April 2006 Volume 2 Issue 3

Trade Science Inc.

Analytical CHEMISTRY

An Indian Journal

Full Paper

ACAIJ, 2(3), 2006 [81-91]

Identification Of Carbonyl Compounds In The Urine Of Healthy Adults: An Evaluation Of Biomarkers Using Pattern Recognition Methods

Warren J.Goux
Department of Chemistry,
The University of Texas at Dallas
P.O. Box 830688, Richardson,
TX 75083-0688 (USA)
Ph.: 972-883-2660; Fax: 972-883-2925
E-mail: wgoux@utdallas.edu

Gabriela Luta
Department of Chemistry,
The University of Texas at Dallas P.O. Box 830688,
Richardson, TX 75083-0688 (USA)

Received: 10th February, 2006
Accepted: 1st April, 2006

Web Publication Date : 7th April, 2006

ABSTRACT

Urinary carbonyls, 8-hydroxy-deoxyguanosine (8-OHdG), and isoprostane 8-epi $F_{2\alpha}$ were measured in healthy smoking and nonsmoking adults. Biomarkers most significant in separating the smoking from the nonsmoking groups were found to be hexanal, octanal, heptanal, 4-methyl-pentanone, butanal, 8-epi $F_{2\alpha}$ and 2-hexenal. Malondialdehyde (MDA), a frequently used oxidative stress marker, was found to be a poor discriminator. The smoking and nonsmoking groups were independently modeled using principal component analysis (PCA) and the validity of the approach tested using simple modeling by class analogy (SIMCA). Using SIMCA, greater than 80% of unknown data could be classified correctly using the 10 most discriminating biomarkers. Our results suggest that the significance of a biological effect may depend on the selection of biomarkers used.

KEYWORDS

Pattern recognition;
Oxidative stress;
Biomarker;
Carbonyl;
Antioxidant.

INTRODUCTION

Biological systems produce a diverse array of polar and apolar carbonyls, the most abundant of which are generated as byproducts of enzymatic and non-enzymatic peroxidation of unsaturated and polyunsaturated fatty acids[1-3]. The initiation of lipid peroxidation is the result of attack by singlet oxygen or hydroxide radical, produced as a byproduct of aerobic respiration[4]. Free radicals, particularly highly reactive hydroxyl radicals, are able to attack DNA, producing oxo- and hydroxylated nucleotide bases[5-

Full Paper

[11], while the generation of lipid peroxides stimulates the production of isoprostanes from membrane-bound arachidonic acid[12-15]. In higher organisms, carbonyl compounds, isoprotanes and enzymatically excised oxidized nucleotides are removed from the circulatory system by the liver and excreted in the urine[16-19]. Under conditions of increased oxidative stress brought about by diet[20-24], exercise[25, 26], exposure to toxins[7, 9, 12-15, 18, 20, 27-34], or disease[16, 34-41], these markers show increased concentrations in the blood and in the urine.

To date there have been more than 15 carbonyls identified in human urine[16,17,19], in addition to the frequently used urinary markers of DNA free radical oxidation, 8-hydroxydeoxyguanosine (8-OHdG), and isoprostane production, 8-epi $F_{2\alpha}$[6-15]. Clearly, some of these markers may be more sensitive to changes in diet or lifestyle than others. In order to identify those markers which are most sensitive, we measured nonpolar urinary carbonyls, 8-OHdG, and 8-epi $F_{2\alpha}$ in a group of subjects having diverse lifestyles. Biomarkers were ranked in their ability to discriminate between smoking and nonsmoking subjects using a pattern recognition approach. Eight carbonyls, in addition to 8-epi $F_{2\alpha}$ and 8-OHdG, were found to be the most discriminatory markers and were used as a basis for the classification of new data into smoking or nonsmoking groups with high reliability. Our results suggest that the significance of the biological effect be determined by treating a set of most discriminating variables in a comprehensive manner.

EXPERIMENTAL

Materials

C18 Sep Packs, Silica Sep Packs and C18 trifunctional packs were purchased from Phenomenex (Torrance, CA). Dowex-50, O-(2,3,4,5,6-pentafluorobenzyl)-hydroxylamine (PFBHA) hydrochloride, formaldehyde (as a 30% formalin solution), acetaldehyde dimethyl acetal, 1,1,3,3-tetraethoxypropane, and other carbonyl standards were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO) and were used without further purification. Acetaldehyde and malondialdehyde (MDA) standard stock solutions were prepared by heating an aqueous solution of 1,1,3,3-tetraethoxypropane and acetaldehyde dimethyl acetal containing a small amount of Dowex-50 ion exchange resin to 80 °C for 20 minutes in a sealed vial.

Study Design

Markers of oxidative stress, 8-OHdG, 8-epi $F_{2\alpha}$ and carbonyl compounds were measured weekly in the urine of 12 healthy adult volunteers (6 men and 6 women, mean age 46±8.9 yr) over a 13-week time period. Half of the volunteers (age 48±8.9 yr) smoked more than 10 cigarettes per day, while the remaining volunteers (age 32.8±8.0 yr) had never smoked. All subjects were healthy, not taking any vitamin or antioxidant other than what was provided to them. Diet was not regulated. First morning urine samples were collected and analyzed for carbonyl compounds, 8-OHdG, 8-epi $F_{2\alpha}$ and creatinine.

Identification of carbonyl compounds in urine

Carbonyl compounds were converted to their PFB-oximes and analyzed following previously published procedures[48,49]. Briefly, to a 10 mL urine sample was added 2 mL of a solution of 0.3 M PFBHA acidified with 40 µL of 37.7% hydrochloric acid. The sample were shaken overnight and then adsorbed onto a C18 Sep-Pack column. The column was washed with 5 mL of deionized water and the PFB-oxime derivatives eluted from the column with 8 mL of hexane. The hexane was dried with anhydrous magnesium sulfate, filtered, and analyzed using a Hewlett Packard 6890 GC equipped with a 5973 mass selective detector (Santa Corita, CA). 2 µL of the sample was injected in splitless mode onto a DB-5MS megabore column (J&W Scientific, 0.25 mm i.d. x 30 m, 0.25 µm film). The injector temperature was 250°C and the column flow was 1 mL/min using helium as a carrier gas. The oven temperature was initially 60°C and was held constant for 1 min, increased to 85°C at a rate of 20°C/min, increased to 300°C at a rate of 8°C/min, and finally held at 300°C for 5 min. In initial runs, spectra were acquired scanning mass to charge ratios (m/z) between 30 and 800 using electron-impact (EI) ionization. Using this method, 19 carbonyl compounds

Full Paper

were identified in urine samples by comparing spectra and retention times to those of PFB-oximes derivatives prepared from aqueous solutions of known carbonyl compounds.

Quantitation of carbonyl compounds by GC/MS

Since the PFB-oxime of furfural was not detected in any of the urine samples examined (12 randomly selected samples), 10 µL of 2 mM furfural (in methanol) was added to 10 mL of urine on which quantitative data was desired. Calibration was carried out using a standard addition method. Accordingly, a stock solution was prepared containing 33 mM formaldehyde, 6.0 mM acetaldehyde, 55 mM acetone, 7.0 mM propanal, butanone and 2-pentanone, 8.0 mM butanal, 1.5 mM 4-methyl-pentanone, hexanal and 2-hexenal, 1.6 mM valeraldehyde, 1.4 mM 2-hexanone, 1.0 mM 4-heptanone, heptanal, octanal, 2-nonenal, decanal, and undecanal, and 56 mM MDA. To each of five 10 mL urine samples already containing the furfural internal standard was added 0 µL, 5 µL, 10 µL, 20 µL and 40 µL of this stock solution. To these urine samples and to a blank containing only the furfural internal standard in 10 mL of deionized water, was added 2 mL of 0.3 M PFBHA acidified with 40 µL of 37.7% hydrochloric acid. The PFB-oxime derivatives of the carbonyl compounds were extracted and subjected to GC/MS, as previously described. Peaks in the chromatograph were identified by selective ion monitoring of the base peak (having a m/z ratio of 181). The relative intensity of each compound, identified by the retention times of its *syn* and *anti* oxime isomer peaks (except for MBA, which had 3 identifiable isomers), was calculated as

$$I_{rel} = \frac{(I_{ave} - I_{blank})}{I_o} \quad (1)$$

where I_{ave} is the average intensity of the *syn* and *anti* oxime isomer peaks (should both isomers be present), I_{blank} is the average intensity of the corresponding peaks in the blank, and I_o is the intensity of the furfural oxime peak. Calibration curves were constructed by plotting the I_{rel} against concentration of the carbonyl compound in the urine sample and carrying out a linear least-squares fit of the data. Fitting to all calibration curves yielded correlation coefficients greater than 0.92. Concentrations of carbonyl compounds (in µmoles/L) in unknown samples were calculated taking into account appropriate response factors.

Quantitation of urinary 8-OHdG and 8-epi $F_{2\alpha}$, and creatinine

Urinary 8-OHdG and 8-epi $F_{2\alpha}$ were measured using an immunoassay kit (Bioxytech immunodiagnostic kit, Oxis International, Portland, OR). Creatinine was measured using a colorimetric diagnostic kit (Sigma-Aldrich Chemical Co.).

Statistical analysis

The original data set consisted of 100 data points in the nonsmoking group and 85 data points in the smoking group. Outliers were eliminated on the basis of a modified z-test[50] to yield 88 and 72 data points in the nonsmoking and smoking groups, respectively. Significant differences between smokers and nonsmokers for a single variable were evaluated using a one-tailed t-test with unequal variances. The null hypothesis was rejected at the 95% confidence level ($p<0.05$).

Principal component analysis (PCA) and SIMCA

A data matrix can be defined such that the number of rows corresponds to the total number of subjects times the number of times each subject was tested (NP = NS x NT) and the number of columns corresponds to the total number of variables (NV).

$$\mathbf{X} = \begin{array}{c} \overset{\longleftarrow \text{NV} \longrightarrow}{} \\ \left\uparrow \atop \text{NP} \atop \left\downarrow \right. \right. \end{array} \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,21} \\ x_{2,1} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ x_{n,1} & \cdots & \cdots & \cdots \end{bmatrix} \quad (2)$$

There were 19 variables corresponding to the 19 carbonyl compounds identified in the GC/MS and two additional variables corresponding to the measured levels of 8-OHdG and 8-epi $F_{2\alpha}$, for a total of 21 variables. Separate data matrices were defined for the smoking and nonsmoking groups. Data matricies were mean-centered and autoscaled to unit variance[51].

The scores are the projections of the data onto

Full Paper

the principal components and are defined as

$$\mathbf{T} = \mathbf{XP} \quad (3)$$

Since the minimum number of principal components are usually retained,

$$\mathbf{X} = \mathbf{TP}^{\mathsf{T}} + \mathbf{E} \quad (4)$$

where E is the error matrix. The Q statistic is the sum of squares of the row elements of E and represents the residual between a sample and its projection onto the principal components retained in the model. The Q statistic is a measure of how well each data point fits the model. Data points that were not able to fit the model for the nonsmokers or smokers were eliminated if they yielded a Q statistic that fell outside the 95% confidence level. After eliminating the first round of outliers the data was once again modeled and the process of elimination was once again repeated. Using this iterative procedure, 78 nonsmoker and 72 smoker data points were retained.

Of the 21 variables used in defining the data, not all are essential for modeling a specific class or for separating the class of smokers from that of nonsmokers. The modeling power was defined as a variable's ability to define a class and was evaluated according to

$$MP(k) = 1 - \frac{s_k^q(error)}{s_k^q(x)} \quad (5)$$

where the residual variance of variable k in class q, $s_k^q(error)^2$, and the meaningful variance of variable k in class q, $s_k^q(x)^2$, are defined as

$$s_k^q(error)^2 = \sum_{i=1}^{NP} \frac{e_{i,k}{}^2}{(NP - NC - 1)} \quad (6)$$

$$s_k^q(x)^2 = \sum_{i=1}^{NP} \frac{(x_{i,k} - \bar{x}_k)^2}{(NP - 1)} \quad (7)$$

and $e_{i,k}$ are the elements of $\mathbf{E}$, $x_{i,k}$ are elements of the data matrix $\mathbf{X}$ and $\bar{x}_k$, the mean value of points defined by variable k. The discriminatory power of a variable, *DP(k)*, corresponds to the ability of the variable to separate the smokers class from the nonsmokers class and is defined as by

$$DP(k) = \left[\frac{[s_k^{s,ns}(error)]^2 + [s_k^{ns,s}(error)]^2}{[s_k^s(error)]^2 + [s_k^{ns}(error)]^2} \right]^{1/2} - 1 \quad (8)$$

where $s_k^{s,ns}(error)^2$ the residual variance of variable k when fitting the nonsmoking data points to the smoking model, and $s_k^{ns,s}(error)^2$ the residual variance of variable k when fitting the smoking data points to the nonsmoking model. Data sets having reduced variables were constructed by eliminating points with the lowest discriminatory power.

Data were plotted in two dimensions by plotting their scores on the two PCs having the largest eigenvalues (scores plot) or by plotting the Q statistic of each data point (residual outside the model) against Hotelling's T^2 statistic (residual within the model), where

$$\mathbf{T}_i^2 = \mathbf{t}_i \boldsymbol{\lambda}^{-1} \mathbf{t}_i^{\mathsf{T}} \quad (9)$$

and $\mathbf{t}_i$ is the score vector of the i^{th} data point and $\boldsymbol{\lambda}^{-1}$ is the inverse of the associated eigenvalue (residuals plot).

As the test to the validity of the model, test points were classified according to K-Nearest Neighbour (KNN) or SIMCA[51]. Using the KNN model, all the points in the data set serve as test points and are classified according to the class of their nearest neighbour in Euclidean space (1KNN) or their 3 nearest neighbours (3KNN), where the majority of neighbours of the same class determines the category. For SIMCA modeling, the test set consisted of all the points used in modeling or, alternatively, of 10 randomly selected points withheld from the modeling set on four consecutive trials (4 x 10 out).

All pattern recognition calculations were carried with MATLAB 6.0 (Mathworks, Natick, MA) and the PLS Toolbox (Eigenvector Research, Manson, WA).

RESULTS

Carbonyl Compounds

Nonpolar carbonyl compounds were determined by GC-MS of their PFB-oximes. Figure 1A shows a typical chromatogram of a urine extract acquired using selected ion monitoring. Peaks arising from the *syn* and *anti* oxime isomers are apparent for many of the shorter chain carbonyls (except for MDA, which

Full Paper

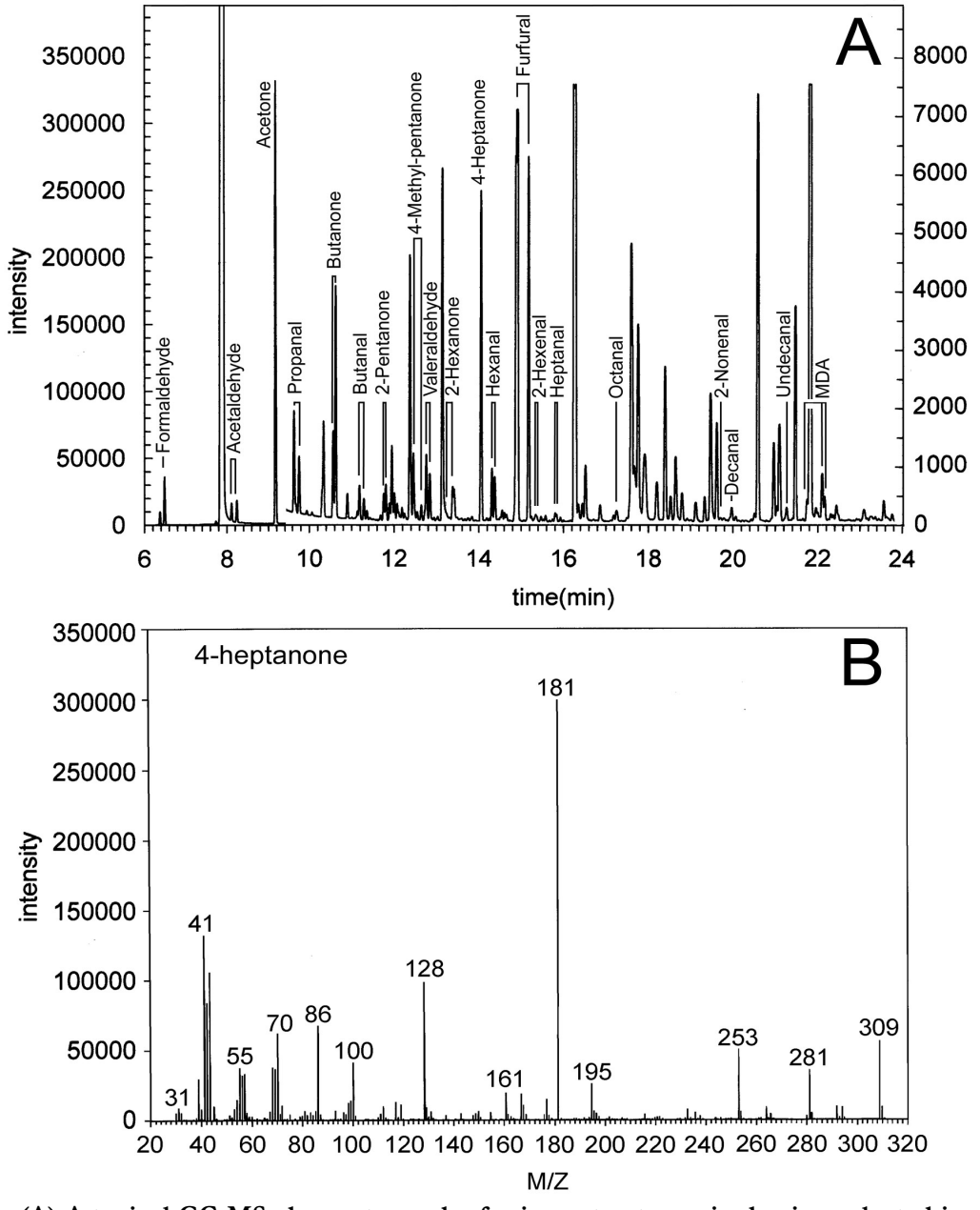

Figure 1: (A) A typical GC-MS chromatograph of urine extract acquired using selected ion monitoring of the parent m/z 181 ion peak, corresponding to pentafluorotropylium ion ($C_6F_5CH_2^+$) [49]. The scale on the left is valid for peaks arising between 0 and 9.5 min, while the scale on the right is for the remaining peaks. The large peak at 7.8 min arises from unreacted PFBHA. (B) The mass spectrum of 4-heptanone, a previously unidentified ketone (retention time of 14.07 min) in human adult urine.

had 3 identifiable isomers). In all, peaks from 19 carbonyl compounds could be identified by comparing their retention times and mass spectrum to those of standards. A summary of retention times and molecular ion peaks is presented in TABLE 1.

A plot of the carbonyl compound concentrations determined in the urine of nonsmokers and smokers is shown in figure 2A. Each of the concentrations shown represents an average of the data collected over a 13-week period for 88 nonsmoker and 72 smoker data sets. Acetone and MDA were the most abundant nonpolar carbonyl compounds found in

Full Paper

TABLE 1: GC-MS retention times and molecular ion data

Carbonyl compound	RT_1[a]	RT_2[a]	RT_3[a]	M^+(oxime)
formaldehyde	6.56			225
acetaldehyde	8.20	8.32		239
acetone	9.26			253
propanal	9.66	9.78		253
butanone	10.59	10.65		267
butanal	11.21	11.31		267
2-pentanone	11.83	11.97		281
4-Me-pentanone	12.48	12.66		292
valeraldehyde	12.78	12.86		281
2-hexanone	13.20	13.40		295
4-heptanone	14.07			309
hexanal	14.33	14.39		295
furfural[b]	14.93	15.19		291
2-hexenal	15.33	15.38		293
heptanal	15.82			309
octanal	17.26			323
2-nonenal	19.50	19.70		335
decanal	19.99			351
undecanal	21.26			365
MDA	21.75	22.11	22.17	430

[a]Retention times for both *syn* and *anti* oxime isomers (except for MDA, which has 3 identifiable isomers).
[b]Furfural was added as an internal standard to urine samples.

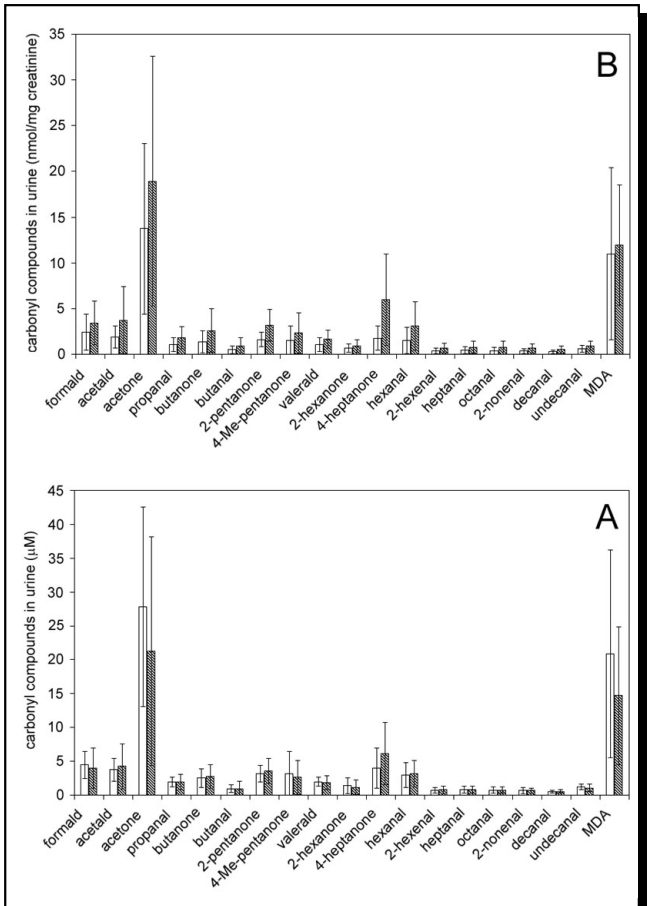

Figure 2: Concentrations of carbonyl compounds in first morning urine of healthy smoking (shaded) and nonsmoking (clear) adults. (A) Plot of concentrations (μM). (B) Ratio of carbonyl:creatinine (nmol/mg creatinine).

urine (14 – 30 μM), and comprise 33.3% and 25.0% of the total carbonyl compounds measured in the nonsmoking group and 28.2% and 19.5% of in the smoking group. These two compounds were found to be significantly higher in the nonsmoking group than in smoking group ($p<0.05$). Other carbonyl compounds appear at 4-10 times lower abundance compared to acetone and MDA (1-5 μM). 4-heptanone, a ketone not previously identified in urine, was one of the few carbonyl compounds that was found to be more abundant in smokers than in nonsmokers. Others included acetaldehyde, butanone, 2-pentanone, hexanal and 2-hexenal. The concentrations of carbonyl compounds normalized to creatinine are shown in figure 2B. All creatinine-normalized concentrations, except those for MDA, were found to be significantly higher in the urine of smokers compared to nonsmokers. This result is surprising since MDA has long been used as a biomarker for oxidative stress[16, 19, 26, 32, 33, 40, 52]. In addition to higher levels of total carbonyl compounds:creatinine (TCC:creatinine), the smoking group was found to have higher levels of 8-epi $F_{2\alpha}$:creatinine and 8-OHdG:creatinine than the nonsmoking group, in agreement with previous work where urinary 8-OHdG or 8-epi $F_{2\alpha}$ were used as biomarkers[6, 7, 8, 9, 12-15].

Principal component analysis (PCA)

Using PCA we modeled the smoking and nonsmoking groups, using as a basis set the 19 carbonyl compounds, 8-OHdG and 8-epi $F_{2\alpha}$ normalized to creatinine. Using this method, each of the principal components (PCs) used in modeling are defined by of a combination of the original 21 biomarkers and point in the direction of greatest variance of the data. Between 65-70% of the variance for either the smok-

Full Paper

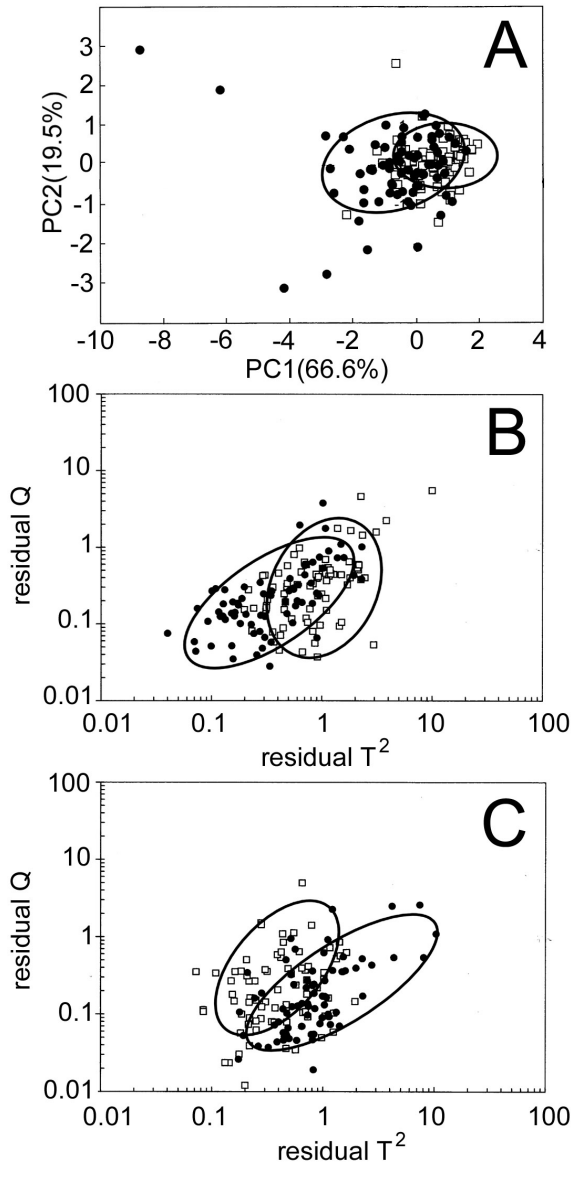

Figure 3: Principal component plots.. Data are shown separately for the smoking (filled circles) and nonsmoking (open circles) group. (A) Scores plot using the two most significant PCs. (B) Residuals plot, where within model residual error (Hotelling's T^2) is plotted against the residual error outside the model (Q statistic). Residuals were calculated using the model for the smoking group. (C) Residuals plot using the model for the nonsmoking group.

ing or the nonsmoking group was captured using either 2 or 3 PCs. Figure 3A shows the combined data in a coordinate system defined by the two most significant PCs. The majority of the nonsmoking group data lies clustered to the right of the origin while the smoking group data appears scattered to the left of the origin and partially overlapping the nonsmoking group data. A more pronounced separation between the two groups is seen when the residual of each data point within the model (Hotelling's T^2) is plotted versus the residual outside the model (Q statistic) (Figure 3B and 3C).

The first two columns of TABLE 2 summarize those variables most important for modeling, where each variable is ranked according to its modeling power. Variables having high modeling power are those with high within group variability. Although the most important variables differ depending on the group considered, it is noteworthy that 8-OHdG and 8-epi $F_{2\alpha}$ rank among the least significant variables defining within group variability. Carbonyl compounds in relatively high concentrations, such as acetone, MDA and 4-heptanone also have low modeling power. This indicates that these variables are reasonably homogeneous within the group.

The ability of a particular variable to separate the smoking from the nonsmoking group is reflected in its discriminatory power (DP), listed in decreasing order in the last column of TABLE 2. We define variables as strongly, moderately and weakly discriminatory based on their DP ranking (7 variables in each group). Variables most significant in separating the smoking from the nonsmoking group were found to be hexanal, octanal, heptanal, 4-methyl-pentanone, butanal, 8-epi $F_{2\alpha}$ and 2-hexenal. MDA, a commonly used oxidative stress biomarker, was found to have poor ability to separate the smoking from the nonsmoking group.

The SIMCA method was used to check the validity of the PC modeling approach. The results for SIMCA classification are summarized in TABLE 3 where they are compared to those of a simple K-Nearest Neighbour (KNN) classification. For SIMCA, modeling with two or three PCs produced nearly equivalent results, with approximately 90% of the data correctly classified using as test points the full data set and 82.5% correctly classified using data originally excluded from the model (4 x 10 out). SIMCA did significantly better than the one nearest neighbour (1KNN, 78.7%) or three nearest neighbour

Full Paper

TABLE 2: Modeling power and disciminatory power of variables[a]

Modeling Power Nonsmokers		Modeling Power Smokers		Discriminatory Power	
heptanal	0.709	formaldehyde	0.724	hexanal	2.460
hexanal	0.666	undecanal	0.694	octanal	2.000
formaldehyde	0.653	decanal	0.667	heptanal	1.637
octanal	0.634	2-hexenal	0.659	4-Me-pentanone	1.370
undecanal	0.630	2-pentanone	0.630	butanal	1.239
butanone	0.623	valeraldehyde	0.626	8-epi $F_{2\alpha}$	1.036
butanal	0.608	hexanal	0.593	2-hexenal	1.034
2-hexenal	0.592	heptanal	0.578	acetone	0.804
propanal	0.587	4-Me-pentanone	0.578	propanal	0.753
decanal	0.580	propanal	0.567	8-OHdG	0.721
acetaldehyde	0.575	acetone	0.556	2-hexanone	0.659
valeraldehyde	0.565	octanal	0.535	decanal	0.626
2-pentanone	0.506	2-hexanone	0.483	acetaldehyde	0.464
4-Me-pentanone	0.497	butanal	0.482	4-heptanone	0.269
acetone	0.395	2-nonenal	0.447	MDA	0.162
2-hexanone	0.393	acetaldehyde	0.442	formaldehyde	0.146
2-nonenal	0.353	8-epi $F_{2\alpha}$	0.366	undecanal	0.127
8-epi $F_{2\alpha}$	0.331	8-OHdG	0.321	butanone	0.115
4-heptanone	0.315	butanone	0.316	2-pentanone	0.079
8-OHdG	0.309	MDA	0.313	valeraldehyde	0.050
MDA	0.268	4-heptanone	0.175	2-nonenal	0.004

[a]**Modeling power and discriminatory power calculated according to eqns. (7) and (10) in experimental section. Horizontal lines in column 3 denote biomarkers having strong, moderate and weak DP.**

TABLE 3: Summary of pattern recognition results

Model	Points	PC Model[a]	Variables	Test Set	%Correct
SIMCA-21	150 (78N+72S)	2N x 2S	21	full set	89.4
SIMCA-21	150 (78N+72S)	3N x 3S	21	full set	90.6
SIMCA-21	140 (73N+67S)	2N x 2S	21	4 x 10 out	82.5
SIMCA-21	140 (73N+67S)	3N x 3S	21	4 x 10 out	82.5
1KNN-21			21	KNN	78.7
3KNN-21			21	KNN	76.7
SIMCA-13	150 (78N+72S)	3N x 3S	13	full set	88.0
SIMCA-13	140 (73N+67S)	3N x 3S	13	4 x 10 out	82.5
SIMCA-10	150 (78N+72S)	3N x 3S	10	full set	85.3
SIMCA-10	140 (73N+67S)	3N x 3S	10	4 x 10 out	80.0
SIMCA-3	150 (78N+72S)	2N x 2S	3	full set	67.3
SIMCA-3	140 (73N+67S)	2N x 2S	3	4 x 10 out	72.5

[a]**The number of PCs used to model the non-smoking and smoking groups**

(3KNN, 76.7%) classification models. Modeling with 3 PCs, the SIMCA results were only slightly decreased when variables having low DP were eliminated from the basis set (13 or 10 variable models). The results were poor when only three variables frequently used to define oxidative stress, TCC:creatinine, 8-epi $F_{2\alpha}$:creatinine and 8-OHdG:creatinine, were used. These results suggest that variables having strong to

moderate DPs are important in discriminating between the smoking and nonsmoking groups.

DISCUSSION

Concentrations of urinary nonpolar carbonyl compounds found using the PFB-oxime GC-MS method are in the same concentration range as those previously determined for healthy adults[16, 17, 19]. The concentrations of hydroxyaldehydic lipid peroxidation products, such as 4- hydroxy-nonenal, were not detected using our procedure due to their high polarity and low volatility. Typically, polar aldehydes are detected as their trimethylsilyl esters[53]. Of the nonpolar carbonyls, acetone was found to be the most abundant carbonyl compound in urine (24±14 μM), followed by MDA (18±13 μM), 4-heptanone (5±4 μM), formaldehyde (4±2 μM) and acetaldehyde (4±2 μM). Using a pentafluorophenyl hydrazine GC-MS method, Stashenko et al.[17] identified 15 nonpolar carbonyls, with the highest concentrations being those of acetone (38±34 μM), hexanal (16±11 μM), propanal (12±8 μM) and acetaldehyde (13±19 μM). Far fewer carbonyl compounds could be detected using HPLC methods, with relative concentrations 2-pentanone>hexanal>butanal>2-butanone> acetone>2-hexenal in nonfasted human subjects[16,19].

Roughly 40% of aldehydes are conjugated to proteins in plasma[54] and amino acid-MDA adducts are known to be present in urine[19, 37]. During initial trials we found that only a fraction of the urinary carbonyls could be detected as PFB-oximes using reaction times in acidic medium shorter than 4 hours. Others have similarly found that oxime formation at ambient temperature requires a minimum of 2 h for the derivatization of simpler aldehydes and as long as 24 h for the reaction of longer chain carbonyls[55]. It is possible that differences between our results and those previously reported arise from differences in reaction times for the hydrolysis of the Schiff base and subsequent derivatization.

In laboratory animals the concentrations of urinary carbonyl compounds, particularly MDA, appear to correlate with the formation of lipid peroxides brought about by exposure to toxins[16, 18, 19, 32, 33], exercise[26], aging[56], and vitamin E deficiency[19, 40, 52]. In human subjects, smoking cigarettes has been shown to increase plasma MDA[6, 38, 57] and MDA excreted in the urine has been correlated with plasma MDA[58]. Hence, it was quite surprising that in the absence of normalization to creatinine, the concentrations of 13 of the 19 carbonyl compounds identified in the urine of nonsmokers were higher than those of smokers. This suggests that non-normalized aldehyde concentrations in urine may not by themselves be used as reliable biomarkers of oxidative stress produced by smoking.

The concentrations of carbonyls normalized to creatinine in the present study (1 - 25 nmol/mg) fall within the range previously reported by others[18, 34, 59]. In human subjects, the urinary MDA:creatinine ratio, determined by a relatively simple colorimetric assay with thiobarbituric acid (TBARS assay), has been shown to reflect changes in oxidative stress status[21-25, 34, 60, 61]. However, it has been established that the TBARS assay overestimates the MDA levels by over 10-fold, possibly resulting from cross-reactivity with other aldehydes[62]. In light of these findings, it is unclear if increases in the TBARS accurately reflect increases in MDA or other carbonyl compounds. In comparing the carbonyl:creatinine levels of smokers to nonsmokers, we found that the smokers had significantly higher levels of all carbonyls except for MDA.

In addition to carbonyl compounds, 8-epi $F_{2\alpha}$ and 8-OHdG in plasma and urine have been used widely as biomarkers of oxidative stress. When creatinine normalized data were taken, it was found that both markers were significantly higher in the urine of smokers compared to nonsmokers, a result consistent with previous studies[6-10,12-15]. We found that while the variance of our data with respect to these two markers was slightly higher using the immunoassay method, the concentrations found were within the range of previously reported values (1.7±0.6 and 2.5±1.1 ng/mg creatinine for 8-OHdG and 281±252 and 506±446 pg/mg creatinine for 8-epi $F_{2\alpha}$ in nonsmokers and smokers, respectively).

We used PCA to determine which of the 21 oxidative stress biomarkers measured were most sensitive. Initial PC plots showed that the smoking group could be differentiated from the nonsmoking group

Full Paper

(Figure 3). Our results suggest that some of the longer chain carbonyls, which occur in relatively low concentrations, such as hexanal, octanal and heptanal, have a greater weight in separating smokers from nonsmokers. Despite its high variance, 8-epi $F_{2\alpha}$ is among the most strongly discriminatory variables while 8OHdG is moderately discriminatory. Urinary MDA, which occurs at a relatively high concentration and is frequently used as oxidative stress biomarker, is calculated to be one of the least most discriminatory variables for distinguishing between smokers and nonsmokers. The SIMCA results suggest that classification of unknown data to a model based on the 10 most discriminatory variables is within error of classification using all 21 variables. However, the much poorer results using only creatinine normalized TCC, 8OHdG and 8-epi $F_{2\alpha}$ as a basis suggests that other strongly and moderately discriminatory variables are important in discriminating smokers from nonsmokers.

One of the most significant aspects of the present work was to reveal the complex manner in which the biomarkers complement one another to provide an integrated picture of oxidative stress. While one set of biomarkers is most appropriate for separating smokers from nonsmokers, quite a different set may be needed for revealing the effects of different vitamins or dietary supplements. For example, concentrations of carbonyls in urine, in the absence of normalization to creatinine concentration, were found to be a quite poor set of features for distinguishing between smokers and nonsmokers. However, non-normalized carbonyl concentrations have been used successfully to determine the effects of toxins in laboratory animals[16, 19]. Hence, the significance of a biological effect may depend on the selection of biomarkers used.

ACKNOWLEDGEMENT

The authors thank Mannatech, Inc. (Coppell, TX) for their financial support.

REFERENCES

[1] H.Esterbauer, R.J.Schaur, H.Zollner; Free Radic.Biol. Med., **11**, 81-128 **(1991)**.
[2] E.N.Frankel; J.Am.Oil Chem.Soc., **61**, 1908-1916 **(1984)**.
[3] H.Esterbauer, K.H.Cheeseman, M.U.Dianzani, G.Poli, T.F.Slater; Biochem.J., **208**, 129-140 **(1982)**.
[4] B.Halliwell, J.M.C.Gutteridge; 'Free Radicals in Biology and Medicine', Clarendon Press, Oxford **(1985)**.
[5] L.Chen, P.E.Bowen, D.Berzy, F.Aryee, M.Stacewicz-Sapuntzakis, R.E.Riley; Free Radic.Biol.Med., **26**, 695-703 **(1999)**.
[6] J.E.Klaunig, X.Yong, C.Han, L.M.Kamendulis, J. Chen, C.Heiser, M.S.Gordon, E.R.Moher; Proc.Soc. Exp.Biol.Med., **220**, 249-254 **(1999)**.
[7] J.Suzuki, I.Yuichiro, S.Suzuki; Free Radic.Biol.Med., **18**, 431-436 **(1995)**.
[8] A.B.Nia, F.J.Van Schooten, P.A.E.L.Schilderman, T.M. C.M.De Kok, G.R.Haenen, M.H.M.Van Herwijnen, E.Van Agen, D.Pachen, J.C.S.A.Kleinjans; Carcinogenesis, **21**, 395-401 **(2001)**.
[9] M.Lodovici, C.Casalini, R.Cariaggi, L.Michelucci, P. Dolara; Free Radic.Biol.Med., **28**, 13-17 **(2000)**.
[10] T.Renner, T.Fechner, G.Scherer; J.Chromatogr.B, **738**, 311-317 **(2000)**.
[11] C.Y.Lee, J.M.F.Wan; J.Nutr., **130**, 2932-2937 **(2000)**.
[12] M.Reilly, N.Delanty, J.A.Lawson, G.A.FitzGerald; Circulation, **94**, 19-25 **(1996)**.
[13] R.Pawlosky, J.Hibbeln, B.Wegher, N.Sebring, N.Salem, Jr.; Lipids, **34**, S287 **(1999)**.
[14] H.Pilz, A.Oguogho, F.Chehne, G.Lupattelli, B. Palumbo, H.Sinzinger; Thromb.Res., **99**, 209-221 **(2000)**.
[15] T.Obata, K.Tomaru, T.Nagakura, Y.Izumi, T. Kawamoto; J.Chromatogr.B, **746**, 11-15 **(2000)**.
[16] S.S.Kim, D.D.Gallaher, A.S.Csallany; Lipids, **34**, 489-496 **(1999)**.
[17] E.E.Stashenko, M.C.Ferreira, L.G.Sequeda, J.R. Martinez, J.W.Wong; J.Chromatogr.A, **779**, 360-369 **(1997)**.
[18] L.L.De Zwart, R.C.A.Hermanns, J.H.N.Meerman, J.N. M.Commandeur, P.J.M.Salemink, N.P.E.Vermeulen; Toxicol.Appl.Pharmacol., **148**, 71-82 **(1998)**.
[19] H.H.Draper, A.S.Csallany, M.Hadley; Free Radic.Biol. Med., **29**, 1071-1077 **(2000)**.
[20] R.Van den Berg, T.Van Vliet, W.M.R.Broekmans, N.H. P.Cnubben, WH.J.Vaes, L.Roza, G.R.M.M.Haenen, A.Bast, H.Van den Berg; J.Nutr., **131**, 1714-1722 **(2001)**.
[21] J.Hallfrisch, K.M.Scholfield, K.M.Behall; J.Nutr. Biochem., **8**, 497-501 **(1997)**.
[22] E.D.Brown, V.C.Morris, D.G.Rhodes, R.Sinha, O.A.

Levander; Lipids, **30**, 1053-1056 **(1995)**.
[23] G.J.Nelson, V.C.Morris, P.C.Schmidt, O.Levander; Lipids, **28**, 757-761 **(1993)**.
[24] V.Rezar, T.Pajk, L.R.Marinsek, J.V.Jese, K.Salobir, A. Oresnik; J.Salobir; Ann.Nutr.Metab., **47**, 78-84 **(2003)**.
[25] S.Demir, G.Turgut, O.Yurtseven, D.Aslan, O.Genc; Acta Medica (Hradec Kralove), **44**, 41-42 **(2001)**.
[26] P.C.Mills, J.C.Ng, J.Thornton, A.A.Seawright, D.E. Auer; Brit.Vet.J., **150**, 53-63 **(1994)**.
[27] A.B.Bridges, N.A.Scott, G.J.Parry, J.J.F.Belch; Eur.J. Med., **2**, 205-208 **(1993)**.
[28] L.A.Mooney, F.P.Perera, A.M.Van Bennekum, W.S. Blaner, J.Karkoszka, L.Covey, Y.Hsu, T.B.Cooper, K. Frenkel; Cancer Epidemiol.Biomarkers & Prev., **10**, 641-648 **(2001)**.
[29] K.M.Brown, P.C.Morrice, G.G.Duthie; Am.J.Clin. Nutr., **60**, 383-387 **(1994)**.
[30] E.A.Meagher, O.P.Barry, J.A.Lawson, J.Rokach, G.A. FitzGerald; JAMA, **285**, 1178-1182 **(2001)**.
[31] M.A.Shara, P.H.Dickson, D.Bagchi, S.J.Stohs; J. Chromatogr.B Biomed.Appl., **576**, 221-233 **(1992)**.
[32] D.Bagchi, M.Bagchi, E.Hassoun, S.J.Stohs; Toxicol., **75**, 81-89 **(1992)**.
[33] M.Bagchi, D.Bagchi, E.A.Hassoun, S.J.Stohs; Int.J. Exp.Path., **75**, 197-202 **(1994)**.
[34] M.Yazdanpanah, X.Luo, R.Lau, M.Greenberg, V. Fisher, D.C.Lehotay; Free Radic.Biol.Med., **23**, 870-878 **(1997)**.
[35] M.V.Novotny, M.F.Yancey, R.Stuart, D.Wiesler, R.G. Peterson; Biochim.Biophys.Acta, **1226**, 145-150 **(1994)**.
[36] A.W.Dohlman, H.R.Black, J.A.Royall; Am.Rev.Respir. Dis., **148**, 955-960 **(1993)**.
[37] H.Mahmoodi, M.Hadley, Y.X.Chang, H.H.Draper; Lipids, **30**, 963-966 **(1995)**.
[38] E.Hoshino, R.Shariff, A.Van Gossum, J.P.Allard, C. Pichard, R.Kurian, K.N.Jeejeebhoy; J.Parenteral Enteral Nutr., **14**, 300-305 **(1990)**.
[39] E.A.Dratz, C.C.Farnsworth, E.C.Loew, R.J.Stephens, D.W.Thomas, F.J.G.M.Van Kuijk; Ann.NY Acad.Sci., **570**, 46-60 **(1989)**.
[40] H.S.Lee, D.W.Shoeman, A.S.Csallany; Lipids, **27**, 124-128 **(1992)**.
[41] M.K.Horwitt; Vita.Hormon., **20**, 541-558 **(1962)**.
[42] P.Patrignani, M.R.Panara, S.Tacconelli, F.Seta, T. Bucciarelli, G.Ciabattoni, P.Alessandrini, A.Mezzetti, G.Santini, M.G.Scuilli, F.Cipollone, G.Davi, P.Gallina, G.B.Bon, C.Patrono; Circulation, **102**, 539-545 **(2000)**.
[43] M.E.Keith, K.N.Jeejeebhoy, A.Langer, R.Kurian, A. Barr, B.O'Kelly, M.J.Sole; Am.J.Clin.Nutr., **73**, 219-224 **(2001)**.
[44] M.P.Habib, L.J.Tank, L.C.Lane, H.S.Garewal; Chest, **115**, 684-690 **(1999)**.
[45] R.B.Weinberg, B.S.VanderWerken, R.A.Anderson, J.E. Stegner, M.J.Thomas; Arterioscler.Thromb.Vasc.Biol., **21**, 1029-1033 **(2001)**.
[46] F.Visioli, D.Caruso, C.Galli, S.Viappiani, G.Galli, A. Sala; Biochem.Biophys.Res.Commun., **278**, 797-799 **(2000)**.
[47] M.E.Tornwall, J.Virtamo, J.K.Haukka, A.Aro, D. Albanes, B.K.Edwards, J.K.Huttunen; Arterioscler. Thromb.Vasc.Biol., **17**, 3475-3480 **(1997)**.
[48] X.P.Luo, M.Yazdanpanah, N.Bhooi, D.C.Lehotay; Anal.Biochem., **228**, 294-298 **(1995)**.
[49] G.F.Hoffmann, L.Sweetman; Clin.Chim.Acta, **199**, 237-242 **(1991)**.
[50] V.Barnett, T.Lewis; 'Outliers in Statistical Data', 3rd ed., John Wiley & Sons, New York, **(1994)**.
[51] M.A.Sharaf, D.L.Illman, B.R.Kowalski; 'Chemometrics Chemical Analysis', John Wiley & Sons, New York, **82**, **(1986)**.
[52] H.H.Draper, L.Polensek, M.Hadley, L.G.McGirr; Lipids, **19**, 836-843 **(1984)**.
[53] A.Loidl-Stahlhofen, W.Kern, G.Spiteller; J. Chromatogr.B, **673**, 1-14 **(1995)**.
[54] G.Quash, K.Maharaj; Clin.Chem.Acta, **29**, 13-16 **(1970)**.
[55] S.M.Breckenridge, X.Yin, J.M.Rosenfeld, Y.H.Yu; J. Chromatogr.B, **694**, 289-296 **(1997)**.
[56] V.De Tata, G.Lorenzini, L.Cecchi, C.Ciuffi, E. Bergamini; Exp.Gerontol., **36**, 507-518 **(2001)**.
[57] F.Nielsen, B.B.Mikkelsen, J.B.Nielsen, H.R.Andersen, P.Grandjean; Clin.Chem., **43**, 1209-1214 **(1997)**.
[58] J.A.Drury, J.A.Nycyk, R.W.I.Cooke; Clin.Chim.Acta, **263**, 177-185 **(1997)**.
[59] C.Feillet-Coudray, F.Chone, F.Michel, E.Rock, P. Thieblot, Y.Rayssiguier, I.Tauveron, A.Mazur; Clin. Chim.Acta, **324**, 25-30 **(2002)**.
[60] A.Rajiv; Am.J.Physiol.Renal Physiol., **284**, F863-F869 **(2003)**.
[61] C.Tagesson, M.Kaelberg, G.Wingren; Int.Arch. Occup.Environ.Health, **69**, 5-13 **(1996)**.
[62] H.C.Yeo, H.J.Helbock, D.Chyu, B.N.Ames; Anal. Biochem., **220**, 391-396 **(1994)**.