

Target-testing factor analysis of Fourier transform infrared spectra as a tool for the determination of the composition of human urinary calculi

Igor Kuzmanovski ^{a,*}, Mira Trpkovska ^a, Bojan Šoptrajanov ^{a,b}, Viktor Stefov ^a

^a *Institut za hemija, PMF, Univerzitet "Sv. Kiril i Metodij", PO Box 162, 91001 Skopje, Macedonia*

^b *Makedonska akademija na naukite i umetnostite, Skopje, Macedonia*

Received 17 July 1998; received in revised form 13 November 1998; accepted 16 November 1998

Abstract

The qualitative interpretation of the infrared spectra of urinary calculi requires expert knowledge, and still it is not always possible to determine all the components, especially those with similar spectra that appear together, e.g., whewellite, $\text{CaC}_2\text{O}_4 \cdot \text{H}_2\text{O}$ and weddellite, $\text{CaC}_2\text{O}_4 \cdot 2\text{H}_2\text{O}$. Instead of empirically assigning the spectra, factor analysis was employed to determine the number of components of urinary calculi. The method was tested on binary and ternary artificial mixtures, and then employed for assigning the composition of the human urinary calculi. The results of this work show that the target-testing factor analysis of the infrared spectra is a promising method, which minimises the role of the person carrying out the analysis. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Urinary calculi; Target-testing factor analysis; Infrared spectroscopy

1. Introduction

The determination of the composition of urinary calculi is important both for the possible prevention of further urolithiasis and for the treatment of the disease. Infrared spectroscopy is probably the instrumental method of analysis which is most frequently used for this purpose [1].

In order to obtain reliable results when determining the calculi composition, it is desirable to minimise the role of the person carrying out the analysis and this is why numerous computerised methods for

calculi analysis have been developed [2–5]. In this paper, the target-testing factor analysis applied to investigate the composition of the calculi is presented. It should, perhaps, be mentioned that factor analysis (a statistical method based on linear algebra) was developed at the beginning of this century and was at first used in sociology and psychology. Presently, factor analysis is used in many exact and applied sciences, including chemistry.

A detailed description of the target-testing factor analysis in chemistry is given in Ref. [6] and here only the essential points will be briefly summarised. The first step of the procedure is the setting of the data matrix **D** followed by premultiplying of this matrix with its transpose to obtain the covariance matrix **C** ($\mathbf{C} = \mathbf{D}^T \cdot \mathbf{D}$). The next step is the decom-

* Corresponding author. E-mail: shigor@iunona.pmf.ukim.edu.mk

position of the covariance matrix into its eigenvalues (these will be written as λ_i) and eigenvectors. When the target-testing factor analysis is applied to spectroscopic data, the data matrix (consisting of the points defining the recorded spectra) is first multiplied by a matrix \mathbf{E}' consisting of primary eigenvectors to obtain the matrix \mathbf{A} ($\mathbf{A} = \mathbf{D} \cdot \mathbf{E}'$) the elements of which define the abstract eigenspectra.¹ The eigenspectra are then converted into real ones using the matrix transformation $\mathbf{T}_i = \mathbf{\Lambda}'^{-1} \cdot \mathbf{A}^T \cdot \mathbf{R}_i$ where \mathbf{T}_i is a column of the transformation matrix, $\mathbf{\Lambda}'^{-1}$ is the inverse diagonal matrix of the eigenvalues matrix consisting of the primary set of eigenvalues and \mathbf{R}_i is a column of the matrix corresponding to a prearranged library of standard spectra. Once the abstract eigenspectra are attained using a primary set of eigenvectors, the real spectra of the constituents are finally obtained using the target-transformation routine [6].

The wide-spread use of Fourier transform infrared instruments makes the application of statistical methods (including factor analysis) practical since the data are stored and processed in digital form. It is, thus, easy to set up the data matrices and to manipulate them as shown above.

The critical step in the application of factor analysis for determining the composition of calculi (or, for that matter, any mixture) is the decision about the number of components present in the system since the number of eigenvectors in the \mathbf{E}' matrix depends on this choice. To solve this key problem, many statistical and empirical criteria have been developed [6,7]. Experience shows that it is advisable to use more than one such criterion since not all of them always indicate the same number of components [6]. In our work several criteria were tested, the use of two of them being presented here. These two criteria are the *minima* of the indicator function IND and the predicted sum of squares PRESS [6,7]. The functions themselves are defined as follows:

$$\text{IND} = \sqrt{\frac{\sum_{j=n+1}^c \lambda_j}{r(c-n)^5}}$$

¹ The abstract spectra, as their name implies, have no physical meaning.

$$\text{PRESS} = \sum_{i=1}^r \sum_{k=1}^c (d_{i,k} - \hat{d}_{i,k}(l))^2$$

where λ_j are the eigenvalues resulting from the decomposition of the covariance matrix \mathbf{C} , n is the predicted number of components, c is the number of columns in the data matrix \mathbf{D} , r is the number of rows in the data matrix, l is the number of eigenvectors estimated as members of the primary set of eigenvectors used for the reconstruction of the data matrix, $d_{i,k}$ is a point in the original data matrix and $\hat{d}_{i,k}(l)$ is the reconstructed point of the data matrix using l eigenvectors.

2. Experimental

The samples and the standards (the latter were needed to set up the library of standard spectra) were prepared as KBr pellets. Each pellet was made up of 1 mg of sample and 250 mg of spectroscopy grade KBr. The digital spectra were recorded at room temperature in the 4000–370 cm^{-1} region on an FTIR Perkin-Elmer System 2000 instrument (the resolution was 4 cm^{-1}), with 32 scans for the samples and 32 scans for the background (atmosphere and KBr). All calculations were carried out using the software package MATHCAD 7.0 Professional [8] for which a suitable program (available on request) has been written.

3. Results and discussion

In order to test the method, a number of binary and ternary mixtures of compounds which are known [2,9–12] to be constituents of urinary calculi were prepared and subjected to the target-testing factor analysis (the mass ratios of the constituents in each mixture were different but exactly known). As an illustration, the results obtained for four binary mixtures (**1**—whewellite and weddellite; **2**—whewellite and cystine; **3**—weddellite and uric acid; **4**—whewellite and whitlockite²) and four ternary mixtures (**1**—whewellite, weddellite and brushite³;

² This is the mineralogical name of $\text{Ca}_3(\text{PO}_4)_2$, one of the calcium phosphates reported to be a constituent of urinary calculi.

³ The formula of this compound is $\text{CaHPO}_4 \cdot 2\text{H}_2\text{O}$.

Table 1
The values for IND and PRESS functions for the artificial mixtures

Mixture	Criterion	Number of constituents			
		1	2	3	4
<i>Binary mixtures</i>					
1	IND	$6.970 \cdot 10^{-4}$	$2.862 \cdot 10^{-4}$	$7.780 \cdot 10^{-4}$	$6.962 \cdot 10^{-4}$
	PRESS	0.556	$2.376 \cdot 10^{-2}$	$7.743 \cdot 10^{-2}$	$9.245 \cdot 10^{-2}$
2	IND	$1.601 \cdot 10^{-3}$	$1.077 \cdot 10^{-3}$	$1.117 \cdot 10^{-3}$	$2.760 \cdot 10^{-3}$
	PRESS	2.628	0.282	0.681	0.671
3	IND	$1.853 \cdot 10^{-3}$	$3.959 \cdot 10^{-4}$	$5.530 \cdot 10^{-4}$	$1.712 \cdot 10^{-3}$
	PRESS	3.521	$3.813 \cdot 10^{-2}$	0.622	0.641
4	IND	$1.519 \cdot 10^{-3}$	$7.978 \cdot 10^{-4}$	$8.719 \cdot 10^{-4}$	$2.172 \cdot 10^{-3}$
	PRESS	2.365	0.154	0.278	0.291
<i>Ternary mixtures</i>					
1	IND	$1.143 \cdot 10^{-3}$	$1.183 \cdot 10^{-3}$	$9.175 \cdot 10^{-4}$	$4.27 \cdot 10^{-3}$
	PRESS	1.339	0.340	$1.824 \cdot 10^{-2}$	$2.697 \cdot 10^{-2}$
2	IND	$8.852 \cdot 10^{-3}$	$4.798 \cdot 10^{-3}$	$2.007 \cdot 10^{-3}$	$1.110 \cdot 10^{-2}$
	PRESS	2.679	0.166	$2.513 \cdot 10^{-2}$	$4.256 \cdot 10^{-2}$
3	IND	$2.690 \cdot 10^{-3}$	$2.898 \cdot 10^{-3}$	$9.261 \cdot 10^{-4}$	$1.668 \cdot 10^{-3}$
	PRESS	7.415	2.042	$2.747 \cdot 10^{-2}$	$5.483 \cdot 10^{-2}$
4	IND	$1.496 \cdot 10^{-3}$	$8.770 \cdot 10^{-4}$	$5.403 \cdot 10^{-4}$	$2.718 \cdot 10^{-3}$
	PRESS	2.294	0.187	$9.352 \cdot 10^{-3}$	$2.921 \cdot 10^{-2}$

2—whewellite, weddellite and whitlockite; **3**—whewellite, weddellite and uric acid; **4**—brushite, cystine and uric acid) are presented below.

The factor analysis was applied to a data matrix consisting of all the recorded spectra of the mixtures in the $1450\text{--}450\text{ cm}^{-1}$ region.

The values for the IND and PRESS functions are shown in Table 1. As expected, both functions reach their respective minima at the expected number of

components ($n = 2$ for the binary and $n = 3$ for the ternary mixtures).

The satisfactory results of the analysis of the artificially prepared mixtures suggested that an analogous procedure could be used for analysing real urinary calculi. This was done for a number of urinary calculi, the results for five of them being given in Table 2. As an illustration, the analysis of one of the studied calculi will be presented in some

Table 2
The values for IND and PRESS functions for the analysed calculi

Calculus	Criterion	Number of constituents				Composition
		1	2	3	4	
1	IND	$3.097 \cdot 10^{-5}$	$1.219 \cdot 10^{-5}$	$7.320 \cdot 10^{-6}$	$5.617 \cdot 10^{-5}$	Whewellite, weddellite and octacalcium phosphate
	PRESS	19.843	0.562	0.329	0.586	
2	IND	$1.786 \cdot 10^{-7}$	$2.740 \cdot 10^{-7}$	$5.862 \cdot 10^{-7}$	$2.304 \cdot 10^{-6}$	Uric acid
	PRESS	$1.101 \cdot 10^{-8}$	$1.826 \cdot 10^{-8}$	$3.270 \cdot 10^{-8}$	$1.072 \cdot 10^{-3}$	
3	IND	$7.253 \cdot 10^{-4}$	$2.876 \cdot 10^{-6}$	$4.931 \cdot 10^{-6}$	$1.832 \cdot 10^{-5}$	Octacalcium phosphate and cystine
	PRESS	0.539	$2.012 \cdot 10^{-6}$	$5.159 \cdot 10^{-3}$	$2.966 \cdot 10^{-2}$	
4	IND	$2.783 \cdot 10^{-3}$	$3.232 \cdot 10^{-4}$	$2.656 \cdot 10^{-4}$	$1.502 \cdot 10^{-3}$	Whewellite, weddellite and calcium carbonate
	PRESS	18.060	$3.283 \cdot 10^{-2}$	$1.023 \cdot 10^{-2}$	$2.952 \cdot 10^{-2}$	
5	IND	$7.474 \cdot 10^{-4}$	$7.777 \cdot 10^{-4}$	$3.215 \cdot 10^{-4}$	$1.582 \cdot 10^{-3}$	Whewellite, weddellite and uric acid
	PRESS	0.573	0.147	$5.271 \cdot 10^{-4}$	$1.095 \cdot 10^{-2}$	

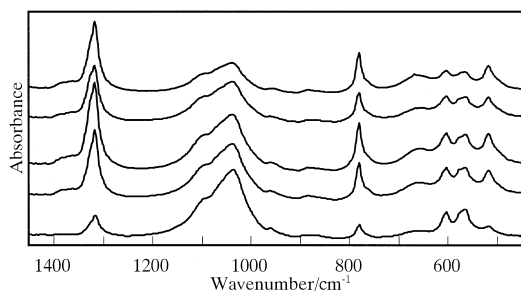


Fig. 1. Spectra recorded from different calculus layers.

detail. The analysed calculus (denoted as calculus **1** in Table 2) had a diameter of 8 mm and several layers could be observed on its cross-section. Five spectra were recorded, each of them corresponding to a different layer (cf. Fig. 1).

On inspecting the spectra shown in Fig. 1, one could perhaps conclude that only *two* components are present in the sample. One of them undoubtedly is an oxalate and the other component must be some phosphate as suggested by the presence of the bands around 1050 cm^{-1} . However, the results obtained by factor analysis showed that the calculus is not composed of two, but of three components as indicated by the minima of both criteria which were used (cf. Table 2). In fact, a more detailed analysis of the spectra could have given the same result since the oxalate peaks at around 1320 cm^{-1} are found at slightly different positions (intermediate of those characteristic for whewellite and weddelite) apparently dependent on the content of the two oxalates in each layer.

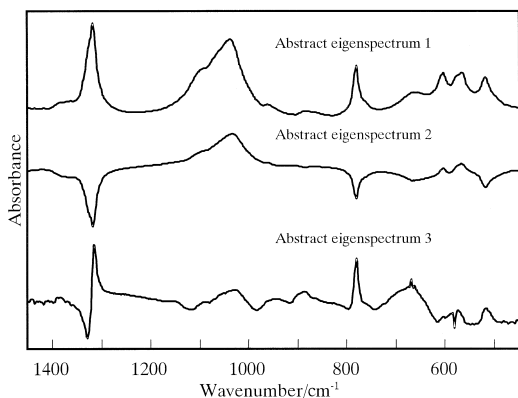


Fig. 2. Abstract eigenspectra computed using three eigenvectors.

In order to find which substances were present in the calculus, a primary set of eigenvectors was used to obtain abstract eigenspectra (Fig. 2). Those spectra were target-transformed into real ones. As shown in Fig. 3, two of the components are indeed oxalates ($\text{CaC}_2\text{O}_4 \cdot \text{H}_2\text{O}$ and $\text{CaC}_2\text{O}_4 \cdot 2\text{H}_2\text{O}$) and the third component is octacalcium phosphate.

The main limitation for the application of this method is that it requires spectra recorded from different calculi layers. Since most of the calculi are layered anyhow, the limitation is not a serious one [1].

The results of this work show that the target-testing factor analysis is a promising method which minimises the role of the person carrying out the analysis (this is important especially when the operator is not highly qualified specialist in infrared spectroscopy). Furthermore, it does not require specialised (and expensive) software which may not be available in each laboratory. At present, we have concentrated on the qualitative analysis of the calculi (determination of the number and nature of their

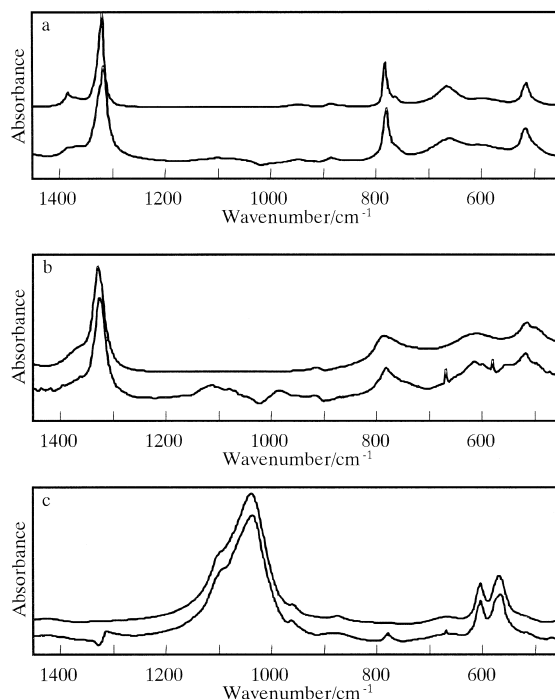


Fig. 3. Original and reconstructed spectra of the components by target-transformation. (a) Whewellite, (b) wheddelite, (c) octacalcium phosphate.

constituents) and have not attempted to apply the method to quantitative or semi-quantitative analysis.

References

- [1] M. Daudon, R.J. Reveillaud, *Presse Med.* 16 (1987) 627.
- [2] S.H. Kandil, T.A. Abou El Azm, A.M. Gad, M.M. Abdou, *Comput. Enhanced Spectrosc.* 3 (1986) 171.
- [3] A. Hesse, M. Gergeleit, P. Schüller, K. Möller, *J. Clin. Chem. Clin. Biochem.* 27 (1989) 639.
- [4] H. Hobert, K. Meyer, *Fresenius J. Anal. Chem.* 334 (1992) 178.
- [5] E. Peuchant, X. Heches, D. Sess, M. Clerc, *Clin. Chim. Acta* 205 (1992) 19.
- [6] E.R. Malinowski, *Factor Analysis in Chemistry*, 2nd edn., Wiley, New York, 1991.
- [7] E.R. Malinowski, *Anal. Chem.* 49 (1977) 612.
- [8] MATHCAD 7.0 Professional, 1986–1997, MathSoft.
- [9] I. Petrov, B. Šoptrajanov, K. Stojanova, *Experientia* 25 (1969) 693.
- [10] K. Stojanova, I. Petrov, B. Šoptrajanov, *Maked. Med. Pregled* 24 (1969) 71.
- [11] K. Stojanova, I. Petrov, B. Šoptrajanov, *Jugoslav. Physiol. Pharmacol. Acta* 5 (1969) 137.
- [12] M. Berthelot, G. Cornu, M. Daudon, M. Helbert, C. Laurence, *Clin. Chem.* 33 (1987) 2070.