Journal of Pharmaceutical & Biomedical Analysis
Vol. 11, No. 11/12, pp. 1295–1301, 1993
Printed in Great Britain

0731–7085/93 $6.00 + 0.00

Liquid chromatography solvent optimization: potential pitfalls when using a black box for developing a quality separation*

P.B. BOWMAN,‡ J.T. HANN,§ J.G.D. MARR,†‡ D.J. SALVAT‡ and B.E. THOMPSON‡

‡*Control Development, The Upjohn Company, 7000 Portage Road, Kalamazoo, MI 49001, USA*
§*Unicam Analytical Systems, York Street, Cambridge CB1 2PX, UK*

Abstract: ICOS and DIAMOND are two commercially available, semi-automated HPLC solvent optimization software packages. The resultant optimized chromatographic separation is dependent on a combination of the operator's objective, the capability of the software system and the appropriateness of the data input. The latter contains components that represent the match between the requirements of the algorithms used and the information content of the data on which those algorithms operated. Knowledge about the sample content, stability and potential sample–solvent interactions can have a significant effect on the quality of the optimal solvent composition that is calculated. The results generated during the optimization of the separation of a mixture of U-83,757 and a variety of related compounds illustrate the need to consider the significance of the contribution to the calculated optimal separation of each of these potential pitfalls, both individually and in combination with the mode of operation of the relevant algorithms. Our results indicate that the quality of the final result is highly dependent on the intelligence content of the data used.

Keywords: *Reversed-phase liquid chromatography; solvent optimization; sample stability; automation; chromatographic peak deconvolution; chromatographic peak tracking; spectral library.*

Introduction

Whether or not a particular chromatographic separation is considered to be optimal is highly dependent on the objective for which the separation will be used. For example, the separation of one component from a group of others, that are not themselves separated, may be an ideal system for use in the preparative chromatographic isolation of that single component, but it would not be useful for implementation as an impurities assay, where all the components need to be quantified. Therefore, prior to and during the development of an optimized separation for a particular series of compounds, one must always consider the ultimate objective of the exercise.

Over the years numerous rules and formulae have been developed to aid the chromatographer to develop appropriate separations. Various practical and philosophical approaches to the general issue have evolved, and some have been commercialized, as either complete instruments or as separate software packages. (References 1 and 2 provide a brief overview of some of these systems.) Two of the most recent commercial systems that have appeared are known as ICOS (for Interactive Computerized Optimization of HPLC Separations) and DIAMOND (which was initially referred to as PU 6100 solvent optimization software). Both systems are similar in that they are stand-alone, but symbiotic, software program suites, employing multivariate interpretive methodologies. They are symbiotic in that both suites will only operate on data derived from a particular diode array-based system and collected by a particular operating system — Hewlett–Packard's 1090M chromatograph and operating system for the ICOS and Unicam's Crystal (or a related detector) system and UICS data collection software. While the multivariate methodologies that form the basis of both suites may be considered to be grossly similar, there are distinct differences both in the background philosophy and mathematical complexity of the algorithms used, the applicability and the individual operation of both systems. These factors have been described in detail elsewhere [1–10].

*Presented at the 'Fourth International Symposium on Pharmaceutical and Biomedical Analysis', April 1993, Baltimore, MD, USA.

†Author to whom correspondence should be addressed.

HPLC optimization involves five steps: (i) definition of the criterion of evaluation; (ii) definition of the parameter space; (iii) data collection; (iv) data analysis and interpretation; and (v) prediction and confirmation of the optimum. Underlying each of these steps are the assumptions that the analyst understands the separation objective; the restrictions of the chromatographic system used; the limitation of any software that might be used; and the composition of the samples being investigated. While the first two assumptions might be considered fundamental, the latter two are equally important but often overlooked.

In a previous paper [1], the performance of the ICOS and DIAMOND systems was investigated and it was determined that the importance of understanding the philosophy behind the development, the mode of operation and the limitations of the software was critical when using either system to develop an optimized separation. Each of these play a role in determining the appropriateness of any data that is input into the system. In this paper we investigate some of the sample-related issues that we have encountered that also have a bearing on the consideration of data appropriateness. The significance of each issue will be examined individually, and in relation to the potential influence it might have on the operation of each of the two software systems. While not unique, many of the issues were unexpected, and serve to highlight some of the pitfalls that await analysts that approach semi- or fully automated software-driven systems with a 'black box' mentality.

Experimental

Two chromatographic systems were used, one system for each software package. In System I, the ICOS software (Version 1.0) was run on an HP 9000 Series 300 (Chemstation) computer and data was collected from a Series II HP 1090M chromatograph using a Pascal-based operating and data collection software (version 5.3), operating on the same computer. In System II, the DIAMOND software was run on a WIN 486 computer, configured with 8 MB RAM and additional operating boards. The Unicam Integrated Chromatography System (UICS version 1.0) control and data collection software was resident on the same computer, and was operated under a Windows 3.0 environment. Data were collected from a Crystal 240 diode array detector. For further details, see ref. 1.

Similar Zorbax SB-phenyl (250 × 4.6 mm i.d.) (MacMod Analytical, Chadds Ford, PA) were used on both systems. Column serial number UU 1345 was used with the HP 1090M system, and column serial number UU 1128 was used with the Unicam system. The columns were operated at ambient temperature with identical mobile phase flow rates of 1 ml min^{-1}. Methanol, acetonitrile, tetrahydrofuran (THF) and water (all HPLC grade) were obtained from Burdick and Jackson (Muskegon, MI). Phenol was obtained from Mallinckrodt (St Louis, MO), the aminopyridine (AP) from Aldrich (Milwaukee, WI) and trifluoroacetic acid (TFA) from Pierce (Rockford, IL). The other chemicals were obtained in-house.

The test mixtures were composed of varying combinations and amounts of U-83,757 (an amine), several isomers and homologues of U-83,757, and other related compounds, including an aminopyridine (AP) and phenol. The sample solvent was composed of acetonitrile–water (1:1) with 0.1% TFA.

The mobile phase combinations used were derived from the various iso-eluotropic planes that were determined using the PLANE software in the DIAMOND system (see ref. 1 for further details). Spectral libraries of the set of reference (test) compounds were created in both systems. For ICOS, the library was created outside of the ICOS software environment using the regular 'Data Editor' software. For DIAMOND, the reference library was created within the DIAMOND software environment.

An HP 8450 spectrophotometer (Hewlett–Packard, Novi, MI) was used to collect static UV spectra.

Results and Discussion

Since the operation of the ICOS and DIAMOND systems have been discussed previously [1], this will not be considered further in this paper. Also, since the same experiments were run on both systems, where similar results were observed on both systems, only those results obtained from the DIAMOND system will be presented and discussed. The same criteria will be applied when considering the significance of the observations. This approach should enable emphasis to be placed on issues

that are sample related and reduce the potential for confusion that might stem primarily from differences in the presentation of the data by the two systems.

The user must have an understanding that is as complete as possible of the samples used for developing the separation. For example, instability of any standard in either the injection solvent or the mobile phase can result in changes in the chromatograms during the runs which make interpretation of the results impossible. Our initial test mixture was made up to contain nine components, including U-83,757. Initially, we assumed that the unexpected optimization results (Fig. 1) were due to our lack of experience with the software. However, while that may have indeed been a factor, upon further investigation we discovered that there were significant sample-related elements that had contributed, unexpectedly, to the results.

While it may have been considered over optimistic to try to develop a separation for nine related compounds, at the time this was believed to reflect the pessimistic reality of the potential composition of future samples. However, this composition highlighted certain latent pitfalls, due to the conflict between the logic of the composition of the standard set of components in the test sample and the similarity of some critical physico-chemical properties of the components. Too many components with very similar spectra can produce results that are un-interpretable. The user cannot keep track of the identity of each component in the standard set. This is seen in Fig. 1, where the sharp peaks in the response surface near the methanol corner of the iso-eluotropic plane were due to an inability to correctly distinguish between spectrally similar components. A further indication of peak misidentification can be obtained by examining the retention surface for the individual peak where, again, sharp discontinuities of the surface are manifestations of a potential misassignment of that component.

While the DIAMOND algorithms can track peaks using a combination of spectra, concentration and retention times, if spectra are similar, concentration differences become critical. However, the significance of this often only becomes obvious with hindsight during or after the optimization process. A similar conclusion occurs with the ICOS where peak identification is left up to the user. If the spectra are similar, different peak heights or

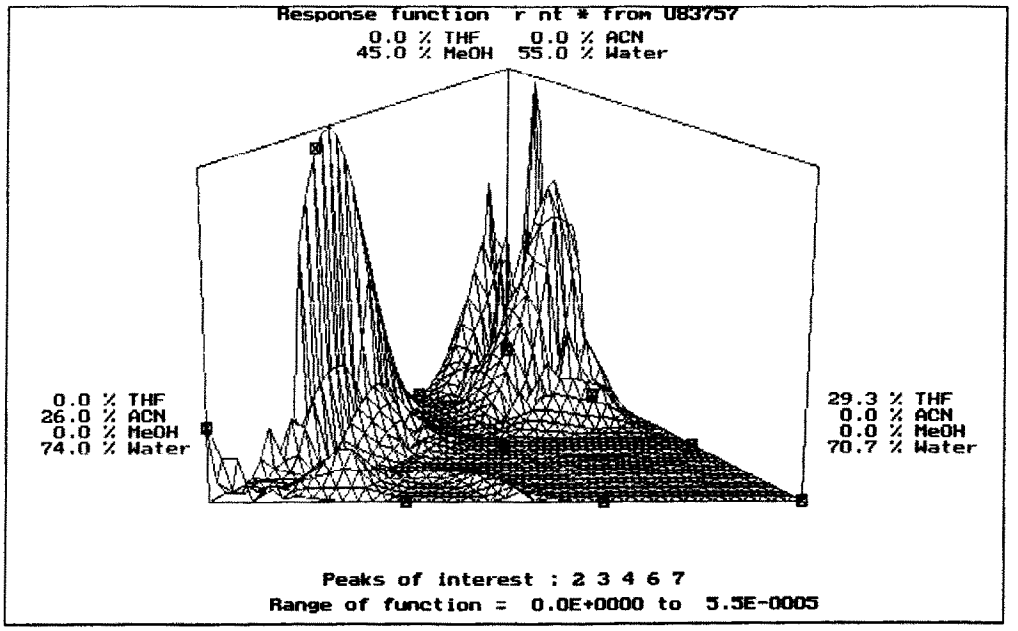

Figure 1
A response map produced following the calculation of an 'optimized' separation. While visual judgement and extrapolation of the results might have indicated an optimal in the region of the methanolic corner, the response surface indicates the 'optimum' to be along the methanol–acetonitrile–water + 0.1% v/v TFA edge. The sharp peaks apparent in the resolution map in the methanolic corner are indicative of possible peak missassignments during the peak tracking phase. (See text for further details.)

areas become the only 'reliable' means of identifying the peaks. Part of the issue related to the concentration of the individual components in the sample is the linear range of the detector. The elution times of most compounds will alter across the iso-eluotropic plane. If the compound elutes rapidly, the potential exists that the peak height may exceed the linear range of the detector. Consequently, even for a pure peak, apical spectra may not match those collected on the sides of the peak, giving the false impression of an impure peak. Likewise, the apical spectra are unlikely to match the reference in the spectral library, further adding to the impression of co-eluting components. Again, the elution profiles of the various compounds is generally only apparent after all the solvent compositions have been run, requiring a degree of judgement by the analyst when making up the initial samples.

Grossly impure 'standards' can provide two peaks where only one was expected. At various points across the iso-eluotropic plane, >10 apparently real spectra could be de-convoluted, despite only nine components being added into the sample (Fig. 2). This can lead to erroneous conclusions, particularly if the levels of the components were varied to facilitate peak tracking.

Another source of extra components in a test sample was more readily observed during a later set of experiments. The methanolic corner of the mobile phase had been determined using the PLANE software (see ref. 1). However, it required a large number of iterations to determine the acetonitrile and THF corners. After completing data collection for all 10 solvent compositions, it was discovered that all the components of interest in the sample had remained unresolved across the plane, with a single peak eluting between 20 and 35 min in each data set (Fig. 3). Investigation of this single peak determined that it was an acid induced degradation component produced as a consequence of the sample diluent that was used (acetonitrile–water, 1:1 + 0.01% TFA). The rate of degradation was such that degradation component had not been present in sufficient quantities to affect the determination of the methanolic corner, but had begun to be detected during the determination of the acetonitrile corner. Hence the numerous iterations that had occurred.

One of the consequences of this observation was the removal of TFA from the sample diluent and the mobile phase. This resulted in poor peak shape for at least one of the amines in the standard mixture. Thus, compromises

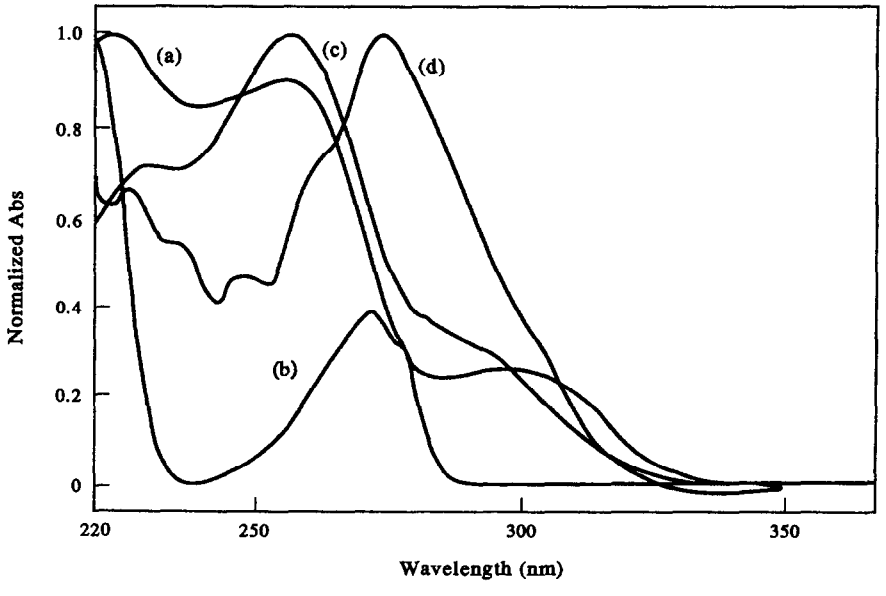

Figure 2
Comparison of two of the reference spectra ((a) and (b)) used for peak tracking with two unidentified spectra ((c) and (d)) that were only detected following peak deconvolution in some, but not all, of the 10 chromatographic data sets collected during an optimization experiment. The unidentified spectra did not correspond to any of the spectra derived from the known sample components, nor were they ever isolated as single component peaks under any of the chromatographic conditions used. (See text for further details.)

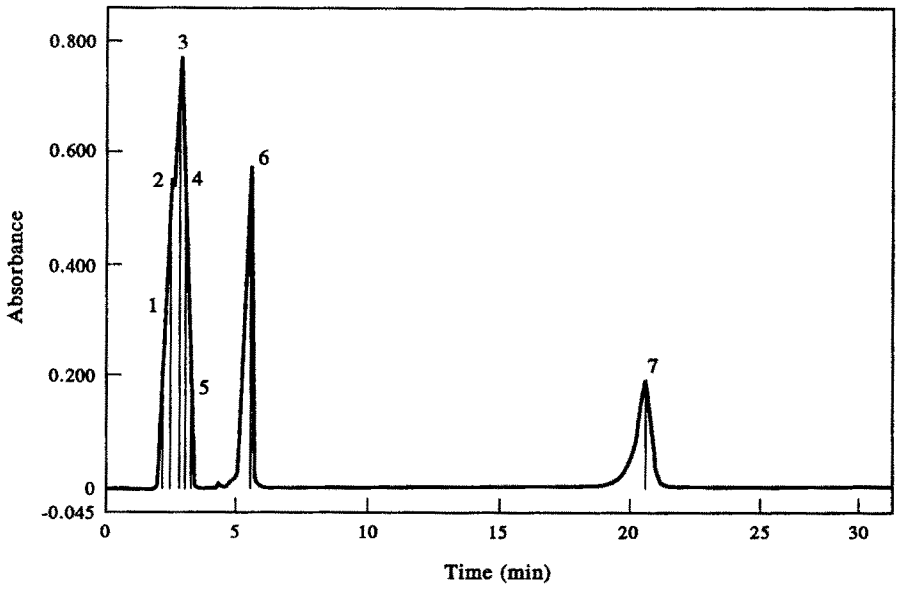

Figure 3
Peak 7 was not present when the methanol corner was defined in the iso-eluotropic plan. This degradation product continued to form during the optimization runs, resulting in a situation where the sample set used for the optimization runs did not correspond to that used to establish the plane. (See text for details.)

may have to be accepted due to the reality of the situation. In this case, it was anticipated that the unstable component was a more likely impurity than the amine. For automated optimization systems, this has another consequence. While many of the algorithms involved do not include peak shape factors in the calculations, when the projected results are presented graphically, as in a predicted chromatogram, a correlation between retention time and traditional peak shape are often assumed. Optimized separations may therefore prove not to be so ideal if there is considerable peak tailing.

Most sample–solvent interactions can be predicted from an understanding of some of the basic physico-chemical principles involved, e.g. pH-induced changes in spectra. Other sample–solvent induced changes that have been reported include ligand exchange [11] or solvent induced forms of degradation [12]. Until an appropriate knowledge base of the compound is developed such interactions may appear unexpectedly. One such example of sample–solvent interaction occurred during the experiments described in a previous paper [1], where only four components were in the test sample. Tracking the aminopyridine peak proved to be 'difficult' for the algorithms, especially when there was THF in the mobile phase. Manual investigation of this resulted in the following observations — the presence of THF in the mobile phase resulted in a change in the peak shape, with significant peak fronting occurring (Fig. 4) accompanied by significant changes in the UV spectrum of the compound (Fig. 5). Unlike those sample–solvent interactions reported elsewhere [11, 12], this interaction was determined to be a reversible phenomenon by re-injecting the sample under the different mobile phase conditions, in a random manner, rather than a time-dependent degradation of the sample. The influence of THF on the spectrum was also confirmed by dissolving the aminopyridine in solvents containing varying amounts of THF, and recording the resultant spectra in a spectrophotometer (Fig. 6). Unless the analyst is aware of the composition of the test sample, and is willing to manually review all the data, it might initially be concluded that the component of interest had not eluted within the timeframe allotted for that particular mobile phase composition. This, in turn, could lead to incorrect information being used by the algorithms, resulting in an inappropriate 'optimized' separation being predicted.

Built into the algorithms involved in both of the automated solvent optimization systems that we investigated, is the underlying assumption that any changes in retention are due solely to changes in the content of the mobile

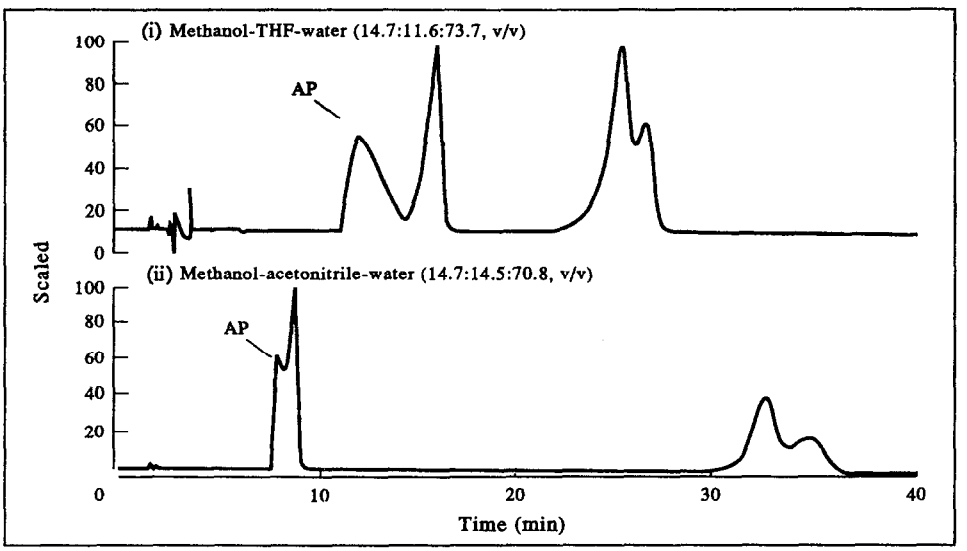

Figure 4
Comparison of the effect on the peak shape of aminopyridine (AP) collected from a sample chromatographed using a mobile phase composition of (i) methanol–THF–water (14.7:11.6:73.7, v/v/v), and (ii) methanol–acetonitrile–water (14.7:14.5:71.8, v/v/v); the detection wavelength was 220 nm. Under the second set of conditions (ii), AP is only partially resolved from a second component in the mixture. Data were collected on the ICOS system; chromatographic data has been normalized. (Sample was dissolved in acetonitrile — see text and ref. 1 for further details.)

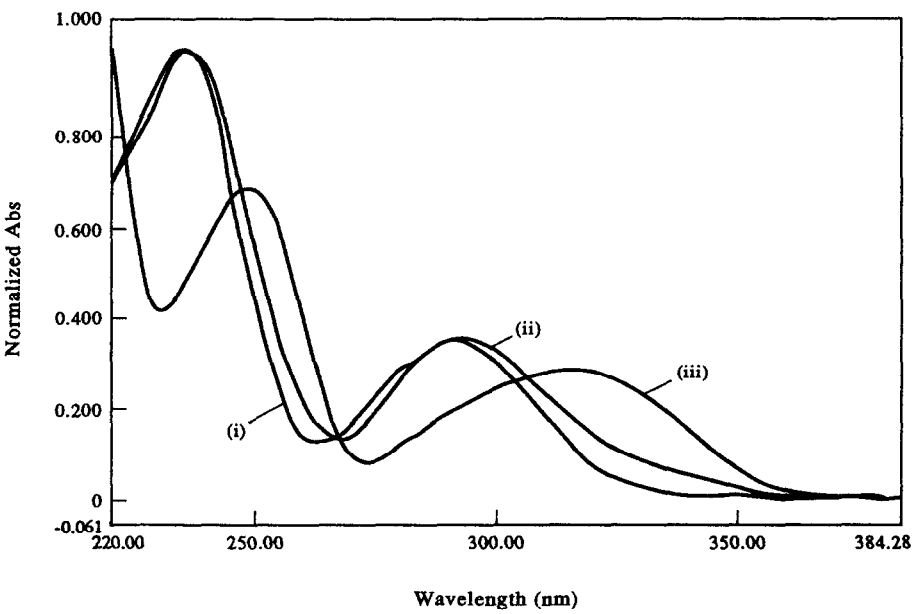

Figure 5
Comparison of the effect on the on-line UV spectrum of an aminopyridine, collected from samples chromatographed using different compositions of methanol–acetonitrile–THF–water in the mobile phase. The mobile phase compositions were (i) 44.2:0:0:55.8, v/v/v/v; (ii) 14.7:7.2:5.8:72.3, v/v/v/v; and (iii) 0:0:17.4:82.6, v/v/v/v. (Sample was dissolved in acetonitrile — see text and ref. 1 for further details.)

phase. Therefore, if the column performance changes during the run, incomprehensible results will be produced. To avoid this pitfall, it is best to verify column stability by repeating run 1 after the last run has been completed.

Conclusions

Automated solvent optimization systems have the potential of significantly impacting the way analysts approach the issue of evaluat-

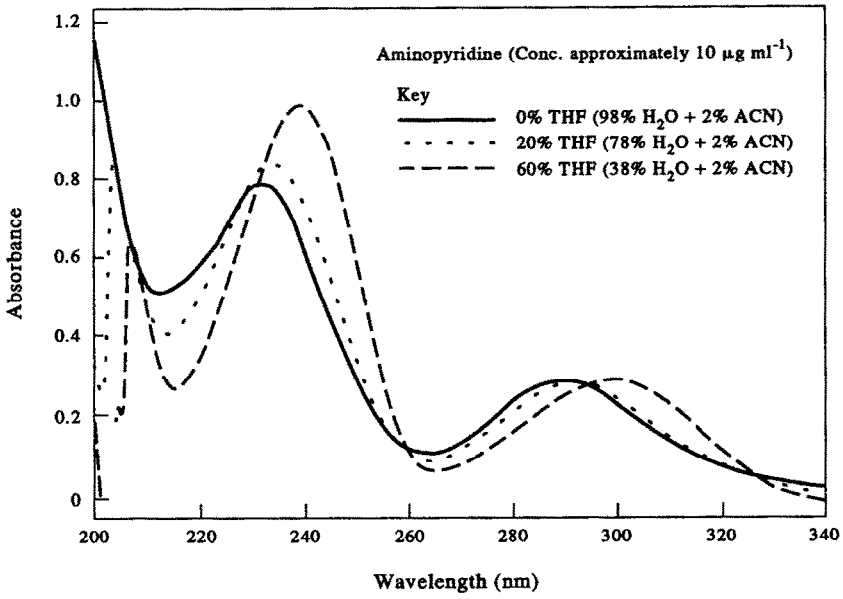

Figure 6
Comparison of the effect on the UV spectrum of an aminopyridine (AP) of changing the THF content of the sample diluent.

ing a chromatographic separation. However, it is important that the user understands the limitations of the particular approach chosen. Naïveté and/or over optimism led the authors to discover a number of 'pitfalls' that await the user of automated solvent optimization systems.

Good information leads to good optimization — the quality of the result depends on the quality of the data input. However, automated systems do not distinguish between the good and bad informational content or the appropriateness (i.e. the 'quality') of the data on which they operate. It is therefore crucial that in deciding which data to use with the system, the analyst has an understanding of the particular basic philosophy, design and limitations of the particular software package. Our work indicates that part of the data quality assessment involves a certain fundamental knowledge of the test sample, at least with respect to the potential number of components, their respective purities, stability and relative concentrations. Sample–solvent interactions can occasionally be foreseen but, again, that often depends on the analyst's knowledge of the physico-chemical properties of the sample components. With new research compounds, it is likely that this information will not be available at the time the assay development is undertaken. The significance of each of our observations will vary, depending on the rates of change, objective behind the separation and the software system used. The amount of intelligence included in the analyst–software interaction and the extent of the reliance that is placed on the final prediction have been found to be critical factors involved in producing a final result, that is optimal in both component separation and quality.

References

[1] P.B. Bowman, J.G.D. Marr, D.J. Salvat and B.E. Thompson, *J. Pharm. Biomed. Anal.* **11**, 1303–1315 (1993).
[2] A. Drouen, J.W. Dolan, L.R. Snyder, A. Poile and P.J. Schoenmakers, *LC–GC* **9**, 714–724 (1991).
[3] A.C.J.H. Drouen, H.A.H. Billiet, P.J. Schoenmakers and L. de Galan, *Chromatographia* **10**, 48–52 (1982).
[4] *Optimisation Software*, Hewlett–Packard, Publication number 12-5091-0325E.
[5] R. Faulstich and T. Catalano, *LC–GC* **9**, 776–779 (1991).
[6] R. Lynch and G. Measures, *Lab. Prac.* **39**, 61–64 (1990).
[7] S.D. Patterson, R.E.A. Escott and R.J. Lynch, *LC–GC Int.* **9**, 54–60 (1990).
[8] P.J. Naish, R.J. Lynch and T. Blaffert, *Chromatographia* **27**, 343–358 (1989).
[9] S.D. Patterson, *J. Chromatogr.* **592**, 43–49 (1992).
[10] P.J. Naish-Chamberlain and R.J. Lynch, *Chromatographia* **29**, 79–89 (1990).
[11] G.S. Baldew, K.J. Volkers, J.J.M. DeGroeij and P.E. Vermeulen, *J. Chromatogr. (Biomed. Appl.)* **491**, 163–174 (1989).
[12] J. Hansen and H. Bundgaard, *Int. J. Pharmaceutics* **6**, 307–319 (1980).

[Received for review 19 April 1993; revised manuscript received 24 June 1993]