COMPARISON BETWEEN AUTOMATED KNOWLEDGE-BASED SYSTEMS IDENTIFYING AND CLASSIFYING AMPHETAMINE ANALOGUES USING VAPOR-PHASE FTIR SPECTRA AND MASS SPECTRA

Inge DIRINCK, Mirela PRAISLER, Jan F. VAN BOCXLAER,
Andreas DE LEENHEER, D. Luc MASSART

Laboratory of Toxicology, Faculty of Pharmaceutical Sciences, University of Ghent, Ghent, Belgium

ABSTRACT: The feasibility of identifying and classifying novel amphetamine analogues according to the substitution pattern of the phenyl ring was explored by principal component analysis (PCA) run for appropriately feature weighed vapor-phase infrared spectra and mass spectra. The appropriately pre-processed FTIR spectral features enabled the automated discrimination and recognition of the subclasses of phenyl nonsubstituted (stimulant) and 3,4-methylenedioxy- (hallucinogenic) amphetamines within higher, statistically determined, reliability limits.

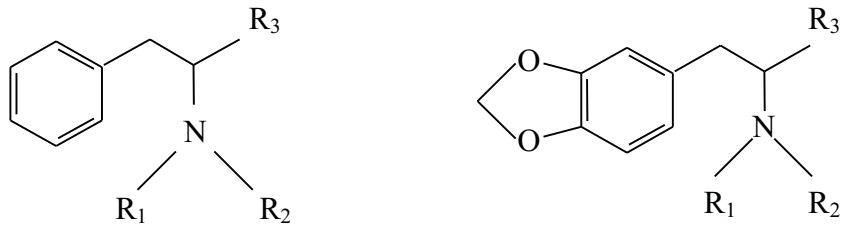
KEY WORDS: Amphetamines; FTIR spectra; Mass spectra.

Problems of Forensic Sciences, vol. XLIII, 2000, 62–69
Received 9 September 1999; accepted 16 May 2000

INTRODUCTION

Many amphetamine analogues have emerged as recreational drugs, so-called "designer drugs", over the last decade. In attempts to circumvent existing controlled substance laws, clandestine laboratories are synthesizing slightly modified chemical structures by adding or changing substituents at various positions on the amphetamine molecule without significantly altering its psychotropic effect [3]. When these changes address the substitution pattern of the phenyl ring, the resulting structures possess hallucinogenic and mood-modifying properties in addition to the CNS stimulation of the phenyl nonsubstituted amphetamines (Figure 1). The latter drugs of abuse have increased toxicity and thus no medical use, being listed under Schedule I of CSA. The unambiguous identification of novel amphetamine analogues is frequently necessary in judicial cases mainly related to deaths from overdose or to illicit drug seizures.

Fig. 1. Modeled amphetamine structures: phenyl nonsubstituted amphetamine analogues (left) and 3,4-methylenedioxy- amphetamine analogues (right).



We explored the feasibility of identifying and classifying novel amphetamine analogues according to the substitution pattern of the phenyl ring by principal component analysis (PCA) [2, 4]. We are reporting a comparative evaluation of the results obtained with automated knowledge-based systems, built using appropriately feature weighed vapor-phase FTIR spectra (GC-FTIR) and mass spectra (GC-MS) of amphetamines.

EXPERIMENTAL

Vapor-phase infrared spectra were imported from a laboratory-made Vapor-Phase FTIR Library, which contains 159 reference spectra of commercially available amphetamines, several in-house synthesized amphetamine analogues, and other related drugs of abuse [1]. The electron impact mass spectra of the same compounds were imported from general MS libraries (NIST Mass Spectral Database, AAFS Spectral Library) and an In-House MS Library.

The knowledge-based classification and identification system for amphetamine analogues was obtained, by using the same computational procedure for both FTIR and MS spectra, as follows: a training set was selected and a feature weight function wk was calculated; PCA was run for the feature weighed spectra of the training set; clusters (subclasses) were identified and characterized using the principal components. The modeled data were mean-centered, to ensure that all results will be interpretable in terms of variation around the mean. The validation method was full cross-validation, as the number of samples in the training set (43) was relatively small. PCA was performed using the software package The Unscrambler® (Camo AS, Sweden).

RESULTS AND DISCUSSION

The training set consisted of class I containing the spectra of interest (amphetamine analogues and their heptafluorobutyric anhydride (HFB) derivatives), and class II containing counter-examples (nonamphetamine compounds). Class I (Table I) was composed of phenyl nonsubstituted amphetamine analogues (class code M),

3,4-methylenedioxy- analogues (class code T), and their HFB-derivatives (class code DM and DT, respectively). Class II contained randomly selected nonamphetamine compounds (class code **N**) with very different molecular structures, spanning a broad range of various toxicologically interesting compounds.

TABLE I. LIST OF THE COMPOUNDS SELECTED IN CLASS I (AMPHETAMINE ANALOGUES AND DERIVATIVES) OF THE PCA TRAINING SET WITH THEIR SPECTRAL LIBRARY ENTRY CODE, ACRONYM, AND SUBSTITUENTS

ID	Name of the compound	Acronym	R_1	R_2	R_3
M4	α-phenylethylamine	APEA	H	H	
M102	*N*-methyl-α-phenylethylamine[a]	MAPEA	$-CH_3$	H	
M7	amphetamine	AMP	H	H	$-CH_3$
M17	β-phenylethylamine	BPEA	H	H	H
M74	methamphetamine	MAMP	$-CH_3$	H	$-CH_3$
M96	*N*-ethylamphetamine	EAMP	$-CH_2-CH_3$	H	$-CH_3$
M114	*N*-*n*-propylamphetamine[a]	PAMP	$-(CH_2)_2-CH_3$	H	$-CH_3$
T79	3,4-methylenedioxyamphetamine	MDA	H	H	$-CH_3$
T82	3,4-methylenedioxy-*N*-ethylamphetamine	MDEA	$-CH_2-CH_3$	H	$-CH_3$
T85	1-(3,4-methylenedioxyphenyl)-2-butanamine	BDB	H	H	$-CH_2-CH_3$
T87	3,4-methylenedioxymethamphetamine	MDMA	$-CH_3$	H	$-CH_3$
T106	*N*-methyl-1-(3,4-methylenedioxyphenyl)-2-butanamine	MBDB	$-CH_3$	H	$-CH_2-CH_3$
T154	3,4-methylenedioxy-*N*-hydroxyamphetamine	N-OH-MDA	-OH	H	$-CH_3$

[a] – Sythesis at the Laboratory of Toxicology, University of Ghent.

All FTIR spectra were reduced in size by eliminating the wavenumber intervals where the compounds in the data base have no IR absorptions. Hence, FTIR data used for multivariate analysis ranged from 3750 to 2550, and from 2000 to 600 cm^{-1}. The 43 samples forming the training set and the remaining 523 wavenumber intervals of 5 cm^{-1} resulted in a data matrix with 43 x 523 entries.

The MS data of all 159 FTIR-library compounds were all normalized (highest abundance set to 100%) and ranged from m/z 12 to 412. A data matrix with 159 rows and 401 columns was built, containing for each sample the relative abundance values (%) for each m/z. The mass spectra of the training set compounds resulted in a data matrix with 43 x 401 entries, which was used for multivariate exploratory analysis.

A feature weight spectrum wk (I, II) [2] was calculated for each wavenumber or m/z from the ratio of the intercategory variances to the sum of the intracategory variances:

$$w_k = \frac{\sum \frac{A_I^2}{N_I} + \sum \frac{A_{II}^2}{N_{II}} - 2\sum\sum \frac{A_I A_{II}}{N_I N_{II}}}{\sum \frac{\left(A_I - \overline{A}_I\right)^2}{N_I} + \sum \frac{\left(A_{II} - \overline{A}_{II}\right)^2}{N_{II}}}$$

where A_I, A_{II} are the absorbances (FTIR) or the relative abundances (MS) at wavenumber |m/z k for the samples of class I and II; A_I, A_{II} the mean absorbances|relative abundances at wavenumber|m/z k for the samples of class I and II; N_I, N_{II} the numbers of samples in the training sets of class I and II. Data scaling improved the separation between the two classes, and the identification of more than two clusters. The feature weights w, w^2, and $(w-1)^2$ were investigated, and the best cluster discrimination was found with $(w-1)^2$.

PCA was run for the feature weighed FTIR and MS spectra of the training set. The explained variance, measured as a percentage of the total variance in the data, is a measurement of the proportion of information (variation in the data) being described by the current PC. The effect of the number of model principal components (PC) on the explained variances was studied to decide how complex the PCA model should be. The

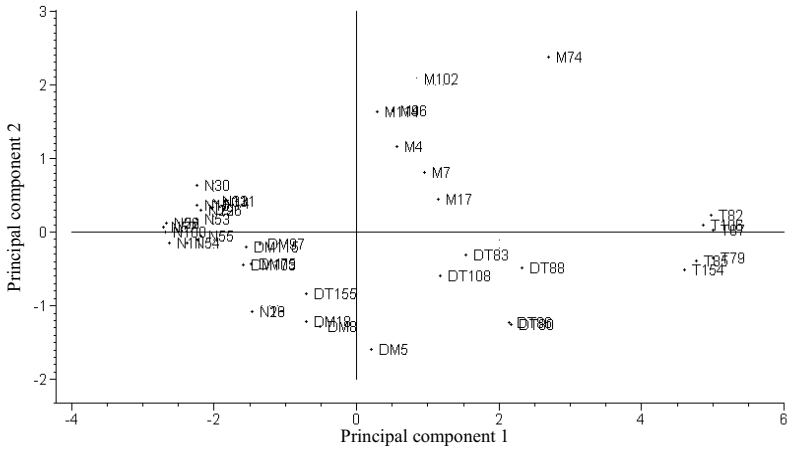


Fig. 2. Discrimination of stimulant (class code M) and hallucinogenic (class code T) amphetamines with vapor-phase FTIR spectra.

exploratory analysis of the FTIR data was done with the first three PCs, which expressed 84%, 10% and 3% of the information respectively. The exploratory analysis of the MS data done with the first three PCs explained 78%, 7% and 4% of the information respectively. Models built with more than three PCs did not improve the explained

variances of the first PCs, nor the discrimination among the stimulant and hallucinogenic amphetamine analogues.

The score plots of the FTIR exploratory analysis proved the feasibility of discriminating among the different (sub)classes of compounds in the training set and put in evidence substitution patterns that could be recognized using this system. The $(w-1)^2$ feature weighed FTIR spectra enabled a reliable discrimination and recognition of the subclasses of phenyl- nonsubstituted and 3,4-methylenedioxy- amphetamines. Indeed, the stimulant (M) amphetamine analogues are very well separated from the rest of the com-

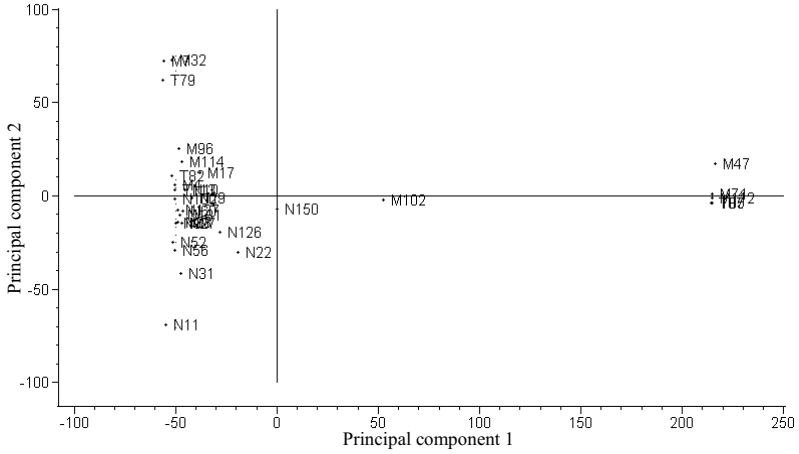


Fig. 3. Discrimination of stimulant (class code M) and hallucinogenic (class code T) amphetamines with MS spectra.

pounds, being the only ones with positive PC1 and PC2 scores (Figure 2). The hallucinogenic (T) amphetamine analogues cluster the best, and are characterized by the highest positive PC1 scores. The phenyl substitution pattern has only a limited influence on the scores of the HFB-derivatives of the amphetamines. In addition, the derivatives are not well separated from nonamphetamine compounds. For example, codeine (code **N**29) lies within the same region as the DM derivatives.

PCA analysis of the MS data matrix proved that the differences in fragmentation pattern are not significantly correlated with the substitution patterns of the phenyl ring (Figure 3). Discrimination according to this structural feature could not be obtained within statistically acceptable reliability limits. The MS spectra of amphetamine analogues can be distinguished from those of nonamphetamines (negative PC1 and PC2) only as a whole class.

The results are due to the fact that, while in the case of the FTIR spectra a much larger number of variables are contributing to the system with information specific to the sub-

stitution pattern of the phenyl ring, for the MS spectra the discrimination relies only on m/z 44 and 91 that cause high positive PC2 scores, and on m/z 58 generating high positive PC1 scores (Figure 4). The molecular structures of the samples forming the clusters in relation to the discriminating mass numbers, indicate that the cleavage always takes place between the first and the second carbon atom of the side chain. The most stabile ion represents the fragment of the side chain containing the nitrogen atom. This frag-

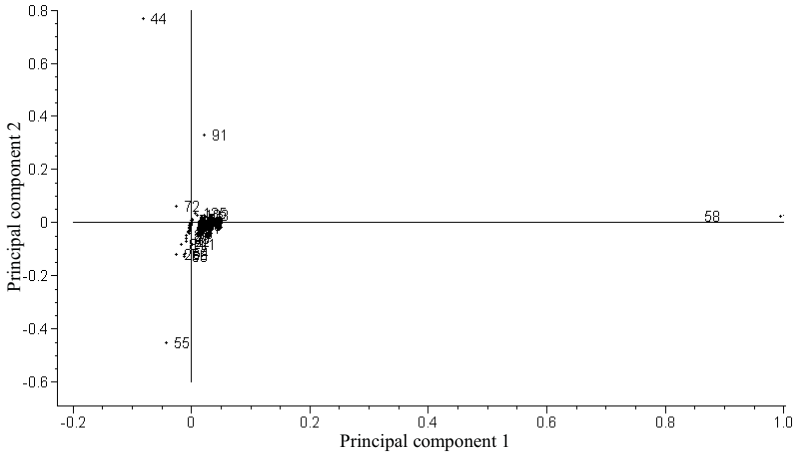


Fig. 4. Loading plot identifying the variables with the most discriminating power in the MS spectra of amphetamines.

mentation pattern explains the overlapping of stimulant and hallucinogenic analogues, the clusters being formed only as a function of the mass of the side-chain fragment. PC1, explaining 78% of the data variation, distinguishes the stimulant and hallucinogenic analogues with a side-chain fragment of 58 amu (C_3H_8N), such as methamphetamine (code M74), 1-phenyl-2-butanamine (code M47), or 1-(3,4-methylenedioxyphenyl)-2-butanamine (code T85). PC2 discriminates, to a much lower extent (explains 7% of the data variation), between nonamphetamines and the stimulant and hallucinogenic amphetamines with a side-chain fragment of 44 amu (C_2H_6N), such as amphetamine (code M7) or 3,4-methylenedioxyamphetamine (code T79). Thus, the mass spectrum of some amphetamine analogues may not provide enough specificity for structural pattern identification. On the other hand, due to amino-dominated fragmentation, MS spectra seem to be more appropriate than FTIR spectra for the classification of stimulant and hallucinogenic amphetamines with an identical side chain and nitrogen substituted fragment.

CONCLUSIONS

The PCA score plots offered evidence that the system built with FTIR spectra can discriminate between amphetamine analogues and nonamphetamine compounds, and thus could act as an automated drug-of-abuse screening test. The best selectivity in discriminating with FTIR spectra among phenyl nonsubstituted (stimulant) amphetamine analogues, 3,4-methylenedioxy- (hallucinogenic) amphetamine analogues, and nonamphetamine compounds was obtained with the FTIR spectra of nonderivatized samples rather than with their HFB-derivatives. The score plots also indicated that more accurate discrimination among amphetamine analogues according to substitution patterns and related biological activity is obtained with their vapor-phase FTIR spectra, rather than with their MS spectra.

Acknowledgements:
This study has been made within a project (BIL 97/75, 174B2198) included in and financed by the Bilateral Scientific and Technological Cooperation Agreement between Flanders and Romania. This work was also supported by the Fund for Medical Scientific Research through grant 9.0024.92.

References:

1. Dirinck I., Meyer E., Van Bocxlaer J., Lambert W., De Leenheer A., Application of gas chromatography-Fourier transform infrared spectrometry to the analysis of amphetamine analogues, *Journal of Chromatography A* 1998, vol. 819, pp. 155–159.
2. Hasenoehrl E. J., Perkins J.H., Griffiths P. R., Expert system based on principal component analysis for the identification of molecular structures from vapor-phase infrared spectra, *Analytical Chemistry* 1992, vol. 64, pp. 656–663.
3. Karch S. B., Drug abuse handbook, CRC Press, New York 1998.
4. Massart D. L., Vandeginste B. G. M., Buydens L. M. C., De Long S., Lewi P. J., Smeyers-Verbeke J., Data handling in science and technology: handbook of chemometrics and qualimetrics (part A), Elsevier, Amsterdam 1997.