Predicting Gas Chromatographic Separation and Stationary-Phase Selectivity Using Computer Modeling

Frank L. Dorman,*,† Paul D. Schettler,‡ Christopher M. English,† and Dinesh V. Patwardhan†

Restek Corporation, 110 Benner Circle, Bellefonte, Pennsylvania 16823, and Juniata College, 1700 Moore Street, Huntingdon, Pennsylvania 16652

A computer modeling technique has been developed which allows for the prediction of chromatographic separation and stationary-phase selectivity. This technique enables development of application-specific gas chromatographic columns by allowing for the simultaneous optimization of physical dimensions, flow and temperature programs, and stationary-phase composition. Stationary-phase selectivity is the most powerful tool available to achieve a separation; however most commercially available columns were not designed to have a selectivity specific to the separations for which they are used. The techniques described in this paper were developed to address this need.

The general resolution equation below describes the terms that impact separation between two analytes.

$$R = (N^{1/2}/4)(\alpha - 1/\alpha)(k/1 + k)$$

where $(N^{1/2}/4)$ is the efficiency (column) factor expressed in terms of theoretical plates, $(\alpha-1/\alpha)$ is the separation (selectivity) factor, and (k/1+k) is the retention factor.

Chromatographically, the most important term in this expression is the selectivity factor $(\alpha - 1/\alpha)$, where α is selectivity, k_2/k_1 and the k's are retention factors. This is the term that can have the greatest impact on separation of two closely eluting analytes, but it is also the term that is least controllable by analysts using a given column.

It is possible to address this term in some limited capacity by combining lengths of individual columns,^{2–5} but the required use of press-fit connections or other connectors limits the utility of

this in a laboratory setting. Adjusting the temperature program also has limited utility for increasing resolution R by adjusting α .

In general, chromatographers do not typically consider the stationary phase as a tool for separation, because it is not something that can be easily varied or adjusted to achieve the optimum separation. Historically, capillary column stationary phases used for gas-liquid chromatographic (GLC) separations were synthesized with little regard for the compounds that would ultimately be separated. Liquid phases were constructed from either poly(ethylene glycol) (PEG, Carbowax, Stabilwax, etc.) or, more commonly, polysiloxanes having phenyl, methyl, trifluoropropyl, or cyanopropyl functional groups as side chains. In many cases, these capillary phases were the same phases that were developed for use in packed-column GLC. Most of the developments in capillary phases have been to address the thermal stability of the polymers, 6 or in the deactivations used prior to coating of the stationary phase. Many of these techniques have been previously summarized.7

Most analysts choose a commercially available phase, in common column dimensions ($\sim\!30~\text{m}\times0.25\text{-mm}$ i.d. \times 0.5 μm d.f.), and then adjust the temperature regime (retention factor) to get the best possible separation. Often, this approach works well enough for common analytes, but it sometimes fails for critical analyte pairs. In these situations, particularly, the ability to "tune" the stationary phase to the particular analysis being performed would be an advantage.

Presently, tuning a stationary phase for a particular separation consists of matching the polarity of the compounds to be separated to the polarity of the stationary phase. Frequently, a "like dissolves like" rationale is used to explain why certain compounds are more strongly retained on certain phases than on others and why separation does not follow simple analyte boiling points.

The key to achieving an optimum separation is to use a column with a stationary phase that is *selective* for the analytes of interest. However, retention and selectivity are controlled by all of the analyte—stationary-phase forces. These forces are categorized as electrostatic (including partial charges on atoms), dipolar, hydro-

^{*} To whom correspondence should be addressed.

[†] Restek Corporation.

[‡] Juniata College.

Modern Practice of Gas Chromatography, 3rd ed.; Grob, R. L., Ed.; John Wiley and Sons: New York, 1995.

Krupcik, J.; Guiochon, G.; Schmitter, J. M. J. Chromatogr. 1981, 213, 189– 201.

⁽³⁾ Krupcik, J.; Mocak, J.; Simova, A.; Garaj, J.; Guiochon, G. J. Chromatogr. 1982, 238, 1–12.

⁽⁴⁾ Takeoka, G.; Richard, H. M.; Mehran, M.; Jennings, W. J. High Resolut. Chromatogr., Chromatogr. Commun. 1983, 6, 145–151.

Ingraham, D. F.; Shoemaker, C. F.; Jennings, W. J. Chromatogr. 1982, 239, 20-50

⁽⁶⁾ Jennings, W.; Lautamo, R. M. A.; Reese, S. L. Advances in Capillary Columns for Gas Chromatographic Applications, Chromatography, InCom: Duesseldorf. 120 1997.

⁽⁷⁾ Sandra, P.; David, F.; Proot, M.; Diricks, G.; Verstappe, M.; Verzele, M. J. High Resolut. Chromatogr., Chromatogr. Commun. 1985, 8 (11), 782-798.

gen bonding, and London (or van der Waals). A change in any one or more of these forces affects retention (and with it, selectivity). For most separations, one or two of these forces will dominate the separation, but rarely is it as simple as matching the polarity of the liquid stationary phase to the compounds to be analyzed.

If each of these forces is not considered in the development and selection of a stationary phase for an analysis, an optimum separation will not be achieved. For many years, method development chemists have attempted to perform analyses using general-purpose columns that were not designed for a specific separation or selectivity. As a result, these columns seldom perform the analysis as well as possible, leaving analysts to settle for a less than ideal separation.

Prediction of Resolution. The science of retention properties has been understood for some time;9 however, only in the past few years have computers been able to apply thermodynamics to the prediction of chromatographic retention at a fundamental level. We now have the ability to predict retention order, retention time, peak width, and therefore, resolution for stationary phases that have yet to be synthesized in a reasonable amount of time. Several new stationary phases designed for specific separations have been previously reported using the modeling process addressed in this paper. 10,11 This approach consists of two different processes: (1) modeling the selectivity of polymers and chromatographic separation of columns on the basis of empirical data and (2) determining new stationary-phase functionalities based on a molecular dynamics approach. Using this approach results in stationary-phase selectivity being "tuned" to the compounds being separated while also optimizing the column dimensions. In addition to tuning for the specific separation, other desirable qualities also can be designed during the modeling process; for example, thermal stability, inertness, etc. This paper will address the first approach, in which the selectivity of the stationary phase is determined using an empirical approach, and the separation is predicted using thermodynamic relationships. The molecular dynamics approach to determining selectivity will be discussed separately.

Previous Methods. The first technique reported that was capable of determining selectivity from mixtures of different polysiloxanes that resulted in commercially available capillary columns utilized "window diagramming". 5,12,13 This technique is capable of predicting separation for analytes on polysiloxane mixtures or a single polysiloxane with functionalities corresponding to the formula of the predicted mixture. Window diagrams are limited, however, because they typically account for polymer variation only and do not also account for the column and retention factors in the general resolution equation shown above. These factors must be modeled using some other technique, such as EZGC (Analytical Innovations, Inc., Beavercreek, OH) or another physical modeling program that allows for optimization of the

oven program and column dimensions simultaneously with optimization of stationary-phase composition. In addition, many modern analytical separations require more than one or two functionalities in the polysiloxane stationary phase in order to achieve an optimal separation. Window diagrams do not lend themselves to simultaneous optimization in multidimensional space, making that approach limited to most current analytical challenges, especially when considering that there is not a simultaneous optimization of physical parameters when using window diagram methods.

Procedure. To calculate separation of compounds on the basis of stationary-phase selectivity, as well as the physical dimensions of the column, a computer-modeling program was written. This program is called Computer Assisted Stationary Phase Design (CASPD). Its output is a set of simulated columns ranked according to quality from user-selected column variables and design control parameters. The program steps through design parameters and calculates new retention times. Each new simulated solution is ranked according to the number of peak overlaps and entered into an output sheet. As each simulated solution is entered, CASPD does a bubble sort so that the best-simulated separations rise to the top and low-ranking solutions are discarded. The final result is a ranked list of the 10 best solutions, or cybercolumns, with column variables and predicted retention times. The criterion for ranking is the number of coeluting peaks. Columns with the same ranking are further sorted by run time. The net result is that the cyber column at the top of the output sheet is the design with the fewest number of coelutions and the minimum total run time. The advantage of this form of data output is that, although the computer may have generated thousands of possible solutions, the designer has immediate and simple access to the best results.

Some column simulations focus on the separation of a set of related compounds or a particular separation. For example, if the design goal is a column for chlorinated hydrocarbons, the ability to separate xylene isomers may not be important. A feature of CASPD is that separation failures can be weighted differently, with the result that CASPD focuses on particular separations. The default is that all overlaps are weighted (and ranked) equally. This would mean, for example, that one solution that exhibits coelution of two xylene isomers has the same ranking as another solution in which two dichlorobenzene isomers coelute. The designer can favor certain solutions by decreasing the xylene-xylene weighting factor from the default value of 1. If, for example, the hydrocarbonhydrocarbon weightings were decreased to 0, the primary ranking of a solution would correspond simply to the number of dichlorobenzene-dichlorobenzene coelutions (0, 1, 2, 3). The run time criteria would then sort all those columns with the same primary ranking according to run time irrespective of the number of xylene-xylene overlaps. In contrast, if the xylene-xylene weightings were set at 0.01, the possible rankings of columns with a single overlap would be 0, 0.01, 0.02, 0.03, 1, 1.01, 1.02, 1.03, 2, 2.01, 2.02, 2.03, 3.00, 3.01, 3.02, 3.03. In this case, the best solutions would be those that could separate both xylenes and dichloro hydrocarbons; only in the case of equivalent xylene and dichlorobenzene rankings would run time be a consideration.

The adjustable column variables include column length, internal diameter, film thickness, and stationary-phase composi-

⁽⁸⁾ Forgacs, E.; Cserhati, T. Molecular Basis of Chromatographic Separation, CRC Press: Boca Raton, FL, 1997.

⁽⁹⁾ Giddings J. C. Unified Separation Science; John Wiley & Sons, Inc.: New York, 1991.

⁽¹⁰⁾ Dorman, F. L.; English, C. M.; Feeney, M. J.; Kowalski, J. American Laboratory 1999, 31 (6), 20–26.

⁽¹¹⁾ Dorman, F. L.; Schettler, P. D.; English, C. M.; Feeney, M. J. LCGC 2000, 18 (9), 928-934.

⁽¹²⁾ Laub, R. J.; Purnell, J. H. Anal. Chem. 1976, 48 (6), 799–803.

⁽¹³⁾ Maier, H. J.; Karpathy, O. C. J. Chromatogr. 1962, 8, 308.

tion. Adjustable run time variables include temperature program variables as well as inlet and outlet pressures. CASPD has three levels, or options, for adjustment. In the first option, useful for finding preexisting possible solutions, the user simply inputs a sequence of complete trial simulations. These might be retention data from off-the-shelf columns, for example. In the second option, which allows for physical parameter optimization, the user inputs values of each variable to be tried; CASPD generates and ranks cybercolumns with each possible combination. For example, if the user input 0.1- and 0.2- μ m film thickness and 0.25 and 0.53 mm for column i.d., CASPD would run all four possible combinations of film thickness and column i.d.. This can quickly evolve into a large number of combinations to be tested; adjusting eight variables over three possible values would result in generation of 38 (6561) possible designs. For stationary-phase composition optimization, a third option is possible in which the computer picks compositions at random, either within some range set by the designer or over the whole range of 0-100% volume fraction. The stationary phase can contain as many as eight components. Here again, CASPD will consider a large number of possible designs.

The program is completely interactive. As a first step, the user inputs retention data. If a two-component stationary-phase column is under development, then retention data is input for two columns of different stationary-phase composition, each under two different temperature programs. If a three-stationary-phase-component column is to be optimized, then data from three columns (each with different stationary-phase compositions and operated under two different temperature programs) must be input, and so forth. An option to actual experimental data is to utilize Kovats Retention Indices to estimate retention times for a specific column.^{14,15} In the second step, CASPD then converts this input data to the partition coefficient and its temperature dependence for each sample component with each stationary-phase component. The user then sets initial program variable options and ranges. CASPD then runs and ranks all combinations within these boundaries. A typical run can rank several thousand combinations within 24 h on a PC. All variables-column, run time, and computer-are on the spreadsheet and, thus, under user control. The user can then change any combination of these at any step during the design process. Upon rerunning, CASPD ranks the new combinations on top of the old. Thus, fresh designer ideas are automatically tested against previous results.

Most of the manipulations of the variables in CASPD are performed by Microsoft's Visual Basic for Applications (VBA) included with EXCEL. Since both viscosity and partition coefficients change with temperature, and the temperature changes during a run in a complicated way, retention time must be calculated numerically. For this reason, the chromatographic calculations are done in a Mathematica, a very flexible and powerful language for this sort of work. MathLink links the Mathematica I/O to EXCEL.

Viscosity for helium is a function of temperature, but not of pressure for the range of pressure associated with gas chromatography. Data from Hinshaw and Ettre¹⁶ was fit to the equation

$$\eta(T) = a' + \frac{b'T}{1 + c'T}$$

$$a' = 5.6159426 \times 10^{-6}$$

$$b' = 0.053286909 \times 10^{-6}$$

$$c' = 0.00026796519$$
(1)

Since the viscosity is not a function of pressure, the pressure profile through the column remains constant (although the flow rate changes) with changes in viscosity.

$$p(z) = p_0 \sqrt{b - cz}$$

$$b = P^2; \quad c = \left(\frac{P^2 - 1}{L}\right) \tag{2}$$

For this case, standard fluid mechanics results in a mean mobile phase velocity in the column as a function of temperature and position within the column as

$$v(z,T) = \frac{a(T)}{\sqrt{b - cz}} = \frac{p_0 a(T)}{p(z)}$$

$$a = \frac{r_0^2 p_0 b}{16\eta(T)}$$
(3)

Each analyte zone velocity is obtained by multiplying this by its retention factor K(T)

$$K = \frac{1}{1 + \beta \sum_{i} \phi_{i} K_{i}(T)} \tag{4}$$

 K_i 's are the partition coefficients for each analyte between each stationary-phase component, i, and the mobile phase. The value of K_i as a function of temperature is determined from

$$K_i(T) = \exp\left(-\frac{\Delta H_i^0 - T\Delta S_i^0}{RT}\right)$$
 (5)

where ΔH_i^0 and ΔS_i^0 are the standard enthalpy and entropy of solution of the analyte in a stationary phase consisting of pure component i.

The oven temperature (as a function of time) is determined by one of several temperature programs entered into CASPD by the designer as a series of linear temperature ramps. (A constant temperature is simply treated as a horizontal ramp for program purposes.) Since for each such temperature program, temperature becomes a function of time (only), v(x,T) and R(T) also become defacto functions of time, v(x,t) and R(t). The retention time is computed numerically by considering the elution to occur by a

⁽¹⁴⁾ Kovats, E. Helv. Chim. Acta 1958, 41, 1915.

⁽¹⁵⁾ McReynolds, W. O. Gas Chromatography Retention Data, Preston Technical Abstracts. Co.: Evanston, IL, 1966.

⁽¹⁶⁾ Hinshaw, X. X.; Ettre, X. X. J. High-Resolut. Chromatogr. 1997, 20, 471-

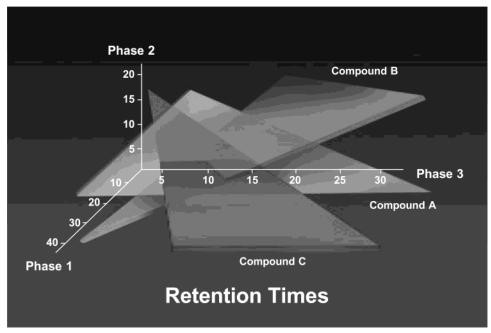


Figure 1. Selectivity surfaces for three compounds as a function of polymer composition. Planes A, B, and C represent the retention index for each compound as a function of changing the percentage of the polymer functionalities 1, 2, and 3.

dividing migration into a number of (nearly) isothermal steps, $\Delta t_i = t_{i+1} - t_i$. The corresponding Δx_i can then be calculated for each time increment using an analytical expression valid for constant temperature.

$$L = \sum_{i=0}^{N-1} \Delta x_i$$

$$\Delta x_i = \left(\frac{3}{2} acR(t) \Delta t_i + (b + cx_i)^{3/2}\right)^{2/3} - (b + cx_i)$$
 (6)

This expression for Δx is obtained by integrating

$$\Delta t_i = \int_{x_i}^{x_i + \Delta x_i} \frac{1}{R(t) v(x, t)} \, \mathrm{d}x \tag{7}$$

analytically and solving the resulting equation for Δx_i . Iteration starts with the analyte at the column entrance at $x_0 = 0$ and $t_0 =$ 0 and ends with analyte at the column exit, $x_n = L$ and $t_N = t_r$. A trial step size, Δt_i , is chosen initially to correspond to some minimum temperature change, ΔT . However, this trial step size may be decreased for two reasons. First, if the trial step size would mean spanning a time-temperature vertex of the temperature program, the step size is decreased so the step endpoint is the vertex. This ensures that the temperature change during any step is a linear function of time. The (isothermal) temperature chosen for each step is that at the segment midpoint, $t_i + \Delta t_i/2$. Cancellation of errors for such a choice means that ΔT can be as large as 1 K or even 5 K without significant error in the computation of Δx_i . The second reason for decreasing Δt_i is at the last step. In this case, Δt_i originally chosen is usually larger than the time remaining to exit the column. Since the last step is unknown a priori, the isothermal time for elution, Δt_e , is calculated at each x_i . Then Δt_i is decreased to Δt_e , and the iteration is terminated when $\Delta t_e \leq \Delta t_i$.

RESULTS

Input data used for this optimization consists of retention times for each compound analyzed under 2 different temperatures, on each stationary-phase functionality. This data is then used to calculate a retention index for each compound on each stationary-phase functionality. The resulting data can be plotted, and a typical resulting three-space graph is shown in Figure 1.

The equation for the selectivity surface is explicitly calculated so that the retention index of the compound can be determined for any polymer composition. It is straightforward to then perform an optimization that determines the polymer composition that yields maximum separation of all compounds from each other in the minimum total run time. This solution provides the best polymer suited for this specific separation. For some analytical methods in which the numbers of compounds are relatively small and the functionalities exhibiting good selectivity are commercially available, this technique is sufficient to produce an optimum column for the analysis. In other cases, however, physical parameters need to be accounted for, such as the column dimensions and the flow and temperature programs, in order to obtain a column with the desired resolution in a reasonable amount of time.

When this case arises, optimization of phase composition only will not yield an acceptable solution. When this occurs, a subroutine in CASPD begins to optimize physical parameters: oven program, flow rate, column internal diameter, film thickness, and column length. Typically, column length is the last variable used because it is the least important and most costly to the end user. Keeping in mind that the equation for the surface as shown in Figure 1 is retention index, not a retention time, it is possible to account for changes in the various physical parameters. Therefore, the optimization routine searches for maximum separa-

Table 1. Target Compound List for the Analytes in USEPA Method 8021

dibromomethane	o-xylene
bromodichloromethane	1,4-dichlorobutane
trichloroethene	1,2,3-trichloropropane
1,2-dichloropropane	isopropyl benzene
2-chloroethylvinyl ether	bromobenzene
cis-1,3-dichloropropene	<i>n</i> -proplyl benzene
trans-1,3-dichloropropene	2-chlorotolune
1,1,2-trichloroethane	4-chlorotoluene
toluene	1,3,5-trimethylbenzene
2-bromo-1-chloropropane	tert-butylbenzene
dibromochloromethane	1,2,4-trimethylbenzene
1,3-dichloropropane	<i>sec</i> -butylbenzene
1,2-dibromoethane	1,3-dicȟlorobenzene
tetrachloroethene	1,4-dichlorobenzene
1,1,1,2-tetrachloroethane	<i>p</i> -isopropyl toluene
chlorobenzene	1,2-dichlorobenzene
bromoform	<i>n</i> -butylbenzene
ethyl benzene	1,2-bromo-3-chloropropane
<i>m</i> -xylene	4-bromo-1-chlorobenzene
<i>p</i> -xylene	1,2,4trichlorobenzene
1-chloro-2-fluorobenzene	naphthalene
styrene	hexachlorobutadiene
1,1,2,2-tetrachloroethane	1,2,3-trichlorobenzene
	bromodichloromethane trichloroethene 1,2-dichloropropane 2-chloroethylvinyl ether cis1,3-dichloropropene trans-1,3-dichloropropene 1,1,2-trichloroethane toluene 2-bromo-1-chloropropane dibromochloromethane 1,3-dichloropropane 1,2-dibromoethane tetrachloroethene 1,1,1,2-tetrachloroethane chlorobenzene bromoform ethyl benzene m-xylene p-xylene 1-chloro-2-fluorobenzene styrene

Table 2. Predicted Retention Times (from CASPD Model) versus Actual Retention Times for 12 of the Compounds Listed in Table 1, as Analyzed on the Optimized Polymer

percentage of each phase column compositions ^a	9.10 100% dimethyl- polysiloxane	13.50 74% trifluoro propyl-methyl- /26% dimethyl- polysiloxane	37.4 8% pentafluro- benzylpropyl- methyl/92% dimethyl- polysiloxane	40.0 5% tridecafluoro- 1,1,2,2,-tetrahydro- octylmethyl/95% dimethyl- polysiloxane	100 10% trifluoro-propyl-methyl 3% pentafluroro-benzypropyl- methyl 2%trideca-fluoro- 1,1,2,2, tetra-hydrooctylmethyl 85% dimethyl-polysiloxane	
compounds	Rtx-1 b	$Rtx-200^b$	exptl $column^b$	exptl column b	predicted	actual bonded polymer
<i>n</i> -propyl benzene	20.46	18.20	20.67	19.17	19.97	20.25
2-chlorotoluene	20.42	18.66	20.80	19.79	20.07	20.35
4-chlorotoluene	20.52	19.04	20.95	19.94	20.25	20.53
1,3,5-trimethylbenzene	20.75	18.50	21.25	20.17	20.40	20.68
<i>tert</i> -butylbenzene	21.31	19.06	21.52	20.73	20.85	21.13
1,2,4-trimethylbenzene	21.32	19.15	21.80	20.70	20.96	21.24
sec-butylbenzene	21.69	19.16	21.81	21.06	21.14	21.41
1,3-dichlorobenzene	21.54	19.87	21.89	20.95	21.21	21.49
1,4-dichlorobenzene	21.64	20.00	21.97	21.00	21.28	21.55
<i>p</i> -isopropyl toluene	21.91	19.30	22.14	21.30	21.45	21.68
1,2-dichlorobenzene	22.13	20.73	22.51	21.51	21.83	22.12
<i>n</i> -butylbenzene	22.53	20.12	22.70	21.88	22.01	22.28

^a Column dimensions: 60 m \times 0.53 mm \times 3.0 df. Dead time, 1.8 min @ 35 °C. Constant head pressure. ^b Bold values are coeluting analytes.

tion in minimum total run time as a function of polymer composition, temperature and flow programs, and column dimensions.

As an example, several volatile compounds, listed in Table 1, were analyzed on a series of different functionalities in order to develope a capillary column for the analysis of compounds listed in the United States Environmental Protection Agency (USEPA) method 8021. A subset of these compounds, listed in Table 2 present a difficult separation challenge, having a 17 $^{\circ}\mathrm{C}$ boiling point difference among them. Additionally, all of these compounds are related in chemical structure, which places considerable selectivity demands on the liquid stationary phase in order to achieve adequate resolution.

These compounds were analyzed on columns employing the following functionalities: dimethyl-, methyl, phenyl-, trifluoropropyl, methyl-, pentafluorophenyl, methyl-, cyanopropyl, methyl, and tridecafluoro-1,1,2,2,-tetrahydrooctyl-methyl, methyl-polysiloxane.

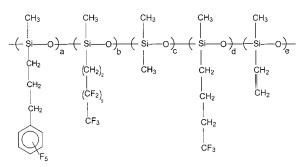


Figure 2. Structure of optimized polymer for the separation of USEPA Method 8021 compounds, as described in the text.

The retention data were loaded into the CASPD program, and the Table 2 shows the predicted retention times for a subset of these compounds seperated on the optimized polymer with the

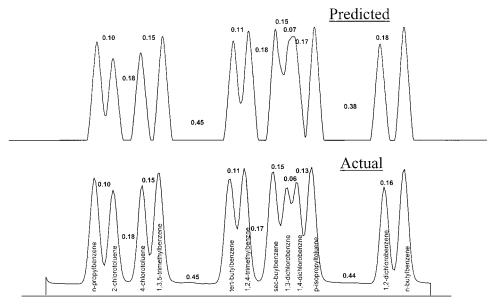


Figure 3. Portion of the chromatogram for USEPA Method 8021 for the 12 compounds listed in Table 2. Values represent the peak-to-peak separation in minutes for both the predicted and the actual separations.

struture shown in Figure 2 for this seperation. This result was reached using the third level of CASPD optimization, where optimization of the stationary-phase composition and the physical variables of separation was performed, as described in the experimental section.

Figure 3 demonstrates the agreement between the predicted separation and the actual separation for the polymer solution listed in Table 2. This polymer was determined to be the best at the separation for the compounds of interest on the basis of any combination of the stationary-phase functionalities, and possible column and GC operation parameters.

CONCLUSIONS

In summary, computer modeling can be effectively employed to design stationary phases that yield optimum resolution for a specific list of compounds. The techniques described in this paper

have successfully developed 8 new commercially available stationary phases and columns over the last two years that provide improved separation performance over previously available column products. This improved performance results from optimization of phase and column designs that yield optimum separation for a particular method. These innovative application-specific stationary phases will certainly improve analyses that were previously difficult by use of general-purpose phases. Now that this technology has been developed, it can also be employed to design applicationspecific phases and columns for a single analysis. Computer modeling will certainly be used for future gas chromatographic column development to achieve optimum separations.

Received for review October 1, 2001. Accepted February 16, 2002.

AC0110496