*Journal of Chromatography*, 485 (1989) 297–309
Elsevier Science Publishers B.V., Amsterdam Printed in The Netherlands

CHROM. 21 863

# KNOWLEDGE-BASED EXPERT SYSTEM FOR TROUBLESHOOTING HIGH-PERFORMANCE LIQUID CHROMATOGRAPHIC ASSAY METHODS

KIYOSHI TSUJI* and KEVIN M. JENKINS
*Pharmaceutical Quality Control Division, The Upjohn Company, Kalamazoo, MI 49001 (U.S.A.)*

## SUMMARY

An expert system has been written to assist in troubleshooting 21 high-performance liquid chromatographic (HPLC) assay methods frequently performed in our laboratories. The knowledge used in this expert system includes the facts, heuristics (rules-of-thumb based on experience that provides short cuts over a strict textbook approach to a problem) and problem-solving strategies for diagnosing problems that occur in HPLC assay methods. The expert system operates by gathering necessary data from a user, analyzing the assay parameter, determining the most probable causes of the problem and displaying its conclusions together with justifications for its diagnosis. Data are quickly and easily entered by a user through the use of a notebook-style form that is displayed on a PC screen. The expert system analyzes the assay problem based on user-supplied chromatographic symptoms and selectivity, capacity factor ($k'$) and resolution ($R_s$) that are derived from the data. Once the system has diagnosed the most probable causes, it prints these out ranked in order of most likely causes and proceeds for further in-depth troubleshooting.

## INTRODUCTION

An expert system is a computer program that performs tasks in a particular domain or subject area, using knowledge gained through years of experience from a human expert. Expert systems are being developed for a number of different reasons: to preserve knowledge within a company or department when an expert retires, to reduce the amount of time required to perform a given task, to allow less experienced people to access easily expert-level knowledge, to act as a training tool, etc.

Expert systems are amenable to high-performance liquid chromatography (HPLC), as HPLC systems are complex, consisting of multiple instruments that require frequent attention. Several tools exist to aid in troubleshooting HPLC assays and system components. Textual information to aid in troubleshooting HPLC systems is available[1,2]; however, often it is not specific to the type of problem encountered in the pharmaceutical industry and, owing to the need for extensive interpretation, it is not readily available in a crisis situation. The HPLC Doctor, a PC-based HPLC troubleshooting program, is being marketed by LC Resources (Lafayette, CA,

U.S.A.). However, the program was found to be limited in scope and not specific to the types of instrumentation and chromatographic conditions utilized in our laboratories specializing in HPLC assays. An expert system, Expert Chromatographic Assistance Team (ECAT), is being developed by Varian (Walnut Creek, CA, U.S.A.) to perform the task of designing, analyzing, optimizing and troubleshooting HPLC methods[3]. ECAT is written in Zetalisp, a specialized artificial intelligence language, and operates on a Symbolic 3670 workstation. This program will require considerable refinement and further input from the scientific community in order to meet its ambitious objectives.

This paper describes the knowledge domain of the knowledge-based expert system program. The knowledge-based programming technique[4,5] was utilized previously by us to construct two expert system programs, the Expert Microbiological Decision Advisory System (XMiDAS)[6] and the Expert Troubleshooting Gel-Clot *Limulus* Amebocyte Lysate (LAL) for bacterial endotoxin assay[7].

## EXPERIMENTAL

### *Operating environment*

The HPLC troubleshooting advisory system was developed on an IBM PC/AT using the M.1 (version 2.1) software (Cimflex-Teknowledge, Palo Alto, CA, U.S.A.). M.1 is rule-based expert system software which uses inference (*modus ponens*), depth first and backward chaining search processes to solve problems. M.1 deals with uncertainty associated with human decision making by use of a certainty factor developed for the MYCIN Expert System by Buchanan and Shortliffe[8]. The capability of M.1 in dealing with uncertainty is utilized in the program to weigh evidence and thereby establish priority for efficient troubleshooting.

In order to collect pertinent information, notebook-style data entry formats were constructed using Windows for Data software (Vermont Creative Software, Richford, VT, U.S.A.). Eight graphic panels were drawn using HALO Graphics version 2.26a and Dr. HALO III software (Media Cybernetics, Silver Spring, MD U.S.A.). A typical graph is presented in Fig. 1. The data entry formats, graphics and database file are linked with M.1 software using an external call function developed by us using C language. The program developed can be distributed through the VAX network or by sending six 360K floppy diskettes to run on an IBM PC under DOS version 3.10 and higher operating environment.

### *Knowledge domain*

The knowledge domain residing in the program was obtained from publications[1,9–12] and chromatographic expertise residing in the Pharmaceutical Quality Control Division of The Upjohn Company (Kalamazoo, MI, U.S.A.). The program contains files specific to 21 HPLC assay methods frequently performed in our laboratories. The program uses four notebook-style data entry formats to facilitate collection of pertinent assay information, such as theoretical plate number, mobile phase flow-rate, column pressure and peak retention time. Selectivity, capacity factor ($k'$) and resolution ($R_s$) are calculated from the data to evaluate peak elution characteristics (Table I). A chromatographer's observations/symptoms, such as the presence of ghost peak, peak shapes, baseline noise, pressure problems, short column life, assay

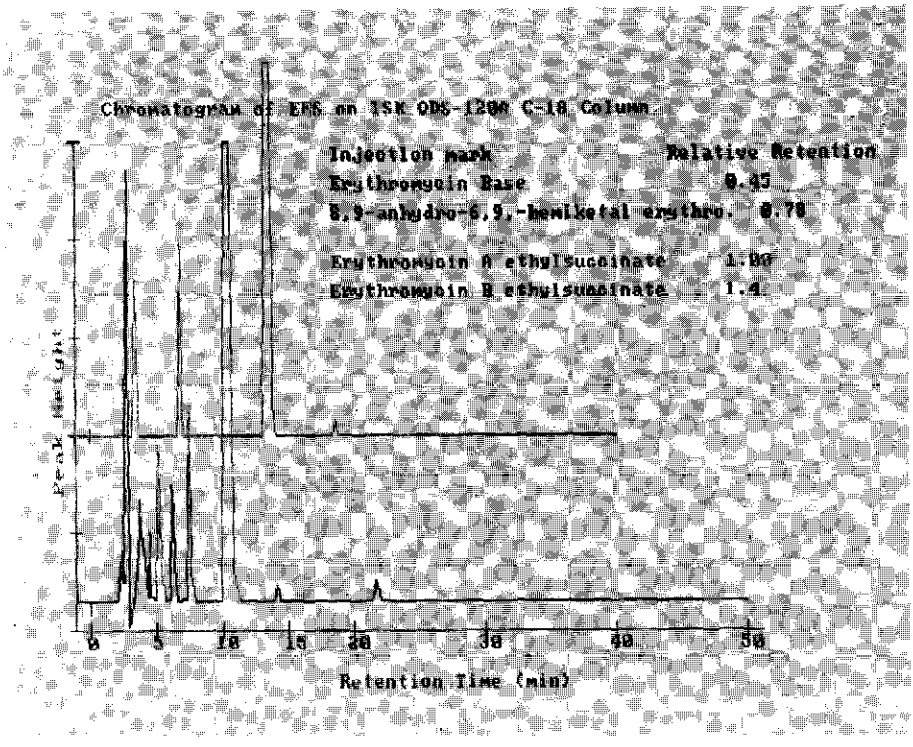


Fig. 1. Graphic illustrating superimposed, actual sample (upper) and impurity sample (lower), reversed-phase HPLC traces for erythromycin ethylsuccinate to aid in peak identification.

relative standard deviation (R.S.D.) and shifting retention times, are also used to analyze the assay problem.

The diagnosed probable causes of the assay problem having a certainty of greater than 30% are listed according to their degree of certainty. Rationale to support the diagnosis is also presented. The two most probable causes are selected for interactive, in-depth troubleshooting. As this is a beta-test version, user's comments are frequently requested to criticize and improve the utility of the program. The style of consultation format used in the program and input of raw data to analyze chromatographic performance, have been selected with the idea of eventually imbedding the program in the LIMS/VAX environment to evaluate automatically and continuously chromatographic performance during each HPLC assay run. This would allow on-line troubleshooting of HPLC assays, provide for multiple users of the system over a network and facilitate sharing of useful troubleshooting experience and knowledge throughout our organization.

## RESULTS AND DISCUSSION

### *Developing an expert system*

Expert systems require that the purpose of the system, the actual application area, potential users and source of expertise be identified before development can

TABLE I

CALCULATED DATA PRINT-OUT FORMAT

*HPLC troubleshooting advisory system*

| | |
|---|---|
| Date: 04/20/89 | Time: 08:13:57 |
| User Name: K. Tsuji | |
| HPLC Assay for: Erythromycin Ethylsuccinate | |
| Analytical Column: Toyo Soda ODS | |
| Column pressure: | 750 psi |
| Column temperature: | 70.0 degree C. |
| Theoretical Plate: 1500 per column | |
| Calculated using the actual value, | |
| alpha: | 1.4 |
| peak 1 capacity factor: | 4.1 |
| peak 2 capacity factor: | 2.9 |
| resolution: | 3.1 |
| Mobile Phase Composition, | |
| percent am-acetate: | 60% |
| percent water: | 30% |
| percent acetonitrile: | 10% |
| pH: | 7.4 |
| Actual Flow-rate: | 1.0 ml/min |

begin. The aim of our project was to construct an expert system that would advise an entry-level laboratory technician to diagnose problems in an HPLC system. The sources of the expertise were publications[1,9–12] and chromatographic experts in the Pharmaceutical Quality Control Division of the Upjohn Company.

Appropriate computer hardware and software must meet our selection criteria. The expert system has to run on computer hardware that is readily available in our laboratories and is fairly inexpensive. Expert systems often require that the following types of software be selected: an expert system shell program, graphics package (for displaying diagrams to the user), programming language and database. The HPLC troubleshooting advisory system was developed using an IBM PC/AT as the computer hardware. Software used by the system includes M.1, Windows for Data (for developing data entry forms), Dr. Halo (graphics), Dr. Halo III (graphics) and DBASE III (database).

As there are varieties of HPLC assay methods and numerous types of problems that can occur with them, the decision was made to focus the development of the system on problems that occur with normal and reversed-phase isocratic chromatography with UV detection. Twenty-one HPLC assay methods frequently performed in our laboratories were selected.

Development of the expert system begins with a knowledge acquisition session. The developer of the expert system, or knowledge engineer, extracts information from the experts, analyzes and organizes the knowledge and represents that information in a computer program. The development of a knowledge system is an interactive

process, one of knowledge acquisition, implementation, testing and expert feedback. This process repeats itself each time the expert system is expanded to handle more and different types of problems in the domain.

*Design of the HPLC troubleshooting advisory system.* The problem-solving approach of the HPLC troubleshooting advisory system is as follows:

(1) gather symptoms and other relevant information from the laboratory technician;

(2) derive selectivity, capacity factor ($k'$) and resolution ($R_s$) from the data;

(3) analyze the assay problem, obtaining more data if necessary;

(4) diagnose the most probable causes of the problem;

(5) print out the probable causes ranked in order of most likely causes and justifications for those decisions;

(6) in-depth troubleshooting of highly ranked causes to identify the source of the problem.

The development of each of these steps is described in more detail below.

*Data entry.* Data are quickly and easily entered by the laboratory technician through the use of four notebook-style forms that are displayed on the PC screen. These forms facilitate the collection of information such as theoretical plate number, mobile phase flow-rate, column pressure, peak retention time, etc. An example of one form is shown in Table II.

*Analyze assay problem.* Assay problems are analyzed based on user-supplied chromatographic symptoms and selectivity, capacity factors and resolutions that are derived from the entered data. The program evaluates the chromatographic parameters using the following rules-of-thumb:

(1) The analytical column in use must be packed with the same column packing material as specified in the procedure.

(2) Unless specified, the accepted column temperature is between 16 and 27°C. Temperatures less than 16°C are considered too low and those higher than 27°C too high. If an assay requires an elevated column temperature, a deviation of ±1°C is permitted from that specified in the procedure.

(3) The program permits a deviation of the mobile phase pH if it is within ±0.1 of that specified in the procedure. Exceptional (out of the limit) pH by itself may not initiate action unless it is coupled with other exceptional symptoms.

TABLE II

CHROMATOGRAPHIC SYMPTOMS/OBSERVATIONS INPUT BY A USER

| | |
|---|---|
| Asymmetric peak | Shifting retention times |
| Broad peak | Spike in baseline |
| Split peak | Cyclic drift of baseline |
| Tailing peak | Continuous drift of baseline |
| Ghost peak | Sudden shift of baseline |
| Negative diffraction | Unsteady pressure |
| Inconsistent internal standard peak height | Pulsation |
| Low detector responses | Sudden loss of pressure |
| Short column life | Sudden increase of pressure |
| High R.S.D. | |

(4) The mobile phase flow-rate is compared with that specified in the procedure. Flow-rate is considered exceptional if it is over 1.15× or less than 0.85× of the flow-rate specified in the procedure. A flow-rate of less than 0.2 ml/min is considered as no flow.

(5) A column pressure of less than 800 p.s.i. is considered too low and that of over 3000 p.s.i. too high. A column pressure of less than 200 p.s.i. is considered as no pressure.

(6) Most HPLC assays performed in our laboratories use a major component peak and an internal standard peak to calculate the resolution factor. As these two peaks are approximately equal in height, a band-width ratio of 1:1 is used to determine peak resolution characteristics from the $R_s$ value. An $R_s$ of over 1.25 is considered acceptable, between 1.25 and 1.0 as fair and less than 1.0 as peak coelution[11,12].

(7) A theoretical plate number ($n$) per column of between 900 and 1200 is considered fair and less than 900 as poor.

(8) The following windows (limits) have been set around the ideal capacity factor ($k'$) to evaluate peak elution characteristics:

| Too fast | Fast | O.K. | O.K. | Slow | Too slow |
|---|---|---|---|---|---|
| $k'-0.25k'$ | $k'-0.1k'$ | $k'$ | $k'+0.1k'$ | $k'+0.25k'$ | |

When the capacity factor of a peak is greater than $k' + 0.25k'$, the peak is considered to elute too slowly. When a peak elutes between $k' + 0.1k'$ and $k' + 0.25k'$, it is considered as eluting slowly, between $k' - 0.1k'$ and $k' - 0.25k'$ as eluting fast and less than $k' - 0.25k'$ as eluting too fast. If elution of a non-retained peak takes over 8.0 min, chromatography is considered too slow. When an internal standard peak elutes but not the sample peak, the program notes "no sample peak eluting". When the solvent peak elutes but not an internal standard or sample peak, the program notes "only solvent eluting". When no peak elutes, the program notes "no peak eluting".

Possible problematic chromatographic symptoms provided by a user may include such items as listed in Table II. The entry of misshaped chromatographic peaks is aided by the graphs typified in Fig. 2.

*Diagnosis.* From the combination of chromatographic parameters and chromatographic observations and symptoms, the program uses inference mechanisms to diagnose the problem. Probable causes of the assay problem are listed according to their certainty and reasons to support the diagnosis are presented. The two most probable causes of the problem will then be interactively examined to troubleshoot the problem.

*Certainty factor.* The certainty factor is used to control flow of logic. For example, when the chromatographic data indicate that no peak is eluting by detecting no column pressure and no mobile phase flow, the program asks if HPLC instruments, *e.g.*, pump, detector and recorder are turned on and detector wavelength is set as specified in the procedure. When the user confirms that all these instruments are on and the wavelength is properly set, the program concludes that there is a 98% certainty that a diagnosis "pump lost prime" is true. When the diagnosis is supported by less evidence then the certainty of the diagnosis becomes low. For example, if "no column

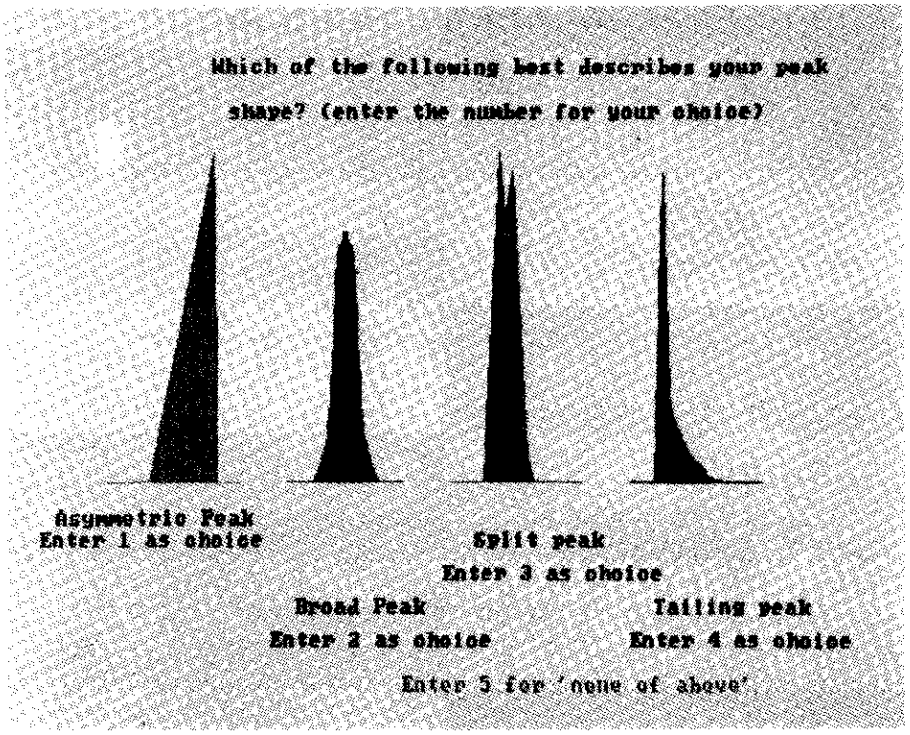


Fig. 2. Graphic to aid in computer entry of misshaped peaks.

pressure" is the only supporting evidence, then a low certainty (40%) accompanies the diagnosis. When additional evidence supports the diagnosis, *e.g.*, (1) no peak elutes, (2) no column pressure and (3) no flow of the mobile phase, the certainty increases to 70%.

*Consultation sub-modules.* Use of an analytical column not specified in the procedure prompts a warning, "use of a wrong column", regardless of whether a chromatographic parameter shows any problem. The interactive consultation continues and the degree of certainty for the diagnosis increases when additional evidence, such as sub-optimum peak resolution and/or low effective plate number, is noted.

When the existence of a dead volume is strongly suspected as the cause of a broad peak, the consultation may use the following four consultation sub-modules: (1) tubing size, (2) tubing connector, (3) tube cutting technique and (4) ferrule. Each sub-module consist of question and answer sessions and recommendations. These sub-modules integrate during consultation and operate in the following manner:

... [When system suspects the existence of dead volume.]

*Are you using a pre-cut 0.007-in. internal diameter (I.D.) or a smaller I.D. tubing between the column and the detector?*

>> Yes.

We suspect that loss of column performance may be due to incorrect plumbing creating dead volume. Dead volume is a gap or space between the connecting tubing

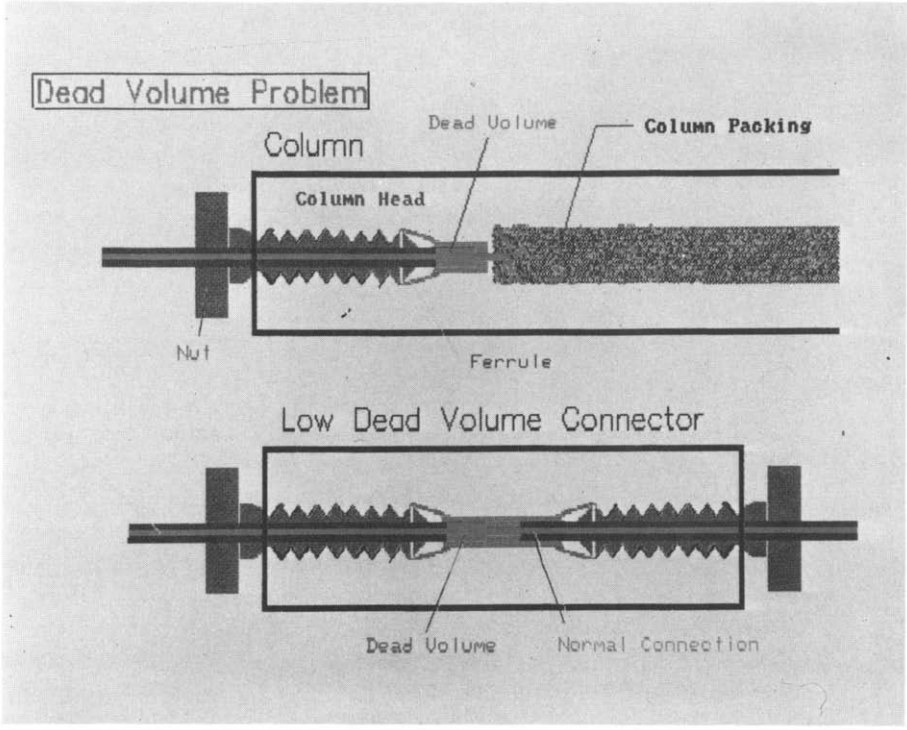


Fig. 3. Cut-away diagram of tubing connections illustrating location of two common dead volume situations.

or between tubing and a column. Extra volume mixes samples and causes turbulence which impedes separation of compounds. Dead volume must be avoided to achieve successful separation of chromatographic peaks.

... [System now displays a diagram (Fig. 3) on the screen].

*RECOMMENDATION:*

(i) We recommend the use of a pre-cut stainless-steel 0.007-in. I.D. connecting tubing. It is difficult to make a clean cut necessary to obtain the maximum peak resolution. However, if you must cut your own tubing, use a tubing cutter specifically designed to produce smooth burr-free cuts. Use a file carefully to remove burr and smooth the surface of the cuts.

(ii) Tubing connecting the analytical column and detector must be as short as possible; it must be one continuous line with an I.D. of 0.007 in. or less. Never use tubing having an I.D. greater than 0.01 in.

*Example session with the HPLC troubleshooting advisory system*

... [System suspects loss of pump prime] (Table III).

*Consultation 1:*

Examine the solvent tubing in the mobile phase reservoir.

TABLE III

EXAMPLE SESSION TO TROUBLESHOOT "LOSS OF PUMP PRIME" PROBLEM

| *System prompts* | *User's input* |
|---|---|
| *CONSULTATION 1:* | |
| Examine the solvent tubing in the mobile phase reservoir. | |
| Does the end of the tubing reach the mobile phase fluid level? | Yes |
| Check the height of the mobile phase reservoir. | |
| Is the mobile phase reservoir located higher than the pump head? | No |
| *Recommendation:* | |
| The mobile phase reservoir must be located higher than the pump head to maintain an adequate pump prime. | |
| *End of consultation* | |
| *CONSULTATION 2:* | |
| Examine the solvent tubing in the mobile phase reservoir. | |
| Does the end of the tubing reach the mobile phase fluid level? | Yes |
| Check the height of the mobile phase reservoir. | |
| Is the mobile phase reservoir located higher than the pump head? | Yes |
| Examine the pump high-pressure limit switch setting. | |
| Is the pump pressure limit switch set at or above 3000 p.s.i. | Yes |
| Are you still losing pump prime? | Yes |
| *Statement:* | |
| Since your pump continues losing prime, the possibility exists that an "air lock" has formed in your pump head assembly. An air lock is created when (1) insufficiently degassed phase is used and/or (2) leakage of air from the low-pressure side of the pump. | |
| *Statement:* | |
| Piston action creates vacuum which draws the mobile phase into the pump head. An insufficiently degassed mobile phase forms air bubbles by vacuum. | |
| *Instruction:* | |
| Apply sufficient vacuum while sonicating for 3–5 min to degas thoroughly your mobile phase. To prime your pump, apply suction at a pump head outlet or a pump bypass tube opening while running the pump at a high flow-rate, 5–10 ml/min. | |
| Has this consultation solved your problem? | No |
| *Instruction:* | |
| Disassemble the pump head assembly. Examine the inlet and outlet check valves and the ball valves under a stereomicroscope for cracks and scratches. | |
| *Instruction:* | |
| Replace the worn out parts. You may call the Instrument Shop to repair the pump. If frequent replacement of check valves is required, the presence of particulate matter in the mobile phase is suspected. It is a good idea routinely to filter the mobile phase containing buffer. | |
| Examine the pump head assembly (Fig. 4 is shown). | |
| *Instruction:* | |
| The pump inlet/outlet check valve may be clogged. Disconnect the pump tubing and connect a syringe (50 ml or larger size) to the pump outlet opening. Start the pump at a flow-rate of 5–10 ml/min. | |
| Type a number to identify leading site (Fig. 5 is shown). | # |
| Interactive consultation continues until the problem is identified and solved or the program exhausts its reasoning. | |

*Does the end of the tubing reach the mobile phase fluid level?*
>> Yes.
... [System requests the user to check the location of the mobile phase reservoir.]

*Is the mobile phase reservoir located higher than the pump head?*
>> No.
... [System determines that the causes for the problem is identified, as loss of pump prime is due to low height of the mobile phase container relative to the pump head].

*RECOMMENDATION:*
The mobile phase reservoir must be located higher than the pump head to maintain an adequate pump prime.
... [This ends the consultation session].

*Consultation 2:*
... [If the answer to the question, "*Is the mobile phase reservoir located higher than the pump head?*" is ">> Yes"]:
Examine the high pressure limit switch setting of the pump.

*Is the pump pressure limit switch set at or above 3000 p.s.i.?*
>> Yes.
... [System tries to make certain that the problem still persists].

*Are you still losing pump prime?*
>> Yes.
Since your pump continues losing prime, the possibility exists that an "air lock" has formed in your pump head assembly. An air lock is created when:
(i) insufficiently degassed mobile phase is used and/or
(ii) leakage of air from the low-pressure side of the pump.

*INSTRUCTION:*
Piston action creates vacuum which draws the mobile phase into the pump head. An insufficiently degassed mobile phase forms air bubbles by vacuum. Apply sufficient vacuum while sonicating for 3–5 min to degas thoroughly your mobile phase. To prime your pump, apply suction at a pump head outlet or a pump bypass tube opening while running the pump at a high flow-rate, 5–10 ml/min.

*Has this instruction solved your problem?*
>> No.

... [System goes on to question leakage of air from the low-pressure side and request disassembly of the pump head].

*INSTRUCTION:*
Disassemble the pump head assembly. Examine the inlet and outlet check valves and the ball valves under a stereomicroscope for cracks and scratches. Replace

the worn out parts. You may call the Instrument Shop to repair the pump. If frequent replacement of check valves is required, the presence of particulate matter in the mobile phase is suspected. It is a good idea routinely to filter mobile phase containing buffer.

... [System now displays a cut away diagram of a pump head assembly (Fig. 4) to assist graphically in repairing a pump.]

... [If pump is still losing prime]:

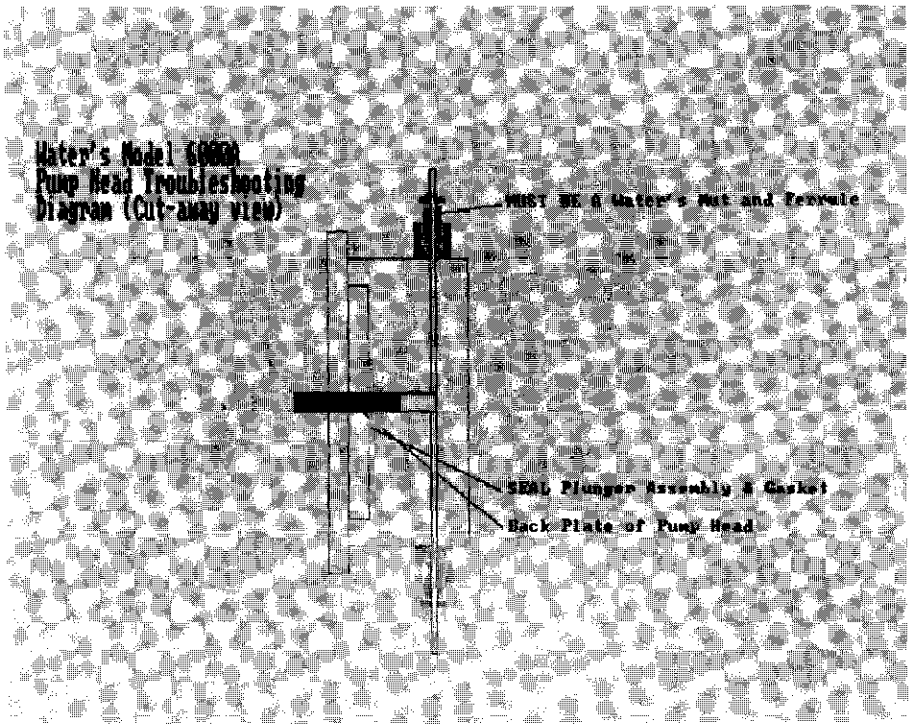


Fig. 4. Cut-away diagram of the Waters Assoc. Model 6000A pump.

*INSTRUCTION:*

Examine the pump head assembly. The pump inlet/outlet check valve may be clogged. Disconnect the pump outlet tubing and connect a syringe (50 ml or larger size) to the pump outlet opening. Start the pump at a flow-rate of 5-10 ml/min.

... [When solvent leakage is suspected, a diagram is displayed to assist in locating the leakage site (Fig. 5)].

... [Consultation continues until the problem is identified and solved].

At the end of a consultation, the program automatically saves the entire consultation and also the comments made by the user. This is to help the critical knowledge domain of the program, examine its utility and improve its performance.

Benefits of this program include increased productivity by minimizing assay down-time, improved assay quality and interactive training for new personnel.

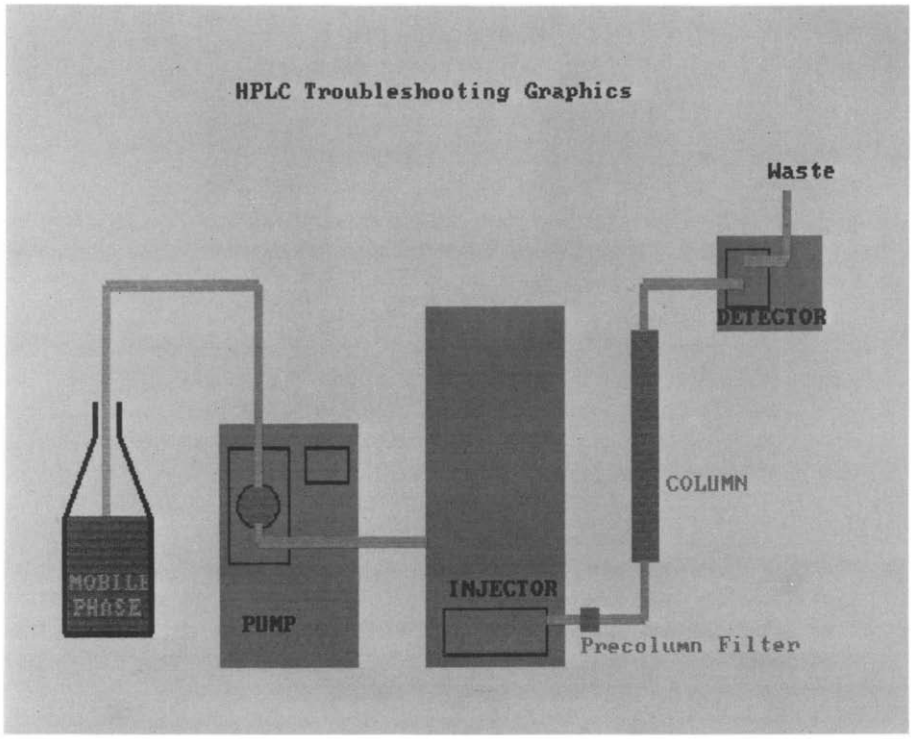


Fig. 5. Graphic to assist in locating solvent leakage site.

ACKNOWLEDGEMENTS

We gratefully acknowledge the technical contributions of W. F. Beyer, C. S. Foran, S. M. Ledden, J. D. Stodola and J. S. Spencer (Kalamazoo College, Kalamazoo, MI, U.S.A.). Shoko Tsuji (Cimflex-Teknowledge, Palo Alto, CA, U.S.A.) is also thanked for revision of the manuscript.

REFERENCES

1 D. J. Runser, *Maintaining and Troubleshooting HPLC Systems*, Wiley, New York, 1981.
2 J. W. Dolan and L. R. Snyder, *Troubleshooting LC System*, Humana, Clifton, NJ, 1989.
3 R. Bach, J. Karnicky and S. Abbott, in T. H. Pierce and B. A. Hohne (Editors), *Artificial Intelligence Applications in Chemistry* (*ACS Symposium Series*, No. 306), American Chemical Society, Washington, DC, 1986, p. 278.
4 P. Harmon and D. King, *Expert Systems*, Wiley, New York, 1985.
5 F. Hayes-Roth, D. A. Waterman and D. B. Lenat, *Building Expert Systems*, Addison-Wesley, Reading, MA, 1983.
6 K. Tsuji, K. M. Jenkins and J. M. Price, *Pharm. Technol.*, 12 (1988) 154.
7 K. Tsuji, *J. Parenteral Sci. Technol.*, 43 (1989) 240.
8 B. G. Buchanan and E. H. Shortliffe, *Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Addison-Wesley, Reading, MA, 1984.

9 P. J. Schoenmakers, *Optimization of Chromatographic Selectivity, a Guide to Method Development*, Elsevier, Amsterdam, 1986.
10 S. B. Schram, *The LDC Basic Book on Liquid Chromatography*, Milton-Roy, St. Petersburg, FL, 1980.
11 L. R. Snyder and J. J. Kirkland, *Introduction to Modern Liquid Chromatography*, New York, 1979.
12 C. L. Olson, in K. Tsuji and W. Morozowich (Editors), *GLC and HPLC Determination of Therapeutic Agents, I*, Marcel Dekker, New York, 1978, Ch. 1, pp. 1–46.