

ACD/ChromManager

Version 5.0 for Windows

Technical Note

Working with Applications Databases and Retention Time Prediction



Advanced Chemistry Development Inc.

Copyright © 2001 Advanced Chemistry Development Inc. All rights reserved.

ACD/Labs is a trademark of Advanced Chemistry Development Inc.

Microsoft is a registered trademark and Windows is a trademark of Microsoft Corporation.

IBM is a registered trademark of International Business Machines Corporation.

All the other trademarks mentioned within this Manual are the property of their respective owners.

All trademarks are acknowledged.

Information in this document is subject to change without notice and is provided "as is" with no warranty. Advanced Chemistry Development Inc. makes no warranty of any kind with regard to this material, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. Advanced Chemistry Development Inc. shall not be liable for errors contained herein or for any direct, indirect, special, incidental or consequential damages in connection with the use of this material.

Working with Applications Databases and Retention Time Prediction

Mike McBrien

*Advanced Chemistry Development,
90 Adelaide Street West, Suite 702, Toronto, Ontario, M5H 3V9*

Introduction

Most method development chromatographers are familiar with the concept of the applications database. Separation conditions for samples are archived, with varying levels of searchability such that chromatographers can retrieve and reproduce these conditions for both identical and new samples. Programs like ACD/ChromManager that offer substructure searchability give this old tool new life. Novel compounds can be separated using procedures that were developed for similar functionality.

With the advent of accurate prediction of physicochemical properties has come retention time prediction in chromatography. The ability of ACD/LC Simulator to predict LogP, pKa, and LogD for compounds offers the method development chemist a tremendous tool. Once likely separations have been retrieved, they can be evaluated and modified based on chemical structures *prior to the first injection*.

In this technical note, we will retrieve an application that seems to contain a reasonable method for our compounds. We will then predict the retention times for our compounds. Based on the predicted retention times, we will make some modifications to the method.

The Tools Involved

ACD/ChromManager Databases are fully searchable databases composed of chromatograms in the .esp format. By comprising our Applications Database of .esp files, we give the searcher access to the original chromatogram with assigned structures, full data collection parameters, column information, and other custom-designed data. This information may take the form of the sample number originally used, or method-specific data that might include the storage or cleanup specifications of the column. All of this information is fully searchable.

ACD/LC Simulator is software designed to predict chromatographic retention times. In this technical note, we will utilize its ability to predict retention times based on structure and the pH-dependent hydrophobicity (LogD).

Opening the Database

In Method Development Technical Note #1, we saw how to search on multiple databases at the same time. For this note, our work will be exclusively based on the ACD/Chromatography Applications Database. Start up ChromManager and move to the database window by clicking on the bottom toolbar. From the database window, choose **File -> Open**, and browse to the file called App_DB.ndb.

This database is a collection of applications consisting of chromatogram with assigned structures and complete separation method information. If we can find a compound in the database that is similar to our compound, this should be a method that is applicable to our compound.

Retrieving a Method

There are many ways in which to search the Applications Databases, but the most likely search technique will be by substructure. Using this search capability, we can find compounds in the database that are similar to the compound that we are working with today.

For example, let's say that we have a sample that contains some impurities as well as our active ingredient, verazide, Figure 1.

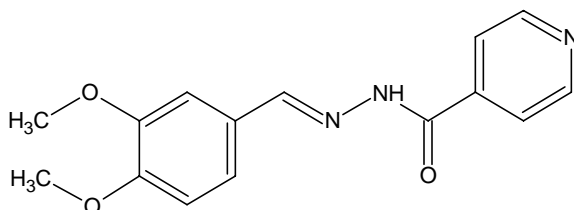



Figure 1. The structure of verazide.

Retrieve verazide from the **Dictionary** . The first thing that we can do is to perform a structure search and see if the compound is in the database. Click and choose search structure in the dialogue box. We can see that there is no hit. From here, we can try to find structures that share functionality.

We can use any part of the structure to do a substructure search. While it is tempting to choose large portions of the structure, depending on the type of compound and the size of the Applications Database, it is often better to simply choose simple functional groups, particularly ones that are common between each component. Let's choose one portion of the molecule by deleting all other atoms (you may want to copy the structure and paste a second version in the window first). Select portions of the structure and delete everything except the fragment shown in Figure 2.

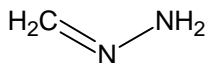



Figure 2. Fragment of structure to use for substructure search.

Now click Substructure Search on the lower toolbar . This search results in 16 hits. Since there are many methods here, we will refine the search further. ChromManager is

designed to do iterative searches. Until we click **Retrieve All** , any subsequent searches will be on the subset that is the result of the search before. Return to ChemSketch via the bottom toolbar, and choose a different part of the structure to target. Let's choose the functionality shown in Figure 3.

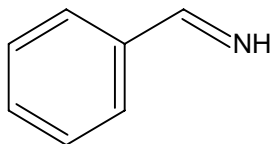


Figure 3. Fragment to be used for secondary search.


The second substructure search leads to seven hits; seven chromatograms have structures that satisfy both criteria. From here, we can choose an application by examination. We can scroll through the chromatograms using the browse toolbar at the top of the window




Select the chromatogram with ID number 35. Double click the chromatogram, carrying it to the chromatogram window.


Generating a Prediction Equation

In the chromatogram window, it is possible to view the structural assignments by moving the

pointer over each peak. In addition, clicking the **HPLC Parameters** button  shows the chromatographic method that produced the chromatogram. Using the structures and retention times from the original method, we can generate an equation that will predict the retention times for any compound using that method.

Click on **Optimize Chromatogram** . This will carry the structures and retention times to LC Simulator.

The Data Input Window should contain all of the pertinent information about the separation. In this case, there are five structures and retention times. Five structures are enough to generate a prediction equation. If there were many more than this, we could browse through and deselect compounds that are not similar to our molecules. This will increase the speed of the calculation and it will help with accuracy. The more similar the training set to the test set, the more accurate we can expect the predictions to be.

Click on  Calculate Predicting Equation. LC Simulator will calculate the LogD and other values for each compound in the training set, and then model the prediction equation on these results. We can now predict retention times for our new compounds.

Predicting Retention Times for Compounds

From the ChemSketch window, draw or import any structures that are involved in the separation. In our case, we will draw the following compounds corresponding to our compound and two expected impurities, Figure 4.

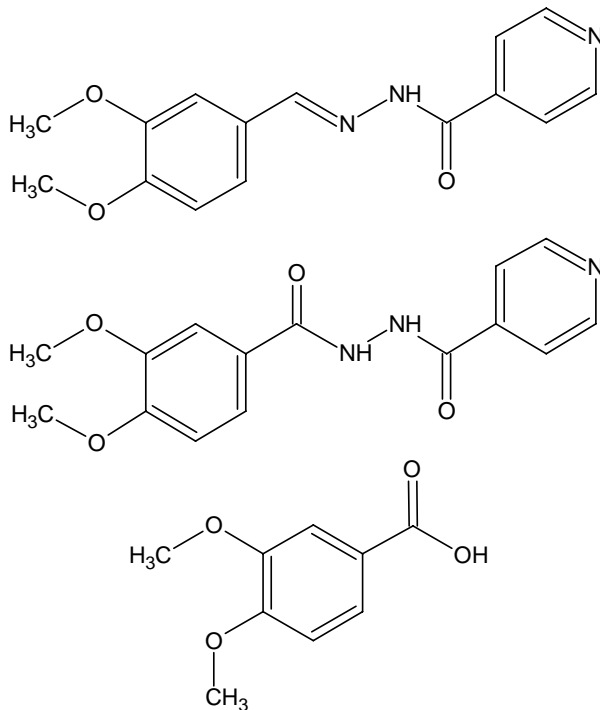


Figure 4. Verazide and expected impurities.

We will transfer our three compounds to LC Simulator one by one. Select the first compound and click **Optimizations**. This will move you to the optimization window and show the predicted retention time, Figure 5. Move back to the ChemSketch window and transfer the impurities in the same way. Note that the resulting chromatogram is composed of both our compounds and the training set. We can turn off the view of the training set by simply deselecting the compounds in the Data Input Window

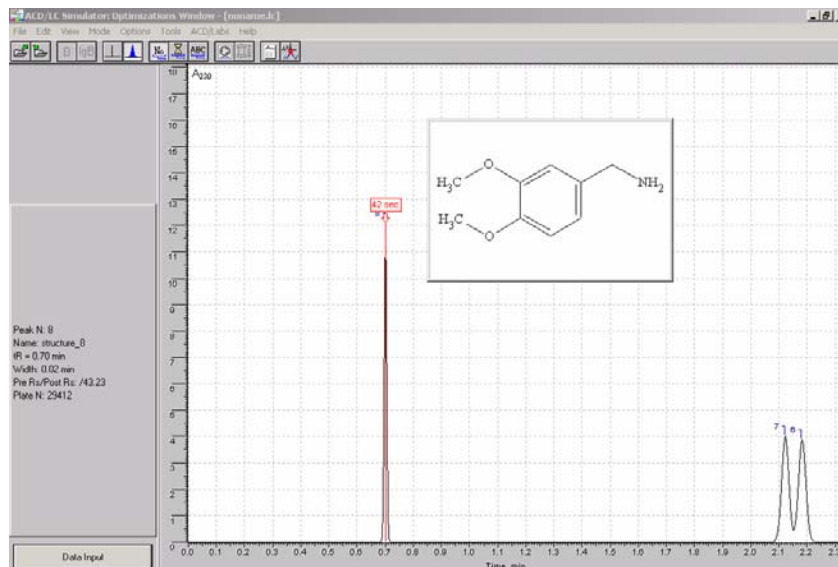


Figure 5. Optimized chromatogram of compounds input from Figure 4.

It is obvious that this is not an ideal chromatogram. Based on the prediction, all of the compounds would elute very early. Ideally, we would like to see the compounds eluting reasonably far apart, with intermediate elution times.

Modifying the Method Prior to Experimentation

One of the greatest benefits of Prediction Mode is that we do not have to attempt an injection in order to assess the value of a method. Upon predicting the retention times of the compounds, we can make some changes immediately. In this case, our compounds are eluting very early. We would be very lucky if we were to get a useful separation from this method, and we can expect that our first impurity will elute close to the solvent front, making it impossible to quantify. However, the column, mobile phase components, etc., may still be good choices. Based on the prediction, we can decide to emulate the method, but simply choose a slightly weaker solvent. Alternatively, we could return to the Applications Database and try to find a method better suited to our system.

If we return to the Applications Database, we can see that we already have a method that corresponds to the modification that we might make. Browsing to entry ID 40, we can see that the method is almost exactly the same, but the solvent ratio is 60:20:20 instead of 60:10:30 water:MeOH:MeCN. The lower MeCN component should mean that our compounds will elute later. Indeed, if we carry the compound to LC Simulator and predict the retention time, it is a more reasonable 4.5 minutes, Figure 6.

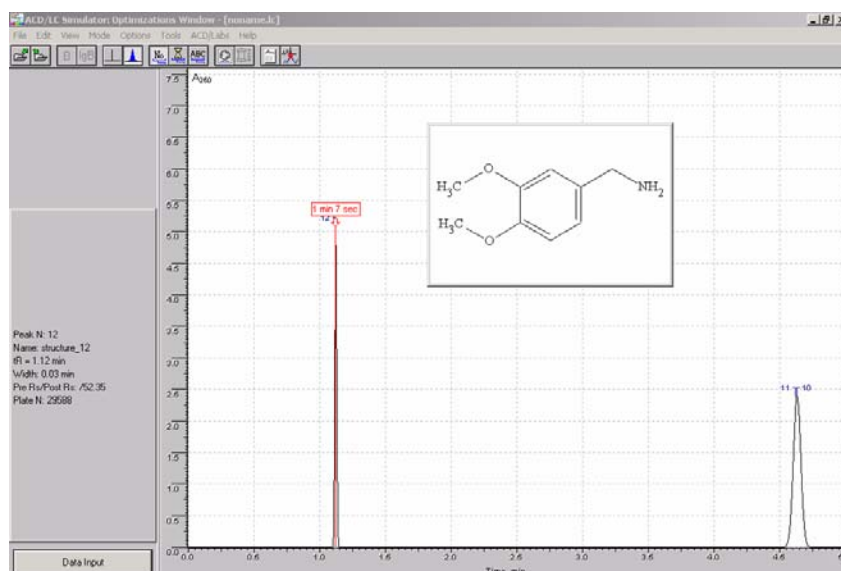


Figure 6. Optimized chromatogram under new conditions.

From at least one perspective, this seems like a better method. Depending on our needs, we might even choose less organic component in the solvent. We should still consider the co-elution of peaks 10 and 11, however.

Predicted Resolution and HPLC Mode

In the above example, neither predicted chromatogram had complete resolution between verazide and one impurity. This should not generally be a reason for rejection of a method, but there are some steps that might be taken when this is the case.

In practice, prediction of retention times depends on a number of different factors, and errors may easily be in the 5% range. This means that predicted co-elution is not necessarily a reason to reject a method outright. If the unresolved components elute very early in the chromatogram, then this could turn out to be a problem. In general, the best solution to this is to use a weaker solvent and see if they are resolved in that way.

There is another avenue, however. Note that the predicted retention times are for a particular method at a specified pH. We have the option to examine the behavior of the compounds as a function of pH using HPLC Mode.

HPLC Mode models the retention behavior of compounds based simply on hydrophobicities, or LogD. It predicts only retention order for the compounds, but it will predict retention order at any pH.

Click **HPLC** and the LogD values for the compounds at that pH will be displayed, Figure 7. Note that the pH for the values is shown in the lower left corner. You can view their LogD values at any pH you wish, by inputting it there.

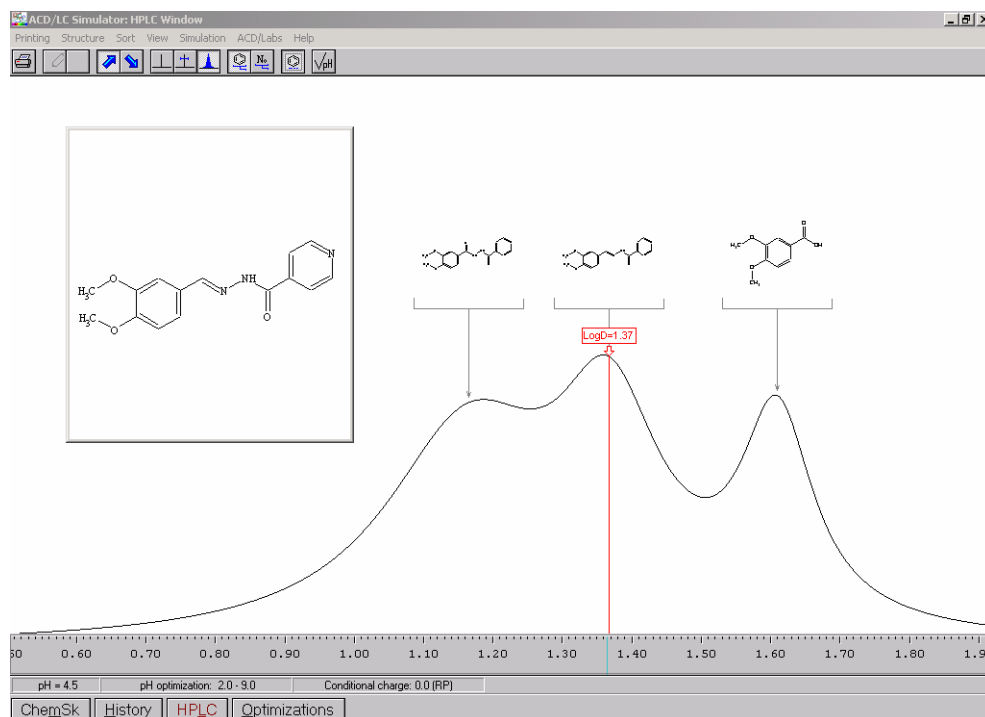



Figure 7. Predicted HPLC for sample compounds.

The LogD values for all three of the compounds are quite similar at this pH (4.5). However, if we can find a pH for which they have different LogD values, then their retention times should move apart as well. Some experimentation should lead you to the conclusion that making the pH more

basic should pull these two components apart. You can also attempt to perform this step automatically, using the **Optimize pH** button, . Keep in mind the parameters of your column; you can input the range at which the column is stable in the box at the bottom of the screen, Figure 8.

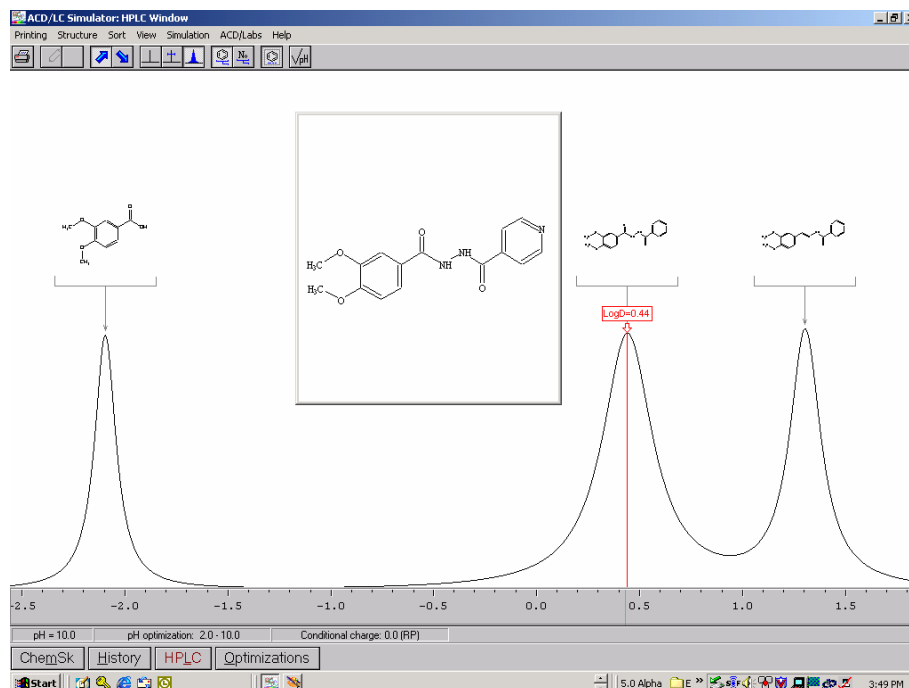


Figure 8. pH optimized predicted HPLC.

You may wish to use HPLC Mode prior to injecting this sample, but often it is better to use this mode after you have experimental retention times to consider. It is a good idea to keep in mind that predicted retention times will not be exactly correct, and often only a small difference will turn borderline resolution into a sufficient method. To see how to further increase the accuracy of both Prediction and HPLC Modes, see Method Development Technical Note #3, *Increasing Accuracy of Predictions By User Training with Experimental Values*.

Conclusion

Prediction Mode is designed to predict retention times for compounds based on a given method, using LogD prediction and a training set of compounds. Using this technique, the chromatographer should be able to evaluate and modify methods very quickly, drastically reducing the number of experiments necessary to obtain a suitable separation method. While ACD/Labs provides methods for your use, it should be noted that most scientists work with similar classes of compounds in their day-to-day work. It is advisable to consider archiving one or two separations for each major class of compounds that you study. Even years later, substructure search can retrieve these methods, and Prediction Mode will enable you to adapt them to new compounds.