP) chromatography as the retention is governed mostly by hydrophobic forces. The prediction of the hydrophobicity of compounds from the structural formula is more or less a solved problem. Several types of software are commercially available for calculation of octanol–water partition coefficients as a measure of hydrophobicity based on Hansch π values[13] or Rekker's[14] fragmental constants (f). The relationships between hydrophocity and reversed-phase chromatographic retention have been also extensively studied[15–18].

The prediction of HPLC retention for metabolites described in this paper is carried out by using a database that contains the retention changes caused by a substituent of the molecule that appears or disappears in the physiological metabolic route. The database can be regarded as a compilation of fragmental constants which refers to the retention changes caused by the hydrophobicity changes similarly to the above-mentioned software.

Several expert systems and other software have already been developed in chro-

matography[19–21]. Their purpose was to compile the knowledge of a well experienced chromatographer who has already analysed hundreds of compounds and developed hundreds of analytical methods. On the basis of this knowledge, solving a new problem is much easier.

In order to reduce the experimental time for the isolation and identification of metabolites, our expert system suggests HPLC measurement conditions and retention times for possible metabolites on the basis of the input experimental HPLC conditions for the parent drug molecule. As the quality of an expert system can be assessed by using it in practice, the predictive capability of HPLC-METABOLEXPERT was investigated by measuring RP-HPLC retention parameters of drugs and their metabolites.

THEORETICAL BACKGROUND

The cluster of approaches underlying the METABOLEXPERT system is based on the simultaneous modelling of the metabolic transformations and pharmacokinetic fate of a substance, as shown in the 'Biotransform graph' (G) in Fig. 1. G is composed of edges corresponding either to transport or metabolism steps, and of vertices corresponding to the structure and concentration of a compound at a definite moment in time. In the realization of the model, the vertices and edges in G are represented by 'functions' of first-order logics, the Horn clauses[22]. Horn clauses have also been used for representing the chemical structure of the compounds and the production rules. These rules describe the conditions of the metabolic transformations.

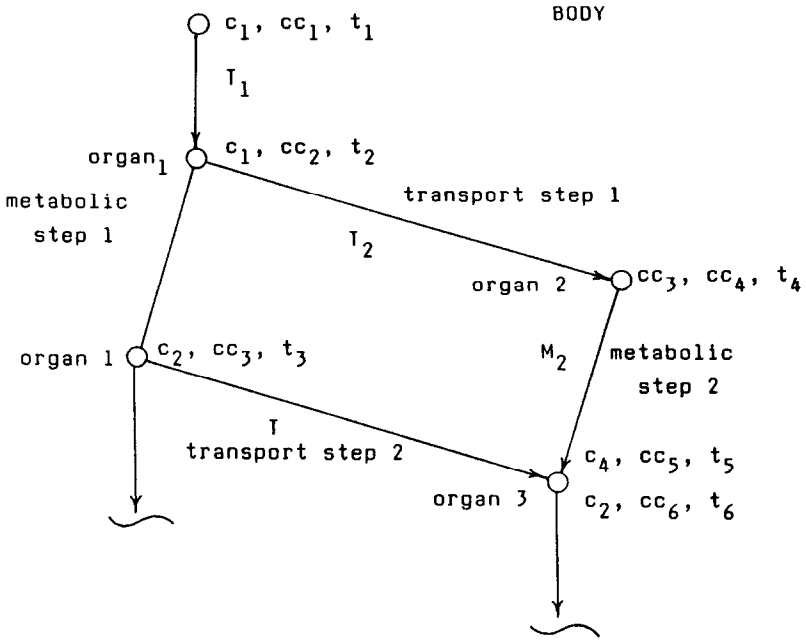


Fig. 1. Biotransformation graph. c_1 is the parent compound at concentration cc_1 at time t_1; c_2–c_6 are metabolites at concentrations cc_2-cc_6 at times $t_2 - t_6$.

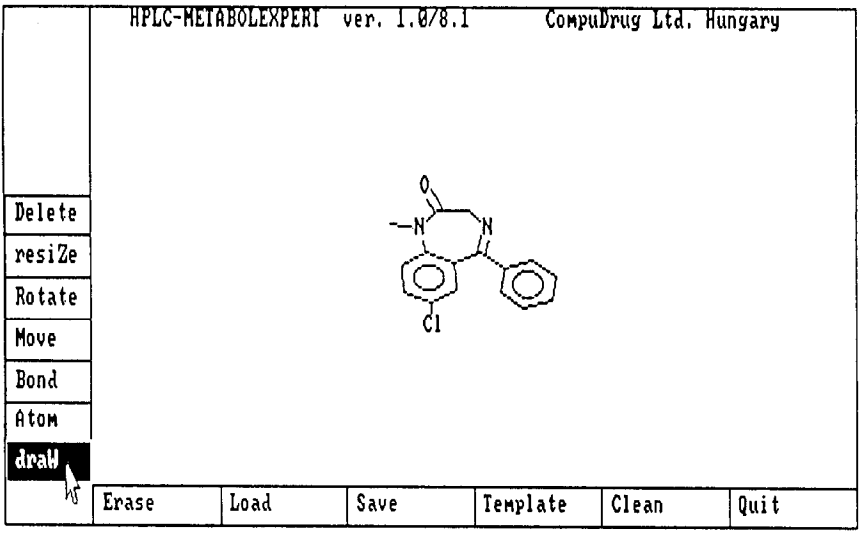


Fig. 2. Input of the compounds into HPLC-METABOLEXPERT.

Version 1.0 of HPLC-METABOLEXPERT[23] serves for predicting metabolic pathways of organic compounds in humans and/or in species commonly used in laboratory investigations. The knowledge base of the system consists of about 120 generalized metabolic transformations taken from Testa and Jenners's book[24]. After

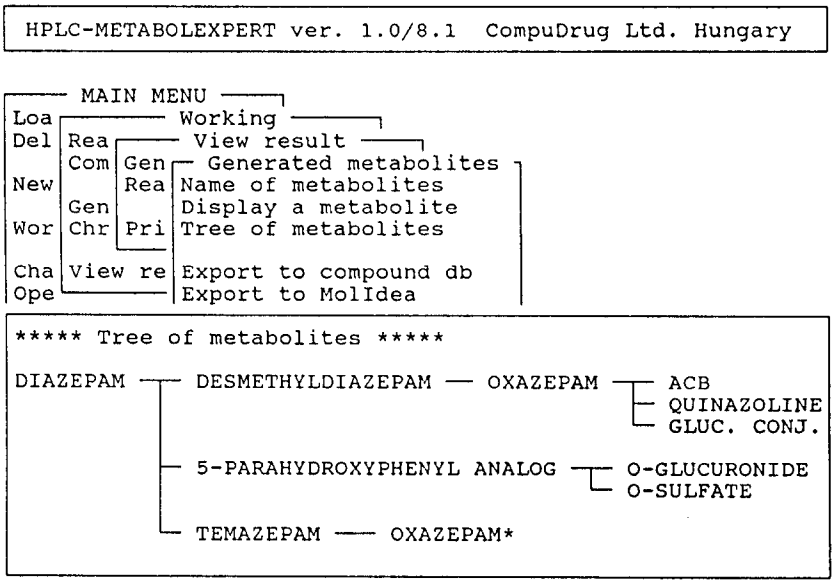


Fig. 3. Tree-like picture for metabolite map in a screen output obtained by the HPLC-METABOLEXPERT for diazepam (compound **4**). Desmethyldiazepam, oxazepam and ACB are compounds **5**, **6** and **8**, respectively.

input of the compounds (Fig. 2), the simulation of the metabolic transformations follows, which results in a tree-like picture and in structures of metabolites (Fig. 3). They might be displayed in three-dimensional form with the attached program MOLIDEA[25]. There is a possibility of calculating hydrophobicity[26] or quantum chemical indices[27] of the metabolites through connected programs. METABOLEXPERT accepts and produces chemical structures also in standard molfile format, which allows interfacing with widely used database managers such as CHEMBASE and MACCS[28,29].

The prediction of HPLC retention for the metabolites is based on the following considerations. The HPLC capacity factor (k') can be calculated from the retention time (t_R) and the dead time (t_0) measured in the given chromatographic system according to the equation

$$k' = (t_R - t_0)/t_0 \tag{1}$$

Log k' is linearly related to the logarithm of the distribution coefficient (log K) of the compound between the mobile and stationary phases according to

$$\log k' = a \log K + b \tag{2}$$

where a and b are constants. As the log K values are linear free-energy related parameters, the log k' values can be regarded as the same. It is also known that log k' values are linearly related to the logarithm of the octanol–water partition coefficient (log P)[15–18]:

$$\log P = c \log k' + d \tag{3}$$

By analogy with the log P predictions (eqn. 4), the log k' values also can be regarded as a sum of the $\Delta\log k'$ values referring to all fragments in the molecule (eqn. 5).

$$\log P = \Sigma\pi_i \text{ or } \log P = \Sigma F_i f_i \tag{4}$$

where π_i is the Hansch π value[13], f_i is the Rekker fragmental constant[14] and F_i represents the number of f_i fragment in the molecule.

$$\log k' = \Sigma\, \Delta\log k' \tag{5}$$

Finally, we can describe a metabolic transformation leading from a structure c_1 to c_2:

$$c_2 = c_1 + S \tag{6}$$

where the structure of c_1 differs from that of c_2 by an S group that appears or disappears during the metabolism. By simple algebraic substitutions, we can resolve that log k' for c_2 is linearly dependent on log P, which again depends linearly on log $P+\pi_s$. Hence, if we know log k' for c_1 (the parent compound) and the contributions of S (which are characteristic of a metabolic transformation type), we can predict log k' for the metabolic c_2:

$$\log k'(c_2) = \log k'(c_1) + \Delta \log k'_s \tag{7}$$

Eqn. 7 was used for predicting retention times of metabolites. The $\Delta \log k'$ values in principle are also dependent on the organic phase concentration in the mobile phase according to eqn. 10, which can be derived by subtracting eqn. 9 from eqn. 8:

$$\log k'_{c_2} = \text{slope}_{c_2} OP\% + \log k'_{o_{c_2}} \tag{8}$$

$$\log k'_{c_1} = \text{slope}_{c_1} OP\% + \log k'_{o_{c_1}} \tag{9}$$

$$\Delta \log k'_s = (\text{slope}_{c_2} - \text{slope}_{c_1}) OP\% + \log k'_{o_{c_2}} - \log k'_{c_1} \tag{10}$$

where $OP\%$ is the organic phase concentration and $\log k'_o$ is the $\log k'$ value extrapolated to the pure water mobile phase. When the change in the contact hydrophobic surface area[30] due to the metabolic transformation is small (*i.e.*, slope_{c_2} is almost equal to slope_{c_1}), the dependence of $\Delta \log k'$ on the mobile phase composition can be disregarded. At the present stage of the expert system the average values of the $\Delta \log k'$ data obtained by any mobile phase composition were considered. The HPLC-METABOLEXPERT contains an extensive database of $\Delta \log k'$ values referring to more than 100 structural changes observed in various metabolic pathways. In order to predict the HPLC conditions for the metabolites, the following input data should be applied to the system.

(1) The HPLC column (stationary phase) on which the parent compound was analysed. Only a reversed-phase column is accepted, and RP-18 is suggested.

(2) The mobile phase composition should also be given, considering expecially the pH and the organic phase concentration.

(3) The detection wavelength used for the analysis of the parent molecule is also an important parameter.

(4) The retention time of the compound and the dead time should also be given in order to be able to calculate the capacity ratio of the parent compound in the given chromatographic system.

When the chromatographic conditions and retention times for the metabolites are predicted, the following considerations were taken into account.

The $\log k'$ values of the compounds should be between -0.5 and 3 from a practical point of view (*i.e.*, not too short or long retention times).

When the predicted $\log k'$ value of the metabolite is outside the above range, the expert system suggests an appropriate change in the mobile phase composition, with an increase or decrease in the organic modifier concentration. According to our previous observations[31], increasing the acetonitrile concentration of the mobile phase by 10% decreased the retention by $\log k' = 0.285$ on average. The observed minimum was 0.129 and the maximum was 0.456. The $\Delta \log k'$ value caused by a change in mobile phase composition can be related to the contact hydrophobic surface area of the molecule[30]. Increasing the methanol concentration of the mobile phase by 10% decreased the retention by $\log k' = 0.298$ on average[32]. The observed minimum was 0.214 and the maximum was 0.379. As the average values for acetonitrile and methanol, did not differ significantly, the same $\log k'$ database was used in both instances.

The detection wavelength is also important because a chromophore can appear

or disappear during the metabolism. Our expert system also makes suggestions for changing the UV detection conditions in these instances, otherwise the predicted retention time can be correct but in certain instances the user cannot see the metabolite peak on the chromatogram.

Changing the pH of the mobile phase can also be important when an acidic or basic group appears or disappears during the metabolism. When the metabolite to be analysed by HPLC is present in partially dissociated form, it can cause a very bad peak shape and makes the peak undetectable or unmeasurable. When acidic or basic groups appear or dissappear during the metabolic transformation, our expert system makes suggestions for changing the mobile phase pH.

Although the change in retention time caused by a given substituent can vary over a wide range, the $\Delta\log k'$ value can be regarded as a constant, neglecting its dependence on the organic phase concentrations. Similarly for the log P predictions, when the π and f values of a molecular fragment can be changed by an adjacent substituent which influences the prediction, the $\Delta\log k'$ values can also be different for the same reason. Moreover, the $\Delta\log k'$ value of a metabolic route can be different on different reversed-phase columns. Several papers[33,34] have reported the significant effects of the free silanol groups usually located on reversed-phase stationary phases on the chromatographic retention of molecules. The present expert system has not yet considered this effect in predicting the changes of retention times. We assume that the error in the prediction is higher than the $\Delta\log k'$ difference obtained using reversed-phase columns produced by different manufacturers and having different amounts of free silanol groups on the surface. In order to check the validity of this assumption, retention measurements were repeated on three different reversed-phase columns.

EXPERIMENTAL

The structures of the investigated pyridopyrimidine derivatives are shown in Fig. 4. The derivatives were synthesized by Fülöp *et al.*[35] and the metabolites were isolated and identified by Jemnitz *et al.*[36].

The structures of the investigated benzodiazepine derivatives are shown in Fig. 5. Their synthesis[37–39] and metabolic pathway[40–42] have been described.

The metabolite of 3-trifluoromethyl-α-ethylbenzhydrol (produced by Gedeon Richter under the trade-name Zixoryn) and its derivative are shown in Fig. 6. The synthesis[43] and the metabolic pathway have been described by Klebovits *et al.*[44] and Ledniczky *et al.*[45].

Fig. 7 shows the structure of Deprenyl and its desmethyl metabolite, for which the synthesis was described by Chinoin[46] and Fowler[47]. The metabolism was studied by Kalász *et al.*[48].

HPLC-grade methanol, acetonitrile and salts for the phosphate buffers were of analytical-reagent grade and purchased from Reanal (Budapest, Hungary),

Fig. 4. Structures of the investigated pyridopyrimidine derivative and its metabolites.

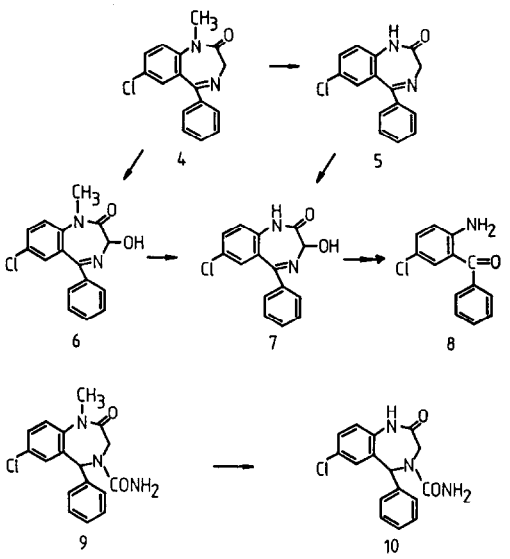


Fig. 5. Structures of the investigated benzodiazepine derivatives (diazepam, oxazepam and Uxepam) and their metabolites.

The HPLC measurements were carried out with a system consisting of a Model 312 Liquopump pump (LaborMIM, Budapest, Hungary) Rheodyne injector valve with a 20-μl loop (Rheodyne, Cotati, CA, U.S.A.), Cecil UV absorbance monitor and Waters 740 Data Module recorder–integrator (Millipore–Waters, Milford, MA, U.S.A.).

Hypersil ODS (5 μm) stationary phase was packed in our laboratory in a stainless-steel column (250 × 4 mm I.D.) (column I). The second column tested (250 × 4.6 mm I.D.) was packed by Bioseparation Technique (Budapest, Hungary) with LiChrosorb RP-18 (5 μm) stationary phase (column II). The third column (250 × 4.6 mm I.D.) contained Sepharon RPS stationary phase purchased from Laboratorní Přístoje (Prague, Czechoslovakia) (column III).

The mobile phases were: methanol–phosphate buffer (0.005 *M* K_2HPO_4, pH 7.8) (50:50) for compounds **1–3**; methanol–buffer (0.005 *M* K_2HPO_4, pH 7.8) (60:40) for compounds **4–10**; methanol–buffer (0.005 *M* KH_2PO_4, pH 3.5) (70:30) for compounds **11–13**; and acetonitrile–ammonium phosphate buffer (pH 3.01) (20:80) for compounds 14 and 15. The flow-rate of the mobile phase was always 1.00 ml/min except on column II (0.6 ml/min). The measurements on compounds **1–4** on column III were carried out by using an increased methanol concentration (60% methanol and 40% 0.005 *M* K_2HPO_4 buffer pH 7.8) because of the long retention times.

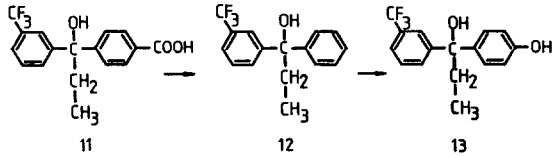
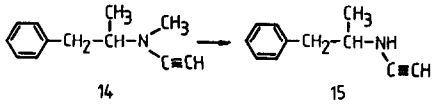


Fig. 6. Structures of the investigated 3-trifluoromethyl-α-ethylbenzhydrol derivatives.



14 15

Fig. 7. Structures of Deprenyl and its metabolites.

Detection was always performed at 260 nm. Compounds were dissolved in the mobile phase at a concentration giving an appropriate UV absorbance at a sensitivity of 0.05. The dead volume of the columns was determined by injecting 2% sodium nitrate solution.

The structures of the metabolites shown in Figs. 4–7 were predicted by HPLC-METABOLEXPERT version 1.0. Running times were of the order of a few seconds.

For retention prediction, first the chromatographic conditions and retention times of the parent compounds were input. By requesting the HPLC retention prediction for metabolites, the necessary mobile phase compositions and predicted retention times can be seen on the display together with a note about the necessary pH and detection wavelength changes.

RESULTS AND DISCUSSION

Figs. 4–7 show the structural formulae of compounds for which HPLC retention measurements were carried out. The indicated metabolic transformations were also predicted by HPLC-METABOLEXPERT version 1.0. Of course, many more metabolites have been predicted by the expert system and many more compounds were identified as real metabolites[36,40–42,44,45] than were investigated in this study. In this work the predictive power of the expert system considering only the HPLC retention was considered by measuring synthetic mixtures of parent compounds and metabolites.

The retention time data and log k' values measured for the compounds in Figs. 4–7 are summarized in Table I. Table I also shows the predicted retention time for each compound under each set of chromatographic conditions. The difference between the measured and predicted retention times was always less than 8 min. The average difference between the measured and predicted retention times was 1.8 min, which can be considerd as an acceptable prediction. Representative chromatograms for each compound are shown in Figs. 8–11.

The $\Delta\log k'$ values obtained for the same structural change on the three types of columns are summarized in Table II. The average $\Delta\log k'$ values are also given in Table II.

In order to reveal the retention changes on the three columns for the same structural change, the differences between the $\Delta\log k'$ values obtained for the C-hydroxylation of compounds **1**, **4**, **5** and **12** and for the N-demethylation of compounds **4**, **6**, **9** and **14** are summarized in Table III. The average $\Delta\log k'$ values and their standard deviations were also calculated. In comparison with the average and the standard deviation for the measured $\Delta\log k'$ values on the three types of columns and for the same structural change obtained on the same columns, but different mobile phase compositions, it can be seen that the standard deviation is much higher

TABLE I

MEASURED (mt_R) AND THE PREDICTED (pt_R) RETENTION TIME DATA AND (LOG k') VALUES FOR THE 15 INVESTIGATED COMPOUNDS

All pt_R values were calculated from the mt_R of the parent compound, except as indicated in the footnotes.

Compound	*Column I (t = 1.4 min)*			*Column II (t = 2.38 min)*			*Column III (t = 1.6 min)*		
	mt_R	pt_R	*log k'*	mt_R	pt_R	*log k'*	mt_R	pt_R	*log k'*
1	11.1	–[c]	0.84	24.4	–	0.97	16.6	–	1.00
2	4.7	6.0	0.37	11.0	12.2	0.56	6.3	8.3	0.50
3	5.2	3.9	0.43	12.3	8.8	0.62	7.3	5.1	0.58
4	13.0	–	0.92	14.3	–	0.70	19.6	–	1.08
5	8.7	8.7	0.71	10.8	9.9	0.55	12.1	13.0	0.85
6[a]	7.5	9.6	0.64	9.4	10.8	0.47	9.2	14.3	0.71
7[b]	5.7	6.6	0.49	8.0	8.3	0.37	7.4	9.0	0.59
7	5.7	5.3	0.49	8.0	6.8	0.37	7.4	6.4	0.59
8	20.8	16.0	1.14	19.6	21.4	0.86	29.8	21.3	1.28
9	5.9	–	0.51	9.4	–	0.47	7.2	–	0.54
10	4.9	4.2	0.40	8.4	6.8	0.40	5.7	5.1	0.41
11	4.4	–	0.32	6.2	–	0.21	3.9	–	0.15
12	12.7	9.9	0.91	10.4	13.1	0.53	11.5	8.1	0.79
13	3.7	5.0	0.22	5.8	4.9	0.16	3.3	4.7	0.02
14	8.4	–	0.70	14.3	–	0.74	13.5	–	0.87
15	6.0	5.3	0.51	9.9	9.1	0.54	7.7	8.3	0.58

[a] pt_R calculated from mt_R of **4**.
[b] pt_R calculated from mt_R of **5**.
[c] The measured retention times of the parent compounds were the input data for the prediction of the retention times of the metabolites.

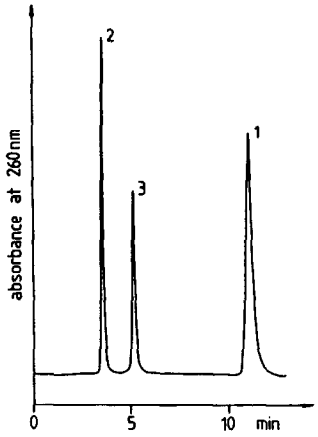


Fig. 8. Chromatogram obtained for the pyridopyrimidine derivative and its metabolites on column I. Mobile phase: methanol–0.005 *M* K_2HPO_4 buffer (50:50, v/v) flow-rate, 1.00 ml/min; column, Hypersil ODS (5 μm) (250 × 4 mm I.D.); detection, 260 nm. Measured (mt_R) and predicted (pt_R) retention times: mt_R = 4.68 min, pt_R = 6.0 min for 2; mt_R = 5.17 min, pt_R = 3.9 min for **3**; mt_R = 11.08 min for **1**.

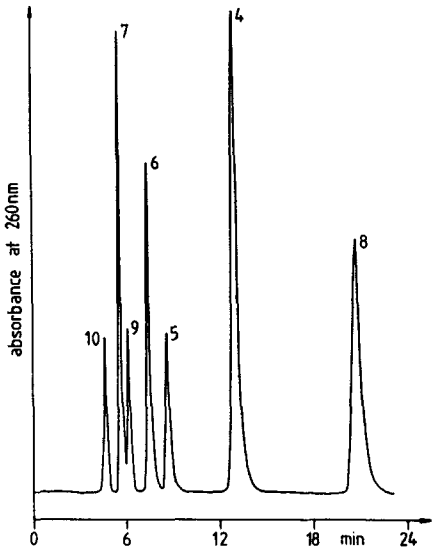


Fig. 9. Chromatogram of the diazepines and their metabolites on column I. Mobile phase methanol–0.005 *M* K_2HPO_4 buffer (60:40, v/v) flow-rate, 1.00 ml/min; column, Hypersil ODS (5 μm) (250 X 4 mm I.D.); detection, 260 nm. Measured (mt_R) and predicted (pt_R) retention times: mt_R = 4.91 min, pt_R = 4.2 min for **10**; mt_R = 5.67 min, pt_R = 5.3 min for **7**; mt_R = 5.91 min for **9**; mt_R = 7.49 min, pt_R = 9.6 min for **6**; mt_R = 8.70 min, pt_R = 8.7 min for **5**; mt_R = 13.02 min for **4**; mt_R = 20.84 min, pt_R = 16.0 min for **8**.

Fig. 10. Chromatogram of 3-trifluoromethyl-α-ethylbenzhydrol and its derivatives obtained on column I. Mobile phase: methanol–0.005 *M* KH_2PO_4 buffer (70:30, v/v); flow-rate, 1.00 ml/min; column, Hypersil ODS (5 μm) (250 × 4 mm I.D.); detection, 260 nm. Measured (mt_R) and predicted (pt_R) retention times: mt_R = 3.70 min, pt_R = 5.0 min for **13**; mt_R = 4.37 min for **11**; mt_R = 12.74 min, pt_R = 9.9 min for **12**.

Fig. 11. Chromatogram of Deprenyl and its metabolite obtained on column I. Mobile phase: acetonitrile–0.1 *M* ammonium phosphate (pH 3.01) (20:80, v/v); flow-rate, 1.00 ml/min; column, Hypersil ODS (5 μm) (250 × 4 mm I.D.); detection, 260 nm. Measured (mt_R) and predicted (pt_R) retention times: mt_R = 5.96 min, pt_R = 5.3 min for **15**; mt_R = 8.36 min for **14**.

TABLE II

ΔLOG k' VALUES MEASURED FOR THE VARIOUS STRUCTURAL METABOLIC CHANGES ON THE THREE COLUMNS AND THE AVERAGE (A) AND VALUES IN THE DATA BASE (DB)

Metabolic route	Δlog k'				
	Column I	*Column II*	*Column III*	*A*	*DB*
+ OH on **1**	−0.470	−0.407	−0.498	−0.458	−0.35
+ OH on **4**	−0.279	−0.230	−0.274	−0.261	−0.15
+ OH on **5**	−0.220	−0.180	−0.260	−0.220	−0.15
+ OH on **12**	−0.723	−0.371	−0.773	−0.622	−0.50
− CH3 on **4**	−0.204	−0.151	−0.231	−0.195	−0.20
− CH3 on **6**	−0.152	−0.105	−0.115	−0.124	−0.20
− CH3 on **9**	−0.110	−0.068	−0.136	−0.105	−0.20
− CH3 on **14**	−0.183	−0.196	−0.290	−0.223	−0.25
− COOH on **11**	+0.598	+0.323	+0.643	+0.521	0.45
Oxidation on **2**	+0.060	+0.060	+0.080	+0.067	−0.13
Ring opening on **7**	+0.650	+0.490	+0.690	+0.610	+0.53

for the Δlog k' values measured for the same structural change but in different molecules. This suggests that the weakest point of the prediction is not due to the use of different reversed-phase columns or conditions.

It should be mentioned that the slopes for pairs of molecules (see eqns. 8–10) can also be different and the organic phase concentration ($OP\%$) in eqn. 10 can be far from zero. Therefore, the higher standard deviation in this instance also reflects to the neglect of the dependence of the Δlog k' values on the mobile phase composition. The average standard deviations are 0.049 (expressed in log k') for demethylation and 0.104 for hydroxylation (see Table III), which are acceptable, especially when expressed as retention time values (1.8 min) (see Table I). The error of the prediction

TABLE III

LOG k' VALUES CAUSED BY THE METABOLIC HYDROXYLATION AND DEMETHYLATION OF THE COMPOUNDS ON THE THREE COLUMNS

Reaction	*Compound pair*	*Column I*	*Column II*	*Column III*	*Average*
Hydroxylation	**1–2**	−0.470	−0.407	−0.498	−0.458 ± 0.038
	4–6	−0.279	−0.230	−0.274	−0.261 ± 0.022
	5–7	−0.220	−0.180	−0.260	−0.220 ± 0.033
	Average	−0.323	−0.272	−0.344	−0.313 ± 0.030
	Standard deviation	0.107	0.097	0.109	0.104
	Phenolic OH: **12–13**	−0.723	−0.371	−0.773	−0.622 ± 0.179
Demethylation	**4–5**	−0.204	−0.151	−0.231	−0.195 ± 0.040
	6–7	−0.152	−0.105	−0.115	−0.124 ± 0.024
	9–10	−0.110	−0.068	−0.136	−0.105 ± 0.034
	14–15	−0.183	−0.196	−0.290	−0.223 ± 0.053
	Average	−0.162	−0.130	−0.193	−0.162 ± 0.026
	Standard deviation	0.047	0.064	0.088	0.049

because of this neglect is under study by measuring the retention of the investigated compounds using a reasonable range of organic modifier concentrations.

Investigating the $\Delta\log k'$ values arising from structural changes on the three types of columns, some interesting features of the reversed-phase columns produced by different manufacturers were observed. Generally the highest $\Delta\log k'$ values were observed on Sepharon RP-18 (column III), which reflects to the high selectivity of the column although the theoretical plate number for this column was the lowest. The significantly higher $\Delta\log k'$ values for the decarboxylation (-0.643) and hydroxylation of the phenyl ring (-0.773) shows that the acidic compounds behaved in a very non-popular manner on the stationary phase. In contrast, these two acidic compounds were fairly polar on LiChrosorb RP-18 (column II). Column I (Hypersil ODS) showed the highest theoretical plate number and it did not show special interactions for either acidic or basic compounds, and its selectivity was between those of columns II and III.

The relatively high standard deviation of $\Delta\log k'$ obtained for the same structural change on various molecules can be explained from a physico-chemical point of view, namely the N-methyl group is not as basic in compounds **4**, **6** and **9** as in **14**, because the former three are carboxamides. For this reason, the change in hydrophobicity cannot be expected to be the same. The effect of C-hydroxylation in $\Delta\log k'$ values cannot be expected to be the same for phenolic (**13**) and aliphatic or ring C-hydroxyl groups (**2**, **6** and **7**). Very different effects with respect to $\Delta\log k'$ values and also hydrophobicity can be assumed for a C-hydroxylation when there is the possibility of the formation of intramolecular hydrogen bonds in the molecule, which is possibly the case for compound **2**. Version 1.0 of HPLC-METABOLEXPERT does not look for adjacent functional groups when taking into account the above-mentioned intramolecular interactions.

Both the knowledge base and the database of the expert system can be continuously enlarged by incorporating more published results and by a better understanding of the relationships between the chemical structure and reversed-phase chromatographic behaviour. In the future, a combined simulation of the metabolic and pharmacokinetic fate of a compound will be solved together with an enhanced precision of predictions. Building up connections with other chromatographic software for easy method development (for *e.g.*, Snyder's DryLab[21]) would also be fruitful for metabolism research. The potential of interfacing HPLC-METABOLEXPERT directly with the chromatograph or through control software (Chrom-A-set) should also be explored. Finally, the HPLC retention prediction module can be connected to other related expert systems.

REFERENCES

1 H. P. Kuemmerle, G. Hitzenberger and K. H. Spitzy, *Klinische Pharmakologie 1*, Ecomed, Munich, 1984.
2 F. Darvas, *J. Mol. Graph.*, 6 (1988) 80.
3 F. Darvas, in K. L. E. Kaiser (Editor), *QSAR in Environmental Toxicology—II*, Reidel, Dordrecht, 1987, p. 71.
4 M. Johnson, *J. Mat. Chem.*, (1989) in press.

5 W. T. Wipke, G. I. Ouchi and J. T. Chou, in L. Goldberg (Editor), *Structure–Activity Correlations as Predictive Tools in Toxicology (PAP Symposium)*, Hemishere Publishing, Washington, DC, 1983, p. 151.
6 A. L. Kidd (Editor), *Knowledge Acquisition for Expert Systems*, Plenum Press, New York, 1987.
7 *METABOLEXPERT Version 4.1, Manual*, CompuDrug Budapest 1986; *METABOLEXPERT Version 5.4, Manual*, CompuDrug Budapest, 1987; *METABOLEXPERT Versions 8.0, 8.1, 8.2, Manual*, CompuDrug, Budapest, and CompuDrug USA, Austin, TX, 1988; *METABOLEXPERT Versions 9.0, 9.1, Manual*, CompuDrug, Budapest and CompuDrug USA, Austin, TX, 1988.
8 I. Molnár and Cs. Horváth, *J. Chromatogr.*, 145 (1978) 371.
9 S. P. Assenza and P. R. Brown, *J. Chromatogr.*, 282 (1983) 477.
10 C. F. Carney, *J. Liq. Chromatogr.*, 8 (1985) 2781.
11 B.-K. Chen and Cs. Horváth, *J. Chromatogr.*, 171 (1979) 15.
12 R. Kaliszan, *Quantitative Structure–Chromatographic Retention Relationships*, Wiley, New York, 1987.
13 C. Hansch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley, New York, 1979.
14 R. F. Rekker, *The Hydrophobic Fragmental Constant*, Elsevier, Amsterdam, 1977.
15 J. F. K. Huber, E. T. Alderlieste, H. Harren and H. Poppe, *Anal. Chem.*, 45 (1973) 1337.
16 E. D. Katz, K. Ogan and R. P. W. Scott, *J. Chromatogr.*, 352 (1986) 67.
17 P. Jandera, *J. Chromatogr.*, 352 (1986) 111.
18 T. Braumann, *J. Chromatogr.*, 373 (1986) 191.
19 P. J. Schoenmakers, N. Dunand, A. Cleland, G. Musch and Th. Blaffert, *Chromatographia*, 26 (1988) 37.
20 A. Peeters, L. Buydens, D. L. Massart and P. J. Schoenmakers, *Chromatographia*, 26 (1988) 10.
21 L. Snyder, *DryLab*, LC Resources, Orinda, CA, 1987.
22 C. L. Chang and R. C. T. Lee, *Logic and Mechanical Theorem Proving*, Academic Press, New York, 1973.
23 *HPLC-METABOLEXPERT Version 1.0/8.1, Manual*, CompuDrug, Budapest, and CompuDrug USA, Austin, TX, 1989.
24 B. Testa and P. Jenner, *Drug Metabolism; Chemical and Biochemical Aspects*, Marcel Dekker, New York, 1976.
25 *MOLIDEA Version 2.0, Manual*, CompuDrug, Budapest, 1988.
26 *PRO-LOGP Version 2.0, Manual*, CompuDrug, Budapest, 1988.
27 *LABSWARE Manual, Version 5.1*, CompuDrug, Budapest, 1987.
28 *CHEMBASE Manual, Version 1.1*, Molecular Design, San Leandro, CA, 1987.
29 *MACCS Manual*, Molecular Design, San Leandro, CA, 1987.
30 K. Valkó, *J. Liq. Chromatogr.*, 10 (1987) 1663.
31 K. Valkó, *J. Liq. Chromatogr.*, 7 (1984) 1405.
32 K. Valkó, *Trends Anal. Chem.*, 6, 8 (1987) 214.
33 H. Colin, A. Krtulovic, G. Guiochon and Z. Yun, *J. Chromatogr.*, 255 (1983) 295.
34 W. Melander, J. Stoveken and Cs. Horváth, *J. Chromatogr.*, 199 (1980) 35.
35 F. Fülöp, I. Hermecz, Z. Mészáros, Gy. Dombi and G. Bernáth, *J. Heterocycl. Chem.*, 16 (1979) 457.
36 K. Jemnitz, G. Dénes, J. Tamás and E. Gács-Baitz, *Biomed. Environ. Mass Spectrom.*, 18 (1989) 308.
37 L. H. Sternbach, R. J. Fryer, W. Metlesics, E. Reeder, G. Sach, G. Saucy and A. Stempel, *J. Org. Chem.*, 27 (1962) 3788.
38 J. Röhricht, I. Kisfaludy, L. Ürögdi, É. Pálosi, Sz. Szeberényi and L. Szporny, *Hung. Pat.*, 171 033 (1974).
39 M. A. Schwartz, B. A. Koechlin, E. Postma, S. Padmer and G. Krol, *J. Pharm. Exp. Ther.*, 149 (1965) 423.
40 S. Garattini, E. Mussini and L. O. Randall, *The Benzodiazepines*, Raven Press, New York, 1973.
41 N. Ya. Golovenko, *Vopr. Med. Khim.*, 24 (1978) 676; *C.A.*, 90 (1979) 48166u.
42 H. C. van der Plas, L. Ötvös and M. Simony, in Zs. Teqyei, L. Vereczkey, J. Tamás, J. Röhricht, L. Kisfaludy and L. Ötvös (Editors), *Bioorganic Heterocycles*, Akadémia, Budapest, 1984, p. 376.
43 E. Tóth, J. Törley, Sz. Szeberényi, É. Pálosi, L. Szporny and S. Görög (Gedeon Richter), *Ger. Pat.*, 2 438 399, 1975; *U.S. Pat.*, 4 039 589 (1977).
44 I. Klebovits, L. Vereczkey, J. Tamás, M. Mák, Gy. Jalsovszky and S. Holly, *Acta Pharm. Hung.*, 58 (1988) 221.

45 M. Ledniczky, I. Szinai, K. Ujszászi, S. Holly, V. Kemény, Gy. Mády and L. Ötvös, *Arzneim.-Forsch./Drug Res.*, 28 (1978) 673.
46 Chinoin, *Fr. Pat.*, FM2635 (1964).
47 J. S. Folwer, *J. Org. Chem.*, 42 (1977) 2637.
48 H. Kalász, L. Kerecsen, J. Knoll, J. Pucsok, R. Dobó and I. Hollósi, in H. Kalász and L. S. Ettre (Editors), *Chromatography 87*, Akadémia, Budapest, 1988, p. 275.