Journal of Chromatography, 452 (1988) 175–189
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 20 730

EXPERT SYSTEM FOR CHROMATOGRAPHY

LU PEICHANG* and HUANG HONGXIN
Dalian Chromatographic R&D Centre of China, Dalian Institute of Chemical Physics, Academia Sinica, Dalian, Liaoning (China)

SUMMARY

A strategy for development of an expert system for chromatography is discussed. A program including the recommendation of the column system, the optimization of operating conditions, peak identification and quantitation on-line and the diagnosis of the hardware system has been developed. Both artificial intelligence and some advanced algorithms have been incorporated.

INTRODUCTION

Since the chromatograph Sentinel[1] with automatic method development was produced by DuPont in 1981, the development of chromatographs with artificial intelligence (AI) has attracted much attention[2–4]. Hardware such as laboratory robotic systems with fully automatic sample preparation and analysis has been produced by Perkin-Elmer[5] and Zymark[6]. To develop hardware to provide the technical support for advanced software is very important, but the key to the development of a chromatograph with AI is the software[7–10]. Research on chromatographic theory and methodology are still the central focus of chromatographers[11–15]. The development of computer programs for solvent optimization, which are heavily dependent on some algorithms, has been achieved by optimizing the separation with respect to specific criteria. Glajch and co-workers[11,12] reported an overlapping resolution mapping (ORM) method which is a powerful tool for aiding the chromatographer in obtaining satisfactory high-performance liquid chromatographic (HPLC) separation, but an human expert is still required to recommend the stationary and mobile phases. Recently, Snyder and Dolan[16] developed a series of software, Drylab, which is based on a set of equations to predict how various changes in chromatographic conditions and column length affect the resolution, plate number, analysis time, solvent consumption, relative peak sensitivity and other chromatographic parameters. Computer simulation can also be used for optimizing gradient-elution separation after two initial experimental gradient separations of the sample[17]. Method development using this procedure, of course, can result in a better separation with much less effort.

The development of expert systems for chromatography using AI techniques has just begun[18–21], *e.g.*, the Expert Chromatographic Assistance Team (ECAT) devel-

0021-9673/88/$03.50

oped by Bach[19] and the Michigan University System (MUS) developed by Wade[20]. The task of ECAT is to perform, at the human expert level, the design, analysis, optimization and troubleshooting of an HPLC separation method. Because there is not a "living" chromatogram base in ECAT, it is difficult to demonstrate the reliability of the reasoning of an expert system and to display chromatograms corresponding to samples analyzed in the literature.

A new generation of chromatographs employ AI[21]. Our basic research in this area has been reviewed[22–24]. In this article we shall discuss our strategy for developing an expert system for chromatography (ESC), the design of the knowledge base and the chromatogram base and some advanced algorithms such as those for separation optimization and data handling. We also discuss how to combine the algorithms with the AI approach to build up the system.

System strategy

The expert system is divided into three parts: knowledge base and chromatogram base, inference engine and user interface. The knowledge base consists of a set of facts and rules on the chromatographic subject. The chromatogram base contains some chromatograms for the samples published in the literature. The inference engine or rule interpreter is responsible for extracting the desired information from the knowledge base and explaining how the answer is obtained. The user interface allows the user to interact with the system through dialogues and to display reasoning results, chromatogram and explanation. The stages involved in the development and application of the ESC are presented in Fig. 1.

The goal of any chromatographic analysis is to separate a mixture into individual peaks for qualitative and quantitative determination. First, it is necessary to select a suitable mode of separation, *i.e.*, gas chromatography (GC) or liquid chromatography (LC). We think that GC should be recommended if two conditions can be met: the sample analyzed is (1) volatile enough to be eluted from the column, (2) stable enough not to change at the operating temperature. The reason is that GC is characterized by high resolving power and low cost. Of course, the volatility of the sample also depends on the detectability of the detector. In addition, derivatization of the sample and column switching techniques can further enlarge the application range of GC. However, for thermally unstable or non-volatile substances only LC can be

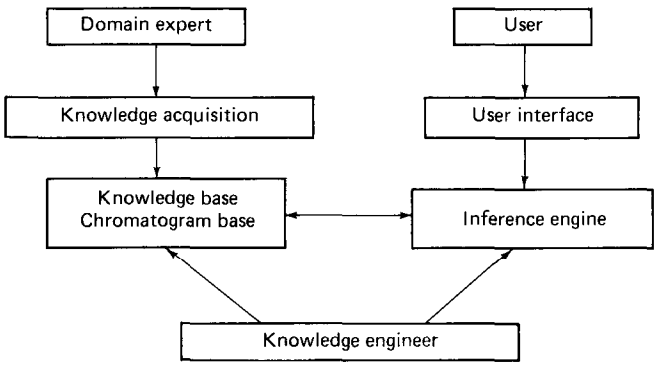

Fig. 1. The stages involved in the expert system for chromatography.

used. After choosing the mode of separation, the column system should be selcted, *i.e.*, the stationary liquid and type of column in GC, and the stationary phase and mobile phase in LC. Secondly, the method of sample pretreatment and detection will be determined by the sample, its matrix and where it is. Thirdly, it is desirable to obtain a good separation within the shortest time with a minimum number of experiments, *i.e.*, the optimum separation. To guarantee good analytical results from the column system, it is necessary to diagnose the instrument hardware. So a program for hardware diagnosis must be included in the system. It is concluded that an expert system for chromatography has to perform at least the following modules:

(1) The recommendation of the mode of separation and column system
(2) The recommendation of the method of sample pretreatment and detection
(3) The optimization of the operating conditions
(4) Peak identification and quantitation on-line
(5) The diagnosis of the hardware system

We have developed the knowledge base and chromatogram base, inference engine, user interface and some algorithms to support task modules in an ESC. In order to implement and use them conveniently, each module is independent of the rest of the system. The flow chart of chromatographic method development is presented in Fig. 2.

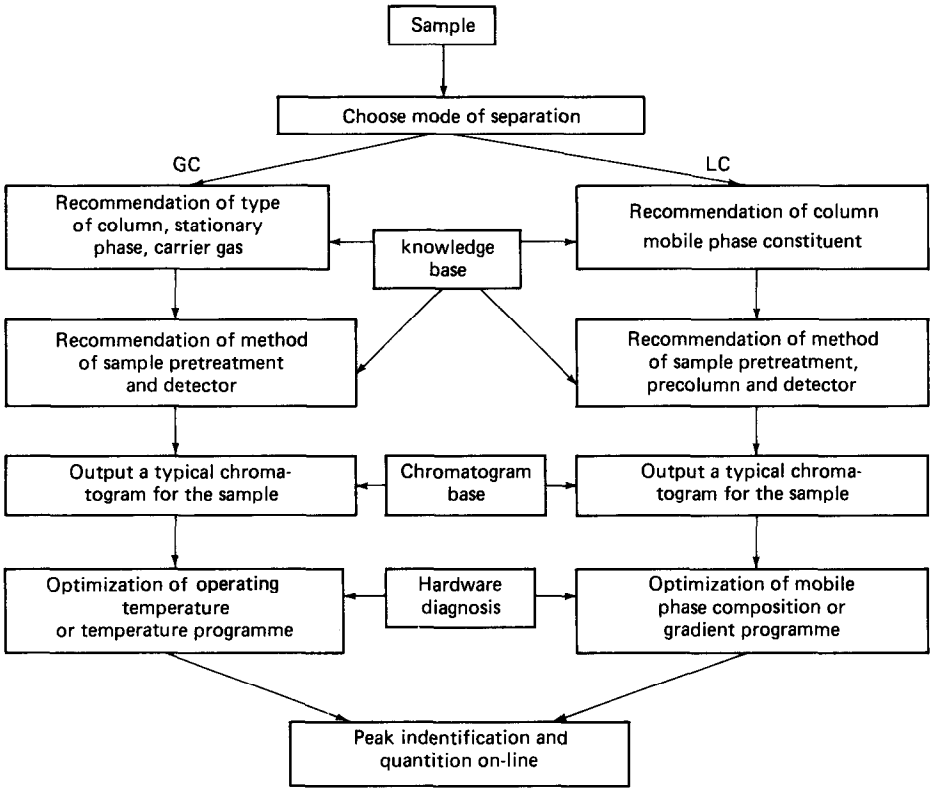

Fig. 2. The flow chart of the chromatographic method development.

As mentioned above, the major feature of an ESC is its knowledge base and chromatogram base. On the other hand, numerical calculation, based on chromatographic theory and algorithms, still plays an important rôle in the development of an ESC. Some problems, such as optimization, chromatographic data handling, should be solved using advanced algorithms. Therefore, we think that an ESC should be based not only on general chromatographic reasoning, but also on advanced algorithms.

Theory and algorithms

Optimization. For the simultaneous experimental optimization for a sample with known constituents, the resolution criterion K1 can be used to select the optimum conditions, considering the effect of the peak height ratio of two adjacent peaks

$$\begin{aligned} \mathrm{K1} &= 2(t_{\mathrm{R}_1} - t_{\mathrm{R}_2})/(w_{\mathrm{h}_1} + w_{\mathrm{h}_2}) \\ &= 1/2.354(\alpha' - 1)\, n \\ &= 1/2.354(\alpha' - 1)[(1 + k')/(\beta + k')] \cdot \sqrt{n_\infty} \\ &= 1/2.354 S \sqrt{n_\infty} \end{aligned} \tag{1}$$

where t_{R} is the retention time, w_{h} the peak width at half height, α' the separation factor, n the plate number, k' the capacity factor, β the coefficient of variation of the plate number with variation of the capacity factor, n_∞ the plate number for a solute having a capacity factor approaching infinity and S is the selectivity criterion of the column system

$$S = (\alpha' - 1)(1 + k')/(\beta + k') \tag{2}$$

For sequential experimental optimization of a sample with unknown constituents, another criterion, the serial chromatographic response function (SCRF), is proposed[25]

$$\mathrm{SCRF} = 100\,000N + 10\,000 \cdot \mathrm{K3} + (100 - t) \tag{3}$$

where N is the peak number, K3 is the peak resolution of the least separated peak pair in a chromatogram and t is the analysis time. Once the SCRF is combined with the simplex optimization method, the optimum conditions can be searched for. The principles of selecting the optimum conditions are:

(1) The higher the number of peaks, the better are the operating conditions. This implies that more information about a sample can be extracted from an analysis.

(2) When the peak number has reached a maximum the resolution of the least-separated peak pair should meet the criterion K1 or K3.

(3) With maximum peak number and satisfactory K1 or K3, the analysis time should be as short as possible.

If the above principles are employed the multi-step linear program optimization for a known sample in HPLC, K1, can be expressed as[26]

$$K1_i = 2[t_0(1 + k_{i+1}^{'x})Q_{i+1}^x - t_0(1 + k_i^{'x})Q_i^x]/(w_{h_{i+1}}^x + w_{h_i}^x) \quad (4)$$

where $k_i^{'x}$, $w_{h_i}^x$ and Q_i^x are the column dead time, capacity factor, peak width at half height and rest fraction of the column length to be moved for the ith component at the x-step mobile phase composition, respectively. The effect of the mobile phase composition on $k_i^{'x}$, $w_{h_i}^x$ can be predicted by:

$$\ln k_i^{'} = A_i + B_i Cb + C_i \ln Cb \quad (5)$$

$$w_{h_i}^x = a + bk_i^{'x} \quad (6)$$

where A_i is related to the adsorption energy of the solute, molecular interactions between the solute, solvent and stationary phase; B_i is only involved in solution interaction; C_i is an entropy function of an adsorbed solute; Cb is the concentration of strong solvent in a binary mobile phase; a and b are constants.

The parameters in eqns. 5 and 6 can be calculated by curve fitting from the data for at least three isocratic experiments. The parameter Q_i^x can be expressed as

$$\left.\begin{aligned} Q &= 1 - \sum_{j=1}^{2x-1} M_{i,j} \\ M_{i,2x-2} &= \int_{C_{x-1}}^{C_x} \frac{dC}{t_0 r_{x-1}[1 + \exp(A_i + B_i Cb + C_i \ln Cb)]} \\ M_{i,2x-1} &= \left[1 - \sum_{j=1}^{2x-2} M_{i,j}\right] \frac{(t_x - t_0)(1 + k_{i-1}^{'x})u}{k_i^x(1 + k_i^{'x})} \end{aligned}\right\} \quad (7)$$

where $M(i,0) = 0$, u is the mobile phase velocity, t_x is the elution time of component i at the x-step mobile phase composition and r_{x-1} is the composition rate from the step $x+1$ to the step x. The $K1_i$ for any adjacent components becomes a function of the mobile phase composition upon inserting eqns. 5–7 into eqn. 4.

Data handling. Data handling is an important topic. The data handling software, which can be used to detect various peaks including shoulders and give information on peak identification and quantitation, has been developed. The flow chart of the software is presented in Fig. 3.

Here, an advanced algorithm, a curve fitting method, for measurement of overlapping peaks is proposed on the basis of an exponentially modified Gaussian (EMG) function.

$$h(t) = \frac{A}{\sqrt{\pi}\tau} \exp\left(\frac{\sigma^2}{2\tau^2} - \frac{t - t_G}{\tau}\right) \cdot \int_{-\infty}^{z} e^{-x^2} dx = A \cdot f(t, t_G, \sigma, \tau) \quad (8)$$

or

$$A = h(t)/f(t, t_G, \sigma, \tau) \quad (9)$$

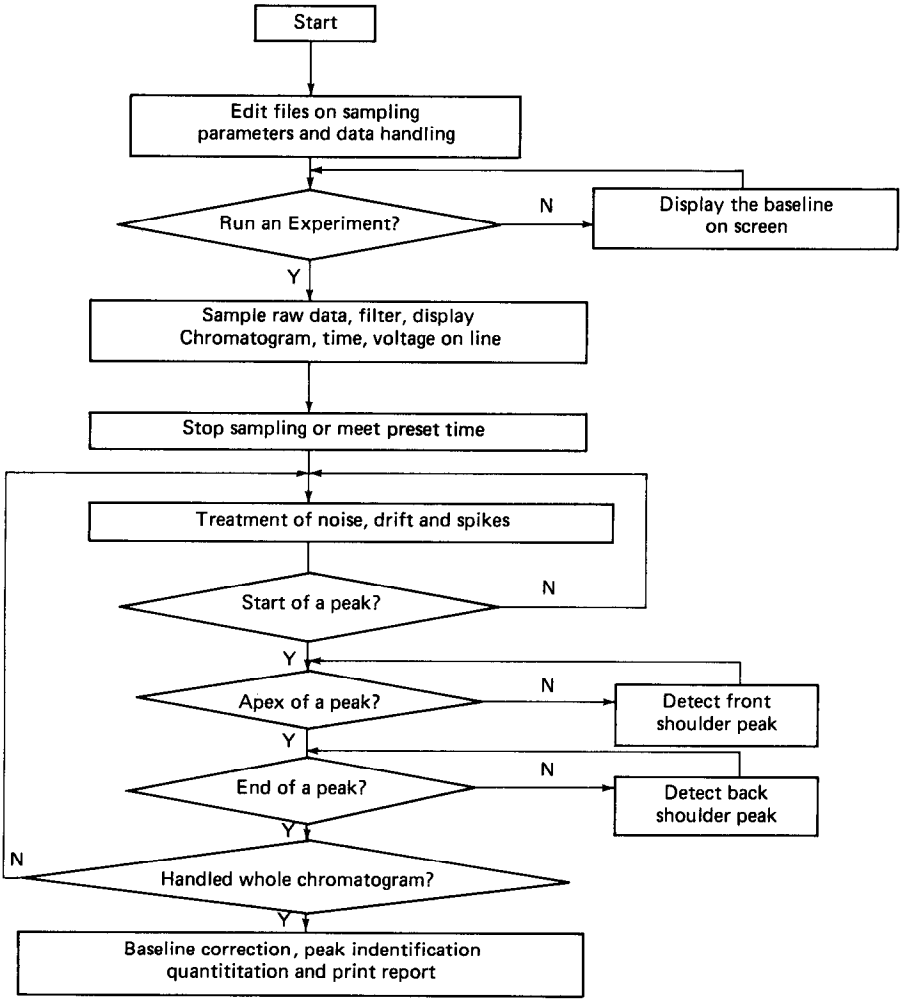

Fig. 3. Flow chart for peak identification and quantitation in the ESC.

where $z = \left(\frac{t - t_G}{\sigma} - \frac{\sigma}{\tau}\right) \Big/ \sqrt{2}$, A is the peak area, t is the time, $h(t)$ is the peak height at t, t_G is the central position of a Gaussian constituent, σ is the standard deviation of a Gaussian constituent and τ is the time constant for exponential decay. When $t = t_R$, $h(t)$ is the maximum peak height. The methods for estimating the parameters are given in ref. 27. We found[28] in practice that the linear relationship among σ, τ and t_R are

$$\sigma = a_1 + b_1 t_R \tag{10}$$

$$\tau = a_2 + b_2 t_R \tag{11}$$

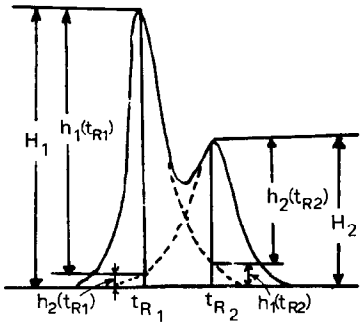

Fig. 4. The resolution of overlapping peaks.

where a_1, b_1, a_2 and b_2 are dynamically related coefficients[29] that can obtained by the regression of several baseline-separated peaks. That means that the function f for all peaks in a chromatogram can be calculated. When we want to resolve the overlapping peaks presented in Fig. 4, the curve-fitting area is represented as

$$\begin{aligned} A_1 &= h_1(t_{R1})/\mathrm{f}(t_{R_1}, \sigma_1, \tau_1) \\ &= [H_1 - h_2(t_{R1})]/\mathrm{f}(t_{R1}, t_{G1}, \sigma_1, \tau_1) \\ &= [H_1 - A_2\mathrm{f}(t_{R1}, t_{G2}, \sigma_2, \tau_2)]/\mathrm{f}(t_{R1}, t_{G1}, \sigma_1, \tau_1) \end{aligned} \tag{12}$$

$$A_2 = [H_2 - A_1\mathrm{f}(t_{R2}, t_{G1}, \sigma_1, \tau_1)]/\mathrm{f}(t_{R2}, t_{G2}, \sigma_2, \tau_2) \tag{13}$$

where H_1, H_2 can be measured from the chromatogram, but A_1, A_2 on the right-hand side of equation are not known before resolving the overlapping peaks. So the peak areas obtained from the perpendicular drop method are first substituted for A_2 and A_1 in eqns. 12 and 13 to get approximate new peak areas A_1, A_2. The new A_1 and A_2 are again substituted for A_1 and A_2 in the right-hand side of eqns. 12 and 13. The iterative calculation proceeds until the error in the peak area is less than desired one (at least two calculations are performed).

Implementation

All software except the numerical processing modules has been written in SCHEME-LISP, a language used to develop programs for symbolic processes. It can call programs written in procedural languages such as BASIC, FORTRAN and C. The decision made by the inference engine can be passed to numerical processing modules via a file or parameters. The results obtained from the numerical processing modules can also be returned to the inference engine to enable further decision making via a file. The combination of the rule-based system with the algorithms approach can thus be realized. The system has been hosted on an IBM PC/XT-286 microcomputer with a 20M hard disk drive, two floppy disk drives and a 20-bit analog-to-digital (A/D) converter to acquire raw data from a detector.

RESULTS AND DISCUSSION

The development of the knowledge base and chromatogram base

Our strategy for building the knowledge base is based on fundamental chromatographic theory. The selection rules for the stationary and mobile phase have been proposed after considering the interaction among solutes, stationary phase and mobile phase. The sample, we think, can be separated into individual peaks if the compounds in the sample have small chemical and physical differences such as functional groups, polarity, solubility, molecular weight. The specific stationary phase and mobile phase should be selected for separation of very similar compounds such as chiral and isomer samples.

An expert system with a chromatogram base has not reported until now, but it is well known that the chromatograms and operating conditions published in the literature are very useful to a novice chromatographer. A brief method for storing a living chromatogram based on an EMG function has been proposed[30]. A chromatogram with n peaks can be stored in a computer by use of $2n+4$ parameters. Based on this principle, we have built a chromatogram base about 500 typical and reliable chromatograms chosen from the published literature. The base can be searched in

```
(define fact)
(set! fact '((fact1
              (if (sample is PTH_amino_acid))
              (then ((PTH_group)
                     (molecular_weight<1000)
                     (moderate_polar)
                     (inactive-form)
                     (slight_aqueous_solubility)
                     (dispersion_force)
                     (orientation_force))))
              ;
              (fact100
                (if (eluent is acetonitrile))
                (then ((methyl)
                       (cyano)
                       (dispersion_force)
                       (orientation_force)
                       (pi-pi_interaction)))
            ;
            ))
;
(define rule)
(set! rule '((rule-mode
              (if ((slight_aqueous_solubility)
                   (moderate_polar)
                   (dispersion_force)))
              (then (reversed_phase)))
;
            (rule-mobile-phase
              (if ((reversed-phase)
                   (inactive-form)
                   (dispersion)
                   (orientation-force)))
              (then (acetonitrile)))
;
            ))
```

Fig. 5. Some examples of facts and rules in the ESC knowledge base.

accordance with compound, analyte class or author(s). So the system can generate a typical chromatogram for a sample analyzed to determine whether the symbolic reasoning over the knowledge base is correct.

The inference engine in the ESC is used to control and coordinate various components of the system. The forward chaining is used to select the mode of separation, mobile phase, stationary phase, method of sample pretreatment and detector.

Now three user entries in the ESC are designed to interface the computer with the user: (1) provision of the molecular structure of the sample; (2) provision of the commercial name of the sample; (3) provision of the analyte class to which the sample belongs. Once one of the entries is inputted, the program can obtain the molecular frame, functional groups and dominant molecular interaction through reasoning over the knowledge base. The column system, method of sample pretreatment and detector can be recommended.

Fig. 5 shows an example of some facts and heuristic rules in the knowledge base. All facts and rules are represented by IF/THEN statements. A rule asserts that the statement at the left-hand side implies the statements on the right-hand side. The inference engine is used to interpret those rules to generate new facts or to answer questions. For example, it can reason the new fact from the rule, fact100, in Fig. 5 that the eluent, acetonitrile, has the cyano, methyl and dominant dispersion force, orientation force and π–π interaction. Fig. 6 shows an example of the application of the ESC to method development in phenylthiohydantoin (PTH)-amino acid separation by HPLC. The user input and system recommendation are automatically processed and generated by the program. After items in Fig. 6 are printed, the system can also

```
                User    Entries

Instrument                      HPLC
Analytical sample               PTH_amino_acid
Sample from                     Synthetic_mixture
Smallest  amounts  (ug)         5
Detector  type                  UV

        Recommendation

Mode of separation:     HPLC

       Analytical column
(Separation mode        Reversed_phase_chromatography)
(Column                 Bonded_phase_C18)

        Mobile phase
(Operating mode         isocratic)
(Eluent_a               acetonitrile)
(Eluent_b               0,01m_NaAC_aqueous_solution)
(Additive agent         0.5%_EtC12)

        Pretreatment method

(Guard column, 5_cm_low_capacity_bonded_phase_C18)

                Detectors

(Detector       UV/254)
```

Fig. 6. The dialogue between the computer and user and method development.

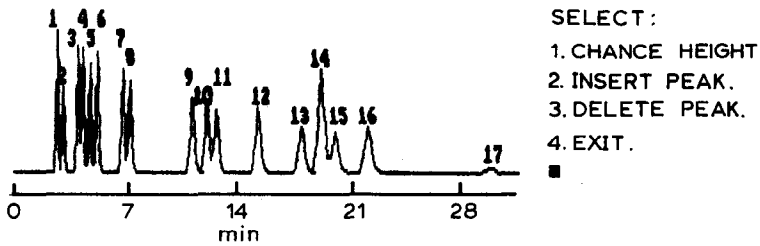

Fig. 7. The chromatogram generated by the chromatogram base of the ESC. Isocratic separation of PTH-amino acids after ref. 37. Column, 25 cm × 4.6 mm I.D. Ultrasphere-ODS; mobile phase, 0.01 *M* sodium acetate (pH 4.9)–acetonitrile (62.2:37.8); flow-rate, 1.0 ml/min; temperature, 55°C; detection, UV at 254 nm. Peaks: PTH derivatives of: 1 = Asp; 2 = Glu; 3 = Asn; 4 = Gln; 5 = Thr; 6 = Gly; 7 = Ala; 8 = Tyr; 9 = Met; 10 = Val; 11 = Pro; 12 = Trp; 13 = Phe; 14 = Lys; 15 = Ile; 16 = Len; 17 = Ser.

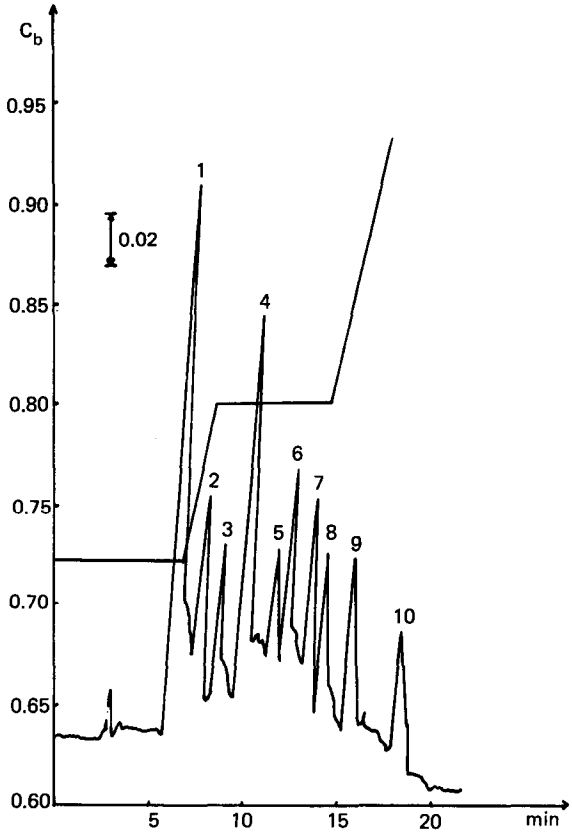

Fig. 8. Separation of ten bile acids with the selected optimum binary mobile phase composition of a multi-step linear programme. Column: 20 cm × 4.6 mm I.D., 10 μm YWG-C_{18}. Eluents: a, pH 5.6, 0.01 *M* KH_2PO_4, methanol–water (73:27); b, pH 5.6, 0.01 *M* KH_2PO_4 methanol–water (98:2). Flow-rate: 1 ml/min. Detector: UV, 210 nm. Solutes: 1 = tauroursodeoxycholic acid, 2 = glycoursodeoxycholic acid, 3 = taurocholic acid, 4 = glycocholic acid, 5 = taurochenodeoxycholic acid, 6 = glycochenodeoxycholic acid, 7 = glycodeoxycholic acid, 8 = taurodeoxycholic acid, 9 = taurolithocholic acid, 10 = glycolithocholic acid.

generate a chromatogram corresponding to an analyte class or compounds. An example of PTH-amino acid separation is presented in Fig. 7 as generated by the ESC. It can be noted that not only are there operating conditions and compounds in Fig. 7, but we can also arbitrarily adjust the peak height and insert and delete peaks in the chromatogram. So we call it a living chromatogram. Comparing the results of the recommendations in Fig. 6 with those in Fig. 7, it is found that the method designed by the ECS is similar to the one from literature. So the knowledge base of the ESC and reasoning are reliable.

The optimization of separation conditions

After choosing the column system, the operating conditions should be optimized. Two strategies are often used in ESC. One is simultaneous experimental optimization. A variety of experimental designs has been adapted to mapping sample retention and to extracting the basic parameters of the column system, and then to predicting the optimum conditions by computation. Another strategy is sequential experimental optimization. The black box approach is used. It is characterized by the lack of knowledge of any chemical information on the sample analyzed, but the drawback of this method is that it can cause local optimization[10].

On the basis of the above-mentioned theory of the ESC, we have proposed optimization of the isothermal operating conditions[31] and of the multi-step temperature programme[32] in GC. In the same way, the optimization of isocratic elution[33] and of a multi-step linear solvent programme[26] in HPLC have been reported. The optimization for samples with unknown constituents has just begun[25].

In order to demonstrate the validity of the theory and algorithm, a sample containing ten bile acids was separated by four different isocratic experiments at constant pH and potassium dihydrogenphosphate concentration initially, and then the coefficient A_i, B_i, C_i and a, b were calculated. Finally, the optimum binary mobile phase composition of the multi-step linear programme was obtained by the overlapping resolution mapping method. The chromatogram obtained from the optimum multi-step programme is presented in Fig. 8. The result shows that the separation is quite good and the optimum programme is successful.

The peak identification and quantitation

To demonstrate the validity of the software, some experiments have been carried out. The results obtained by a C-R1B integrator, where the area measurement is based on the perpendicular drop method, and by our software based on the above-mentioned principle are presented in Table I. When the eluent is methanol–water (80:20) as shown in Fig. 9A, the resolution of the peaks 11 and 12 is about 85%. The peak areas of the pair are 15.87 and 10.92% according to the C-R1B, and 17.92 and 9.20% according to our software. However, when the eluent is methanol–water (90:10) as shown in Fig. 9B, the peak resolution of pair 11, 12 becomes about 0.2; the peak areas are 9.09 and 16.73% from C-R1B, but 17.12 and 9.04% from our software. This means that the error of the peak area measurement by the normal line method increases greatly with decrease in resolution, but the precision of the curve fitting method is very good. In addition, a shoulder, 15 in Fig. 9B, can also be detected and be identified and quantitated by our software.

TABLE I

COMPARISON OF QUALITATIVE AND QUANTITATIVE RESULTS FROM THE SAME CHROMATOGRAMS BY USE OF DIFFERENT SOFTWARE

MK represents the mark of the peak: T = baseline separated, V = overlapping, B = back shoulder peak.

Fig. 9A					*Fig. 9B*				
Peak No.	t_R *(min)*	*MK*	*ESC Area %*	*C-R1B Area %*	*Peak No.*	t_R *(min)*	*MK*	*ESC Area %*	*C-R1B Area %*
1	1.07	T	0.21	0.16	1	1.01	T	0.02	*
2	1.30	V	0.49	0.38	2	1.23	V	0.27	0.07
3	1.89	T	0.05	*	3	1.91	T	0.05	0.06
4	2.12	V	0.05	0.07	4	2.15	V	0.55	0.05
5	4.24	T	5.68	5.55	5	3.03	T	5.50	5.51
6	5.49	T	6.49	6.43	6	3.42	V	6.57	6.44
7	6.94	T	1.81	1.62	7	3.80	V	1.85	1.35
8	7.53	V	3.57	3.81	8	3.98	V	3.55	3.65
9	9.52	T	6.49	5.80	9	4.32	V	6.44	6.09
10	10.27	V	0.82	1.70	10	4.68	V	1.06	2.28
11	13.70	T	17,92	15.87	11	5.44	V	17.23	9.10
12	14.99	V	9.20	10.92	12	5.68	V	9.15	16.73
13	19.86	T	7.14	7.27	13	6.82	V	7.04	5.84
14	27.14	T	0.38	0.34	14	7.48	V	40.94	42.83
15	29.09	T	39.62	40.07	15	7.97	B	0.90	*
					16	11.60	T	0.03	*
Total			100.00	99.99				100.00	100.00

* Not detected.

The diagnosis of hardware

So-called intelligent instruments imply that the system must possess self-diagnosis. However, almost all chromatographs do not possess self-diagnosis at the heart of the system, the column. Though the evaluation of column efficiency has been reported[34], there are few methods to diagnose the chromatograph.

In the expert system for chromatography under development, the diagnosis of hardware is necessary, especially for the column in HPLC. It was demonstrated that the column, especially a small-bore column in HPLC, can exhibit high efficiency only if the column system is suitable[35]. In addition, the extra-column dead volume of the column system has a contribution to the retention time[36]. We have proposed a general strategy for diagnosing hardware. First, to ensure that the column system is suitable for chromatographic peak identification and quantitation, a standard column with a set of parameters, such as efficiency, asymmetry, retention time, selectivity etc., must be connected to the instrument system and be operated once. The results from this experiment can be compared to the set of parameters obtained from the standard column system. If great deviation is found, the column system must be improved, otherwise the system is adopted. On the other hand, monitoring changes in the chromatographic parameters, such as efficiency, asymmetry, retention time, selectivity and operating pressure on-line, allows the verification of column performance.

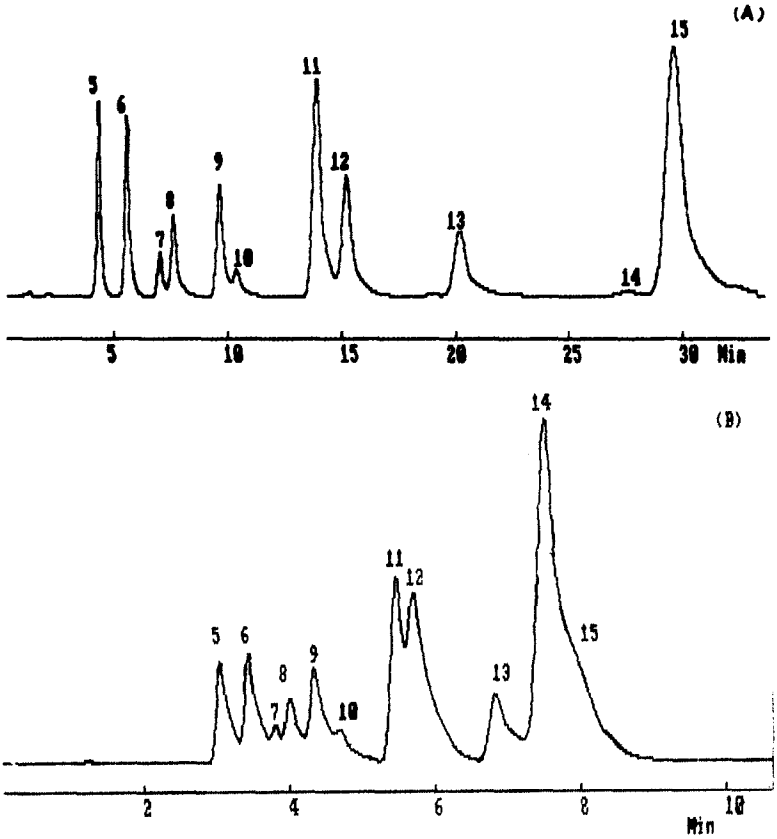

Fig. 9. The chromatograms which are handled by different software. Analysis: mixture of aromatic hydrocarbons. Column, 15 cm × 4.6 mm I.D. 10-μm Nucleosil C_{18}; flow-rate, 1.0 ml/min; detection, UV at 254 nm; mobile phase, (A) methanol–water (80:20), (B) methanol–water (90:10). Peaks: 5 = benzene, 6 = toluene, 7 = naphthalene, 8 = *p*-xylene, 9 = biphenyl, 10 = 1,3,5-trimethylbenzene, 11 = phenanthrene, 12 = anthelene, 13 = fluoranthene, 14 + 15 = *m*-diphenylbenzene + impurity (the large peak is diphenylbenzene).

TABLE II

COMPARISON OF RESULTS FROM A STANDARD SYSTEM WITH THOSE FROM A WATERS SYSTEM

n_1 and $(\tau/\sigma)_1$ are the plate number and the ratio (τ/σ) when $k' = 1$. n_∞ and $(\tau/\sigma)_\infty$ are the plate number and the area ratio (τ/σ) when k' approaches infinity; $(b/a)_{0.1}$ is the asymmetry at 1/10 peak height.

		k'	w_h	$(b/a)_{0.1}$	n	
Standard system	Benzene	0.60	10.69	1.47	7280	n_1 7760
	Naphthalene	1.32	14.78	1.20	8010	n_∞ 8210, 1.07
	Biphenyl	1.85	18.08	1.11	8080	$(\tau/\sigma)_1$ 1.00
	Phenanthrene	2.92	25.19	1.05	7950	$(\tau/\sigma)_\infty$ 0.21
Waters system	Benzene	0.58	10.71	1.56	7030	n_1 7180
	Naphthalene	1.30	15.04	1.30	7490	n_∞ 8420. 1.16
	Biphenyl	1.87	18.21	1.25	7730	$(\tau/\sigma)_1$ 1.17
	Phenanthrene	2.87	24.90	1.15	7800	$(\tau/\sigma)_\infty$ 0.20

Troubleshooting faults in a chromatograph is an high-level task requiring professional knowledge. The results of diagnosis can be passed to the inference engine for decision making in the ESC. For example, a standard column with a set of parameters is connected to an instrument system for evaluation. If a considerable decrease in column efficiency compared with standard parameters is found, our experience shows that it results from three causes, *i.e.*, (1) the dead volume of the pre-column is larger than the standard system, (2) the dead volume of the post-column is larger, (3) the dead volume of both the pre- and post-columns is larger. There are three rules in the ESC to judge what causes the decrease in column efficiency:

A considerable decrease in column efficiency with smaller k' and a constant column efficiency with larger k' is caused by factor (1)

An almost constant column efficiency with smaller k' and a considerable decrease in column efficiency with larger k' is caused by factor (2)

A decrease in column efficiency with various k' is caused by factor (3)

The Waters system has been evaluated by the diagnosis software. The results are presented in Table II. It is noted that the column efficiency and asymmetry were worse than with the standard system. The program indicates that the cause may be factor (3) and it is recommended that a user check the injection system, connecting tube and detector.

CONCLUSION

We have presented a system strategy for developing an expert system of chromatography. The knowledge base and chromatogram base, optimization of operating conditions, peak identification and quantitation on-line and the diagnosis of the hardware system have been discussed. To enlarge the range of application, the knowledge base and chromatogram base should be further expanded. The ESC is still under development.

REFERENCES

1 DuPont's Sentinel, *Am. Lab.*, 6 (1982) 15.
2 J. P. Bonnine and G. Guiochon, *Analusis*, 12 (1984) 175.
3 J. C. Berridge and E. G. Morrissey, *J. Chromatogr.*, 316 (1984) 69.
4 J. W. A. Klaessens, G. Kateman and B. G. M. Vandeginste, *Trends Anal. Chem.*, 4 (1985) 114.
5 M. G. Cirillo, *J. Liq. Chromatogr.*, 9 (1986) 3185.
6 J. N. Little, *J. Liq. Chromatogr.*, 9 (1986) 3197.
7 Cs. Horváth (Editor), *High Performance Liquid Chromatography, Advances and Perspective*, Vols. 1–3, Academic Press, New York, 1980–1983.
8 C. M. White and R. K. Houck, *J. High Resolut. Chromatogr. Chromatogr. Commun.*, 9 (1986) 4.
9 M. A. Quarry, M. A. Stadalius, T. H. Mourey and L. R. Snyder, *J. Chromatogr.*, 358 (1986) 17.
10 C. E. Goewie, *J. Liq. Chromatogr.*, 9 (1986) 1431.
11 J. L. Glajch, J. J. Kirkland, K. M. Squire and J. M. Minor, *J. Chromatogr.*, 199 (1980) 57.
12 J. L. Glajch, J. J. Kirkland and L. R. Snyder, *J. Chromatogr.*, 238 (1982) 269.
13 J. J. Kirkland and J. L. Glajch, *J. Chromatogr.*, 255 (1983) 27.
14 P. J. Schoenmakers, A. C. J. H. Drouen and L. de Galan, *Chromatographia*, 15 (1983) 688.
15 H. A. H. Billiet, A. C. J. H. Drouen and L. de Galan, *J. Chromatogr.*, 316 (1984) 231.
16 L. R. Snyder and J. W. Dolan, *Am. Lab.*, 8 (1986) 37.
17 J. W. Dolan, R. L. Snyder and M. A. Quarry, *Chromatographia*, 24 (1987) 261.
18 R. E. Dessy, *Anal. Chem.*, 56 (1984) 1200A, 1312A.
19 R. Bach in J. H. Pierce and R. A. Hohne (Editors), *Artificial Intelligence Application in Chemistry*, Maple Press, York, PA, 1986, p. 279.

20 "Expert System in HPLC", *Anal. Chem.*, 58 (1986) 1192A.
21 P. Lu and H. Li, *Chin. J. Sci. Instrum.*, 4 (1983) 237.
22 P. Lu, *Proceedings of the 4th Chinese National Conference on Chromatography, Shanghai*, 1985, P1.
23 P. Lu and X. Lu, in Y. Dong (Editor), *Advance in Science of China*, Science Press, Beijing, 1985.
24 P. Lu and X. Lu, *J. Chromatogr.*, 292 (1984) 169.
25 H. Huang, Y. Zhang and P. Lu, *Proceedings of 2nd Beijing Conference and Exhibition on Instrumental Analysis (BCEIA), Beijing, 1987*, p. 937.
26 M. Zou, Y. Zhang and P. Lu, *Proceedings of 2nd Beijing Conference and Exhibition on Instrumental Analysis (BCEIA), Beijing, 1987*, p. 871.
27 J. P. Foley and J. G. Dorsey, *Anal. Chem.*, 55 (1983) 730.
28 Y. Zhang, L. Dong, M. Bao, G. Zhou and P. Lu, *Fenxichesin Tonbao*, 3 (1984) 16.
29 H. Li, B. Li, C. Luo and P. Lu, *Kexue Tongbao*, 5 (1986) 449.
30 B. Lin, H. Li and P. Lu, *LC · GC, Mag. Liq. Gas Chromatogr.*, 4 (1987) 1206.
31 B. Lin, H. Li, C. Luo and P. Lu, *Chin. J. Chromatogr.*, 4 (1986) 193.
32 P. Lu, B. Lin, X. Chu, C. Luo, G. Lai and H. Li, *J. High Resolut. Chromatogr. Chromatogr. Commun.*, 9 (1986) 702.
33 P. Lu, R. Liu and L. Dou, *LC · GC, Mag. Liq. Gas Chromatogr.*, 4 (1986) 354.
34 J. D. Stuart and J. S. Strokis, *Pittsburgh Conference and Exposition on Analytical Chemistry and Applied Spectroscopy*, Atlantic City, NJ, 1985, Abstr. No. 785.
35 H. Chen, H. Bao, S. Fang and P. Lu, *Chromatographia*, 13 (1981) 12.
36 H. Zou, Y. Zhang and P. Lu, *Chin. J. Chromatogr.*, 6 (1986) 334.
37 C. M. Noyes, *J. Chromatogr.*, 266 (1983) 451.