Minimal Invasive Gender Determination of Birds by Means of UV-Resonance Raman Spectroscopy

ARTICLE *in* ANALYTICAL CHEMISTRY · MARCH 2008

Impact Factor: 5.64 · DOI: 10.1021/ac702043q · Source: PubMed

6 AUTHORS, INCLUDING:

Thomas Bartels
University of Leipzig
61 PUBLICATIONS **245** CITATIONS
SEE PROFILE

Kerstin Cramer
University of Leipzig
9 PUBLICATIONS **59** CITATIONS
SEE PROFILE

Juergen Popp
Friedrich Schiller University Jena
409 PUBLICATIONS **6,089** CITATIONS
SEE PROFILE

Anal. Chem. **2008,** *80,* 1080−1086

Minimal Invasive Gender Determination of Birds by Means of UV-Resonance Raman Spectroscopy

M. Harz,[†] M. Krause,[†] T. Bartels,[‡] K. Cramer,[‡] P. Rösch,[†] and J. Popp*[,†,§]

Institut für Physikalische Chemie, Friedrich-Schiller-Universität Jena, Helmholtzweg 4, 07743 Jena, Germany, Klinik für Vögel und Reptilien, Universität Leipzig, An den Tierkliniken 17, 04103 Leipzig, Germany, and Institut für Photonische Technologien, Albert-Einstein-Strasse 9, 07745 Jena, Germany

The identification of avian gender is important for prosperous breeding of birds. Since birds do not possess external genital organs, endoscopic investigations, blood analysis, and molecular biological methods are applied to determine the gender in monomorphic species. However, anesthesia and blood sampling impose stress on the examined bird and should be avoided in terms of animal protection. Here we report on the application of UV-resonance Raman spectroscopy as a minimal invasive method for gender determination of birds via an evaluation of feather pulp samples. Sample preparation for this investigation method is simple and facilitates a quick and easy analysis. The UV-resonance Raman spectra of the feather pulp sample extracts are dominated by DNA and protein signals. The different DNA content in male and female chicken allows for gender differentiation via its characteristic Raman fingerprint. The classification either to male or female chicken is ideally accomplished by support vector machines due to the fact that no unknown classes are involved. Recognition rates of about 95% were compared to less effective results of the unsupervised hierarchical cluster analysis. Within the scope of our investigations, principal component analysis was also applied to determine the important spectral regions for the classification of chicken's feather pulp samples.

The knowledge of a bird's gender is vital to bird keepers and aviculturists for the establishment of breeding pairs or breeding flocks. Furthermore, the poultry industry has an enormous interest in identifying the gender of poultry at a very early stage of body development. Since, for example, the feeding conditions are different for either gender, this has economic reasons.[1] In contrast to mammals, birds do not possess external genital organs. Thus, secondary sexual characteristics such as body size, plumage color, or other phenotypic sexual characteristics have to be used for a gender determination. If a distinct sexual dimorphism is missing, ethological parameters as mating behavior or phonation can provide further information for gender diagnosis. However, in numerous monomorphic bird species no reliable gender determination can be accomplished on the basis of external or behavioral characteristics. Furthermore one is also dependent on alternative methods for gender determination of immature or sexually inactive birds. Among endoscopic evaluations of the gonads of anesthetized birds[2−4] and investigations of the hormone status, molecular biological methods (e.g., DNA−PCR)[5−11] in particular provide reliable results in these cases. In the poultry industry, vent sexing is commonly used. Here a manual sorting of newly hatched male and female poults on the basis of differences in cloacal morphology is accomplished with a high accuracy of 95%.[12] Despite their good reliability, the described methods are time-consuming and stressful for the birds and should be avoided in terms of animal welfare. Furthermore, birds can be injured during blood sampling, and both sampling procedure and analysis presuppose trained specialists as well as special laboratory equipment. Last, anesthesia proposes an increased risk for the birds' health. Fatalities due to anesthetic accidents are extremely unacceptable, not least if the examined species is rare and expensive, such as certain parrots. Hence the use of alternative, less impacting methods would be preferable to preserve the bird's health.

A promising minimal invasive option is the application of Raman spectroscopy for the analysis of complex biological samples. Depending on the Raman laser excitation wavelength, a selective enhancement of different sample constituents can be achieved.[13−22] The application of deep UV excitation wavelengths

* Corresponding author. Phone: (+49)3641 948320. Fax: (+49)3641 948302. E-mail: juergen.popp@uni-jena.de.

[†] Friedrich-Schiller-Universität Jena.

[‡] Universität Leipzig.

[§] Institut für Photonische Technologien.

(1) Shalev, B. A.; Pasternak, H. *Poult. Sci.* **1998**, *77*, 859−863.
(2) Satterfield, W. C. *Vet. Clin. North Am.: Small Anim. Pract.* **1990**, *20*, 1353−1367.
(3) Korbel, R. *Tieraerztl. Prax.* **1993**, *21*, 41−46.
(4) Prus, S. E.; Schmutz, S. M. *Avian Dis.* **1987**, *31*, 420−424.
(5) Bermudez-Humaran, L. G.; Garcia-Garcia, A.; Leal-Garza, C. H.; Riojas-Valdes, V. M.; Jaramillo-Rangel, G.; Montes-De-Oca-Luna, R. *J. Exp. Zool.* **2002**, *292*, 677−680.
(6) Huynen, L.; Millar, C. D.; Lambert, D. M. *Mol. Ecol.* **2002**, *11*, 851−856.
(7) Clinton, M.; Haines, L.; Belloir, B.; McBride, D. *Br. Poult. Sci.* **2001**, *42*, 134−138.
(8) Griffiths, R.; Double, M. C.; Orr, K.; Dawson, R. J. G. *Mol. Ecol.* **1998**, *7*, 1071−1075.
(9) Nomura, O.; Nishimori, K.; Nakabayashi, O.; Yasue, H.; Mizuno, S. *J. Steroid Biochem. Mol. Biol.* **1998**, *67*, 143−148.
(10) Gill, D. V.; Robertson, H. A.; Betz, T. W. *Gen. Comp. Endocrinol.* **1983**, *49*, 176−186.
(11) Tell, L. A.; Lasley, B. L. *Zoo Biol.* **1991**, *10*, 361−367.
(12) Bramwell, R. K. *Avian Advice* **2003**, *5*, 4−5.
(13) Frosch, T.; Schmitt, M.; Popp, J. *J. Phys. Chem. B* **2007**, *111*, 4171−4177.
(14) Schmitt, M.; Popp, J. *J. Raman Spectrosc.* **2006**, *37*, 20−28.

10.1021/ac702043q CCC: $40.75
Published on Web 01/16/2008

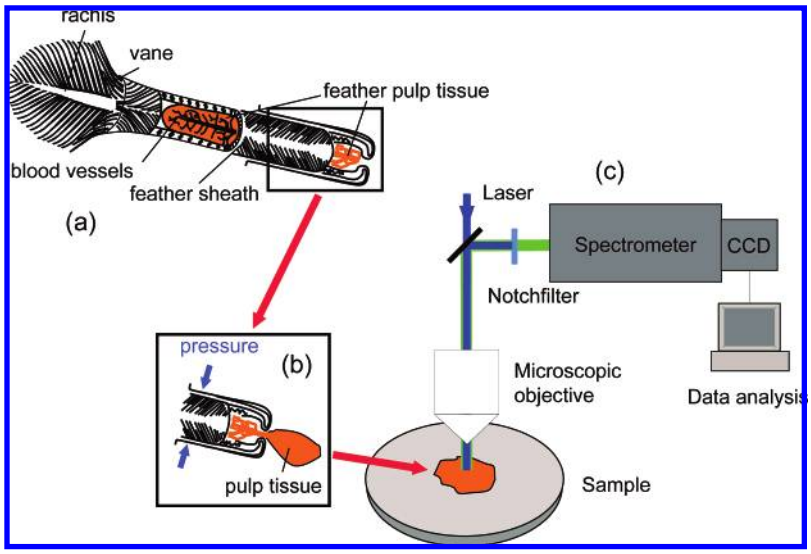

Figure 1. Principle scheme of minimal invasive gender determination of birds by using a feather (a): easy extraction of DNA-rich cell material of the feather pulp by pressing the feathers from tip to base (b) followed by direct recording of a Raman fingerprint spectrum and data analysis (c).

yields a selective enhancement of macromolecules such as aromatic amino acids in proteins as well as DNA bases. The scattering intensity increases up to 10^6 by the electronic resonance enhancement thus improving the signal-to-noise ratio significantly.[23,24] Hence, UV-resonance Raman (UVRR) spectroscopy with excitation in the deep UV region is a powerful approach to study samples with high protein and/or DNA content.

Here we report on the successful application of UVRR spectroscopy to determine the gender of birds. This approach is based on the fact that genome size is larger in male birds as compared to females. In chickens, the difference relative to the respective DNA content amounts to about 2%.[25] This difference should be detectable by means of UVRR spectroscopy since other sample components such as carbohydrates or lipids contribute negligible Raman signals in the deep UV region. After extraction of the DNA-rich cell material by pressing the thawed feathers from tip to base, no further extensive sample preparation is required.

EXPERIMENTAL SECTION

Sample Preparation. Feather samples were collected in a flock of 6-week-old layer chickens (Lohmann Brown) of both sexes. Growing contour feathers containing pulp tissue were manually picked from the skin and were separately stored in vials at −20 °C.

The extraction of cell material containing DNA was carried out by pressing the thawed feathers from tip to base and smearing the obtained liquid sample directly onto a fused-silica surface for drying.

Spectroscopic Instrumentation. UVRR spectra were recorded using a micro-Raman setup (HR800, Horiba Jobin-Yvon, Bensheim, Germany) with a focal length of 800 mm and a 2400 lines/mm grating. The wavenumber resolution of the HR800 spectrometer was 4 cm^{-1}. As excitation wavelength, a frequency-doubled argon-ion laser (Innova 300, MotoFReD, Coherent, Dieburg, Germany) at 244 nm was applied. The laser power on the sample was ca. 1 mW. The laser beam was focused onto the sample by a 40× broad-band antireflection coated UV objective (LMU UVB) with a numerical aperture of 0.5. The entrance slit was set to 300 μm. A video camera, which is sensitive in the UV and in the visible spectral range, was used for positioning the samples under the microscope. The Raman scattered light was detected by a nitrogen-cooled CCD camera with an accumulation time set to 60 s. The samples were rotated and moved in the x,y direction after each rotation to obtain an average spectrum over a large sample area in order to minimize possible photodegradation by UV radiation.

Figure 1 summarizes this quick and reliable approach for minimal invasive gender determination of birds: the procedure includes the easy extraction of DNA-rich cell material of the feather pulp (a) by pressing the feathers from tip to base (b) followed by the direct recording of a Raman fingerprint spectrum and data analysis by the application of chemometrical methods for gender classification (c).

Data Preprocessing and Classification. A total of 18 male and 16 female chicken samples were analyzed by recording about 10 Raman spectra of each sample. Overall, a number of 326 UVRR spectra were recorded.

(15) Gaus, K.; Rösch, P.; Petry, R.; Peschke, K. D.; Ronneberger, O.; Burkhardt, H.; Baumann, K.; Popp, J. *Biopolymers* **2006**, *82*, 286−290.
(16) Neugebauer, U.; Schmid, U.; Baumann, K.; Holzgrabe, U.; Ziebuhr, W.; Kozitskaya, S.; Kiefer, W.; Schmitt, M.; Popp, J. *Biopolymers* **2006**, *82*, 306−311.
(17) Harz, M.; Claus, R. A.; Bockmeyer, C. L.; Baum, M.; Rösch, P.; Kentouche, K.; Deigner, H. P.; Popp, J. *Biopolymers* **2006**, *82*, 317−324.
(18) Krause, M.; Radt, B.; Rösch, P.; Popp, J. *J. Raman Spectrosc.* **2007**, *38*, 369−372.
(19) Neugebauer, U.; Schmid, U.; Baumann, K.; Ziebuhr, W.; Kozitskaya, S.; Deckert, V.; Schmitt, M.; Popp, J. *ChemPhysChem* **2007**, *8*, 124−137.
(20) Rösch, P.; Harz, M.; Schmitt, M.; Popp, J. *J. Raman Spectrosc.* **2005**, *36*, 377−379.
(21) Rösch, P.; Harz, M.; Peschke, K.-D.; Ronneberger, O.; Burkhardt, H.; Schuele, A.; Schmauz, G.; Lankers, M.; Hofer, S.; Thiele, H.; Motzkus, H.-W.; Popp, J. *Anal. Chem.* **2006**, *78*, 2163−2170.
(22) Rösch, P.; Harz, M.; Peschke, K. D.; Ronneberger, O.; Burkhardt, H.; Popp, J. *Biopolymers* **2006**, *82*, 312−316.
(23) Carey, P. R. *J. Biol. Chem.* **1999**, *274*, 26625−26628.
(24) Turpin, P. Y.; Chinsky, L.; Laigle, A. *Stud. Phys. Theor. Chem.* **1987**, *45*, 369−385.
(25) Tiersch, T. R. *World's Poult. Sci. J.* **2003**, *59*, 24−31.

Prior to the chemometric analysis of the Raman spectra, the first derivative of the spectra (Savitzky−Golay, 13 smoothing points) was calculated to realize both a background correction and a signal-to-noise ratio enhancement. The final data preprocessing step consisted in a vector normalization of the spectra. Out of a large variety of classification methods, two methods, namely, the supervised nonlinear classification method support vector machine (SVM) and the unsupervised classification method hierarchical cluster analysis (HCA), are used to characterize the data set. These are well-established methods to classify biological and medical samples.[26,27] The preprocessed data set was employed by SVM using radial basis function kernels.[28] This algorithm allows for differentiating all background information from important signals. No restriction to a special wavenumber region was done since every small spectral difference might be of importance to achieve the best differentiation results. The results were obtained by performing leave-one-out tests and testing validation data.

The results were compared to the analysis of the HCA. Therefore, the program package "Opus Ident 3.1" (Bruker Optics, Ettlingen, Germany) was used accounting the factorization method and the Ward's algorithm for calculating the spectral distances.

Furthermore, the evaluation of essential wavenumber regions for classification could be performed by accomplishing a principal component analysis (PCA). Therefore, the software "The Unscrambler" (CAMO Process AS; version 9.2, Reutlingen, Germany) was applied. This software uses a cross validation to verify the results. Here a subsample set containing the original Raman spectra recorded from each sample was created for all 34 samples. The Raman bands of interest in the wavenumber range between 1140 and 1730 cm^{-1} were compared to the loadings receiving by PCA.

RESULTS AND DISCUSSION

The aim of the investigations presented in the following was a gender determination of chickens. The analysis of feather pulp containing mostly DNA is possible by means of UVRR spectroscopy in combination with chemometric data analysis.

Qualitative Signal Assignment. The prominent bands of the feather pulp UVRR spectra are assigned by comparing those with the UV Raman spectra of well-known constituents.

Figure 2 shows the average UVRR spectra of the extract from male (black) and female (gray) feather pulp samples of chickens. The spectra are baseline corrected and vector normalized. The two Raman spectra show slightly different intensity patterns. Therefore, the Raman difference spectrum by subtracting the female from the male UVRR spectrum was calculated identifying the most pronounced differences between female and male feather pulp UVRR spectra in the wavenumber region between 600 and 1800 cm^{-1}. The difference signals in positive direction correspond to higher band intensities in the averaged male feather pulp spectrum. The main differences appear around 1648, 1615, 1594, 1500, 1480, 1452, 1336, 1172, and 1007 cm^{-1}. Since DNA and

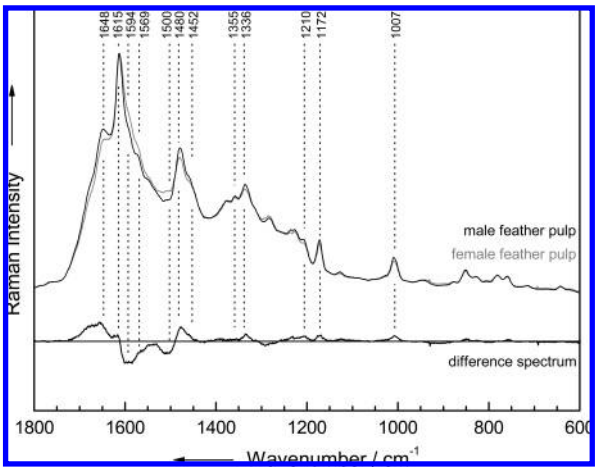

Figure 2. Averaged UVRR spectra of male and female chickens' feather pulp samples (baseline correction, vector normalization) and difference spectrum by subtracting the female from the male UVRR spectrum.

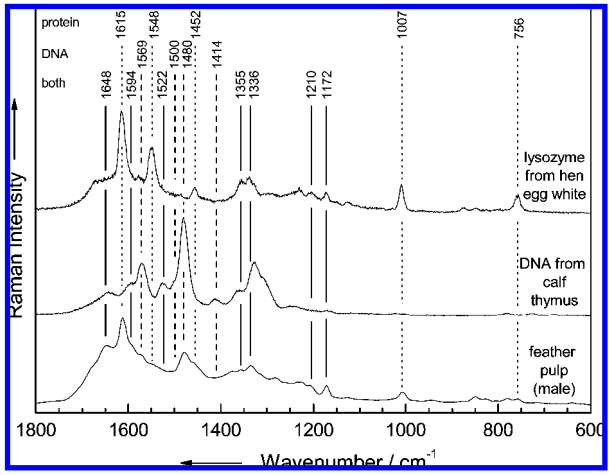

Figure 3. UVRR spectrum of a male chickens' feather pulp sample compared to a DNA spectrum of calf thymus and a protein spectrum of hen egg white lysozyme.

protein components are enhanced by UV excitation, the signal assignment can be focused to these components. Figure 3 compares an averaged UVRR spectrum of a male feather pulp sample with those of pure calf thymus DNA and the hen egg white lysozyme protein being characteristic for a chicken environment. The wavenumber labeling of the Raman bands in Figure 3 is ordered as follows: the protein bands are labeled on top and marked with a dotted line, the DNA bands are in the middle and marked with a dashed line and the superposition of DNA and protein information is found at the end of the legend with a continuous line. A complete band assignment is summarized in Table 1.

The Raman bands of the feather pulp at 1648, 1480, 1569, 1355, and 1172 cm^{-1} can be assigned to the DNA building blocks (nucleic acid bases, nucleosides, and nucleotides). The band at 1648 cm^{-1} can be identified as the CO and CC stretching vibration of the nucleic acid thymine. The two purine bases guanine and adenine contribute to the Raman band at 1480 cm^{-1} with its triene vibration and the CN stretching vibration, respectively. Adenine and guanine also contribute with a NH_2 deformation to the band at 1569 cm^{-1}. The two pyrimidine bases, thymine and cytosine,

(26) Harz, M.; Rösch, P.; Peschke, K. D.; Ronneberger, O.; Burkhardt, H.; Popp, J. *Analyst* (*Cambridge, U.K.*) **2005**, *130*, 1543−1550.
(27) Wood, B. R.; Caspers, P.; Puppels, G. J.; Pandiancherri, S.; McNaughton, D. *Anal. Bioanal. Chem.* **2007**, *387*, 1691−1703.
(28) Rösch, P.; Harz, M.; Schmitt, M.; Peschke, K.-D.; Ronneberger, O.; Burkhardt, H.; Motzkus, H.-W.; Lankers, M.; Hofer, S.; Thiele, H.; Popp, J. *Appl. Environ. Microbiol.* **2005**, *71*, 1626−1637.

Table 1. Raman Band Assignment of DNA and Protein in Feather Pulp Spectra

bands (cm^{-1})	protein	DNA	ref
756	tryptophan, phenylalanine: symmetric benzene/pyrrole in-phase breathing mode		31, 33
1007	phenylalanine, tryptophan: symmetric benzene/pyrrole out-of-phase breathing mode		31, 33, 38
1172	tyrosine: in-plane C−H bending	C, T: CC and CN stretching	33, 38, 39
1210	tyrosine, phenylalanine: ring C−Cβ stretching		33, 38
1336	tryptophan: Fermi resonance between N1−C8 stretching in pyrrole ring and combination bands of out-of-plane bending	A: C5N7, N7C8 stretching G	31, 38, 40
1355	tryptophan: Fermi resonance between N1−C8 stretching in pyrrole ring and combination bands of out-of-plane bending	G: N7C8, N1C6, N5N7 stretching	33, 38, 40
1369		C: C4N, N2C2 stretching T: C6H bending, ring stretching	40
1414		A, G: C4N9 stretching, C8H deformation	39, 40
1452	tryptophan: minor ring mode		
1480		G: C8H deformation, N9C8 and C8N7 stretching A: C8H deformation, N9C8 stretching	31, 32, 39, 40
1500			
1522		C: N3C4 stretching	39, 40
1548	tryptophan: C−C stretching vibration of the pyrrole ring		33, 38
1569	tryptophan: in-plane C−C stretching mode of pyrrole ring	A, G: NH2 deformation vibration	32, 33, 38
1594		G: NH2 bending, C2N stretching C: NH2 bending, C4N stretching	40
1615	tyrosine, tryptophan: in-plane C=C ring stretching		33, 38
1648	amide I: C−O stretching vibration and N−H in-plane bending	C: C2=O stretching T: C4=O−C4C5 stretching G: N2H2 scissoring	31, 32, 34, 39, 40

generate the bands at 1355 and 1172 cm^{-1}, which can be assigned to CC and CN stretching motions.[29−32]

Furthermore, protein bands are found at 1648, 1615, 1569, 1452, 1355, 1336, 1210, 1172, 1007, and 756 cm^{-1} and can be assigned to three aromatic amino acids phenylalanine, tryptophan, and tyrosine. The band at 1648 cm^{-1} contributes to the CO stretching vibration and NH in-plane bending vibration of the amide I mode. The band around 1615 cm^{-1} is due to the in-plane CC ring stretching mode of tryptophan and tyrosine. The bands at 1569, 1355, 1336, and 756 cm^{-1} also arise from tryptophan, more precisely from CC and CN pyrrole stretching and bending vibrations, whereas the band at 1452 cm^{-1} contributes to a minor ring mode of tryptophan. The band at 1172 cm^{-1} finds its origin from an in-plane CH bending vibration of tyrosine. Tryptophan and phenylalanine exhibit a characteristic band at 1007 cm^{-1} due to breathing modes of the benzene and pyrrole rings.[31,33,34]

In summary it can be said that the main differences between female and male feather pulp UVRR spectra result from different DNA contents of both genders. The Raman difference signals (see Figure 1) being not characteristic for either female or male feather pulp and therefore not useful for a gender determination arise in the range of 1450 and 1350 cm^{-1}, between 1150 and 1010 cm^{-1}, and below 1000 cm^{-1}.

Chemometric Analysis. The results presented above demonstrate that female and male feather pulp Raman spectra exhibit subtle differences. However, for a quantitative gender determination a chemometric analysis is required. Prior to this analysis the first derivative of the 326 UVRR spectra from male and female feather pulp was calculated and vector normalized afterward. Different chemometric methods were applied for the quantification of variances between male and female chicken feather pulp samples.[35,36]

Support Vector Machine. Supervised classification methods like nonlinear SVMs require additional information about the sample which can be added to the spectral information as attributes. The quantity of attributes determines the number of classes for the classification. By means of mathematical algorithms, the best separation area of the multidimensional problem is calculated. The problem for differentiating two classes (male and female) can be ideally solved by SVM since no class of unknown features has to be taken into account.

The analysis uses higher dimensional vector spaces because each Raman spectrum contains 2109 data points. The result is a classification model in terms of a multidimensional function. Beside the relevant DNA signal information, the spectra are influenced by the background, protein variations, or pyrolysis effects. These spectral artifacts are taken into account to achieve a closer reality simulation and make the model more robust. To estimate the calibration model, the validation results of both the leave-one-out validation test set and an independent data set of unknown spectra are calculated. The independent data set of

(29) Neugebauer, U. Ph.D. Thesis, Friedrich-Schiller University Jena, Jena, Germany, 2007.
(30) Neugebauer, U.; Schmid, U.; Baumann, K.; Simon, H.; Schmitt, M.; Popp, J. *J. Raman Spectrosc.* **2007**, *38*, 1246−1258.
(31) Overman, S. A.; Bondre, P.; Maiti, N. C.; Thomas, G. J., Jr. *Biochemistry* **2005**, *44*, 3091−3100.
(32) Thomas, G. J., Jr. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 1−27.
(33) Asher, S. A.; Ludwig, M.; Johnson, C. R. *J. Am. Chem. Soc.* **1986**, *108*, 3186−3197.
(34) Kitagawa, T.; Hirota, S. In *Handbook of Vibrational Spectroscopy: Applications in Life, Pharmaceutical and Natural Sciences*; Chalmers, J. M., Griffiths, P. R., Eds.; John Wiley & Sons: Chichester, U.K., 2002; pp 3426−3446.
(35) Beebe, K. R.; Pell, R. J.; Seasholtz, M. B. *Chemometrics—A Practical Guide*; John Wiley & Sons, Inc.: New York, 1998.
(36) Henrion, R.; Henrion, G. *Multivariate Datenanalyse—Methodik und Anwendung in der Chemie und verwandten Gebieten*; Springer-Verlag: Berlin, 1995.

Table 2. SVM Results by Leave-One-Out Method[a]

data set		classification result		estimation (%)	
	real no. of samples	male chicken	female chicken	correct	wrong (error type)
male chicken	171	168	3	98.3	1.7 (type I)
female chicken	155	10	145	93.6	6.4 (type II)
				95.9	4.1

[a] The recognition rate for male and female chickens' samples is demonstrated for all 326 spectra.

Table 3. SVM Results by Test Set Method[a]

data set		classification result		estimation (%)	
	real no. of samples	male chicken	female chicken	correct	wrong (error type)
male chicken	57	56	1	98.2	1.8 (type I)
female chicken	52	4	48	92.3	7.7 (type II)
				95.4	4.6

[a] The original data set of male and female chickens' samples is split into a data set for calibration (217 spectra) and a data set for validation (109 spectra).

Table 4. Number of Correctly and Wrongly Classified Spectra of Male and Female Chickens' Feather Pulp Samples of the Full Data Set (326 Spectra) Resulting by HCA

data set		classification result		estimation (%)	
	real no. of samples	male chicken	female chicken	correct	wrong (error type)
male chicken	171	123	48	71.9	28.1 (type I)
female chicken	155	30	125	80.6	19.4 (type II)
				76.3	23.7

unknown spectra was achieved by dividing randomly the data set into the calibration set of 217 spectra (two-thirds of all spectra) and the validation set of 109 spectra. Therefore, two classification results can be discussed. The classification rates of the leave-one-out validation test set are shown in Table 2. When taking all spectral and attributed information into account, 98.3% of the male and 94.8% of the female feather pulp samples can be classified correctly. The validation results of the data set of unknown spectra are summarized in Table 3. The classification is possible with a probability of 98% for male and of 92% for female feathers. These recognition rates of both the validation data set and the unknown feather pulp samples have comparable error variances as manual vent sexing with an accuracy of about 95%.

These results can be compared to those obtained for an unsupervised HCA. For such an analysis only the spectral data set is required, and no further information about class affiliation (features) of the objects (spectra) are included into the similarity calculation. Table 4 presents the HCA classification results in detail. The classification rates for female decreases to 81% and for male to 72%. These results demonstrate that the variances of the spectra are too high for an acceptable classification. Thus, supervised classification methods such as SVM are advantageous, also in consideration of the possibility for the identification of unknown samples.

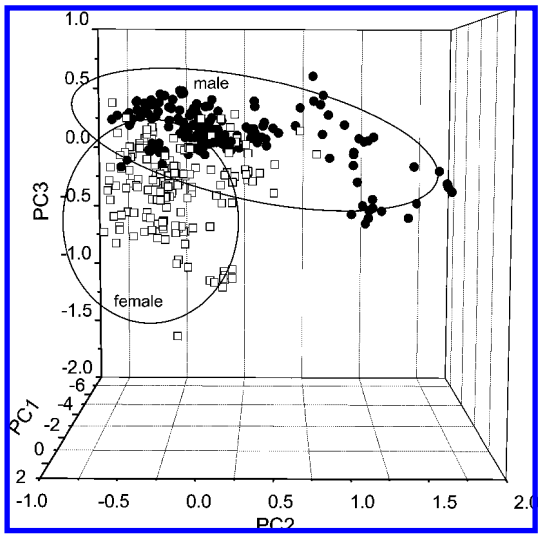

Figure 4. Three-dimensional diagram of the PCA-scores plot of male and female feather pulp samples spectra in the spectral range between 1140 and 1730 cm^{-1}.

Principal Component Analysis. The SVM showed to be a practical separation tool between female and male chicken samples with recognition rates of about 95%. In order to determine the decisive spectral regions for the classification of chicken's feather pulp samples a PCA was applied to the original normalized data set. The PCA is used to model the data and to determine the spectral variances associated with the feather pulp of male and female chickens. The variation of the data set is described via a few basis vectors or principal components (PC). More precisely, the first PC describes the greatest part of the data set variance, whereas higher PCs describe only noise. The coordinates of the data according to the new axes are called scores, and the contributions of the original variables to the PCs are called loadings. After the calculation, similar spectra group together in the scores plot. The loadings plot provides information of the variables (wavenumbers of the spectrum) that are important for group separation.

Figure 4 shows the three-dimensional diagram of a scores plot of the first three principal components of 326 male and female chicken feather pulp samples' UVRR spectra for the wavenumber range between 1140 and 1730 cm^{-1}. The scores plot reflects a clear separation of male and female feather pulp samples. It is possible to separate both genders with a few principal components, although the groups are disperse and sometimes overlap each other. The diffuse grouping originates from several reasons; for example, fluctuations in excitation power can influence the Raman intensity but are corrected by scaling methods. The spectral background resulting from variations in the sample thickness can be corrected by calculating the first-derivative spectra. Furthermore, a different sample composition, e.g., the ratio between DNA and protein content, can influence the spectrum quality. This effect is illustrated in Figure 5 A. An enlarged protein content which can be seen at the 1452 cm^{-1} band showing almost the same intensity as compared to the amount of DNA represented by the 1480 cm^{-1} mode (compare Table 1) is demonstrated in spectrum

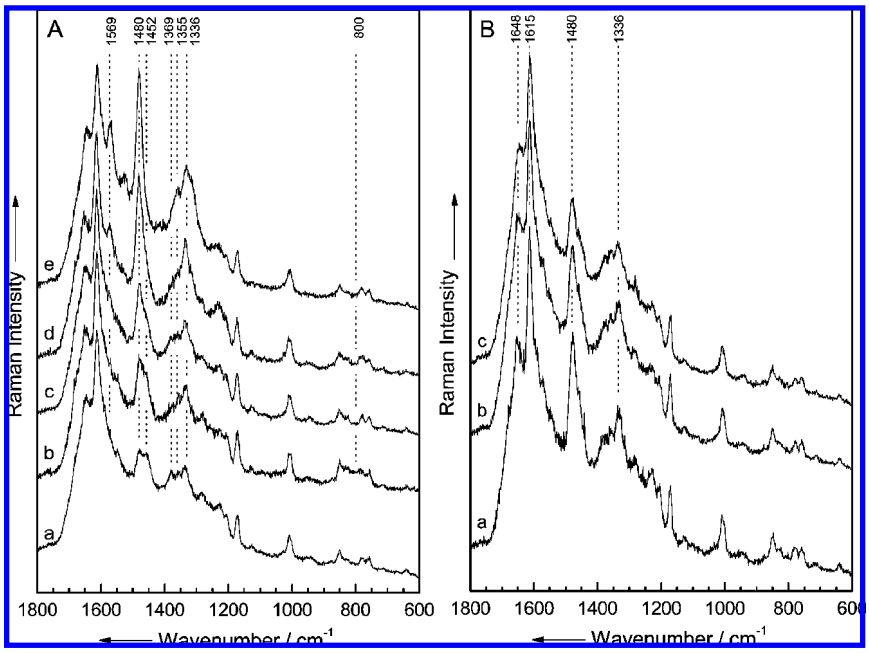

Figure 5. (A) Influence of different cell compositions on the Raman spectra: spectrum a with a heightened protein content, spectrum b with a dominant background signal of quartz, spectra c and d as mixture spectra of variations between protein and DNA composition, and spectrum e with an increased DNA content. (B) Influence of photochemical degradation during a measurement (curves a–c: increasing pyrolysis).

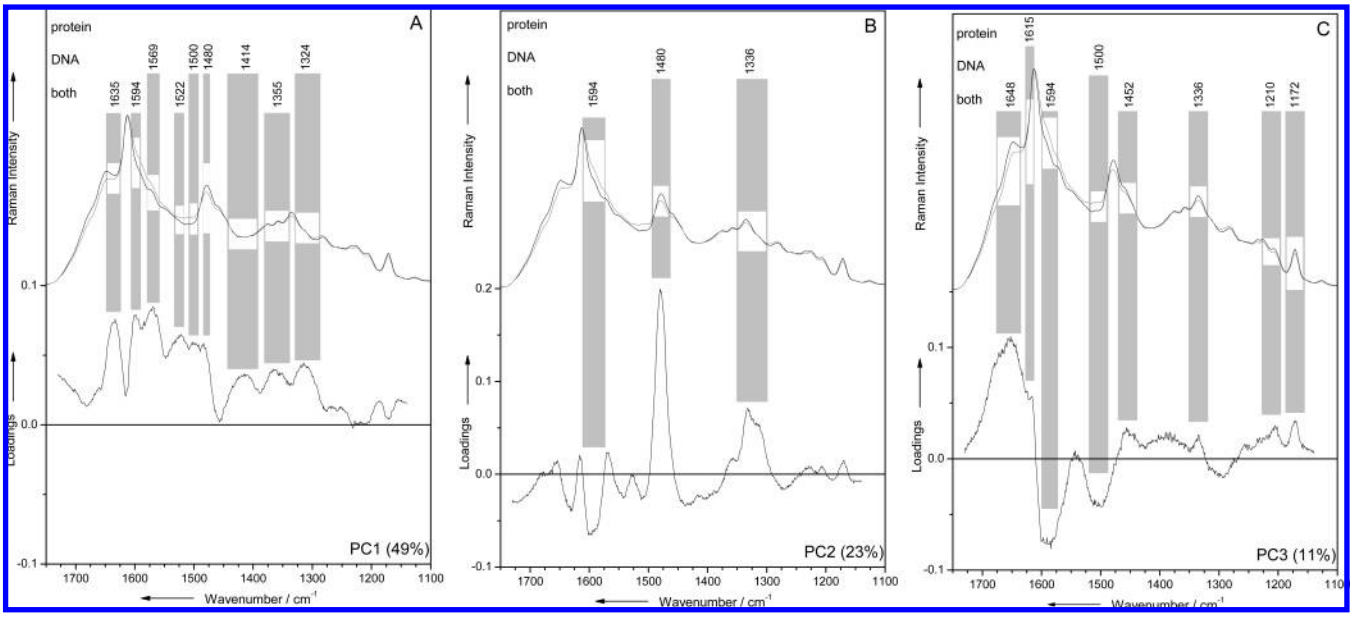

Figure 6. Loading plots of the first three principal components highlighting regions associated with large loadings compared to average spectra of the first derivative of the male and female chickens' feather pulp samples.

a. In contrast, a high DNA content discernible by the enhanced band intensities at 1569, 1480, and 1355 cm^{-1} is represented by spectrum e. During the Raman measurements, a spectral superposition between spectra, which are dominated by protein (spectrum a) or DNA (spectrum e) contents (spectra b, c, and d) occurred. The mode of sample preparation can also influence the Raman spectrum when too thin sample smears are used. In spectrum c, this effect is visualized where a broadened band at 800 cm^{-1} occurs due to the background spectrum of the quartz slide. Furthermore, small variations in signal intensities can be attributed to inhomogeneous surface tension as well as to possible photodegradation of the sample due to UV light irradiation. Figure 5B illustrates the influence of a progressive photodegradation of the samples during the measurement. In comparison to the nondegraded spectrum a, spectrum b shows a less and spectrum c a stronger degradation effect noticeable by decreased signal intensities at 1480 and 1336 cm^{-1} relative to the signal at 1615 cm^{-1}.

In order to determine the data set variances associated with feather pulp of male and female chickens, the loading plots need to be interpreted. In Figure 6A–C the average spectra of the male and female chickens' feather pulp samples (above) and the loading plots of the first three principal components PC1, PC2, and PC3 (below) are compared. Signal regions reflecting information of DNA, proteins, or both are marked with a rectangle and labeled with the respective wavenumber positions. The PC1 loading plot

(Figure 6A) shows nine different signal regions with high loadings. The spectral regions of 1569, 1500, 1522, and 1480 cm^{-1} as well as the region of 1414 and 1324 cm^{-1} show positive loadings and are correlated to DNA information. The wavenumber regions of 1635, 1594, 1522, and 1355 cm^{-1} represent both DNA and proteins bands.

The loading plot of PC2 (Figure 6B) includes three spectral regions with intensive loadings containing DNA information (1480 and 1336 cm^{-1} bands). Bands originating from both proteins and DNA are found in the wavenumber region around 1594 cm^{-1}.

The loading plot of PC3 (Figure 6C) shows eight strong loading regions. The first region around 1615 cm^{-1} represents protein information, and the second region at 1500 cm^{-1} yields DNA information. The regions at 1648, 1594, 1452, 1336, 1210, and 1172 cm^{-1} are correlated to both DNA and protein information.

These first three PCs can clarify 83% of the data set variation and focus with strong loadings on 13 different wavenumber regions in the range between 1660 and 1336 cm^{-1}. Due to the deep UV excitation wavelength leading to a selective enhancement of aromatic amino acids in proteins and DNA bases, the different wavenumber regions can be correlated to a superposition of DNA and proteins signals.

CONCLUSIONS AND OUTLOOK

Gender determination of birds is important for poultry industry as well as for bird keepers. Current invasive analysis methods imply increased risk for the bird's health. The work presented within this paper nicely demonstrates that UVRR spectroscopy of feather pulp is a capable alternative and minimal invasive method for gender determination.

The sample collection can be performed without distressing the animals in contrast to other methods like endoscopic surgery or blood analysis. The risk of anesthetic problems is excluded. The sample preparation is simple, and the sample extracts can directly be analyzed to achieve a Raman fingerprint spectrum within 1 min. For gender determination, the spectrum will be compared to a setup database by chemometrical methods. Thus, this approach presents a quick and easy method for gender determination.

In this contribution the influence of the selectively enhanced feather pulp constituents of the feather extracts to the feather pulp Raman spectrum was discussed. For that purpose the Raman spectra of the extract material were compared to those of pure DNA and proteins.

Different chemometrical methods were used to elucidate the spectral characteristics of male and female chicken samples. For classifying the recorded spectra, in this approach the supervised classification method SVM was applied. A classification model was created and applied to unknown samples for gender determination. The SVM showed to be a practical separation tool for female and male chicken samples with recognition rates of about 95%. The influence of spectra with increased photodegradation or with enhanced protein signals rates did not influence the classification results significantly.

In order to determine the important spectral regions for the classification of chicken's feather pulp samples a PCA was applied. The first three principal components can elucidate 83% of the data set variation and are necessary for a satisfactory differentiation between male and female feather pulp spectra. The loadings plots showed seven different wavenumber regions in the range between 1660 and 1336 cm^{-1} that are important for group separation. Since an excitation wavelength in the deep UV was applied a selective enhancement of aromatic amino acids in proteins and DNA bases is achieved. Therefore, the different wavenumber regions are correlated to a superposition of DNA and proteins signals.

The principle application of UVRR spectroscopy for a gender determination of birds has been demonstrated so far for chickens. In future work these investigations need to be extended to other birds like turkey. Here, a quick and simple gender determination has to be provided for optimizing the feeding conditions to reach better economic profits in poultry industry.

In the future the exploratory focus is also on better classification accuracy. The value of 98.5% which is the commercial accuracy acceptance of automated gender sorting of avian embryos[37] must be the goal. It can be reached by optimizing the sample preparation and sample handling to generate more homogeneous spectra. The minimization of photodegradation will be one of the main aspects in this respect.

ACKNOWLEDGMENT

This study was supported by the Hessen Ministry for Environment, Agriculture and Consumer Protection and the Lohmann Tierzucht GmbH (Cuxhaven). We gratefully acknowledge the Fond der Chemischen Industrie for financial support.

Received for review October 2, 2007. Accepted November 19, 2007.

AC702043Q

(37) Phelps, P. *World's Poult. Sci. J.* **2003**, *59*, 32−37.
(38) Cho, N.; Song, S.; Asher, S. A. *Biochemistry* **1994**, *33*, 5932−5941.
(39) Fodor, S. P. A.; Spiro, T. G. *J. Am. Chem. Soc.* **1986**, *108*, 3198−3205.
(40) Wen, Z. Q.; Thomas, G. J., Jr. *Biopolymers* **1998**, *45*, 247−256.