Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1086 (2005) 185–192
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Evaluation of the DotMap algorithm for locating analytes of interest based on mass spectral similarity in data collected using comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry

Janiece L. Hope, Amanda E. Sinha, Bryan J. Prazen, Robert E. Synovec *

Center for Process Analytical Chemistry, Department of Chemistry, Box 351700, University of Washington, Seattle, WA 98195-1700, USA

Available online 27 June 2005

Abstract

Comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry (GC × GC–TOF-MS) is a highly selective technique ideal for the analysis of complex mixtures. The instrument yields an abundance of data, with complete mass spectral scans at every time point in the GC × GC separation space. The development and application of appropriate tools for data mining is essential in making sense of the wealth of information available. An algorithm for locating analytes of interest based on mass spectral similarity in GC × GC–TOF-MS data, called DotMap, has been previously reported and is rigorously evaluated herein. A thorough investigation into the performance characteristics of DotMap, including the performance near the limit of detection and dynamic range of the algorithm as well as the capacity of the algorithm to deal with peak overlap, is investigated using jet fuel as a complex sample matrix. For instance, the algorithm can successfully identify a spiked compound at the single μg/ml level in a jet fuel sample with an overlapping interferent. The performance of the DotMap algorithm in situations with very limited mass spectral selectivity, specifically in the evaluation of spectra from isomer compounds, as well as the ability to tune DotMap results to provide the location of a specific analyte or of a class of compounds is demonstrated. The DotMap algorithm is demonstrated to be a sensitive tool that is useful in the analysis of complex mixtures and which possesses the capacity to be easily "tuned" to discern the location of analytes of interest.

Keywords: GC × GC–TOF-MS; Chemometrics; Data mining; Selectivity; Data analysis

1. Introduction

Two-dimensional comprehensive gas chromatography coupled with time-of-flight mass spectrometry (GC × GC–TOF-MS) has been demonstrated as a highly selective instrument, perfectly suited for complex mixture analysis [1–15]. Recent complex sample analyses in the literature include fuels [1], pesticide residues in food [3,4], cigarette smoke [6,7], airborne particulate matter [11], metabolite extracts of plants [14], etc. This instrument creates an immense amount of data, which requires the development of tools to assist the user in mining the data to give information pertinent in answering a question. The algorithm, called DotMap, studied herein is a tool designed to locate analytes or classes of analytes that are of interest.

Often in complex mixture analysis, there is a need to locate specific analytes of interest [1,4,16–20], and in many complex mixtures, matrix effects may make peak locations inconsistent from one mixture separation to the next [21–24]. With current software approaches, peak locations are found by generating a list of all of the peaks present in the data set, and matching each peak to a library of spectra. This process can be excessively time consuming, even when automated, and still requires the user to scroll through lists of perhaps hundreds of peaks to find the compound(s) of interest. This approach is very inefficient in situations where the analyst may only be interested in a small subset of the compounds present.

* Corresponding author. Tel.: +1 206 685 2328; fax: +1 206 685 8665.
E-mail address: synovec@chem.washington.edu (R.E. Synovec).

0021-9673/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2005.06.026

Sometimes to, it is important to locate an entire class of compounds in a complex mixture [25], and in some cases, to more fully describe the chemical characteristics of a sample it is necessary to quantify the percent composition of different chemical classes. In cases such as these the DotMap algorithm can also be of assistance. These reasons demonstrate a need for the development of an algorithm such as DotMap, which was recently reported and demonstrated on an organic acid metabolite sample of human infant urine [26].

The DotMap algorithm searches an entire GC × GC–TOF-MS data set for a target analyte of interest. In some applications only a relevant portion of the entire data set needs to be searched. A correlation value is calculated, via a dot product calculation, at every point in the GC × GC separation space between the spectrum at that point and the target analyte spectrum. A contour plot is then generated by the algorithm showing only the location(s) of the highly correlated mass spectra. The mass spectrum at the location of maximum similarity is extracted from the GC × GC–TOF-MS data set and searched against a mass spectral library to confirm the identification.

In the current study a thorough investigation into the performance characteristics of the DotMap algorithm for a jet fuel sample is reported. *O*,*O*,*O*-Triethylphosphorothioate (TEPT) was spiked into the jet fuel at varying concentrations to determine the effective limit of detection (LOD) and dynamic range for this algorithm. Two isomers, 1-methylnaphthalene and 2-methylnaphthalene, were analyzed with DotMap to determine the algorithm performance with very little mass spectral selectivity. The ability to tune the implementation of DotMap to provide single analyte information and class information was also investigated using the alkane *n*-dodecane and the alkane compound class in the jet fuel sample.

2. Theory

The DotMap algorithm has been discussed in detail in a previous publication [26], and will be described only briefly here to provide clarity. DotMap is an algorithm that searches an entire GC × GC–TOF-MS data set to find those spectra that best match the target spectrum of interest. The DotMap algorithm calculates the dot product of the baseline corrected, scaled, weighted, and normalized target analyte mass spectrum with the mass spectral signal in the data set for each point in the GC × GC plane as follows:

$$\left(\frac{m\sqrt{A_\mathrm{d}}}{\sum m\sqrt{A_\mathrm{d}}}\right)\cdot\left(\frac{m\sqrt{A_\mathrm{u}}}{\sum m\sqrt{A_\mathrm{u}}}\right) \quad (1)$$

where A_u is the abundance or ion counts of *m/z* signals in the target spectrum of interest, A_d are the abundances of *m/z* signals at each point in the two-dimensional GC × GC "data" space, and *m* is the vector containing values used for weighting the signals such that higher *m/z* signals contribute more to the classification. The symbol "·" represents the dot product. The mass spectral signal intensities in the data set and spectrum of interest are also scaled to the one-half power to enhance the small intensity signals relative to the high intensity signals.

The result of the algorithm is displayed as a contour plot indicating the location of signals above a threshold. The empirically determined and applied threshold used in all DotMap analyses reported herein is defined as 90% of the maximum dot product value above the median value of all dot products in the raw DotMap data.

$$\text{Threshold} = (0.9 \times (\text{maximum} - \text{median})) + \text{median} \quad (2)$$

The algorithm then extracts the spectrum from the GC × GC–TOF-MS data that corresponds to the location of the maximum dot product value, presumably the peak maximum of the analyte peak, and performs a traditional mass spectral similarity search, to confirm the identification.

3. Experimental

Two standard solutions of compounds commonly found in fuels were analyzed to obtain analyte spectra, Proposed O_PONA System Validation Mixture (AccuStandard, New Haven, CT, USA) and Gasoline Refinery Aromatics Standard (AccuStandard). A sample of jet fuel diluted 10:1 with acetone was analyzed to generate a searchable data matrix for the DotMap Algorithm. This sample was spiked with *O*,*O*,*O*-triethylphosphorothioate, ranging in concentration from 1.6 to 324 μg/ml. The analysis of the fuel sample was performed using a LECO Pegasus 4D GC × GC–TOF-MS instrument (LECO, St. Joseph, MI, USA). The first column was a 40 m × 250 μm I.D. capillary column with a 0.5 μm 5% diphenyl/95% dimethyl polysiloxane (DB-5; J & W Scientific, Alltech, Deerfield, IL, USA). The second column was a 2.5 m × 180 μm I.D. capillary column with a 0.2 μm film of 100% polyethylene glycol (AT-Wax, Alltech). These columns were joined using a Vu2 union (Restek, Bellefonte, PA, USA). Thermal modulation was used to transfer first column effluent onto the second column. Cold nitrogen gas was used to concentrate the column 1 effluent at the head of column 2. At the start of each second column separation the cold nitrogen was shut off and heated air jets (40 °C above the oven temperature) were switched on for a length of time equal to 0.4 s. For the remainder of the column 2 separation, the cold nitrogen gas was applied at the modulator. The total column 2 modulation time was 2 s. Half a microliter of the jet fuel sample was injected using a 50:1 split. The column 1 initial oven temperature was 40 °C with a hold time of 0.5 min then increased at 5 °C/min to 220 °C. Column 2 was housed in a separate oven and held at a constant 5 °C higher than the column 1 oven temperature. The carrier gas was ultra high purity helium and operated at a constant flow rate of

1.5 ml/min, corresponding to a pressure program starting at 1.39×10^5 Pa above ambient room pressure and extending to 2.51×10^5 Pa at the end of the temperature program. No mass spectra were collected during the solvent delay for the first 4 min of each run. The transfer line was maintained at 260 °C and the ion source set point was 200 °C. The detector voltage was −1600 V and the filament bias was −70 V. Mass spectra were collected at 100 spectra/s in the range from *m/z* 41–300 unless otherwise noted. Data were then exported as a comma separated value (.csv) file and loaded into Matlab 6.0 R12 (The Mathworks, Natick, MA, USA) for data processing. Mass Spectral similarity searches were performed using NIST MS Search 2.0 (NIST/EPA/NIH Mass Spectral Library; NIST 02).

4. Results and discussion

The DotMap algorithm studied herein is designed to take a user chosen target spectrum and search a given GC × GC–TOF-MS data set for mass spectra that are similar to the target spectrum, highlight those regions with similar spectra in a GC × GC contour plot, extract the mass spectrum from the raw data in the highlighted region and search that extracted analyte spectrum against a library of spectra using the NIST MS Search program. This complete process is automated and requires approximately 30 s of analysis time. In the first report on this algorithm [26], complex metabolite extracts were explored as a challenging application of the DotMap algorithm. In this report, the capabilities and limitations of the algorithm are studied in greater detail in a different, yet very complex matrix, jet fuel.

The jet fuel sample used for this study is shown in Fig. 1A as a total ion current (TIC) GC × GC chromatogram. To study the ability of the DotMap algorithm to cope with decreasing S/N, a test analyte, TEPT, was spiked into aliquots of the jet fuel sample at a range of concentrations. The concentrations of TEPT in 10 spiked jet fuel samples ranged from 1.6 to 324 μg/ml. These spiked jet fuel samples, in addition to the pure component spectrum of TEPT obtained in-house, were used for the DotMap analysis. The sub-region of the entire GC × GC separation space containing TEPT is shown in Fig. 1B. The specific sample shown in Fig. 1B is that of the lowest concentration spike evaluated in this study (1.6 μg/ml). Despite the small size of the analyte peak in Fig. 1B, it can be seen that there are multiple modulations across the first column elution profile for the TEPT analyte in this system. The minimum number of modulations (i.e., modulation number) needed to produce "comprehensive" two-dimensional chromatographic separations is considered to be three to four modulations per peak [27,28]. It can be seen that even the lowest concentration spike of TEPT qualifies as comprehensive.

The resultant matrix from the DotMap analysis of TEPT spiked at 324 μg/ml in jet fuel (i.e., highest concentration of range), projected on to the column 1 dimension, is shown in Fig. 2A. The threshold, shown as a dotted line, was automatically calculated by DotMap as 90% of the difference between the maximum point and the median of the DotMap raw data, as defined in Eq. (2) of Section 2. An analogous figure showing the raw data in the column 2 dimension was omitted for brevity. Only one signal (i.e., analyte peak) is above the threshold in this raw DotMap data indicating that the similarity between the analyte spectrum and the spectra corresponding to this location in the raw data is high. Further, the presence of only one signal location indicates that the DotMap algorithm shows a high degree of selectivity for the analyte (TEPT) in this matrix.

DotMap analysis was performed on additional data sets with decreasing concentrations of TEPT. At concentrations below 8 μg/ml (signal (S)/noise (N) ∼100), the DotMap analysis was not able to pinpoint the location of TEPT while searching against the full data set. The signal to noise was calculated using *m/z* 65, the most abundant ion in the library spectrum for TEPT, where signal is defined as the maximum peak height and noise is defined as three times the standard deviation of the noise in a representative section of baseline signal. Below 8 μg/ml TEPT and using the full data set, the location of 1-methylnaphthalene is identified instead due to the very high S/N and degree of mass spectral similarity of this compound compared to that of TEPT, the analyte of interest. If prior knowledge of the expected retention time of an analyte exists, this information can be used to reduce the chromatographic time window within the full data matrix that is searched. Thus, only a portion of the full data set would be searched using DotMap.

As an example, the TEPT-containing region of the jet fuel chromatogram spiked with 1.6 μg/ml TEPT was used as the data matrix for a DotMap analysis. This is the region that was shown in Fig. 1B, and it can be seen that there is an interfering analyte close to the TEPT peak. The S/N for TEPT in this sample is 33, calculated using *m/z* 65 as described above. The raw data from this DotMap analysis is shown in Fig. 2B as the column 1 projection of the raw DotMap results. This image is of the same form as was Fig. 2A. It can be seen that using DotMap in a selected region of the data set, wherein the analyte is expected to elute, yields one peak that matches the analyte of interest in this region. The task of searching for a target analyte mass spectrum in a selected region of the chromatographic separation space has the added advantage of speeding up the already rapid DotMap analysis process, with total analysis times for reduced data sets being on the order of 1 s.

The DotMap algorithm outputs the results of the DotMap search in a user-friendly contour plot of the separation space analyzed. A contour plot of the DotMap raw data for the analysis of the 324 μg/ml TEPT spiked sample is shown in Fig. 3A with a single contour line drawn at the threshold indicated in Fig. 2A. This view is referred to as the "DotMap" for the analysis. The result of DotMap analysis on the selected region used to study the low concentration spike (1.6 μg/ml) was successful as indicated in Fig. 2B, and the generated

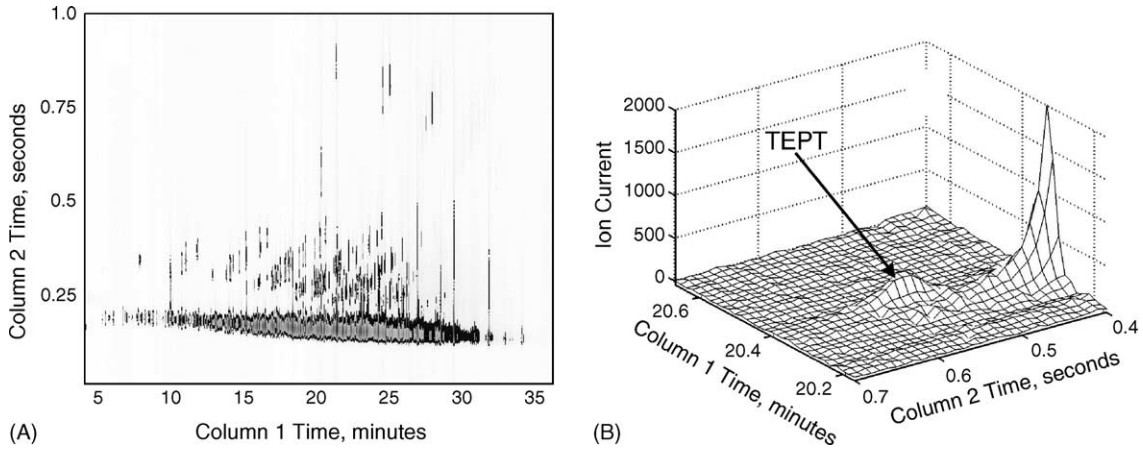

Fig. 1. GC × GC–TOF-MS analysis of jet fuel. (A) A selected region of the GC × GC separation space containing the majority of the analyte peaks is shown as a TIC plot. (B) The submatrix (*m/z* 65) of the full separation space that was used for the DotMap analysis of *O,O,O*-triethylphosphorothioate (TEPT) spiked into jet fuel (1.6 μg/ml). The location of TEPT is indicated and there is evidence of overlap with an interfering component in the jet fuel sample that shares the same selected ion.

DotMap for this analysis was essentially a zoomed-in version of that shown in Fig. 3A (not shown for brevity). From the contour plot in Fig. 3A, the retention indices on both column 1 and column 2 were determined. At those coordinates a spectrum was extracted from the original baseline corrected GC × GC–TOF-MS data matrix and sent to the NIST MS Search program where a similarity search was conducted. The NIST MS Search program returns a match factor on a scale of 0–999 describing the quality of the match between the spectrum being searched and the spectra of the library hit to which it is matched. A match factor of 999 indicates a perfect match or total spectral similarity. A reverse match factor is also calculated in a similar fashion, but in the case of the reverse match, peaks that are present in the sample spectrum but not present in the library best match spectrum are not penalized when computing the reverse match factor [29]. For each DotMap analysis performed in this study, the extracted mass spectrum from the region of the 2D separation space selected by DotMap was searched against the user built library containing the mass spectra for TEPT. For the highest concentration TEPT spiked sample (324 μg/ml) the reverse match factor (818) was much higher than the similarity match factor (716). This difference in the reverse and similarity match factors can be used as an indicator of peak overlap. A similarity and reverse match that are nearly identical indicates a pure peak and a discrepancy as described above for similarity and reverse match factors can indicate peak overlap due to the effect of interfering, overlapping peaks. In the case of the lowest concentration TEPT spiked sample (1.6 μg/ml), the extracted spectrum did match to the library spectrum for TEPT (Fig. 3B), but the match factors were reduced relative to the match factors for the higher concentration TEPT samples (similarity: 523 and reverse: 554). Reasons for the reduced match factor values include the effect of the interfer-

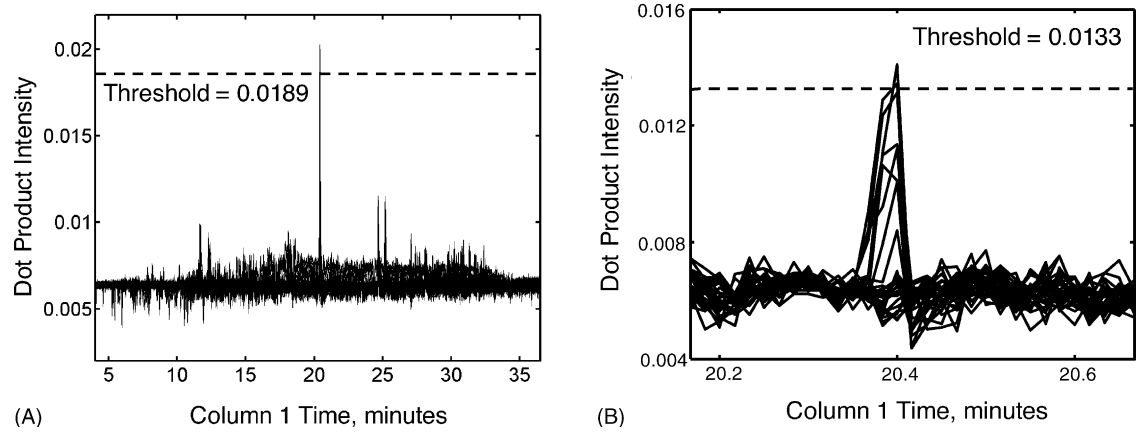

Fig. 2. DotMap analysis of TEPT spiked into jet fuel: column 1 projections of DotMap results. (A) Raw DotMap results projected onto column 1 for TEPT at a spiked concentration of 324 μg/ml. The maximum peak indicates the location of TEPT on column 1. The threshold shown is 90% of the maximum dot product value above the median value of all dot products in the raw DotMap data (Eq. (2)). The location on column 2 is found in an analogous fashion, not shown for brevity. (B) DotMap result on the submatrix of the TEPT spiked jet fuel at a concentration of 1.6 μg/ml is shown. The raw data from the DotMap analysis projected onto column 1 and the peak that crosses the threshold indicates the column 1 location of the analyte.

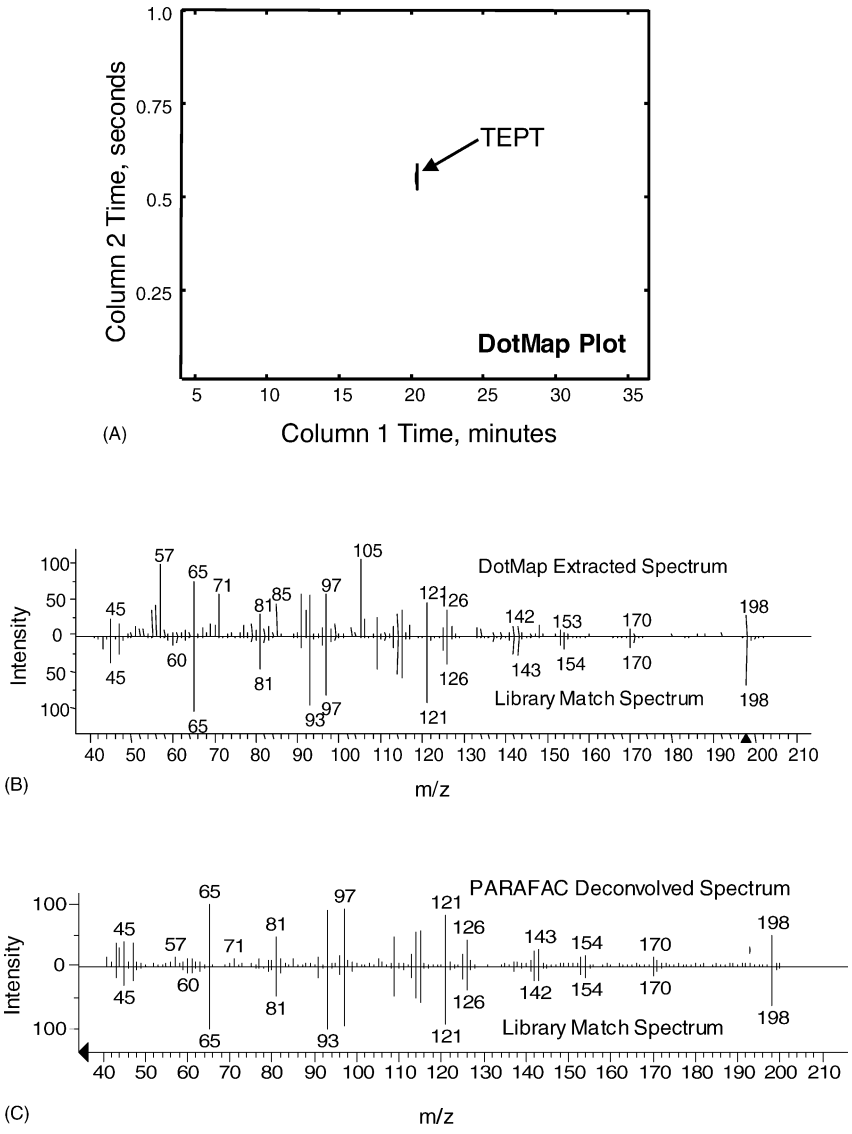

Fig. 3. DotMap analysis of TEPT spiked into jet fuel. (A) The contour plot output, or "DotMap," from the algorithm with the location of the analyte indicated. This is the DotMap for the high concentration TEPT spike (324 μg/ml). (B) The extracted mass spectrum from the DotMap analysis of TEPT (1.6 μg/ml) in the submatrix shown in Fig. 1B is displayed. The spectrum was extracted from the original GC × GC–TOF-MS data at the maximum point in the DotMap analysis. The spectrum was matched to TEPT using the NIST MS Search program and the NIST02 library. The library spectrum is also shown in the lower portion of the figure. (C) The PARAFAC deconvoluted mass spectrum for TEPT (1.6 μg/ml). The spectrum was also matched to TEPT using the NIST MS Search program and the NIST02 library.

ing analyte (Fig. 1B) as well as the effect of the lower S/N of TEPT at a concentration of 1.6 μg/ml. The presence of peaks at *m/z* 57, 91 and 105 in the extracted spectrum but not in the library spectrum (Fig. 3B) provides further evidence that TEPT is overlapped with an interferent. Indeed, it is noteworthy that even though the match factor for TEPT was quite low, the DotMap algorithm was able to locate TEPT in the presence of the interferents.

While DotMap can be used to find an analyte that is highly overlapped with interferent signals, Parallel Factor Analysis (PARAFAC) can be subsequently used to deconvolute the chromatographic and mass spectral profiles from each analyte and interferent in that region thereby refining the mass spectrum of the analyte of interest, in this case TEPT. PARAFAC provides a deconvoluted mass spectrum, generally excluding mass peaks from the overlapping interferent compound(s) and which therefore provides better match values for TEPT at 1.6 μg/ml in the sample (similarity: 739 and reverse: 859), as demonstrated in Fig. 3C [26].

Since the DotMap algorithm is strongly influenced by mass spectral similarity, it is important to investigate the results of DotMap analyses of isomers, which generally have very similar spectra. In this case, the spectra for 1-methyl and 2-methylnaphthalene standards were obtained in-house to reduce spectral variations due to differences in instrumentation. Spectra in commercial libraries are produced from

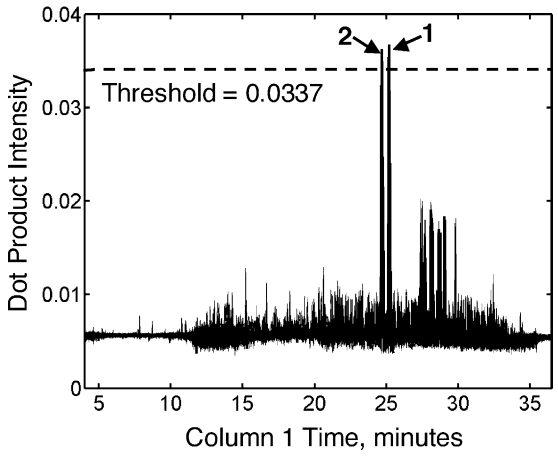

Fig. 4. DotMap analysis of 1-methylnaphthalene present in jet fuel. The raw DotMap results projected onto column 1 are shown. The results give the locations of both the isomers 1-methylnaphthalene and 2-methylnaphthalene, shown as "1" and "2" respectively. The algorithm notes both peaks due to the mass spectral similarity of the two compounds' spectra.

many instruments with instrumental parameters that inflict differences in the spectra and could potentially reduce the accuracy of DotMap. The initial jet fuel sample, diluted in acetone and analyzed by GC × GC–TOF-MS was used as the data set. The raw DotMap data projected onto column 1 using 1-methylnaphthalene as the search analyte is shown in Fig. 4. There are clearly 2 peaks that have virtually identical dot product intensities. While the dot product intensity is slightly greater for the true 1-methylnaphthalene, the algorithm also identifies 2-methylnaphthalene as a potential match. Similarly, when the data matrix is searched against 2-methylnaphthalene both peaks are found with the peak for 2-methylnaphthalene being slightly larger (not shown for brevity). This indicates that the mass spectral similarity between these spectra is too high for DotMap to distinguish the individual isomers and therefore, that the contour plot generated indicates the location of both isomers, not shown for brevity. This does, however, bring up the potential ability of DotMap to be "tuned" to give the locations of individual analytes or classes of compounds that share spectral features.

This concept of tuning DotMap to find individual analytes of a class of compounds was explored using the spectrum, collected in the laboratory, of a dodecane standard and the jet fuel data set studied previously. The raw DotMap results projected onto column 1 from the DotMap generated using the entire dodecane spectrum is shown in Fig. 5A. While there is clearly one peak that falls above the threshold, the extracted spectrum was found to be that of octane rather than dodecane. If Figs. 5A and 2A are compared, it can be seen that there are many more compounds which are similar to the search compound (dodecane) in Fig. 5A than to the search compound (TEPT) in Fig. 2A.

To optimize the DotMap analysis for the search analyte, dodecane, several approaches were investigated including a method using a variety of selected ions (not all approaches shown for brevity), and the best results were obtained using a portion of the analyte's mass spectrum that distinguishes the spectrum of the desired analyte from that of similar compounds. In the case of dodecane, that selected region was found to be *m/z* 86–200, thus omitting *m/z* 45–85 was required. The raw results of the DotMap analysis using *m/z* 86–200 of the dodecane spectrum and the jet fuel data is shown in Fig. 5B. When comparing the column 1 projections of the DotMap results in Fig. 5A and B, it can be seen that the dot product intensity for dodecane is greatly enhanced by using a selective portion of the dodecane spectrum. The extracted spectrum that corresponds to the analyte location indicated in the DotMap in Fig. 5B was matched successfully to the library spectrum collected in-house of dodecane.

To optimize the DotMap analysis to identify a class of compounds, in this example the alkanes, several approaches were also investigated using selected ions (again not all approaches shown for brevity), but the best results were obtained using a portion of the dodecane spectrum that is common to all alkanes in the sample. The region *m/z* 45–85 of the dodecane spectrum was used as the search spectrum for the DotMap analysis. The raw DotMap data for column 1 and the ensuing contour plot are shown in Fig. 6A and B. The raw data indicate a high degree of mass spectral similarity between many compounds in the sample and the portion of the dodecane spectrum that it was searched against. While the highest dot product intensity in Fig. 6A corresponds to dodecane, all compounds that reach the threshold are shown in the DotMap contour plot (Fig. 6B). All of the major alkanes from octane to hexadecane were found in this jet fuel sample using this approach. Although peaks for heptane (C7) and heptadecane (C17) have similar mass spectra to that of other alkanes, it can be seen in Fig. 6A that the DotMap intensity for these compounds does not cross the threshold. The threshold chosen by the algorithm is set at 90% of the difference between the maximum point and the median of the DotMap results. This is an empirically determined set point chosen because it afforded the best compromise between specificity and completeness in its representation of DotMap results. The threshold selection is a parameter that could be modified to fit user needs for a specific application or sample matrix. Specifically in the case of the alkane group identification it is noteworthy that the bulk of the peak intensity in the separation space is concentrated in the region from 10 to 30 min of separation time (Fig. 1A), the peaks for heptane and heptadecane are outside this region (Fig. 6A), and therefore would have lower S/N ratios and correspondingly lower matches to other alkane compounds. This is a similar phenomenon to the effect of decreased S/N on the match factors for TEPT discussed earlier.

It is important to emphasize that DotMap provides an alternate method for the elucidation of compound groups. Selected ion monitoring is the typical method used for locating groups of compounds in a separation space, but there are problems and limitations with this approach. Detailed lists of mass spectral peaks common to the target members com-

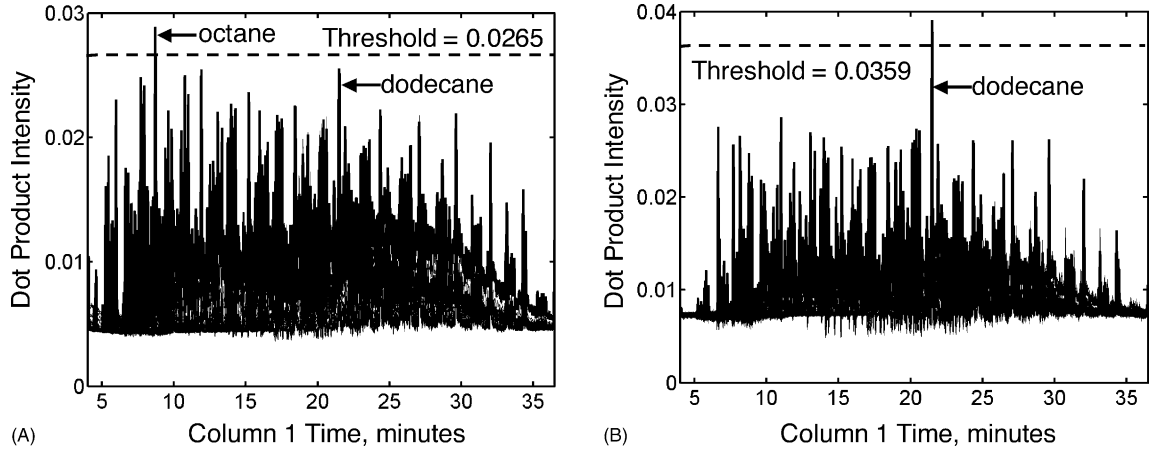

Fig. 5. DotMap analysis of *n*-dodecane present in jet fuel. (A) The raw DotMap results projected onto column 1 are shown. The DotMap analysis was performed using the entire mass spectrum (*m/z* 45–200) of dodecane to search the entire separation space. The mass spectrum of the peak that exceeded the algorithm's threshold, however, matches to that of octane. (B) The raw DotMap results projected onto column 1 for the analysis of *n*-dodecane using the portion of the spectrum unique to dodecane among alkanes (*m/z* 86–200) is shown. In this case the mass spectrum of the peak that exceeded the algorithm's threshold matches to that of dodecane.

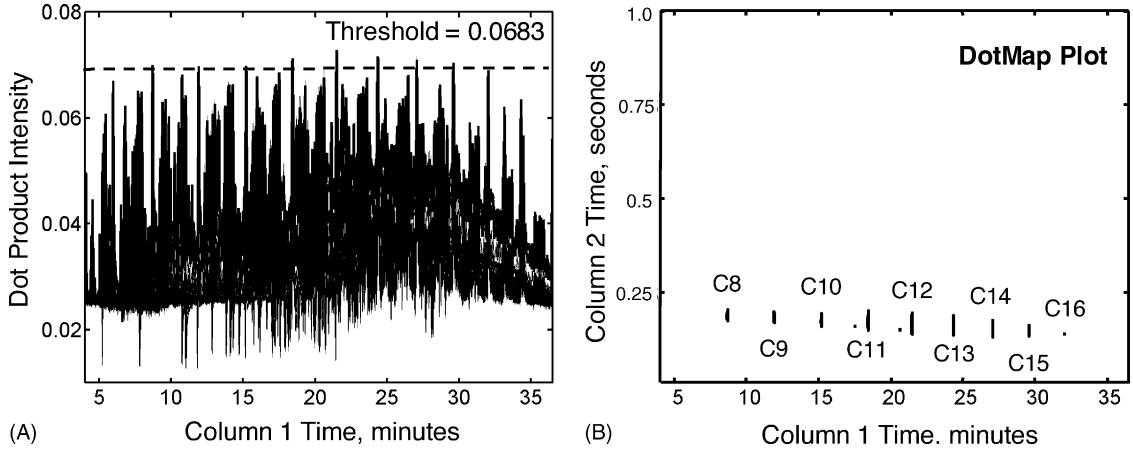

Fig. 6. DotMap analysis of *n*-dodecane present in jet fuel using the portion of the spectrum common to alkanes (*m/z* 45–85). (A) Raw DotMap results projected onto column 1. (B) Contour plot output from DotMap. The location of the alkanes as a class of compounds is depicted in this analysis. The peak labels indicate the identification of the *n*-alkanes located using the DotMap algorithm. Labels refer to the number of carbons in the straight chain *n*-alkane molecule.

pounds must be constructed and optimized and a detailed understanding of the matrix is also needed in some cases to avoid the selection of matrix masses. In Fig. 1B we see an example of a situation in which selected ion monitoring would fail due to matrix interferents. The image shown is that of a selected ion (*m/z* 65) and it is obvious that there is a large interferent overlapping with TEPT in that location. Were selected ion monitoring used to locate such a low concentration of TEPT in jet fuel, a false identification would be likely. It has been observed by van Deursen et al. that when doing selected ion monitoring in GC × GC separations that "Unique masses that were used were not quite as unique as it seems" and that unique masses "have to be chosen with great care to show the differences between two groups" [1]. In many cases, such as in the aforementioned study, it is necessary to sum multiple selected ions to obtain reliable group-type information. Furthermore, DotMap incorporates more information than a simple selected ion approach. Information about the relative intensities of target ion peaks is retained and used in the searching algorithm thus providing enhanced selectivity.

5. Conclusions

The DotMap algorithm is demonstrated to be a sensitive tool that is useful in the analysis of complex mixtures to elucidate the location of analytes of interest. This is particularly useful for analytes in samples exhibiting substantial matrix effects. The results of DotMap can potentially be tuned to provide specific analyte locations or compound class locations. Spectra of interest can be from either a commercial database or from a user's collected library. The algorithm performs successfully under a large range of signal intensities. Conditions of peak overlap can at best have little detrimental effect on DotMap results for two analytes with very different spec-

tra. The degree of spectral similarity and chromatographic resolution will ultimately dictate the success or failure of finding and identifying analytes of interest that are overlapped with interfering components. Additional study is warranted to investigate more deeply the impact of chromatographic overlap on the performance of DotMap.

Acknowledgements

Support for this project was provided by the United States Department of Defense (DOD) and the Center for Process Analytical Chemistry, University of Washington, Seattle, WA, USA.

References

[1] M. Van Deursen, J. Beens, J. Reijenga, P. Lipman, C. Cramers, J. Blomberg, J. High. Resolut. Chromatogr. 23 (2000) 507.
[2] R. Shellie, P. Marriott, P. Morrison, Anal. Chem. 73 (2001) 1336.
[3] J. Dalluge, M. van Rijn, J. Beens, R.J.J. Vreuls, U.A.Th. Brinkman, J. Chromatogr. A 965 (2002) 207.
[4] J. Zrostlikova, J. Hajslova, T. Cajka, J. Chromatogr. A 1019 (2003) 173.
[5] J. Dalluge, R.J.J. Vreuls, J. Beens, U.A.Th. Brinkman, J. Sep. Sci. 25 (2002) 201.
[6] J. Dalluge, L.L.P. van Stee, X. Xu, J. Williams, J. Beens, R.J.J. Vreuls, U.A.Th. Brinkman, J. Chromatogr. A 974 (2002) 169.
[7] X. Lu, J. Cai, H. Kong, M. Wu, R. Hua, M. Zhao, J. Liu, G. Xu, Anal. Chem. 75 (2003) 4441.
[8] A.E. Sinha, B.J. Prazen, C.G. Fraga, R.E. Synovec, J. Chromatogr. A 1019 (2003) 79.
[9] J.-F. Focant, A. Sjodin, D.G. Patterson Jr., J. Chromatogr. A 1019 (2003) 143.
[10] P. Korytar, L.L.P. van Stee, P.E.G. Leonards, J. de Boer, U.A.Th. Brinkman, J. Chromatogr. A 994 (2003) 179.
[11] W. Welthagen, J. Schnelle-Kries, R. Zimmermann, J. Chromatogr. A 1019 (2003) 223.
[12] M. Adahchour, L.L.P. van Stee, J. Beens, R.J.J. Vreuls, M.A. Batenburg, U.A.Th. Brinkman, J. Chromatogr. A 1019 (2003) 157.
[13] A.E. Sinha, C.G. Fraga, B.J. Prazen, R.E. Synovec, J. Chromatogr. A 1027 (2004) 269.
[14] J.L. Hope, B.J. Prazen, E.J. Nilsson, M.E. Lidstrom, R.E. Synovec, Talanta 65 (2004) 380.
[15] S.M. Song, P. Marriott, A. Kotsos, O.H. Drummer, P. Wynne, Forensic Sci. Int. 143 (2004) 87.
[16] V.G. van Mispelaar, A.C. Tas, A.K. Smilde, P.J. Schoenmakers, A.C. van Asten, J. Chromatogr. A 1019 (2003) 15.
[17] R.X. Hua, Y.Y. Li, W. Liu, J.C. Zheng, H.B. Wei, J.H. Wang, X. Lu, H.W. Kong, G.W. Xu, J. Chromatogr. A 1019 (2003) 101.
[18] M. Harju, C. Danielsson, P. Haglund, J. Chromatogr. A 1019 (2003) 111.
[19] M. Harju, A. Bergman, M. Olsson, A. Roos, P. Haglund, J. Chromatogr. A 1019 (2003) 127.
[20] T.T. Truong, P.J. Marriott, N.A. Porter, R. Leeming, J. Chromatogr. A 1019 (2003) 197.
[21] R.B. Poe, S.C. Rutan, Anal. Chim. Acta 283 (1993) 845.
[22] E.A. Dietz, J. High. Resolut. Chromatogr. 19 (1996) 485.
[23] C.G. Fraga, C.A. Bruckner, R.E. Synovec, Anal. Chem. 73 (2001) 675.
[24] C.G. Fraga, B.J. Prazen, R.E. Synovec, Anal. Chem. 72 (2000) 4154.
[25] J.L. Hope, K.J. Johnson, M.A. Cavelti, B.J. Prazen, J.W. Grate, R.E. Synovec, Anal. Chim. Acta 490 (2003) 223.
[26] A.E. Sinha, J.L. Hope, B.J. Prazen, E.J. Nilsson, R.M. Jack, R.E. Synovec, J. Chromatogr. A 1058 (2004) 209.
[27] R.E. Murphy, M.R. Schure, J.P. Foley, Anal. Chem. 70 (1998) 1585.
[28] P. Schoenmakers, P. Marriott, J. Beens, LC GC Eur. 16 (2003) 335.
[29] S.Y. Mirokhin, D. Tchekhovskoi, G. Mallard, NIST (2002).