



ToF-SIMS Investigations on Dental Implant Materials and Adsorbed Protein Films

Falk Bernsmann

► To cite this version:

Falk Bernsmann. ToF-SIMS Investigations on Dental Implant Materials and Adsorbed Protein Films. Biomaterials. Technische Universität Kaiserslautern, 2007. English. NNT: . tel-00602449

HAL Id: tel-00602449

<https://theses.hal.science/tel-00602449>

Submitted on 22 Jun 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ToF-SIMS Investigations on Dental Implant Materials and Adsorbed Protein Films

Diploma Thesis

Falk Bernsmann

Group for Interfaces, Nano Materials and Biophysics
Department of Physics
Technische Universität Kaiserslautern

Supervised by
Prof. Dr. Christiane Ziegler
and
apl. Prof. Dr. habil. Hubert Gnaser

July 2007

Contents

1	Introduction	3
2	Previous work on protein analysis by ToF-SIMS	5
3	Theoretical aspects	7
3.1	Formation of biological films in the oral cavity	7
3.1.1	Saliva	7
3.1.2	Acquired enamel pellicle	7
3.2	Proteins	8
3.2.1	Serum albumin	10
3.2.2	Lysozyme	11
3.2.3	Amylase	11
3.3	Time-of-flight secondary ion mass spectrometry	12
3.3.1	Creation of primary ions	12
3.3.2	Sample interaction	13
3.3.3	Time-of-flight mass analysis	14
3.4	Fluorescence microscopy	16
3.5	Scanning force microscopy	19
3.6	Scanning electron microscopy	20
3.7	Principal component analysis	21
3.7.1	Change of basis	22
3.7.2	Variance and covariance	22
3.7.3	Diagonalisation of the covariance matrix	23
3.7.4	Graphical representation	24
3.7.5	Discriminating principal component analysis	26
3.7.6	Evaluation	28
4	Experimental aspects	29
4.1	Chemicals	29
4.2	Sample preparation	29
4.2.1	Dental material	29
4.2.2	Protein films on silanised substrates	30
4.2.3	Protein films on silicon substrates	32
4.2.4	Protein films on dental implant materials	32
4.2.5	Samples for fluorescence microscopy	33
4.3	Time-of-flight secondary ion mass spectrometry	33
4.4	Fluorescence microscopy	34
4.5	Scanning force microscopy	35
4.6	Determination of amylase activity	35
4.7	Data analysis	35

5	Results	37
5.1	ToF-SIMS of dental implant materials	37
5.2	ToF-SIMS of protein films on silanised substrates	38
5.3	ToF-SIMS of protein films on silicon substrates	44
5.3.1	Proteins solved in water	45
5.3.2	Proteins solved in 100 millimolar buffer solution	47
5.3.3	Proteins solved in 10 millimolar buffer solution	48
5.4	Fluorescence microscopy of protein films on silicon	55
5.5	ToF-SIMS of protein films on dental implant materials	57
5.5.1	Adsorption from binary protein mixtures	60
5.5.2	Sputtering of single protein layers	62
5.5.3	Consecutive adsorption of two proteins	64
6	Summary and Outlook	69
A	Matlab functions	72

1 Introduction

Biofilms play an important role in the health sector, in bioanalytics, in the food industry and in engineering science [24] because the adsorption of organic molecules can alter the physical, biological and chemical properties of a surface. This work deals with the formation of biofilms on dental implant materials.

When a dental implant is placed in the oral cavity, within seconds its surface is covered by a biofilm called pellicle consisting mainly of proteins and other macromolecules [15]. Since the adsorption of proteins is a highly selective process, the proportions of proteins found in the pellicle differ significantly from the ones found in saliva [36]. The pellicle is of great physiological importance because it serves as lubricant, as diffusion barrier to demineralising agents and as reservoir for remineralising electrolytes [15]. Furthermore, proteins in the pellicle play an important role in the colonisation of the surface by bacteria and thus in the formation of dental plaque. On the one hand there are proteins, like amylase, that exhibit specific binding sites for bacterial adsorption [13]. On the other hand enzymes, like lysozyme, immobilized in the pellicle have anti-bacterial properties [14]. Since the adsorption process of proteins is a subject not yet fully understood, this work shall further investigate the adsorption of the proteins amylase, lysozyme and serum albumin on two experimental dental implant materials.

The chosen method is Time-of-Flight Secondary Ion Mass Spectrometry (ToF-SIMS) because it offers the following advantages [2]: The mass distribution of molecules adsorbed on a sample's surface can be measured with a high mass resolution and a high surface sensitivity. It is possible to create depth profiles with a depth resolution of below one nanometre. And the analysis of non-conducting samples is possible without further preparation steps. ToF-SIMS has already been used to analyse adsorbed protein films on different substrates (see section 2) but up to now there are no ToF-SIMS studies of protein films on dental implant materials.

The difficulty in interpreting mass spectra of proteins is their complexity [10]. Every protein consists of a combination of the same twenty amino acids which dissociate within the ToF-SIMS analysis to numerous fragments. Hence one has to take into account the intensities of many different masses for analysis. To reduce the number of variables (i.e. masses), Principal Component Analysis (PCA) is used. This multivariate technique concentrates the variance of the spectra onto only a few variables, called Principal Components (PC).

This work is subdivided into the following sections: In this first section an introduction to the subject is given. The second section deals with previous work on the analysis of adsorbed protein films by ToF-SIMS and multivariate data analysis. Theoretical aspects of the examined systems and the applied techniques are detailed in the third section. The experiments are described in the fourth section. The fifth section contains the following results: First, mass spectra of dental implant materials are examined to determine their elemental surface composition. Then mass spectra of proteins adsorbed to silane coated silicon substrates are analysed to develop the methods and programmes necessary for distinguishing different proteins by their mass spectra. Since this does not work very well, the examined system is simplified to proteins adsorbed directly to silicon substrates. Here the different proteins can be recognized by their mass spectra and the developed statistical

models perform well in evaluation tests. The adsorption conditions are varied to obtain the best results. Additionally the mutual influences of two proteins adsorbed at the same time or consecutively to the same silicon substrate are studied. The results obtained are confirmed by enzymatic activity measurements. Finally the mass spectra of proteins adsorbed to dental implant materials are examined. With only little modifications in data pre-treatment, the programmes developed to analyse the spectra of proteins on silicon can be used to distinguish between different proteins adsorbed to dental implant materials. Again, information on the mutual influence of the proteins upon adsorption is obtained. The sixth section gives a summary and an outlook on possible future investigations.

2 Previous work on protein analysis by ToF-SIMS

In this section a brief overview of the literature available on multivariate analysis (MVA) of time-of-flight secondary ion mass spectra (ToF-SIMS) of adsorbed protein films is given.

In 2001 Wagner and Castner published their article “Characterization of Adsorbed Protein Films by Time-of-Flight Secondary Ion Mass Spectrometry with Principal Component Analysis” [31]. For analysis of single component protein films of various proteins adsorbed to poly(tetrafluorethylene) (PTFE), mica or silicon substrates with principal component analysis (PCA), several peaks of the mass spectra of positively charged ions were selected. The selection was based upon the work of Mantus and others [22] who had developed a spectral interpretation protocol for protein spectra based on strong peaks in the spectra of amino acid homopolymers adsorbed to glass substrates. Wagner and Castner were able to distinguish between several proteins by their scores on the first two principle components. Furthermore, the PCA model developed with the mass spectra of single component protein films allowed qualitative insight into the composition of a complex adsorbed protein film from bovine plasma.

In another article [32] published in 2002, Wagner and others described their attempt to quantitatively characterise multicomponent adsorbed protein films by ToF-SIMS. As before, only peaks related to amino acid fragments were selected from the mass spectra of positively charged ions for analysis. For binary protein films composed of fibrinogen and immunoglobulin G adsorbed on mica or PTFE, a good agreement between surface concentrations predicted from the mass spectra by a partial least squares regression (PLSR) model and radio labelling experiments was observed. PLSR is a method of multivariate analysis closely related to PCA. It is described for example in [8]. For ternary films composed of fibrinogen, immunoglobulin G and bovine serum albumin, only major trends in the surface composition could be traced by a PLSR model. As in the preceding article, qualitative information about the composition of complex protein films adsorbed from bovine serum or bovine plasma was obtained with a PCA model developed with the mass spectra of single component adsorbed protein films.

Still in 2002, Wagner and others compared the interpretation of static ToF-SIMS mass spectra of adsorbed protein films on mica or PTFE by different methods of multivariate pattern recognition [33]. The unsupervised technique PCA and the two supervised techniques discriminant principal component analysis (DPCA) and linear discriminant analysis (LDA) were used. An improved discrimination between the mass spectra of different proteins was observed when comparing the supervised techniques to the unsupervised one. Furthermore, Wagner and others introduced a method to classify unknown spectra to the previously examined proteins using a PCA model developed by the mass spectra of these proteins. A successful classification was possible using PCA but it could be improved by the use of DPCA and especially LDA. Yet the very good classification results of LDA went along with a high risk of spurious discrimination.

In their article “Classification of adsorbed protein static ToF-SIMS spectra by principal component analysis and neural networks” [26], Sanni and others compared the performance of PCA and the artificial neural network (ANN) “NeuroSpectraNet” applied to the mass spectra of proteins adsorbed to silicon substrates. An introduction to neural networks is given for example by Kriesel [19]. Sanni and others concluded that a

discrimination of different proteins using PCA with peak selection was possible but the classification of unknown spectra to the known proteins was difficult due to numerous outliers. On the other hand “NeuroSpectraNet” allowed a full classification of unknown spectra using the complete mass spectra of positively and negatively charged ions. According to the authors the major drawbacks of neural networks lie in the complexity of their algorithms and of data interpretation.

Xia and Castner published the article “Preserving the structure of adsorbed protein films for time-of-flight secondary ion mass spectrometry analysis” [35] in 2003. They wanted to preserve the structure of fibrinogen layers on gold coated silicon substrates upon dehydration. The samples were fixated with trehalose or glutardialdehyde and ToF-SIMS spectra of positively charged ions were analysed with PCA. It was found that unfolding and exposure of hydrophobic domains induced by drying could be prevented by both methods.

In 2003 Belu and others published a review on techniques and applications for characterisation of biomaterial surfaces by ToF-SIMS [2]. They discuss the ToF-SIMS technique with regard to biomaterial samples and give examples of applications and data interpretation.

Michel and Castner reviewed the “Advances in time-of-flight secondary ion mass spectrometry analysis of protein films” [23] in 2006. The article deals mainly with characterisation and classification as well as conformation and orientation of proteins, quantitative studies, spatial distribution of proteins, cluster ion sources and matrix-assisted desorption techniques.

Also in 2006 Graham and others gave an overview of current techniques and future needs in ToF-SIMS data interpretation by multivariate analysis in the article “Information from complexity: Challenges of ToF-SIMS data interpretation” [10].

3 Theoretical aspects

3.1 Formation of biological films in the oral cavity

3.1.1 Saliva

In the oral cavity, saliva fulfils the following tasks:

- The regeneration of dental enamel is enabled by ions solved in the saliva.
- Enzymes like amylase allow the pre-digestion of food.
- The oral cavity is cleaned by removal of nutrition residues.
- Saliva buffers acids either supplied by food or produced by bacteria.
- Mucines form a lubricating film on the tooth surfaces to reduce the mutual abrasion.

Human saliva consists to over 99 % of water. The residue is composed to two thirds by organic and to one third by inorganic compounds [27]. The most abundant inorganic materials are the anions hydrogen carbonate, chloride and phosphates as well as the cations of potassium, sodium and calcium. A large amount of the organic material is formed by proteins. The most frequent of these are albumin, amylase and lysozyme [27]. These proteins analysed in this work are described in detail in section 3.2.

3.1.2 Acquired enamel pellicle

On the surface of teeth as well as dental implants, proteins and other macromolecules are selectively adsorbed to form a film called pellicle. Salivary proteins form an initial layer of 10 to 20 nanometres thickness within a couple of minutes [15]. According to the work of Hannig and Joiner [15], the adsorption is governed by ionic interactions between the proteins' charged groups and calcium and phosphate ions of the enamel surface assisted by van der Waals interactions. In a second phase, proteins and protein aggregates are continuously adsorbed from the saliva. Its thickness reaches a plateau after 30 to 90 minutes and increases further within 60 minutes to reach 100 to 1 000 nanometres. Afterwards the pellicle attains a dynamic equilibrium of adsorption and desorption.

The major salivary components of in-vivo formed pellicle are proteins and glycoproteins [15]. Their proportions are not the same as in whole saliva indicating that the adsorption is a selective process [36]. The three proteins albumin, amylase and lysozyme studied in this work are abundantly found in pellicle.

The main functions of pellicle formed on enamel surfaces are:

- Lubrification of the tooth surface.
- Formation of a diffusion barrier for acidic agents to protect the enamel from erosion.
- Inhibition of mineral precipitation from the tooth surface.
- Modulation of bacterial adherence onto the surface.

Proteins in the pellicle can influence the adhesion of bacteria in different ways. Some of them, like amylase, offer specific binding sites for bacteria [13] while others, like lysozyme, can decompose bacteria by enzymatic processes [14].

Bacteria and other micro-organisms colonise the pellicle and form dental plaque which is clearly distinguished from pellicle [15]. It is a whitish layer difficult to wipe off the dental surface and it can lead to caries and gingivitis. By formation of phosphate crystals the plaque can mineralise to form dental calculus [27].

3.2 Proteins

Proteins are the most complex known molecules. Due to their manifold structure they can fulfil various tasks. They serve the immune system as antibodies, they transport and store metabolic materials (as haemoglobin does), they allow signal transmission as hormones, they catalyse metabolic reactions in form of enzymes, they form supporting structures (e.g. collagen) and allow the movement of muscles (actin and myosin). An introduction to function and structure of proteins can be found for example in the book by Light [21] and in the diploma thesis of Schmitt [27] which have served as main sources for this section.

Every protein is composed of a combination of the twenty proteinogenic amino acids. An amino acid consists of a carboxylic acid ($-\text{COOH}$) with an amino group ($-\text{NH}_2$), which is usually bound to the α carbon atom of the carboxylic acid, and a side chain bound to the same carbon atom. Different amino acids are discriminated by the composition of their side chain which will be denoted R in the following. The structures of the amino acid residues as they are found in proteins are shown in figure 1.

The carboxylic group of one amino acid can react with the amino group of another by dissection of water. Thereby an amide bond ($\text{O} = \text{C} - \text{NH}$) is created between the two amino acids (see figure 2). A protein is a chain of many amino acids linked by amide bonds. It is unambiguously determined by its amino acid sequence called *primary structure*.

Each amide bond contains a partially negatively charged group (CO) and a partially positively charged one (NH). This favours the building of hydrogen bonds ($\text{C} = \text{O} \cdots \text{H} - \text{N}$) which stabilise the polypeptide chain in certain conformations. The most common of these *secondary structures* are the α helix and the β sheet.

In an α helix the polypeptide chain winds around a central axis with 3.6 amino acid residues per turn and a translational distance along the axis of 5.4 Ångströms per turn [21] (see figure 3). This way every carbonyl oxygen is hydrogen bonded to the amide hydrogen of the fourth peptide further along the chain. The side chains of the amino acid residues mostly point away from the helical axis.

In a β sheet some polypeptide chains are closely aligned side by side. To allow the maximum number of hydrogen bonds between the chains, these have to be shorter than a fully extended chain resulting in a conformation that resembles a pleated sheet (see figure 3). The side chains of the amino acid residues are located alternately above or below the plane of the sheet. Regions of the polypeptide chain showing an extended form without one of the secondary structures are called *random coil* regions.

The overall spatial structure (*tertiary structure*) of a protein is built by interactions

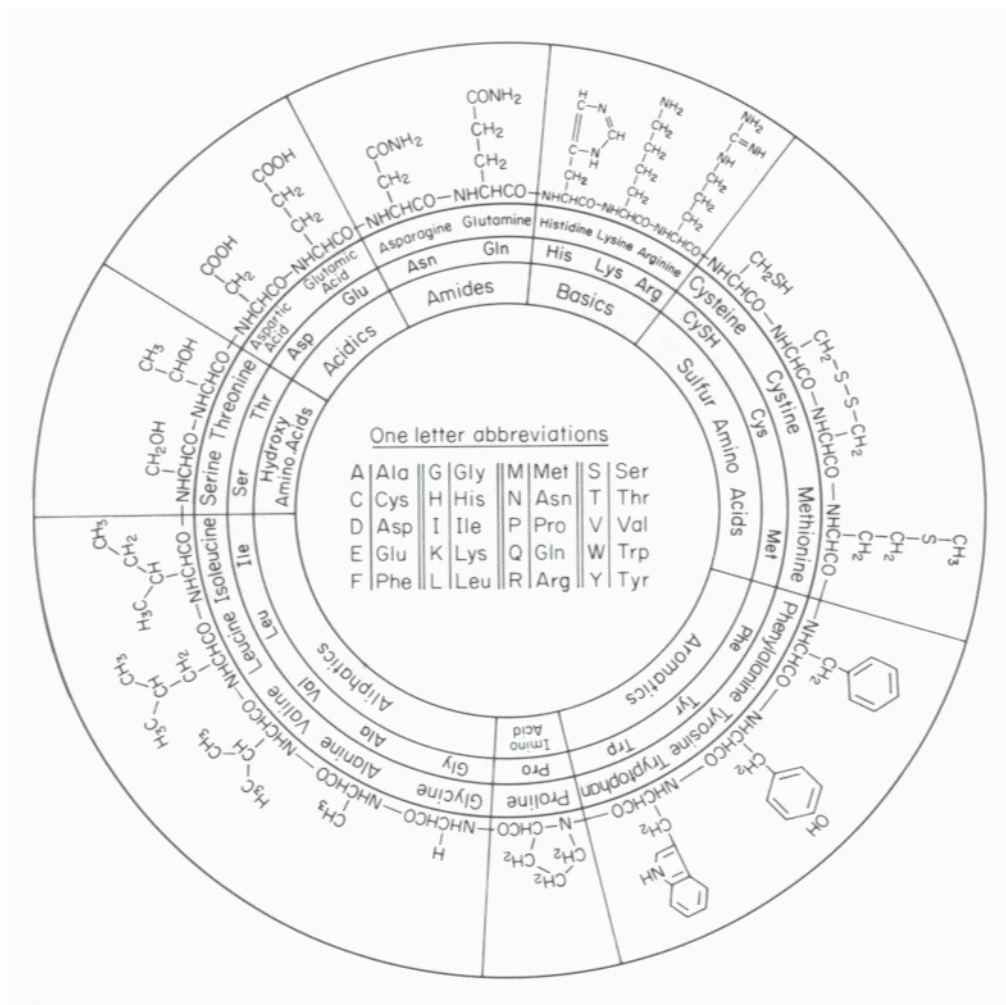


Figure 1: Structures, names and abbreviations of the amino acid residues found in proteins grouped by the properties of their side chains [21]

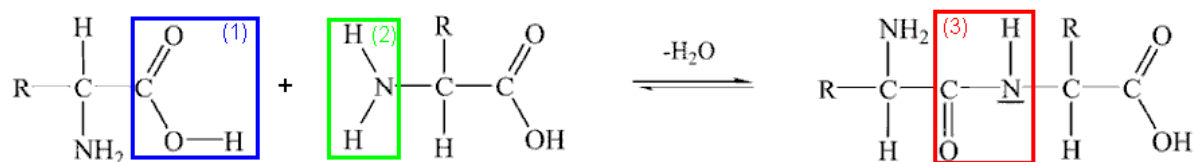


Figure 2: Formation of an amide bond: (1) = carboxylic group; (2) = amino group; (3) = amide bond [27]

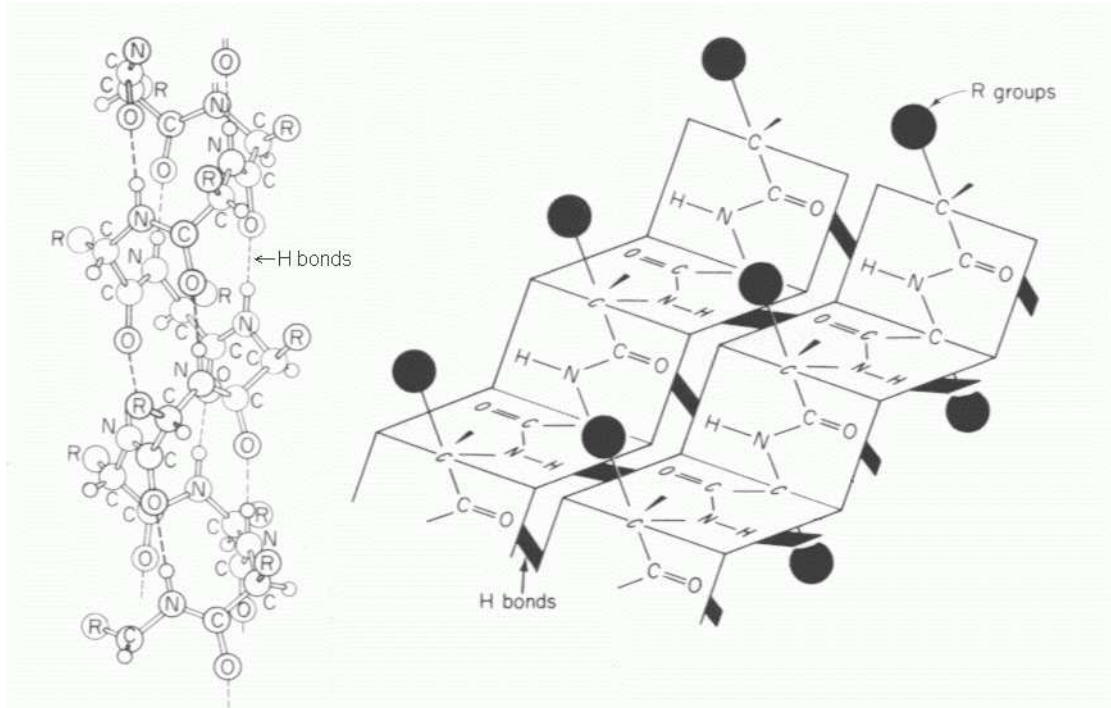


Figure 3: Models of an α helix (left) and a β sheet (right) (modified from [21])

between the side chains of amino acid residues. These can be salt bridges, hydrogen bonds, van der Waals interactions or disulfide bonds ($S - S$). The tertiary structure is vital for the functioning of a protein. If this conformation is changed for example by dehydration or changes in temperature or pH, the protein usually can not fulfil its tasks any more. One speaks of a *denatured* protein in this case.

A protein contains acidic as well as basic amino acid residues that are partly dissociated depending on the pH of the surrounding medium. The dissociation processes create electrically charged residues. Since the numbers of positive and negative charges are usually not equal, the protein carries a non zero net charge. Anyway, there exists for a given protein a pH at which its net electrical charge is zero. This pH is called the proteins *isoelectric point* (pI).

3.2.1 Serum albumin

Serum albumin is the most abundant plasma protein in human blood. It accounts for roughly 60% of the protein mass in the plasma [27]. Its major task is to maintain the colloidal osmotic pressure. Since the concentration of albumin in the blood vessels is higher than in the surrounding tissue, a leakage of water from the vessel is prevented by the osmotic pressure. Furthermore, albumin serves as a transport molecule and buffers the pH-value [6]. Besides in the blood plasma, serum albumin can also be found in the skin, in muscles, in the saliva and in the cerebrospinal fluid.

In this work bovine serum albumin (BSA) is used because its amino acid sequence is to 76% identical to the one of human serum albumin (HSA) and it is significantly cheaper [27].

BSA consists of 607 amino acids and has a molecular weight of 69 kg/mol [1]. In

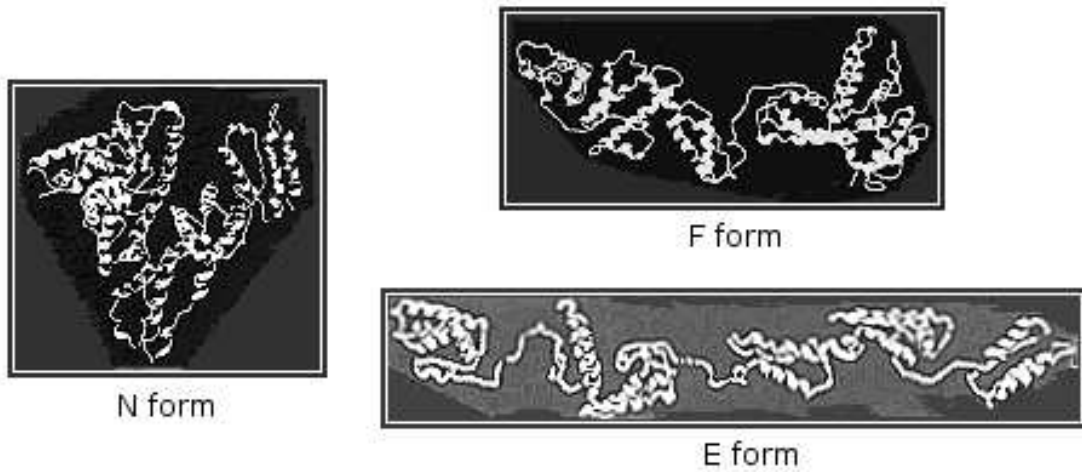


Figure 4: Models of different conformations of BSA. α helices are represented by helices and random coil regions by cords [6]

physiological conditions it shows a heart-shaped tertiary structure called normal form (N form) with a size of roughly 11 nm by 8 nm by 8 nm [1]. Its isoelectric point is at pH 4.7 [36]. Dependent of the pH-value the shape of BSA is reversibly changed. Above pH 8 one finds the basic form (B form), between pH 2.7 and pH 4.3 the fast migrating form (F form) and below pH 2.7 the extended form (E form) [6]. Figure 4 shows simplified models of different spatial conformations of BSA. Instead of the atoms of the polypeptide chain only the different secondary structures and their relative position are drawn.

3.2.2 Lysozyme

Lysozyme can be found abundantly in most body liquids like saliva, plasma and tears. It is the main pellicle-bound bacteriolytical component [14]. Its antibacterial effect is caused by the ability to enzymatically dissolve the cell wall of bacteria and to activate bacterial autolysis [14].

The lysozyme used in this work is extracted from chicken egg white. It consists of 129 amino acids and has a molecular weight of approximately 14.3 kg/mol [27]. Its shape is globular with a size of roughly 4 nm by 3 nm by 2.5 nm [1] as visualised by the model shown in figure 5. Lysozyme is a basic protein with its isoelectric point at pH 9.3 [36]. It shows optimum efficiency at pH 4.5 [27].

3.2.3 Amylase

Amylase is a digestive enzyme that breaks down starch. It is produced in the salivary glands and in the pancreas. In human saliva it is the most abundant enzyme and it is a major component of pellicle as well [12]. Like lysozyme it keeps its enzymatic activity in the pellicle [13]. Amylase complexes in the pellicle are binding sites for pioneer bacteria and play thus an important role in plaque formation [13]. Salivary amylase also called ptyalin can break polysaccharides down into maltose and glucose. It works only at a pH of about 7 and is inactivated in the stomach by gastric acid [27]. The isoelectric point of amylase is at pH 6.3 [36].

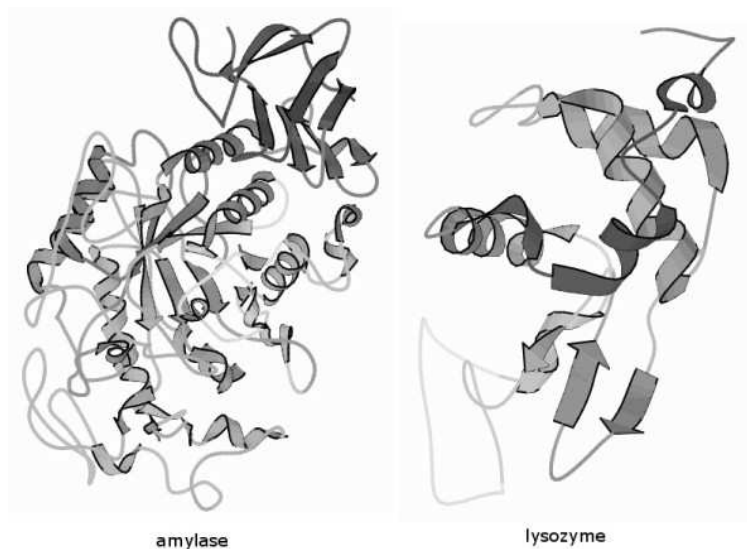


Figure 5: Conformation models of amylase and lysozyme. α helices are represented by helices, β sheets by flat arrows and random coil regions by cords [1]

Human salivary amylase consists of 496 amino acids and has a molecular weight of 56 kg/mol. Its size is approximately 7.5 nm by 4.5 nm by 4.5 nm [1]. A model of the conformation is shown in figure 5.

3.3 Time-of-flight secondary ion mass spectrometry

In secondary ion mass spectrometry (SIMS), a beam of fast primary ions forms secondary particles on a bombarded surface from which they are ejected. Only few of the secondary atoms and molecules are ionized by the collision and can be examined by a mass analyser. In the case of time-of-flight secondary ion mass spectrometry (ToF-SIMS), the ions are accelerated to a certain energy. Then ions of different masses are separated by the time they need to fly a given distance.

The most important condition for the examined samples is their ability to tolerate ultra high vacuum conditions with pressures below 10^{-8} mbar. A vacuum is necessary to prevent collisions of the primary and secondary ions on their drift trajectories with air molecules. Furthermore, it prevents a contamination of the sample surface during analysis. Electric charging of the sample must be prevented because it may degrade or even suppress the secondary ion signal. To examine electric insulators they are flooded with slow electrons between the primary ion pulses to compensate charging of the sample.

3.3.1 Creation of primary ions

The primary ions can be created for example by electron collision ionisation, plasma ionisation, surface ionisation or in a liquid metal ion source (LMIS). The latter is employed in the ToF-SIMS apparatus used for this work. The source is a metal tip in a reservoir of liquid metal that forms a film on the surface of the tip. Between the tip and an aperture lying in front of it, a strong electric field (for a gallium source about 10^{10} V/m [9]) is applied to create ions from the metal film by field ionisation and accelerate them towards

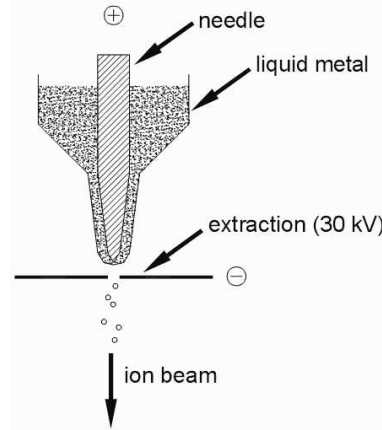


Figure 6: Sketch of a liquid metal ion source [28]

the sample (see figure 6). Compared to other sources [9] a liquid metal source generates a very intense beam and has a small lateral source size of about 50 nanometres. This creates large space charge effects near the tip of the source causing the emitted ions to have a relatively high energy spread of more than five electron volts. Thus the spot size on the surface is limited by chromatic aberrations. Anyway, liquid metal ion sources are the ones that offer smallest spot size compared to other ion sources [9].

3.3.2 Sample interaction

Before their arrival on the sample surface, the primary ions are accelerated to energies of several kilo-electron volts. Since these energies are significantly larger than typical binding energies, most of the molecular bonds near the impact site are broken and mainly atomic secondary particles are emitted. A part of the primary ion's energy propagates into the sample in form of a collision cascade. Thus particles can be emitted from sites further apart from the impact, too. Since the available energy is smaller at these sites, molecular bonds are not necessarily broken and molecules can be emitted in larger fragments or as an integrated whole. The diameter of the collision cascade is typically smaller than five nanometres in organic compounds and smaller than twenty nanometres in metals [28]. Only particles from the uppermost monolayers have enough energy to overcome the surface binding energy and to escape from the sample. A small fraction of them ($10^{-6} - 10^{-1}$ [2]) is charged and can be mass analysed.

To limit the analysis to the uppermost monolayers and to minimize sample damage, an experiment has to be done in static mode (Static Secondary Ion Mass Spectrometry: SSIMS). This is the case if the probability of any sample site to be hit by more than one primary ion is negligible. The mean number $\bar{n}$ of primary ions hitting an area A is related to the ion dose D by

$$DA = \bar{n}. \quad (1)$$

Assuming for static mode that $\bar{n} < 0.01$ and $A = 10 \text{ nm}^2$, one can calculate the maximum ion dose in SSIMS as

$$D < 10^{13} \text{ cm}^{-2}. \quad (2)$$

The secondary ion yield Y is defined as the number of detected secondary ions per

primary ion. It is strongly dependent on the chemical environment at the emission site. This *matrix effect* complicates the analysis of ToF-SIMS spectra because the varying yield for different secondary ions as well as for different sample sites leads to an intensity distribution in the mass spectrum that does not necessarily reflect the sample surface composition.

3.3.3 Time-of-flight mass analysis

The principle of time-of-flight (ToF) mass analysis is as follows: All secondary ions are accelerated to the same kinetic energy and drift over a field free distance. At the end of it, the ions are detected and one can discriminate different masses by their differing times-of-flight. To measure a time interval, one needs a well-defined starting time. Therefore the primary ion beam is pulsed and it is blanked after each pulse until all secondary ions have arrived at the detector. The advantage over other mass analysers like quadrupole or magnetic sector field systems lies in the parallel acquisition of signals from all masses. In addition, ToF mass analysers offer a greater transmission [2] which allows a better exploitation of the limited number of available secondary ions.

Usually the secondary ions are accelerated to kinetic energies of several kilo-electron volts. Thus their kinetic energy is negligible against their rest energy, which amounts already to nearly one giga-electron volt for a single proton. Hence the following calculations are done in non-relativistic approximation.

To calculate the time-of-flight, the following variables are defined: From the sample surface secondary ions of mass m and electric charge q are accelerated on a distance d by an electric field of strength E . Thus they traverse a voltage $V = Ed$. Afterwards they drift on a field free distance D and are detected at the time t . Within the acceleration the force F on the ions and their acceleration a are related by

$$F = Eq \quad \text{and} \quad F = ma \quad (3)$$

$$\Rightarrow a = Eq/m. \quad (4)$$

The final velocity u and the time t_a to traverse the acceleration range are thus

$$u = \int_{t_0}^{t_a} a dt = u_0 + Et_a q/m \quad (5)$$

$$\Rightarrow t_a = \frac{(u - u_0)m}{Eq}. \quad (6)$$

u_0 is the initial velocity of the ions at the starting time t_0 . From the kinetic energy T at the end of the acceleration one can calculate the drift velocity u_d

$$T = qV = qEd = 1/2 m u_d^2 \quad (7)$$

$$\Rightarrow u_d = u - u_0 = \sqrt{\frac{2qV}{m}}. \quad (8)$$

Assuming a small initial velocity ($u_0 \ll u$), it follows that

$$u \approx u_d = \sqrt{\frac{2qV}{m}}. \quad (9)$$

Hence the drift time t_D is

$$t_D = D/u = \frac{D}{\sqrt{2qV/m}}. \quad (10)$$

The observed time-of-flight t is the sum of the preceding times and the reacting time t_R of the detector

$$t = t_0 + t_a + t_D + t_R. \quad (11)$$

Since the acceleration occurs on a distance of typically some millimetres and the drift distance measures about one metre, the drift time is with some 100 μ s much longer than all the other terms and one can use equation (10) as good approximation of the time-of-flight.

The charge of the secondary ions amounts to $q = ze$. z is the charge number and e the elementary charge. Thus it follows from equation (10) that

$$t^2 = \frac{D^2}{2V} \frac{m}{q} \propto \frac{m}{z}. \quad (12)$$

To derive the mass to charge ratio m/z from the time-of-flight, one has to determine by calibration the proportionality constant n in

$$\frac{m}{z} = nt^2 \quad (13)$$

The mass of an ion of given charge is determined by $m = ct^2$ with a proportionality constant c . By deriving, one gets $dm = 2cdt$. Hence the mass resolution $m/\Delta m$ is inversely proportional to the relative uncertainty in the measurement of the time-of-flight $\Delta t/t$:

$$\frac{m}{\Delta m} = \frac{t}{2\Delta t}. \quad (14)$$

Reasons for this uncertainty are the spread of initial velocities, energies and positions as well as the spread in the formation times of the secondary ions. Furthermore non-ideal accelerating fields and jitter of the detection system contribute to the uncertainty.

In case of a solid sample all secondary ions are created at its surface and traverse the same acceleration distance. With an initial energy T_0 , their drift energy T_D and their drift time t_D are

$$T_D = T_0 + qV \quad (15)$$

$$t_D = D \sqrt{\frac{m}{2(T_0 + qV)}}. \quad (16)$$

Assuming a long drift distance ($t = t_D$),

$$dt = dt_D = \frac{\frac{D}{2} \sqrt{\frac{m}{2}}}{(T_0 + qV) \sqrt{T_0 + qV}} dT_0 \quad (17)$$

$$\Rightarrow \frac{dm}{m} = \frac{2dt}{t} = \frac{dT_0}{T_0 + qV} \quad (18)$$

holds.

Since the acceleration energy usually is much larger than the initial kinetic energy ($qV \gg T_0$), the mass resolution is

$$\frac{m}{\Delta m} = \frac{T_0 + qV}{\Delta T_0} \approx \frac{qV}{\Delta T_0}. \quad (19)$$

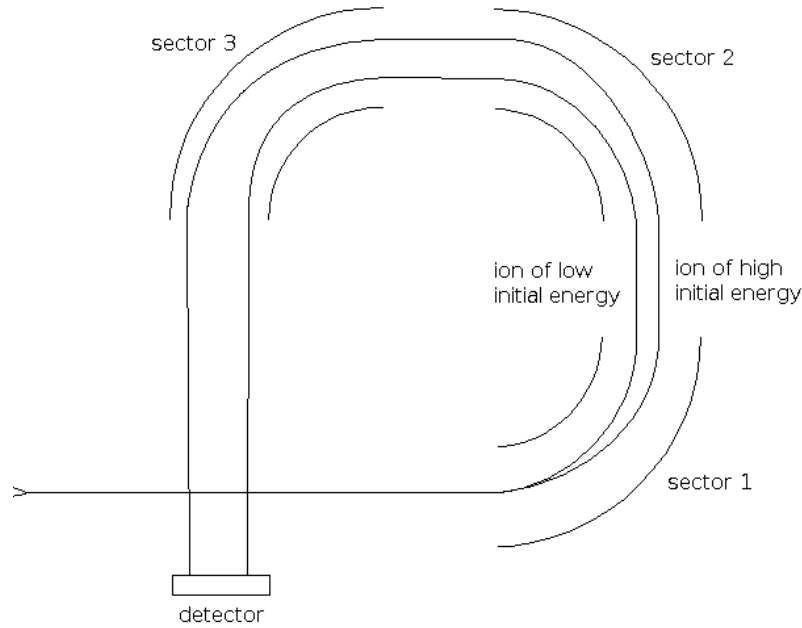


Figure 7: Effect of the initial energy on the trajectories of ions of equal mass

To achieve a high mass resolution, the spread in the initial energy has to be compensated. This is usually done by letting high energy ions traverse a longer distance than low energy ions. In the mass spectrometer used for this work, this is realised by a system of three electric 90° sector fields. Therein the trajectories of high energy ions have larger radii than the ones of low energy ions as sketched in figure 7.

The uncertainty in the time of ion creation t_0 is mainly determined by the length of the primary ion pulse. For the apparatus employed for this work, it amounts to roughly ten nanoseconds.

Because of their high sensitivity, electron multipliers and micro channel plates are used as detectors. In the ideal case, all electrons created by the impact of one secondary ion should arrive at the same time at the detector anode. Furthermore an amplifier with a short time constant and an exact determination of the starting time t_0 are needed to prevent signal broadening. Actually none of these conditions is perfectly met, so that a single secondary ion usually creates a signal pulse of up to ten nanoseconds width [11].

Figure 8 shows the layout of the triple focusing time-of-flight (TRIFT) mass spectrometer used in this work. Due to a system of electrostatic lenses, it offers not only energy focusing but also direction focusing. Hence on the channel plate detector an image of the ion distribution on the sample's surface is created. In the figure, the ion images of aluminium and copper from an aluminium copper grid are shown to illustrate this.

3.4 Fluorescence microscopy

Fluorescence microscopy is a kind of optical microscopy allowing the visualization of objects containing a fluorescing molecule, a so-called fluorophore.

Absorption of light in the visible or ultraviolet range by a molecule can lead to an

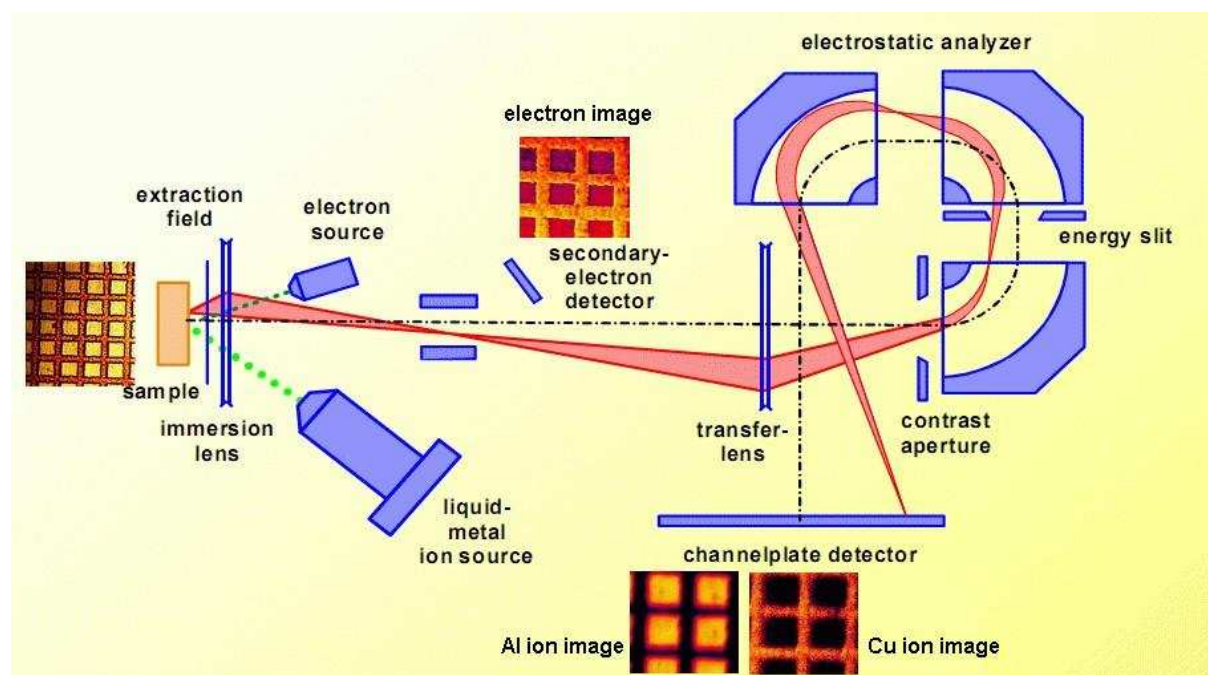


Figure 8: Layout of the TRIFT mass spectrometer (modified from [9])

electronically excited state. Afterwards there exist several different ways for the molecule to return to its ground state. If this relaxation is accompanied by the emission of radiation, one speaks of luminescence. Due to electron pairing, the electronic ground state of a molecule with an even number of electrons is usually a singlet state called S_0 with vanishing total spin. This electronic state can be subdivided into several vibrational states (denoted $\nu_0, \nu_1, \nu_2, \dots$), but at room temperature mainly the vibrational ground state is occupied. By absorption of a photon with the matching energy, the molecule can be excited into a vibrational state of the first (S_1) or higher ($S_2, \dots$) excited electronic singlet states. The principles of excitation and relaxation in a fluorophore are sketched in figure 9. An excitation into a triplet state is forbidden by quantum selection rules because it would require a spin-flip which cannot be caused by a photon [5].

Then the molecule relaxes by radiation free transitions to the vibrational ground state of the first excited electronic state. The energy is dispersed to inner vibrational states of the molecule or to vibrational states of neighbouring molecules [7]. Anyway, the molecule does not relax by radiation free transitions from the first excited electronic state to the electronic ground state. Since the energetic difference among the states S_1 and S_0 is larger than between the others, more vibrations would have to be excited simultaneously reducing the probability of a radiation free transition to occur. Thus the molecule relaxes by emission of a photon.

Once the molecule has arrived in the S_1 vibrational ground state *intersystem crossing* can occur with little probability. In this case the transition from the singlet system to the triplet system takes place. Another possible way to leave the S_1 state is relaxation by fluorescence *quenching* if the excitation energy is transferred to a so-called quencher molecule.

Fluorescence occurs upon the radiating transition from the S_1 vibrational ground state

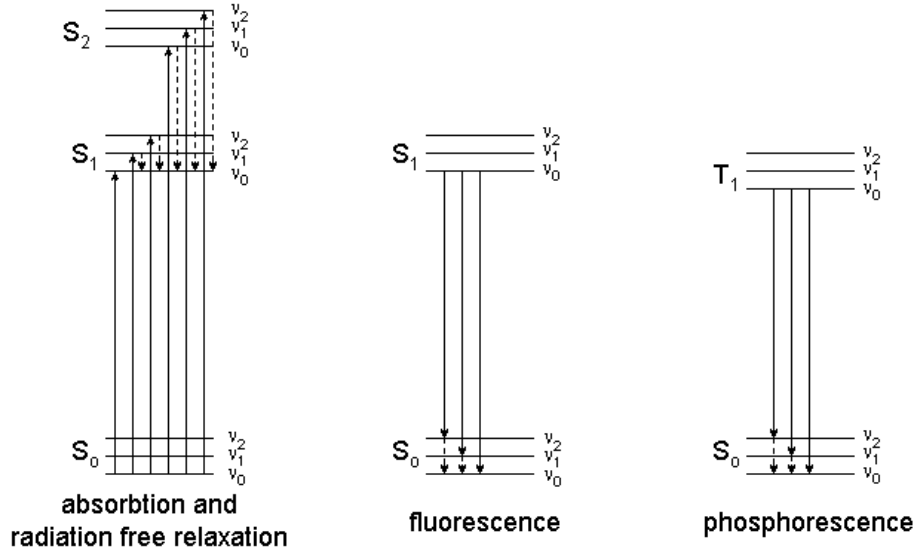


Figure 9: Scheme of excitation and relaxation in a fluorophore. Radiation transitions are symbolised by solid lines, radiation free transitions by dashed lines

into a vibrational state of the electronic ground state S_0 . Typical transition rates are 10^7 to 10^8 per second [7].

Phosphorescence is generated by the radiating transition from the vibrational ground state of the first excited electronic triplet state T_1 to a vibrational state of the electronic ground state S_0 . Its transition rate is much smaller with typical values ranging from less than one up to 10^4 per second [7].

The setup of a fluorescence microscope resembles the one of an ordinary light microscope. Typically a mercury lamp or a laser is used to illuminate the samples. In the first case the wavelength for exciting the fluorophore is selected by a set of filters. A dichroic mirror reflects the light through the observation optics onto the sample surface. The fluorescing molecules emit light which is collected by the observation optics. Since the wavelength of the emitted light is longer than the one of the exciting light, it can pass the dichroic mirror to be guided to the ocular pieces, a camera or a photomultiplier where it is detected.

The advantages of fluorescence microscopy over normal optical microscopy are the following:

- Objects showing little contrast with respect to one another can be distinguished by marking them with different fluorophores emitting at different wavelengths.
- Since they act as light sources, fluorescing objects much smaller than the optical resolution of the microscope can be detected.

A drawback of the method is the possible modification of the sample's chemical or physical properties by the introduction of a fluorophore. Additionally, the properties of the fluorophore, especially its fluorescence yield, can be strongly dependent on its chemical environment making comparisons of different samples difficult.

Using normal light microscopy the protein layers dealt with in this work are not visible. So it is necessary to modify them with fluorescence markers for imaging.

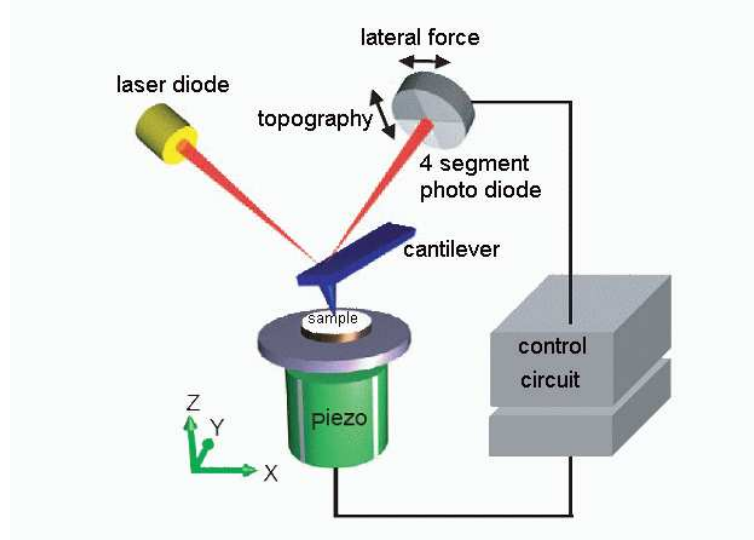


Figure 10: Layout of a scanning force microscope (modified from [27])

3.5 Scanning force microscopy

Scanning force microscopy (SFM) is a scanning probe microscopy method. It allows imaging of the topography of non conducting samples in air as well as in liquid with a typical lateral resolution of some nanometres and a height resolution of less than one Ångström [3]. Since it is possible to achieve a resolution of atomic scale, this method is also called atomic force microscopy (AFM).

Figure 10 shows the layout of a scanning force microscope. The very fine tip of a cantilever serves as probe to explore the surface of a sample placed on a stage which can be moved in three dimensions by piezoelectric actuators. A laser spot is reflected from the cantilever onto the centre of a four segment photo diode. By measuring the light intensities on the four segments and calculating the differences between the intensity arriving on the upper and lower half and between the right and left half of the photo diode, movements of the cantilever can be detected. If the cantilever is bent, the first difference is non-zero; if it is twisted, the second difference is non-zero.

In the so-called *contact mode* the sample is approached to the cantilever until its tip feels the repulsive force caused by the sample surface. This force causes an upwards bending of the cantilever and can thus be detected via the photo diode. There are two possibilities to scan the sample surface:

In the *constant height mode* the z position of the sample is held constant while scanning in x and y directions. Here x , y and z are the axes of a Cartesian coordinate system with the z axis perpendicular to the sample surface and the x and y axes in the sample surface plane (see figure 10). Any height differences of the sample surface cause a change in the force acting on the cantilever tip which is recorded to calculate an image of the sample surface. Due to the limited flexibility of the cantilever, this mode can only be used on very flat surfaces.

In the *constant force mode* the z position of the sample is modified by a control circuit to maintain the force acting on the cantilever tip constant while scanning. This way the distance between cantilever tip and sample surface remains constant. Hence the variations

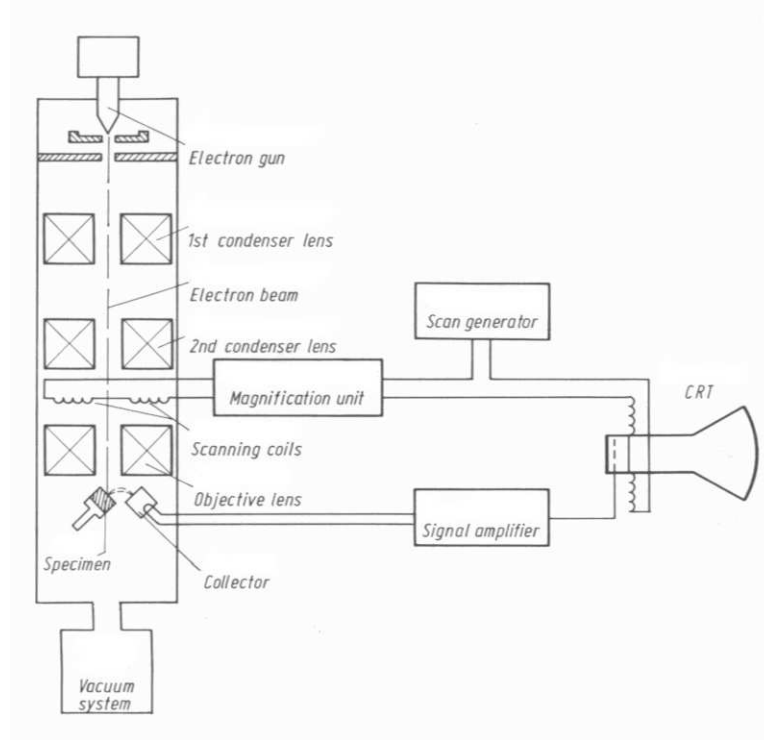


Figure 11: Block circuit diagram of a scanning electron microscope (modified from [25])

of the position of the z piezoelectric drive correspond directly to the height variations on the sample surface and can be used to create an image. This mode can be used on flat as well as on rough surfaces but it is slower than the constant height mode.

Another possibility for imaging the sample surface is the *dynamic mode* (also called tapping mode or AC mode). In this case the cantilever is set to oscillation by a piezoelectric element at a frequency near its resonance frequency. The amplitude of the cantilever oscillation is detected via the photo diode. If the cantilever approaches the sample surface and forces act, the amplitude changes. Similar to the constant force mode, the z position of the sample is modified while scanning in the x and y directions to maintain the amplitude change constant. Hence the distance between the cantilever tip and the surface remains constant and the position variations of the z piezoelectric drive can be used to create an image of the surface topography. Since in dynamic mode the force acting on the sample surface is smaller than in contact mode, the former is especially useful for imaging soft and delicate samples. The advantage of the contact mode lies in its higher lateral resolution [27].

3.6 Scanning electron microscopy

The principle of a scanning electron microscope (SEM) is represented by figure 11. The whole system has to be evacuated to below 10^{-5} millibar [25] to achieve long enough mean free pathways for the electrons. In an electron gun a beam of so-called *primary electrons* (PE) with typical energies of about 10^4 electron volts [25] is produced by thermal emission or field emission. By a column of lenses it is focused on the investigated specimen. Scanning coils scan the primary beam over the surface of the specimen. The scan generator

synchronizes the scanning of the primary beam to the formation of an image on a computer screen (in former times on a cathode ray tube (CRT)). This way the scanning image is faithfully reproduced on the screen. When the primary electrons hit the surface of a specimen, several interactions can occur. For imaging only the produced secondary electrons of low energies (0 - 50 eV [25]) and backscattered electrons of high energies (about 10^4 eV [25]) are used. A laterally placed, positively biased detector collects the electrons. The intensity of its output signal is used to modulate the brightness of the point on the screen corresponding to the point on the sample surface being scanned. Due to their small energy, secondary electrons can only be emitted from a specimen if they are produced very close to the surface. Backscattered electrons can also leave the specimen from deeper layers. They contribute to image formation mainly by releasing more secondary electrons on their way back to the surface. The main cause of contrast in the created image is the number of secondary electrons created close to the surface. It is influenced by the following properties of the specimen:

- Tilting: Surfaces non-perpendicular to the primary beam are brighter than surfaces perpendicular to it.
- Topography: Spikes and edges appear brighter than plane surfaces.
- Electric charge: Negatively charged regions are brighter than positively charged ones.
- Chemical composition: Heavy elements appear brighter than light ones.

If the latter two effects are negligible, the image creates a three dimensional impression of the specimen.

The main advantages of SEM imaging over optical microscopy imaging are an about 300 times greater depth of focus [25] due to the small illumination aperture (0.06° to 0.6°) of the primary beam and a point resolution of down to two nanometres [25].

3.7 Principal component analysis

Mass spectra of proteins are very complex, because every protein can decompose into several different fragments visible as peaks in the mass spectra. Considering only the masses from 1 amu/z (atomic mass unit per charge number) to 200 amu/z, every measured spectrum can be interpreted as a point in a two hundred dimensional space spanned by unit vectors corresponding to the different masses. The intensity of a mass peak determines its projection onto the corresponding vector. In this space, spectra of different proteins should be grouped at different sites due to their similarity. Unfortunately it is not possible to visualize such a high dimensional space.

Principal component analysis (PCA) is a method of linear algebra with the goal to concentrate the information contained in a data set onto a small number of variables called principal components (PC).

Comprehensive information on PCA can be found in the books by Jackson [17] and Jolliffe [18]. The following explanation is based upon the tutorial by Shlens [29].

3.7.1 Change of basis

Let a mass spectrum be represented by a column vector $\vec{x}$ of length m containing as elements the intensities at the masses from 1 amu/z to m amu/z. Out of n spectra one can construct an $m \times n$ data matrix X :

$$X = [\vec{x}_1 \cdots \vec{x}_n]. \quad (20)$$

PCA searches a basis to represent the data set in a way that concentrates the information onto just a few of the new basis vectors. The new basis vectors should be linear combinations of the old ones, to allow the use of the well established techniques of linear algebra. Thus a transformation matrix P is needed to transform the data matrix into a new $m \times n$ matrix Y :

$$Y = PX. \quad (21)$$

The row vectors $p_1 \cdots p_m$ of P are the new basis vectors and by definition the principal components of X .

3.7.2 Variance and covariance

A well transformed matrix Y should fulfil mainly two criteria: Firstly, it should contain as much relevant information as possible. Assuming the measured data are not too noisy, directions in the data space showing a large variance are the interesting ones. By assumption, the effects of interest show stronger dynamics than the measurement noise. Secondly, there should not be any redundancy. Since the number of relevant variables should be as small as possible, mutual dependencies between them have to be eliminated.

The mutual dependency of two variables is characterised by their covariance. Let a and b be two mean centred row vectors of the data matrix X . Their elements are thus the intensities at one mass in different measurements minus the mean intensity at the respective mass. Their covariance σ_{ab}^2 is defined as

$$\sigma_{ab}^2 := \frac{1}{n-1} ab^T. \quad (22)$$

b^T is the transposed vector of b . The covariance fulfils the following:

- $\sigma_{ab}^2 \geq 0$. The covariance is zero if and only if the variables are uncorrelated.
- $a = b \Rightarrow \sigma_{ab}^2 = \sigma_a^2$. The covariance of a variable with itself is its variance.
- $\sigma_{ab}^2 = \sigma_{ba}^2$. The covariance is symmetric with respect to the variables.

Further explanations concerning the concept of covariance can be found in the tutorial on PCA by Smith [30].

Since the x_i are the rows of the data matrix X , one can generalise the definition of the covariance to a covariance matrix C_X :

$$C_X := \frac{1}{n-1} XX^T. \quad (23)$$

It has the following properties:

- C_X is a symmetric $m \times m$ matrix.
- The diagonal elements are the variances of the variables contained in X .
- The off-diagonal elements are the covariances between the variables contained in X .

To minimize covariance, the covariance matrix of the data in the transformed basis C_Y has to be diagonal. This way all covariances are zero which is their smallest possible value.

3.7.3 Diagonalisation of the covariance matrix

A matrix P of basis vectors $p_1 \cdots p_m$ is searched to diagonalise the covariance matrix C_Y of $Y = PX$. The basis vectors are ordered by their importance by assuming that vectors with the highest variance are the most important ones. As a further assumption, PCA postulates orthonormality of the new basis vectors. This is to make the solution achievable by methods of linear algebra. By this assumption, P is an orthogonal matrix ($P^T = P^{-1}$: The transposed matrix of P is its inverse). One can express the new covariance matrix C_Y as a function of the searched matrix P :

$$C_Y = \frac{1}{n-1} Y Y^T \quad // \quad Y = PX \quad (24)$$

$$= \frac{1}{n-1} P X X^T P^T \quad // \quad C_X = \frac{1}{n-1} X X^T \quad (25)$$

$$= P C_X P^T. \quad (26)$$

The covariance matrix C_X is symmetric by its construction:

$$C_X^T = \left(\frac{1}{n-1} X X^T \right)^T = \frac{1}{n-1} X^{TT} X^T = \frac{1}{n-1} X X^T = C_X. \quad (27)$$

Every symmetric matrix can be diagonalised by an orthogonal matrix O . The columns of O are the eigenvectors of C_X . The elements of the resulting diagonal matrix D are the eigenvalues of C_X (a proof can be found in [29]):

$$C_X = O D O^T. \quad (28)$$

A matrix C_X of rank $r \leq m$ has exactly r orthonormal eigenvectors. If C_X is degenerate (i. e. $r < m$), one can freely choose $m - r$ additional orthonormal vectors to fill up O . These do not influence the solution because their eigenvalues are zero.

Let the row vectors p_i of P be the eigenvectors of C_X . Consequently $P = O^T$ holds. Hence from equation(28) it follows that

$$C_X = P^T D P. \quad (29)$$

With this, one can substitute C_X in equation (26):

$$C_Y = P P^T D P P^T \quad // \quad P^T = P^{-1} \quad (30)$$

$$= D \quad (31)$$

Obviously the covariance matrix is diagonal by employing this P . Its diagonal elements are the variances of the respective variables. The solution of PCA can be summarised as follows:

- The mean centred $m \times n$ data matrix X is transformed by the orthogonal matrix P to a new $m \times n$ matrix Y ($Y = PX$). Its elements are the so-called *scores* of the n measurements on the m new basis vectors, the principal components. The scores link the measurements to the principal components.
- To calculate the principal components, the covariance matrix C_X of X is used: $C_X = \frac{1}{n-1}XX^T$.
- The eigenvectors of C_X are the principal components of X and the row vectors of P . The elements of the PC, called *loadings*, show how strongly the initial variables (masses) influence the PC. Loadings link the initial variables to the principal components.
- The eigenvalues of C_X are the diagonal elements of C_Y . Their values are the variances of the respective principal components.
- The eigenvectors and -values are ordered by descending variance, that is by descending eigenvalues.
- This way, the relevant information is concentrated onto the first principal components.

3.7.4 Graphical representation

To visualise the results of principal component analysis, one usually plots the first two or three scores vectors and loading vectors. In the scores plot, the positions of the measurements in the space spanned by the first principal components are visible. Under favourable circumstances the measurements are grouped according to their sample type in this space. The corresponding loadings plot shows which of the initial variables are responsible for the positioning of the measurements in the scores plot.

This shall be clarified by an example. Here, the “measurements” are ten simulated mass spectra each of acetone, ethanol and a 1:1 mixture of the two. The scores and loadings on the first two PC are given in figure 12. The percent values at the axes are the relative variances explained by an axis.

The scores plot shows that the first principal component separates the three sample types. The spread on the second PC is caused by the equally simulated noise of the measurements. In the loadings plot one can see that the first PC is dominated by the masses 58 amu/z (acetone) in the negative direction and by 46 amu/z (ethanol) in the positive direction. Accordingly, in the scores plot acetone shows negative values on the first PC and ethanol shows positive values.

In the representation of a two dimensional scores plot, it is possible to draw a probability ellipse around a group of data points. Therefore one has to do once again PCA on these data represented in the space of the first two principal components of the prior analysis. The centre of the ellipse is given by the mean values of the data points on the two axes. The two new principal components define the directions of the axes of the ellipse. The first principal component gives the direction of highest variance, the major axis of the ellipse. The second principal component is by definition orthonormal to the

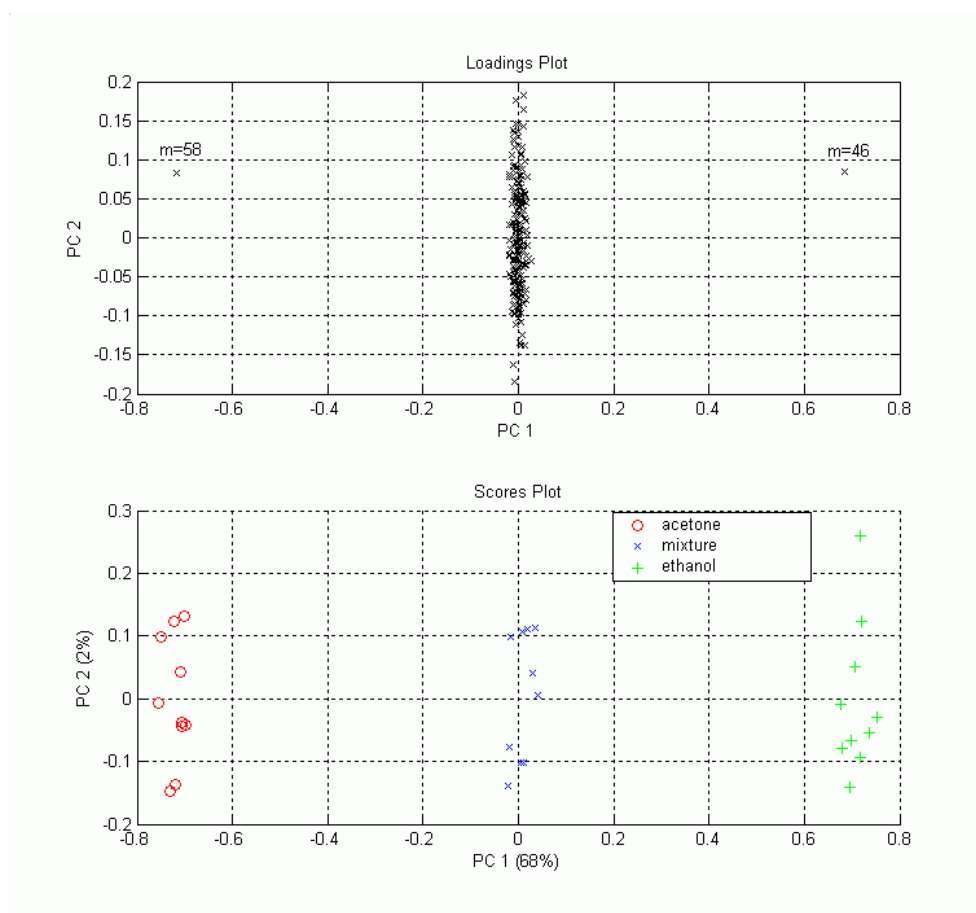


Figure 12: Results of PCA on simulated data

first one and gives the direction of the minor axis. The eigenvalues λ_1 and λ_2 associated to the principal components are used to calculate the lengths of the axes. The length of the semimajor axis is given by

$$l_1 = \sqrt{\lambda_1 T_\alpha^2}. \quad (32)$$

The length of the semiminor axis is

$$l_2 = \sqrt{\lambda_2 T_\alpha^2}. \quad (33)$$

Here T_α^2 is the critical value of the T^2 distribution which is related to the critical value $F_{p,n-p,\alpha}$ of the F distribution by

$$T_{\alpha,n,p}^2 = \frac{p(n-1)}{n-p} F_{p,n-p,\alpha}. \quad (34)$$

p is the number of variables (here $p = 2$) and n is the number of measurements in the data set. The T^2 statistic predicts that a data point of the considered sample can be found with a probability of $P = 1 - \alpha$ in the so defined ellipse [17]. In this work a level of significance $\alpha = 0.05$ is used. The F distribution is a probability density function of a continuous positive random variable x with two independent degrees of freedom a and b as parameters. In the case of the probability ellipses the parameters a and b correspond to the number of variables p and the reduced number of measurements $n - p$. The F distribution is described by the following formula [16]:

$$F(x|a, b) = \frac{\Gamma(\frac{a}{2} + \frac{b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \frac{x^{\frac{a}{2}-1}}{(1 + \frac{a}{b}x)^{\frac{a+b}{2}}}. \quad (35)$$

$\Gamma(x)$ is the gamma function given by

$$\Gamma(x) = \int_0^\infty t^{x-1} \exp(-t) dt. \quad (36)$$

Further properties of the F distribution can be found in the book “Statistik” by Hartung [16]. The critical value of the F distribution $F_{a,b,\alpha}$ is defined by

$$\int_0^{F_{a,b,\alpha}} F(x|a, b) dx = 1 - \alpha. \quad (37)$$

Therefore the random variable has a probability of $1 - \alpha$ to be smaller than the critical value of the F distribution. Figure 13 illustrates this for the case $a = 2$ and $b = 14$. The hatched area represents the probability of 95% for x to be smaller than the critical value $F_{2,14,0.05} = 3.74$. Critical values of the F distribution are tabulated for example in the “User’s guide to principal components” by Jackson [17].

3.7.5 Discriminating principal component analysis

The goal of this work is to recognize different adsorbed proteins by their mass spectra. Therefore a representation of the data stressing the differences between the proteins is necessary. Yet, PCA searches the components with largest variance regardless of whether this is the variance between spectra of one protein (“internal” variance) or between the spectra of different proteins (“external” variance).

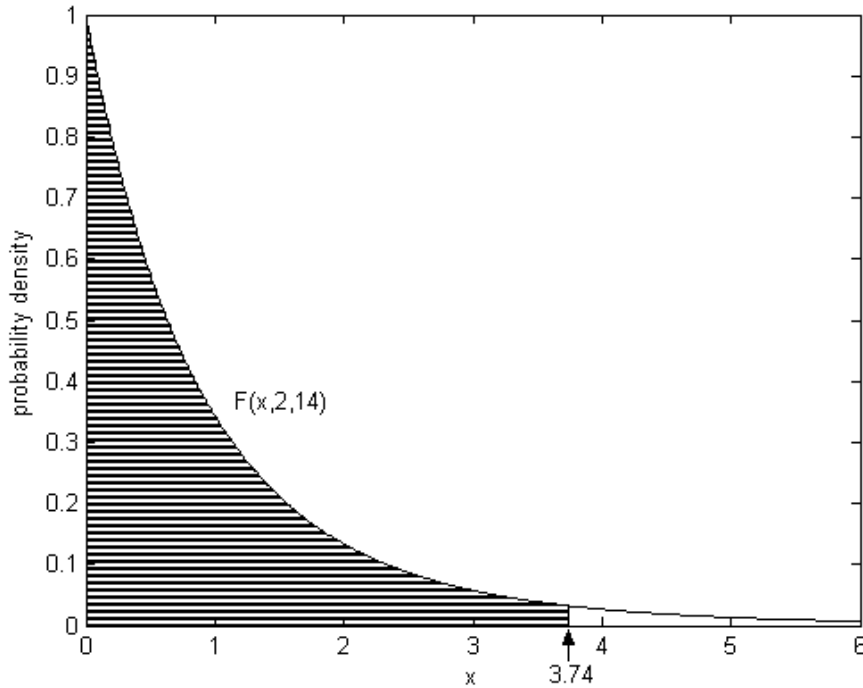


Figure 13: Example of an F distribution with the parameters $a = 2$ and $b = 14$ and its critical value $F_{2,14,0.05} = 3.74$ for a significance level of $\alpha = 0.05$

To achieve a good discrimination between the different sample types, it would be better to maximise the relation of external variance to internal variance. This is the goal of discriminant principal component analysis (DPCA) as described by Yendle [37].

A simple approach to DPCA would be to replace the spectra of one protein by the mean spectrum of this protein. Thus the internal variance would be zero and normal PCA would find the axes of highest external variance. With the resulting principal components one could transform the initial data matrix.

However, in the initial data the axes of highest external variance could be the ones of highest internal variance as well. Hence the simple approach would not maximise just the distance between different proteins but also the spread within the protein groups. Thus the internal variance has to be taken into account to calculate the principal components.

Therefore the initial variables (masses) are scaled by their total internal standard deviation. That is, first one calculates the internal variances of each variable for all the different proteins. These are summed to form the total internal variance of each variable. By root extraction one obtains the total internal standard deviations to scale the values of the respective variables. This way variables with a small internal variance are scaled up and vice versa.

In the resulting scaled data set all the spectra of a protein are replaced by its mean spectrum. Afterwards a normal PCA is done. The resulting principal components are the discriminating principal components (DPC) for transforming the scaled data set. The transformed data can be used to create scores plots that maximise the external variance while minimising the internal variance, to allow the best possible discrimination between the sample types.

The general disadvantage of DPCA is the need of a training data set with known

discrimination of the measurements into groups to calculate the principal components. Therefore DPCA is called a *supervised* technique of multivariate data analysis while PCA is an *unsupervised* technique. After having calculated the DPC, one can use these for transforming new measurement data, to see to which of the known groups these belong. The only condition is that the differences between training data and unknown data must not be too large.

Since the goal of this work is not the recognition of unknown patterns in the data but the attribution of measurements (mass spectra) to known groups (proteins), a training phase is also necessary in normal PCA to know, where in the scores plot the proteins can be found. Hence in this case the named disadvantage of DPCA does not matter.

3.7.6 Evaluation

In this work the *Leave-One-Out-Technique (LOO)* shall be used to evaluate how well a (D)PCA projection fits the data. As suggested by the name, one measurement is left out of the data set before performing (D)PCA. Afterwards this measurement is projected into the space spanned by the principal components using the loadings matrix.

Now the measurement can be assigned to one of the sample groups. This can be done by calculating the euclidean distances on the first principal components to the centres of gravity of the sample groups and assigning the left out measurement to the closest group. The great advantage of this method is its simplicity but on the other hand it cannot deal with outliers. Even samples that do not belong to any group are assigned to one. Another possibility is to draw probability ellipses around the sample groups on the first two principal components and to assign the left out measurement to a group if its projection lies in the corresponding ellipse. With this method outliers can be identified because they should not lie in any probability ellipse. The problem lies in the fact that projections can be found in the overlap of more than one ellipse. In this case no unambiguous assignment is possible.

The procedure is repeated with leaving out different measurements until all of them have been left out once. If most of the measurements are assigned to the correct group, the (D)PCA projection fits well the data and can be used to project unknown data and assign them to the known sample groups. If on the other hand many measurements are assigned to the wrong groups, the (D)PCA projection cannot be used to represent the data.

4 Experimental aspects

4.1 Chemicals

- Bovine serum albumin fraction V, approximately 99%, Sigma-Aldrich, Germany.
- Lysozyme from chicken egg white, approximately 95%, Sigma-Aldrich, Germany.
- α -Amylase from human saliva, Fluka BioChemika, USA.
- Bovine serum albumin fluorescein conjugate, Invitrogen, Germany
- Glutardialdehyde solution, 50% in water, Merck-Schuchardt, Germany.
- 3-Aminopropyl-tri(ethoxy)silane, minimum 98%, Sigma-Aldrich, Germany.
- Sodium di(hydrogen)phosphate monohydrate, minimum 99.5%, Fluka BioChemika, Switzerland.
- Disodium hydrogenphosphate, minimum 99%, Riedel-de Haën, Germany.
- Bidistilled water with a resistivity of 182 k Ω m from a “Milli-Q A10” water purification system, Millipore, USA.
- Dehydrated, denatured ethanol, Department of Chemistry, Technische Universität Kaiserslautern, Germany.
- Hydrogen peroxide, 35%, Department of Chemistry, Technische Universität Kaiserslautern, Germany.
- Sulphuric acid, 95 – 97%, Department of Chemistry, Technische Universität Kaiserslautern, Germany.
- Sodium hypochlorite solution, approximately 13% of active chlorine, Department of Chemistry, Technische Universität Kaiserslautern, Germany.
- Dehydrated toluene, work group of Professor Thiel, Department of Chemistry, Technische Universität Kaiserslautern, Germany.

4.2 Sample preparation

4.2.1 Dental material

The dental implant materials and the samples of bovine enamel are provided by the work group of Professor Matthias Hannig, Universitätsklinikum Homburg, in pieces of roughly five millimetres by five millimetres size. The dental implant materials are polymer matrices containing apatite particles. The two examined types are called FAT and FAW for fluoroapatite with a more transparent or more whitish appearance. They are made of the following substances:

- silanized (for FAW) or unsilanized (for FAT) fluoroapatite ($\text{Ca}_5(\text{PO}_4)_3\text{F}$) particles

- bis-phenol-A-glycidyl-di(methacrylate)
- tri(ethylene glycol)-di(methacrylate)
- poly(methacryl)oligo(maleic acid)
- camphor quinone
- strontium

The surfaces are polished. Before being used, the substrates are ultrasonically cleaned for five minutes in a solution of one percent of sodium hypochlorite. To remove a possible deposit of sodium hypochlorite, they are ultrasonically cleaned for another five minutes in bidistilled water and rinsed three times in bidistilled water.

4.2.2 Protein films on silanised substrates

Approximately ten millimetres by five millimetres large silicon wafer pieces serve as substrate in this case. First they are cleaned for fifteen minutes in a solution consisting to two thirds of concentrated sulphuric acid and to one third of a thirty-five percent solution of hydrogen peroxide. This cleaning solution is called *piranha solution* for its ability to remove most organic matter. Afterwards the substrates are rinsed with bidistilled water and dried with nitrogen. Then they are placed for one hour under protective gas in a mixture of 20 ml of dry toluene and 0.5 ml of 3-aminopropyl-tri(ethoxy)silane (APTES). First the APTES hydrolises with the residual water in the solution to form a silanol. The latter can covalently bind to the oxide atoms of the silicon oxide to form a monolayer as shown in figure 14. At the same time the silane can also polymerise. To remove polymerised silane from the surface, the samples are given into an ultrasonic bath for fifteen minutes in ethanol. Then they are rinsed with bidistilled water. Silanised substrates are imaged with a scanning electron microscope (SEM) by Dr. Stefan Trellenkamp, Nano+BioCenter Kaiserslautern. When the samples are not sonicated in ethanol, many aggregates of polymerised silane can be found on the surface (see figure 15). Their number is strongly reduced after sonication.

Glutardialdehyde serves as link between the silane and the protein. It shall bind to the silane as shown in figure 16. Therefore the samples are placed for one hour in a five percent (volume / volume) solution of glutardialdehyde in pH 7 buffer solution. The pH-value of the buffer solution is adjusted by mixing a 0.1 molar solution of sodium di(hydrogen)phosphate (NaH_2PO_4) and a 0.1 molar solution of disodium hydrogenphosphate (Na_2HPO_4). It is controlled with a pH electrode "Testo 252" by "Testo", Germany. After the adsorption of glutardialdehyde the samples are rinsed with buffer solution. The actual protein adsorption also takes place in pH 7 buffer solution for one hour with protein concentrations of two grammes per litre. All adsorption steps are effectuated at room temperature. To remove loosely bound protein, the samples are rinsed with fresh buffer solution. In a last step they are rinsed another three times with bidistilled water to remove buffer salts and dried with argon.

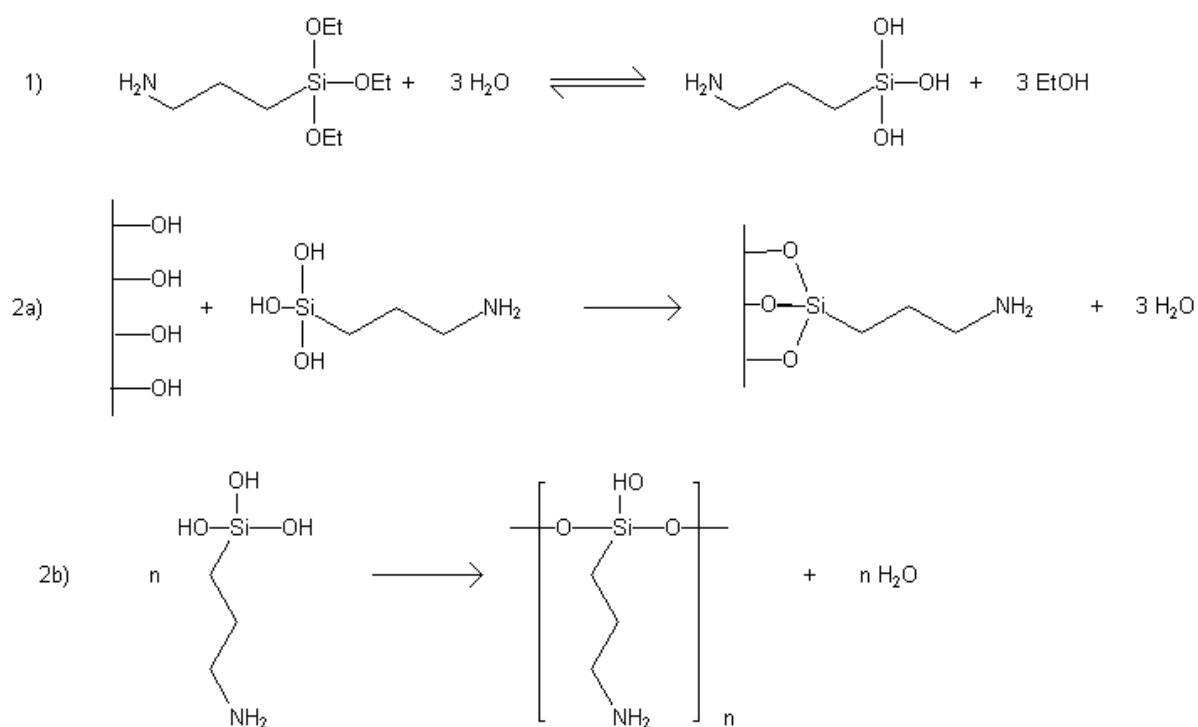


Figure 14: Hydrolysis (1), binding (2a) and polymerisation (2b) of APTES

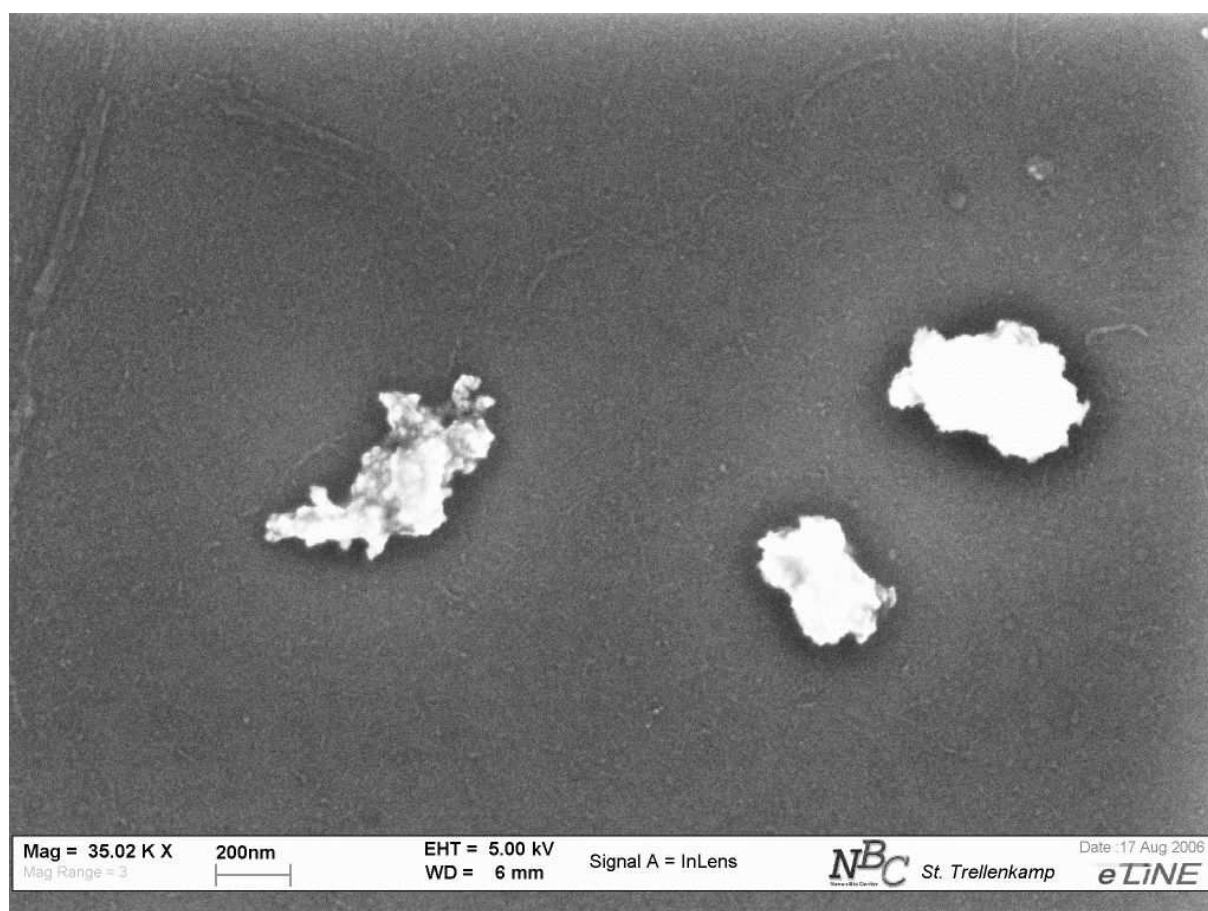


Figure 15: SEM image of a silanized substrate with aggregates of polymerised silane

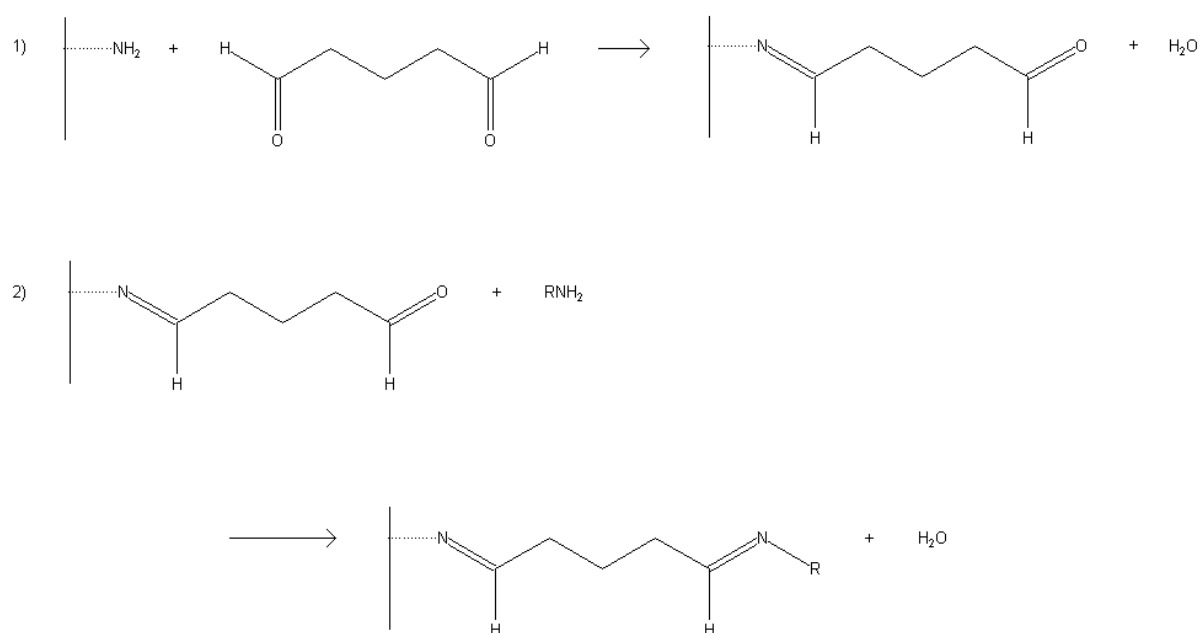


Figure 16: Binding of glutardialdehyde to the silane (1) and of a protein R to glutardialdehyde (2)

4.2.3 Protein films on silicon substrates

To prevent the possible influence of polymerised silane on the protein mass spectra, protein films are prepared directly on silicon. As in the previous section ten by five millimetre silicon wafer pieces are used as substrates. These are cleaned for fifteen minutes in piranha solution. Afterwards the substrates are rinsed with bidistilled water and given into protein solutions with a concentration of one to five grammes per litre. The solvent for the proteins is either bidistilled water or a pH 7 phosphate buffer as described above. After two hours of adsorption the samples are rinsed three times with bidistilled water to remove loosely bound protein and buffer salts and dried with argon.

4.2.4 Protein films on dental implant materials

Substrates of the dental implant materials FAT and FAW are ultrasonically cleaned for ten minutes in a two percent solution of sodium hypochlorite (NaOCl). Afterwards the substrates are given for another five minutes in an ultrasonic bath in bidistilled water and rinsed with bidistilled water to remove remainders of sodium hypochlorite. The proteins are solved with molarities of about 10^{-4} mols of protein per litre pH 7.4 buffer solution, which is prepared from 0.01 molar solutions of disodium hydrogenphosphate and sodium di(hydrogen)phosphate as described above. For protein adsorption the cleaned substrates are given into protein solutions for two hours at room temperature. Then the samples are rinsed three times with bidistilled water to remove buffer salts and loosely bound protein and dried with argon.

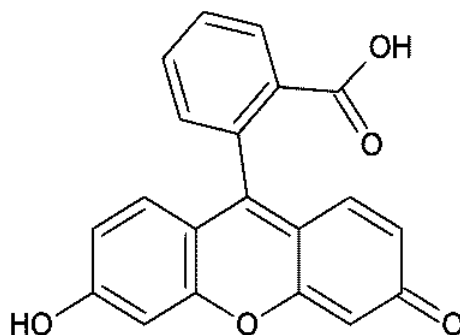


Figure 17: Chemical structure of fluorescein

4.2.5 Samples for fluorescence microscopy

Protein layers are prepared on silicon substrates for fluorescence microscopy. The substrate pieces of ten by five millimetres size are cleaned for twenty minutes in piranha solution and rinsed with bidistilled water. Afterwards proteins are adsorbed from a phosphate buffer solution with pH 7.4. Either only BSA conjugated with the fluorophore fluorescein ($C_{20}H_{12}O_5$, see figure 17 for chemical structure) or a 1:1 (weight:weight) mixture of BSA fluorescein conjugate and non-marked lysozyme are adsorbed for two hours at room temperature. According to the manufacturer Molecular Probes (USA) the absorption and emission maxima of the BSA fluorescein conjugate are at wavelengths of 494 nm and 520 nm. After protein adsorption the samples are either rinsed with bidistilled water and air-dried or air-dried without prior rinsing.

On one sample the protein is adsorbed onto a silane layer. Therefore the cleaned substrate is given into a solution of 0.5 ml of 3-aminopropyl-tri(ethoxy)silane (APTES) in 20 ml of dry toluene at 80 °C to 90 °C for two hours under protective gas. Then the sample is baked out at 150 °C for another two hours to favour the formation of an APTES monolayer. Next the sample is given into a 1% glutardialdehyde solution in pH 7.4 phosphate buffer for one hour. This way a monolayer of glutardialdehyde should build on the silane layer as described above. Before protein adsorption the sample is rinsed with bidistilled water.

4.3 Time-of-flight secondary ion mass spectrometry

All measurements are made with a “TRIFT II” (TRIPLE Focusing Time-of-flight) apparatus by “Physical Electronics”, USA. The use of this instrument was kindly allowed by its owner, the “Institut für Oberflächen- und Schichtanalytik, Kaiserslautern (IfOS). The ion source is a liquid metal gallium gun that produces primary ions with an energy of 25 kilo-electron volts. An electron flood gun compensates charging effects with electrons of an energy of 20 electron volts. The initial energy spread of the secondary ions is compensated by a system of three electrostatic 90° sector fields.

Since the gallium stock of the source ran out during the measurements, the source had to be replaced. Before the replacement the mass resolution was determined on a silicon wafer as $m/\text{FWHM} \approx 300$ at $m = 28$ amu/z. The mass resolution is defined as ratio between the mass m and the full width at half maximum (FWHM) of the corresponding

peak in the mass spectrum. With the new source the mass resolution is $m/\text{FWHM} \approx 420$ at $m = 28 \text{ amu/z}$.

Of all the samples mass spectra of cations and anions in the mass range of 1 amu/z to 400 amu/z are acquired at several sample sites. The scanning size is $A = (120 \text{ } \mu\text{m})^2$. The extractor current of the ion source is maintained at $1.5 \text{ } \mu\text{A}$ corresponding to a primary ion current of below one nanoampère in unpulsed mode. The acquisition time is usually five minutes.

The primary ion dose D can be calculated with the primary ion current I , the acquisition time t , the scanning size A , the repetition frequency f , the pulse length Δt and the elementary charge $e = 1.6 \cdot 10^{-19} \text{ C}$ by

$$D = \frac{Itf\Delta t}{eA}. \quad (38)$$

Here it is $I \approx 10^{-9} \text{ A}$, $t = 300 \text{ s}$, $f \approx 10^4 \text{ Hz}$ and $\Delta t = 12 \cdot 10^{-9} \text{ s}$. Hence the primary ion dose

$$D \approx 2 \cdot 10^{12} \text{ cm}^{-2} \quad (39)$$

lies well in the regime of static analysis (see equation (2)).

To see whether the surfaces of the dental implant materials are contaminated, they are also examined after sputtering with the unpulsed primary ion beam. For sputtering the scanning size is augmented to $A = (240 \mu\text{m})^2$ to exclude the influence of gating effects on the measurements. When sputtering a surface one does not obtain a crater with vertical walls. Instead, there is a transition region at the edges from the crater ground to the unsputtered surface. To obtain a measurement only from the crater ground, one has to choose the scanning size for analysis smaller than the one for sputtering.

In the protein films, depth profiles are acquired. Therefore the following two steps are alternated:

1. Acquisition of a mass spectrum as described above.
2. Unpulsed sputtering of the same sample site for five seconds with 1.44 square millimetres scanning size. This is the largest possible scanning size. It was chosen to obtain a slow removal of the protein layer.

Here only mass spectra of cations are acquired, because according to Wagner et al. [31] the spectra of anions do not allow a discrimination between different proteins, since they contain only signals of the unspecific protein backbone (mainly CN^- and CNO^-).

4.4 Fluorescence microscopy

The samples are imaged using an epifluorescence microscope “Axioskop 2 mot” (Carl Zeiss, Germany) equipped with a CCD camera “AxioCam HRc” (Carl Zeiss, Germany). The samples are illuminated by a mercury lamp with a filter transmitting light with wavelengths of 450 nm to 490 nm. Fluorescence emission of the sample is detected for wavelengths longer than 515 nm. This setup suits well the fluorescence properties of the BSA fluorescein conjugate, because its maximum exciting and emission wavelengths are 494 nm and 520 nm (according to the manufacturer Molecular Probes, USA).

4.5 Scanning force microscopy

Measurements are carried out on a “Molecular Force Probe MFP-3D” scanning force microscope (Asylum Research, USA). The samples are imaged in air in contact mode with constant force acting on the cantilever tip. Scanning areas are varied between fifteen micrometres by fifteen micrometres and two micrometres by two micrometres.

4.6 Determination of amylase activity

The activity of amylase immobilized on silicon substrates is determined by the group of Dr. Christian Hannig, Universitätsklinikum Freiburg, Germany, with a method described in [13]. For these experiments silicon substrates with 25 square millimetres surface area are used. They are cleaned with piranha solution for fifteen minutes and rinsed with bidistilled water. Proteins are solved in ten millimolar phosphate buffer solutions with pH 7.4 at molarities of about $5 \cdot 10^{-5}$ mol per litre. On one sample only amylase is adsorbed for two hours. A second sample is given into a lysozyme solution for two hours, rinsed with buffer solution and given into an amylase solution afterwards for another two hours. The samples are rinsed three times with bidistilled water and dried with argon after the last adsorption step. Then they are sent to Freiburg in micro centrifuge tubes for activity measurements.

The determination of amylase activity is based upon its ability to directly hydrolyse the synthetic trisaccharide 2-chloro-4-nitrophenyl-4-O- β -D-galactopyranosylmaltotrioside (GalG2CNP) without any auxiliary enzymes. One product of this reaction is aglycone 2-chloro-4-nitrophenolate (CNP). According to [13], the formation of CNP is stoichiometric with respect to incubation time and occurs at a constant rate. CNP is detected photometrically by its absorption at a wavelength of 405 nanometres. The samples are incubated for 10 minutes in 300 microlitres test solution at a temperature of 25 °C. The test solution contains five millimol GalG2CNP per litre and is buffered at pH 6.0. Immediately after removal of the sample, the absorption is read against reagent blank at 405 nanometres. One unit of amylase activity is defined as hydrolytic production of 1 micromol CNP per minute. The absorbance A_0 of one micromol CNP in the given experimental setup can be determined independently. Thus the activity a can be calculated from the measured change in absorbance within ten minutes ΔA :

$$a = \frac{\Delta A}{10A_0}. \quad (40)$$

With the known surface S (25 mm²) of the samples, the immobilized activity per square centimetre of sample surface can be calculated as

$$\frac{a}{S} = \frac{\Delta A}{2.5 \text{ cm}^2 A_0} \quad (41)$$

4.7 Data analysis

The mass spectra are acquired and calibrated with the software CADENCE (version 2.0, Physical Electronics, USA). A newer version of the same software (WinCadence 3.7.1)

is used to export the data to unit mass mass spectra in ASCII format. These are analysed in a MATLAB 6.5 environment (The MathWorks, USA) with functions written for this purpose, which can be found in appendix A. First the mass spectra are extracted from the WinCadence export files with the functions “einlesen”, “voll” and “matrix”. A background spectrum can be subtracted (“silanweg”) and peaks related to amino acid fragments can be selected (“auswahl”) from the mass spectra. Principal component analysis with an algorithm based upon the PCA tutorial by Shlens [29] or discriminant principal component analysis with an algorithm based upon Yendle’s article [37] are performed using the functions “pca” or “dpca”. The results are plotted by the function “plotpc2d”. In addition to the data points, “plotdpc2dmitellipse2” also plots probability ellipses around groups in the scores plot. Therefore the function “ellipse” is used. To calculate the size of the probability ellipses, it needs the critical value of an F distribution provided by the function “Fdistribution”, which contains critical values taken from the “User’s guide to principal components” [17]. The quality of the DPCA models is evaluated by leave-one-out tests. The left out measurements are assigned to a group either by their projection’s position with respect to the probability ellipses of the groups (function “loo”) or by the euclidean distance of the projections to the groups’ centres of gravity (function “loo3”). New spectra can be projected into existing scores plots using the functions “pcaprojektion” and “dpcaprojektion”. They can be assigned to existing groups with the function “zuordnen”.

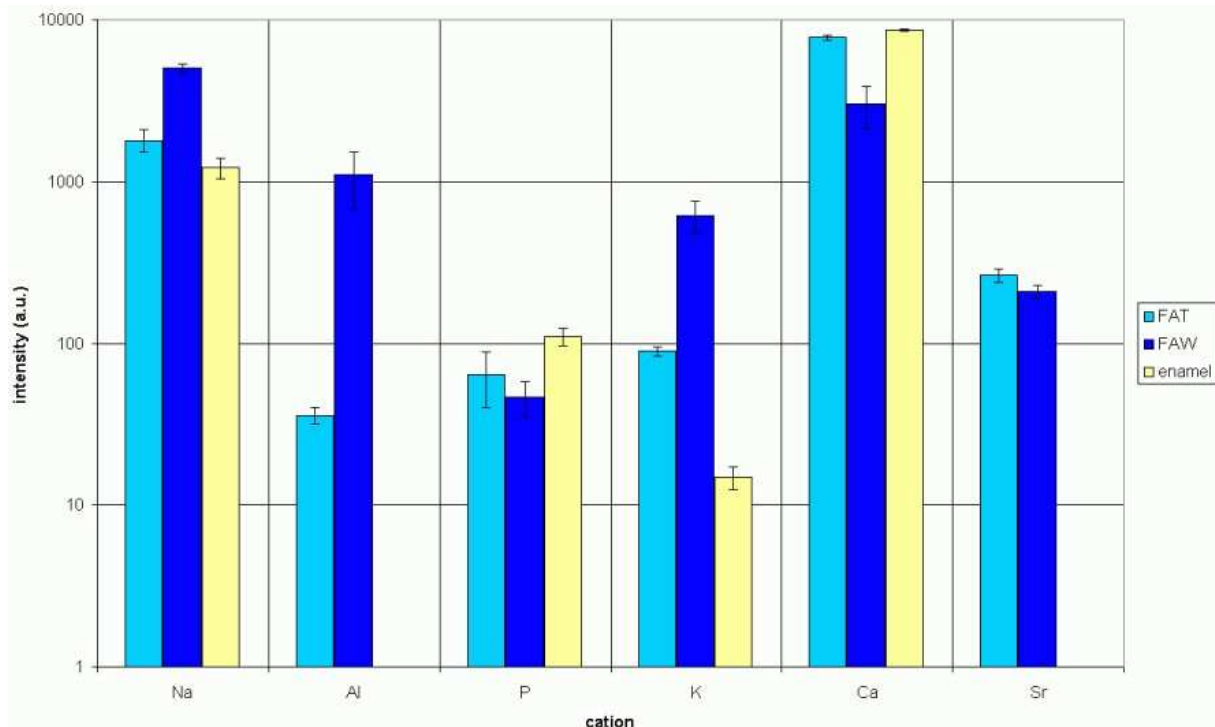


Figure 18: Intensities of cations in mass spectra acquired on dental materials

5 Results

5.1 ToF-SIMS of dental implant materials

Comparison of the mass spectra of the dental implant materials FAT and FAW before and after sputtering shows a significant decrease in intensity of the peaks assigned to organic compounds. Hence spectra acquired without prior sputtering are distorted by organic surface contaminants. These can be compounds not removed by the cleaning process or molecules adsorbed between cleaning and the analysis by ToF-SIMS. In the following, only spectra acquired with prior sputtering are used for analysis.

To compare the surface composition of the dental implant materials and bovine tooth enamel, mass spectra of cations and anions are acquired at three sites on samples of the implant materials FAT and FAW as well as bovine tooth enamel. The peaks showing the strongest signals in the mass spectra are selected and scaled by the total intensity of the selected peaks to compensate for a possible shift in the total secondary ion yield. Using their masses and isotope patterns, the peaks are assigned to the ions that caused them.

In the mass spectra of positively charged ions, six peaks show a noteworthy intensity. These are shown in figure 18. Comparing the two implant materials, FAW shows much higher intensities for the aluminium and potassium signals. The other four cations show comparable intensities for the two materials with a slightly stronger sodium signal and a little weaker phosphorus, calcium and strontium signals for FAW than for FAT. In contrast to the implant materials the enamel sample does not show any detectable signal for aluminium or strontium and a much weaker one for potassium. The signals of calcium and phosphorus are a little stronger and the one of sodium is slightly weaker than for the implant materials.

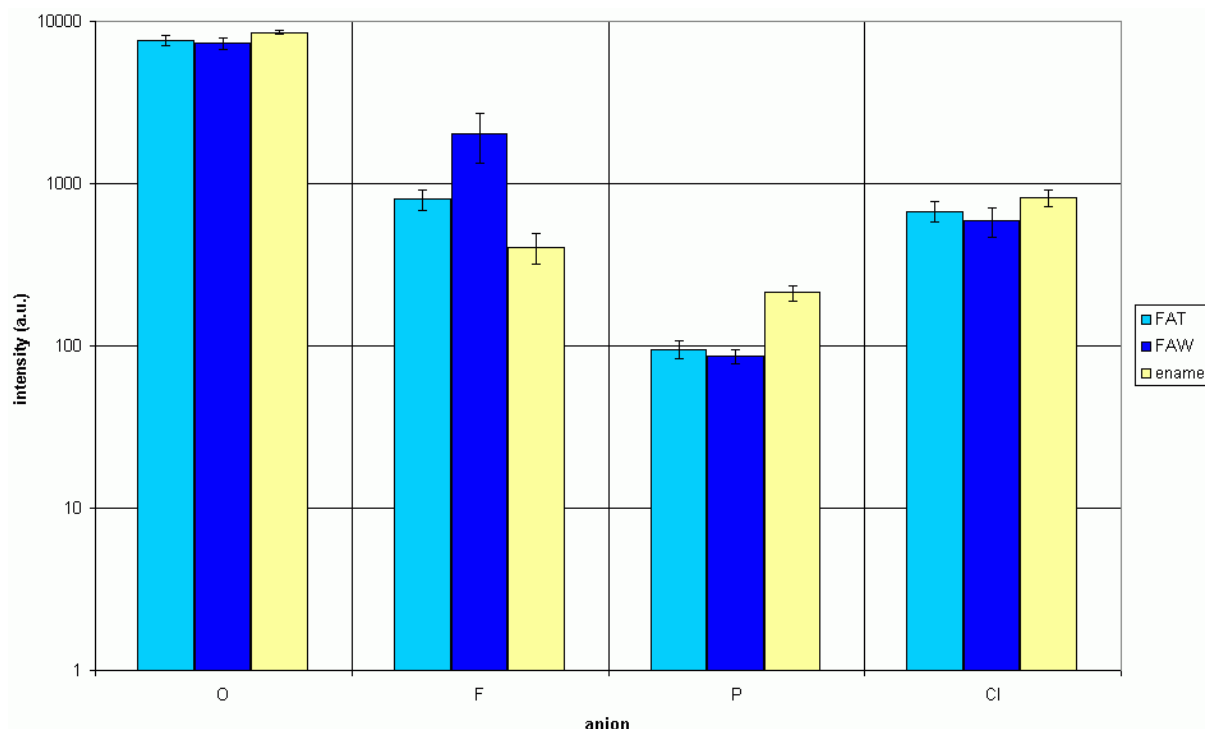


Figure 19: Intensities of anions in mass spectra acquired on dental materials

In the mass spectra of negatively charged ions, the peaks of four components of the samples can be identified. Their intensities are shown in figure 19. Here the two implant materials show quite similar intensities. The fluorine signal is a little stronger and the phosphorus (as in the spectra of cations) and chlorine signals are a little weaker for FAW than for FAT. The bovine enamel sample does not show any strongly differing behaviour either. Its fluorine signal is slightly weaker while the phosphorus and chlorine signals are stronger than for the implant materials. All the samples show similar intensities for the oxygen signal.

Due to the different matrix effects and ionisation probabilities, no exact conclusions to the surface composition of the samples can be drawn from these data. Anyway it can be stated that all three materials resemble in their main components. They contain principally sodium, phosphorus, potassium, calcium, oxygen, fluorine and chlorine. In the implant materials aluminium and strontium are also abundantly present. Strontium is incorporated in the implant materials to make them visible in X-ray imaging. The aluminium originates probably from the polishing of the samples.

5.2 ToF-SIMS of protein films on silanised substrates

Regarding the mass spectra of different protein films, there are no clearly visible differences between the different films. In all the spectra of negative ions the oxygen (16 amu/z), hydroxide (17 amu/z), carbon (12 amu/z) and hydrocarbon anions (13 amu/z) as well as the cyanide anion (CN^- : 26 amu/z) and C_2H^- (25 amu/z) with much weaker intensities are the only clearly visible peaks (see figure 20).

In the spectra of positive ions the peak of the silicon cation (28 amu/z) is the most

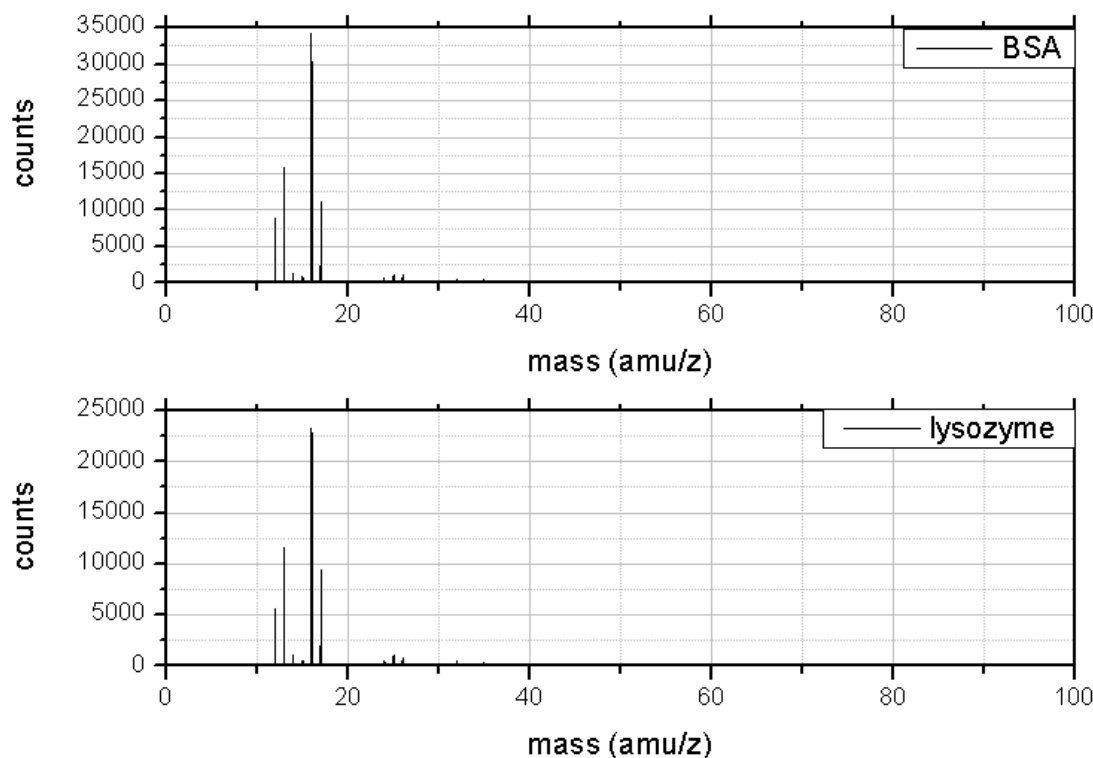


Figure 20: Exemplary mass spectra of anions of protein films adsorbed to silane

intense one. This is probably due to polymerised silane that was not removed from the surface by the cleaning step in ethanol. In the mass range from 27 amu/z to 70 amu/z nearly every mass shows a peak of noteworthy intensity. Furthermore up to 200 amu/z the density of peaks stays large (see figures 21 and 22).

Some sample sites exhibit a strong peak at 23 amu/z due to sodium from the buffer solution. At these sites no spectra for analysis are acquired because according to Wagner et al. [31] sodium causes a strong matrix effect that might distort the measurements.

For principal component analysis only the mass range between 1 amu/z and 200 amu/z is used. The spectra are scaled by their total intensity to compensate for changes in the total number of ions detected per spectrum.

By doing DPCA on the spectra of cations a discrimination of samples coated with lysozyme or BSA by their scores is hardly possible. Figure 23 shows the loadings and scores plots of three samples each coated with lysozyme or BSA on a silane and glutardialdehyde layer. On each sample mass spectra of cations were acquired at four or six sites. In addition to these data, spectra of cations acquired on three samples treated with a 1:1 (weight:weight) mixture solution of the two proteins are projected into the scores plot. Around the data points of lysozyme and BSA spectra, 95%-probability ellipses are drawn. As the ellipses show a large overlap, these results cannot be used to discriminate between the two proteins. In the loadings plot the masses showing strongest influence on the first discriminant principal component are labeled with their mass to charge ratio in atomic mass units per charge number (amu/z). One of these is 28 amu/z, the mass to charge

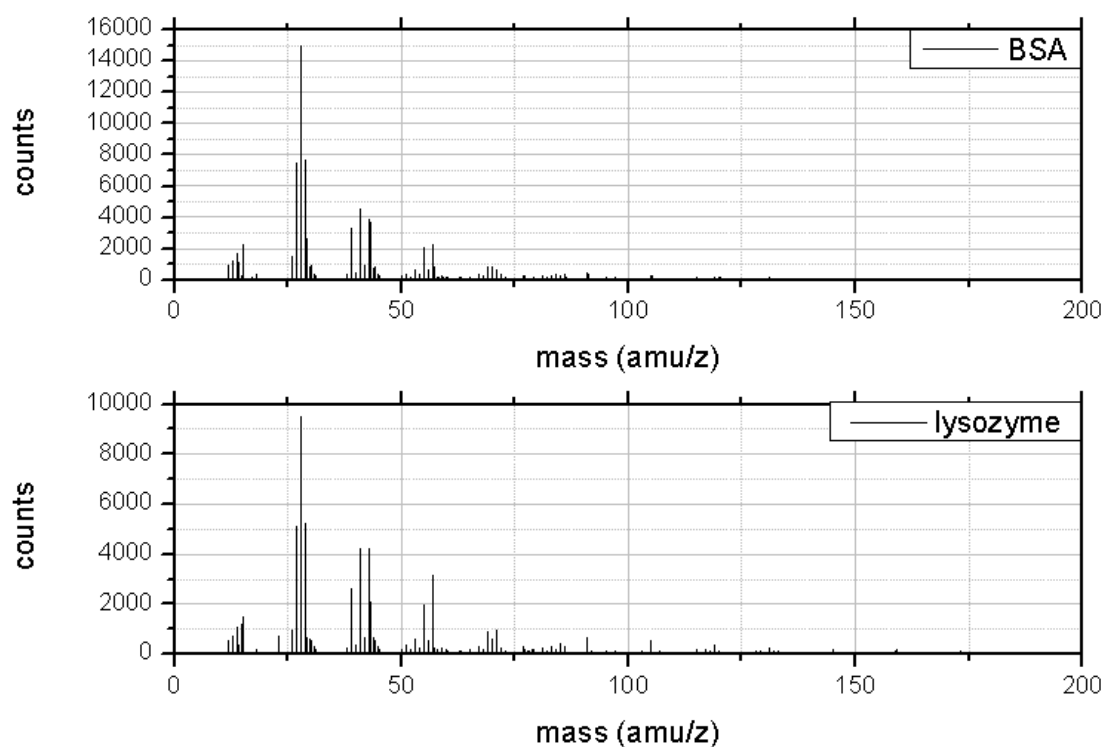


Figure 21: Exemplary mass spectra of cations of protein films adsorbed to silane

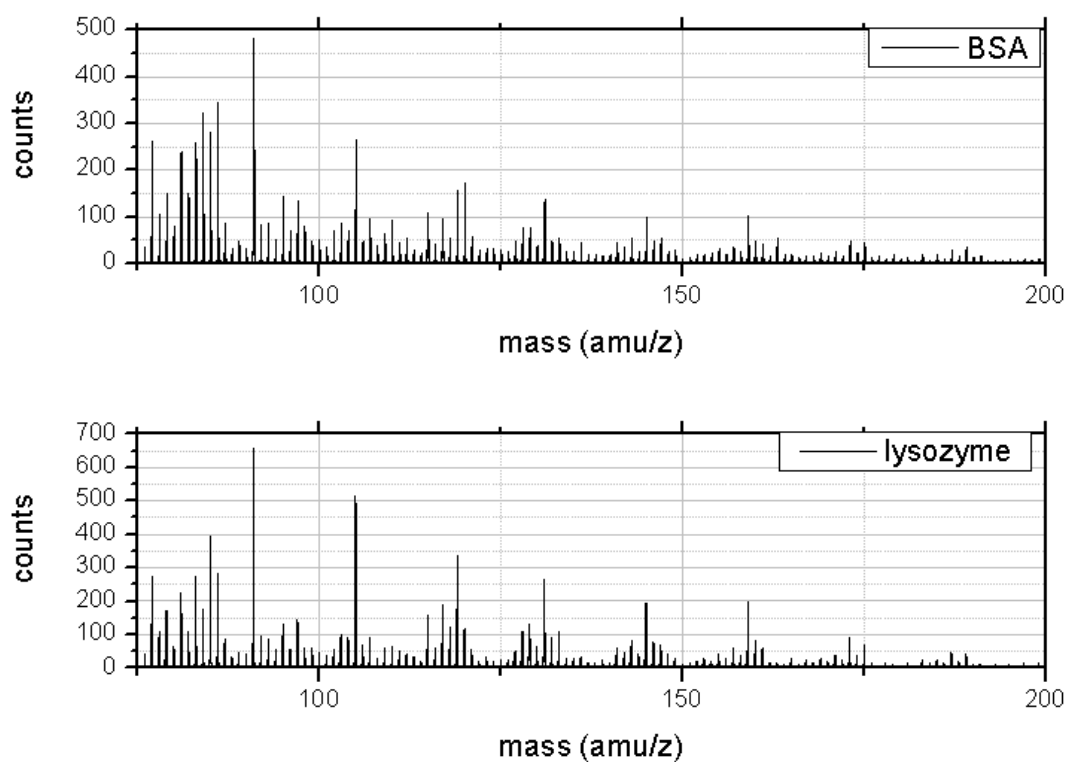


Figure 22: Exemplary mass spectra of cations of protein films adsorbed to silane

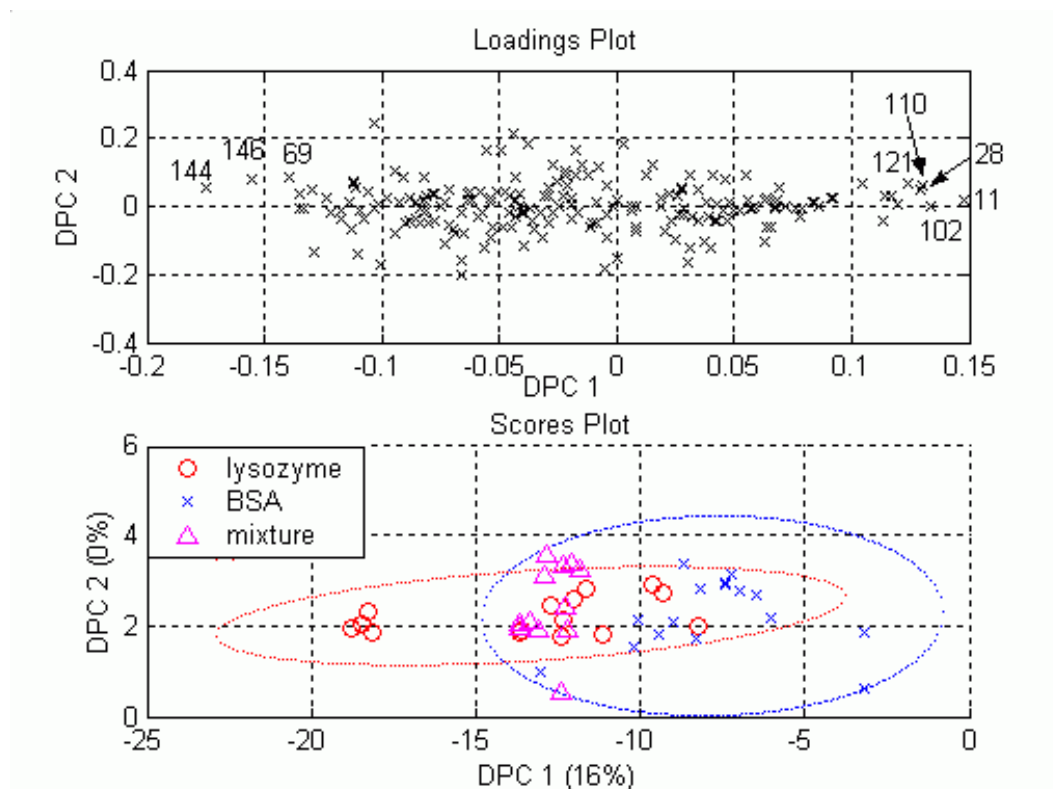


Figure 23: DPCA results of two proteins adsorbed to silane from buffer solution

ratio of the silicon cation. To exclude the influence of silicon from polymerised silane or from the substrate on the analysis, a mean mass spectrum of cations acquired on a sample coated only with silane is subtracted from the protein mass spectra.

The new results after subtraction of the silane mass spectrum are shown in figure 24. There still exists a large overlap of the probability ellipses.

To improve the results, in the next step only masses corresponding to intense peaks of amino acid mass spectra are taken into account. The selection was taken from the work of Lhoest et al. [20] and is listed in table 1. At the masses 30 amu/z and 45 amu/z there are also peaks visible in the mass spectra caused by the substrate, namely $^{30}\text{Si}^+$ and $^{28}\text{SiOH}^+$ that overlap with the peaks of amino acid fragments. Hence these masses are not used for analysis of protein films on silicon substrates.

In the new scores plot the overlap of the probability ellipses is a lot smaller than before (see figure 25) but there still is a large spread within the groups. Furthermore, the projections of the mass spectra of samples treated with a mixture of both proteins are spread all over the probability ellipses of the two proteins. This suggests that the composition of these samples is very variable. In the loadings plot the masses 102 amu/z and 110 amu/z show the strongest positive influence on the first discriminant principal component, while the masses 60 amu/z, 130 amu/z and 87 amu/z show the strongest negative influence. The former are assigned to the amino acids glutamic acid and histidine and the latter are assigned to serine, tryptophane and asparagine (see table 1).

The relative abundances of the different amino acids in lysozyme, BSA and amylase are plotted in figure 26. The values were taken from the work of Lhoest et al. [20] for lysozyme and BSA and from the Protein Data Bank [1] for amylase.

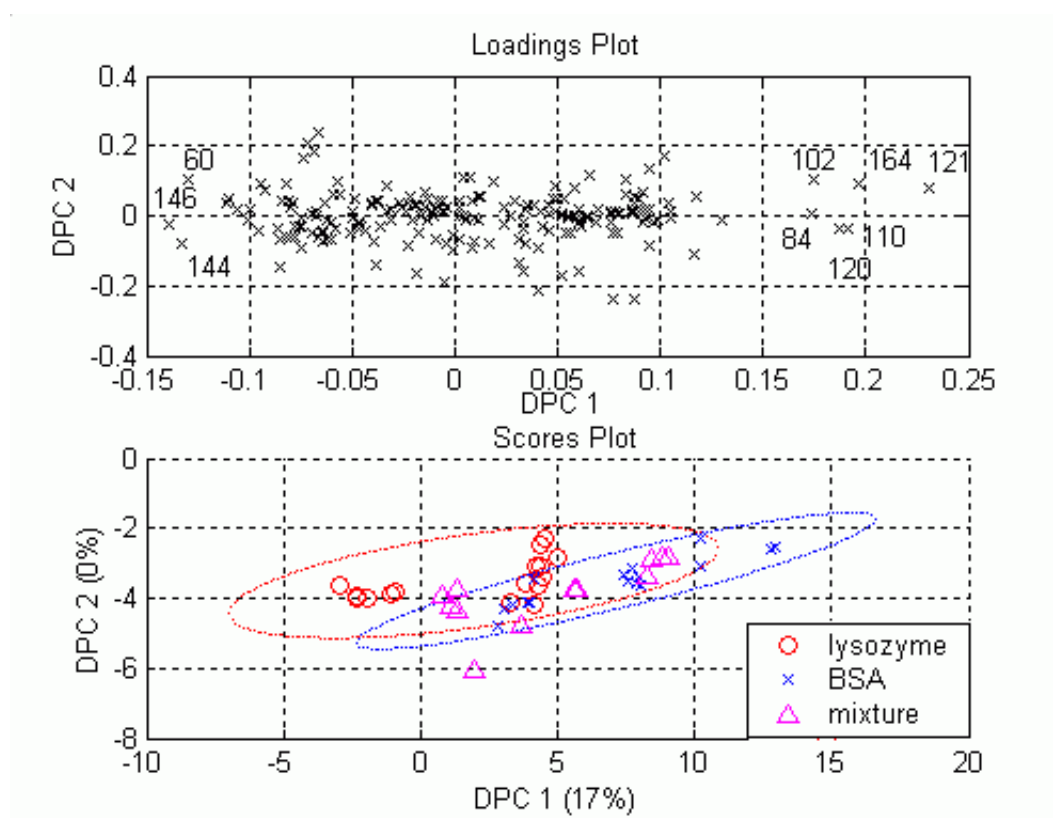


Figure 24: Results of DPCA after subtraction of a silane spectrum of two proteins adsorbed to silane from buffer solution

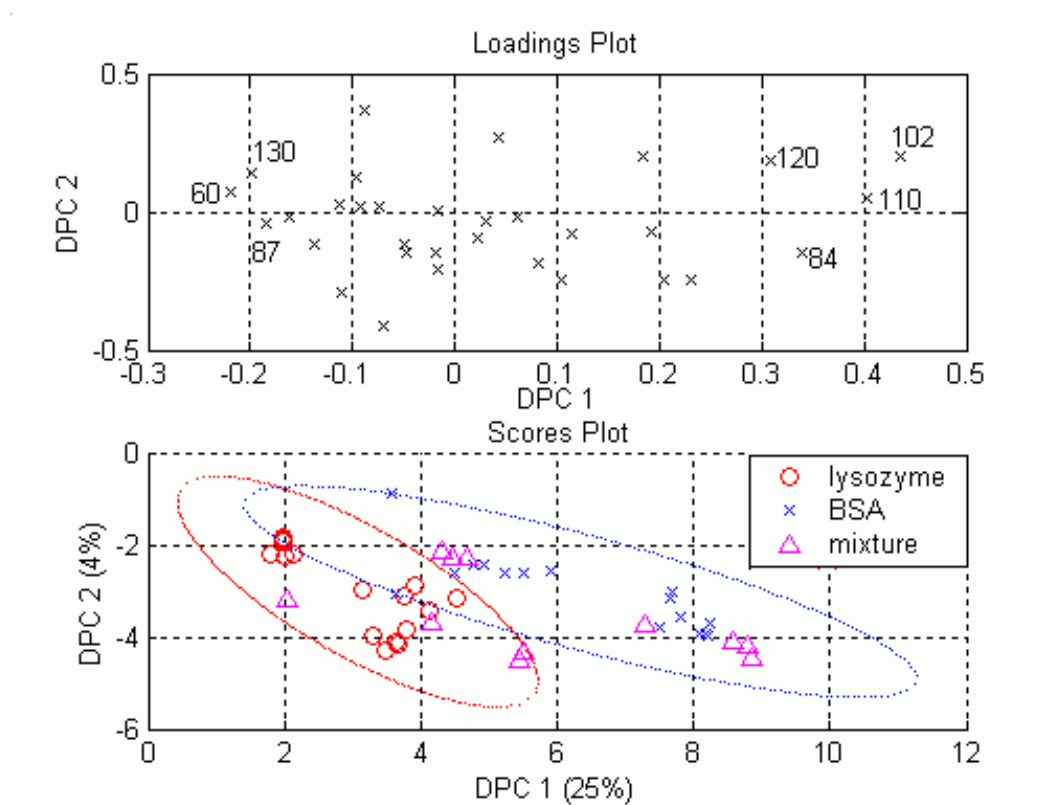


Figure 25: Results of DPCA after peak selection of two proteins adsorbed to silane from buffer solution

mass in amu/z	fragment	corresponding amino acid
30*	CH ₄ N	glycine
43	CH ₃ N ₂	arginine
44 [†]	C ₂ H ₆ N	alanine
45*	CHS	cysteine
60	C ₂ H ₆ NO	serine
61	C ₂ H ₅ S	methionine
68	C ₄ H ₆ N	proline
69 [†]	C ₄ H ₅ O	threonine
70	C ₃ H ₄ NO	asparagine
	C ₄ H ₈ N	proline
71	C ₃ H ₃ O ₂	serine
72	C ₄ H ₁₀ N	valine
73	C ₂ H ₇ N ₃	arginine
74	C ₃ H ₈ NO	threonine
81	C ₄ H ₅ N ₂	histidine
82	C ₄ H ₆ N ₂	histidine
83	C ₅ H ₇ O	valine
84	C ₄ H ₆ NO	glutamine, glutaminic acid
	C ₅ H ₁₀ N	lysine
86 [†]	C ₅ H ₁₂ N	leucine, isoleucine
87 [†]	C ₃ H ₇ N ₂ O	asparagine
88 [†]	C ₃ H ₆ NO ₂	asparagine, aspartic acid
98	C ₄ H ₄ NO ₂	asparagine
100	C ₄ H ₁₀ N ₃	arginine
101	C ₄ H ₁₁ N ₃	arginine
102	C ₄ H ₈ NO ₂	glutaminic acid
107	C ₇ H ₇ O	tyrosine
110	C ₅ H ₈ N ₃	histidine
112	C ₅ H ₈ N ₃	arginine
120	C ₈ H ₁₀ N	phenylalanine
127	C ₅ H ₁₁ N ₄	arginine
130	C ₉ H ₈ N	tryptophane
131	C ₉ H ₇ O	phenylalanine
136	C ₈ H ₁₀ NO	tyrosine
159	C ₁₀ H ₁₁ N	tryptophane
170	C ₁₁ H ₈ NO	tryptophane

Table 1: Selected amino acid fragment peaks. The peaks marked with * are omitted for samples prepared on silicon and the ones marked with [†] for samples prepared on FAT or FAW because of their overlap with substrate caused peaks.

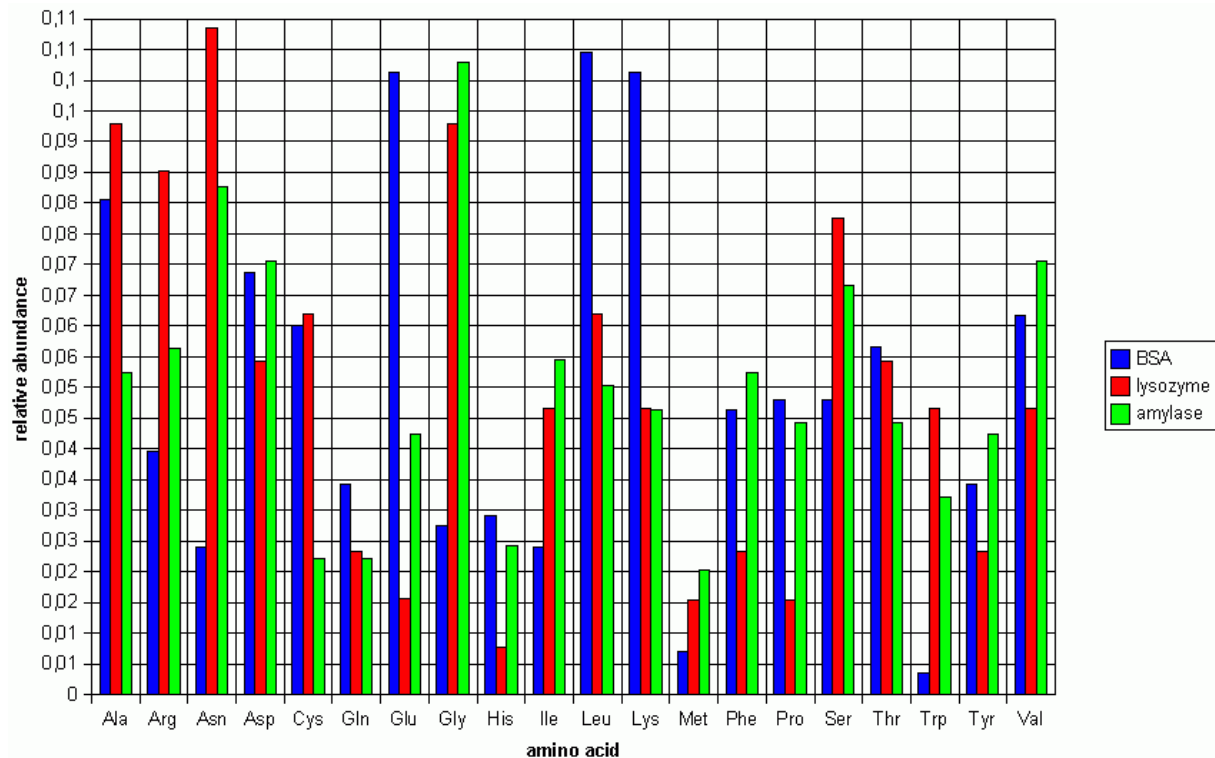


Figure 26: Relative abundances of amino acids in lysozyme, BSA and amylase

It can be seen that glutamic acid and histidine are more abundant in BSA than in lysozyme while the opposite is true for serine, tryptophane and asparagine. The amino acids associated to BSA show the strongest positive loadings on the first DPC, whereas the amino acids highly abundant in lysozyme show the strongest negative loadings on the first DPC. Therefore the loadings are consistent with the fact that BSA can be found on the right side of the scores plot and lysozyme on the left side.

To quantify the quality of the DPCA model of the spectral data, leave-one-out-tests are performed. When the samples are assigned to one of the two proteins by their euclidean distance on the first DPC to the centres of gravity of the two groups, 26 out of 32 spectra (81%) are correctly assigned. Using the probability ellipses, only 18 spectra (56%) are correctly assigned because many spectra are projected into the overlapping region of the two ellipses and thus cannot be assigned to one group. Hence a way for better discrimination of the sample types by reducing the spread between the mass spectra of one protein, has to be found.

Probably the spread within the sample types is due to different amounts of polymerised silane on the sample surfaces. To check this, protein films shall be adsorbed to unsilanised silicon wafers.

5.3 ToF-SIMS of protein films on silicon substrates

To the naked eye the mass spectra acquired from protein films on silicon do not significantly differ from those acquired from protein films on silane. The intensity of the silicon peak at 28 amu/z is even stronger than in the spectra of protein films on silane (see figure 27). Now it can only be caused by the silicon from the substrate. Since the sampling depth

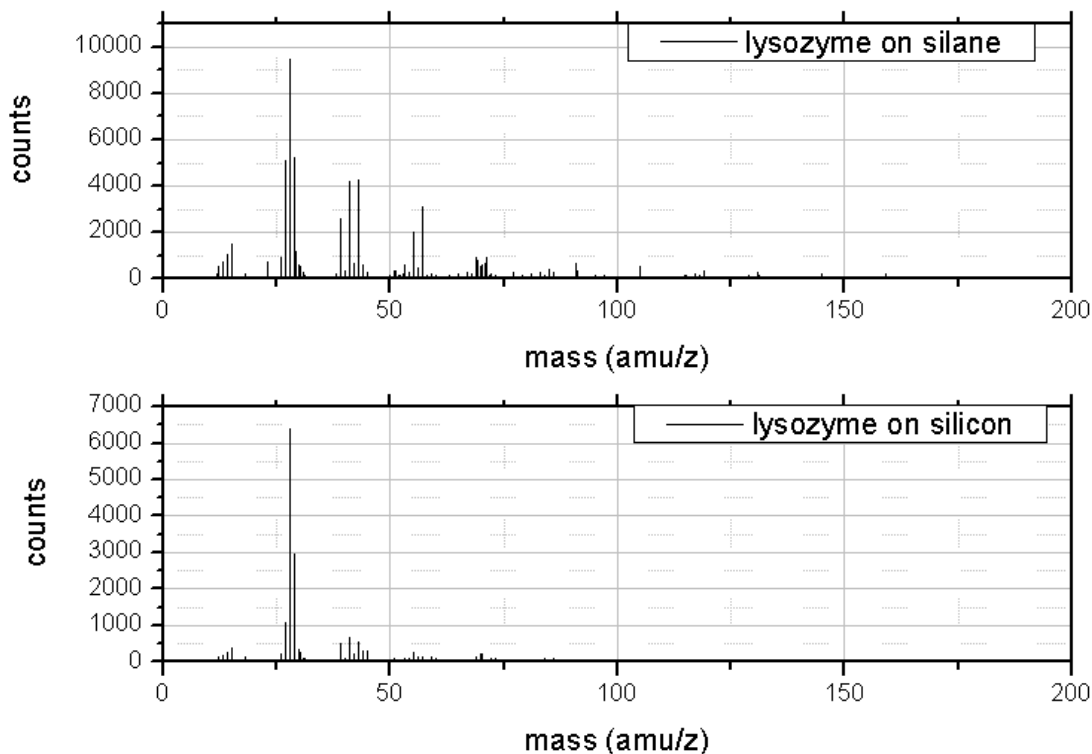


Figure 27: Comparison of mass spectra of cations from lysozyme films on silane or silicon

of ToF-SIMS is only a few Ångströms, no silicon should be detected through a compact protein layer of several nanometres thickness. Probably the protein film is porous due to removal of part of the protein in the rinsing step. In the higher mass region different peaks can be found in the spectra obtained from protein adsorbed to silicon, but still there are no characteristic peaks for the different proteins (see figure 28).

5.3.1 Proteins solved in water

First the spectra of proteins solved in water are treated. On four samples coated with lysozyme, BSA or amylase, mass spectra of cations are acquired at four sites. The spectra are unit-mass binned and amino acid related peaks (see table 1) are selected before doing discriminant principal component analysis. Additional silicon substrates are treated with solutions of 1:1 (weight:weight) mixtures of two of the three proteins or with a solution of a 1:1:1 (weight:weight:weight) mixture of all three proteins solved in water. On two samples of each mixture type, cation mass spectra are acquired at four sites. These undergo the same pretreatment as the spectra of pure protein films.

The results of DPCA are shown in figure 29. In the scores plot there is only little overlap of the 95% probability ellipses of lysozyme and amylase, while BSA is completely separated from the two other sample types. The mass spectra acquired on mixture treated samples are projected into the scores plot. The position of the ternary mixture and the mixture of lysozyme and BSA suggest that all proteins have adsorbed in these cases because the projections are found between the regions of the pure protein spectra. On

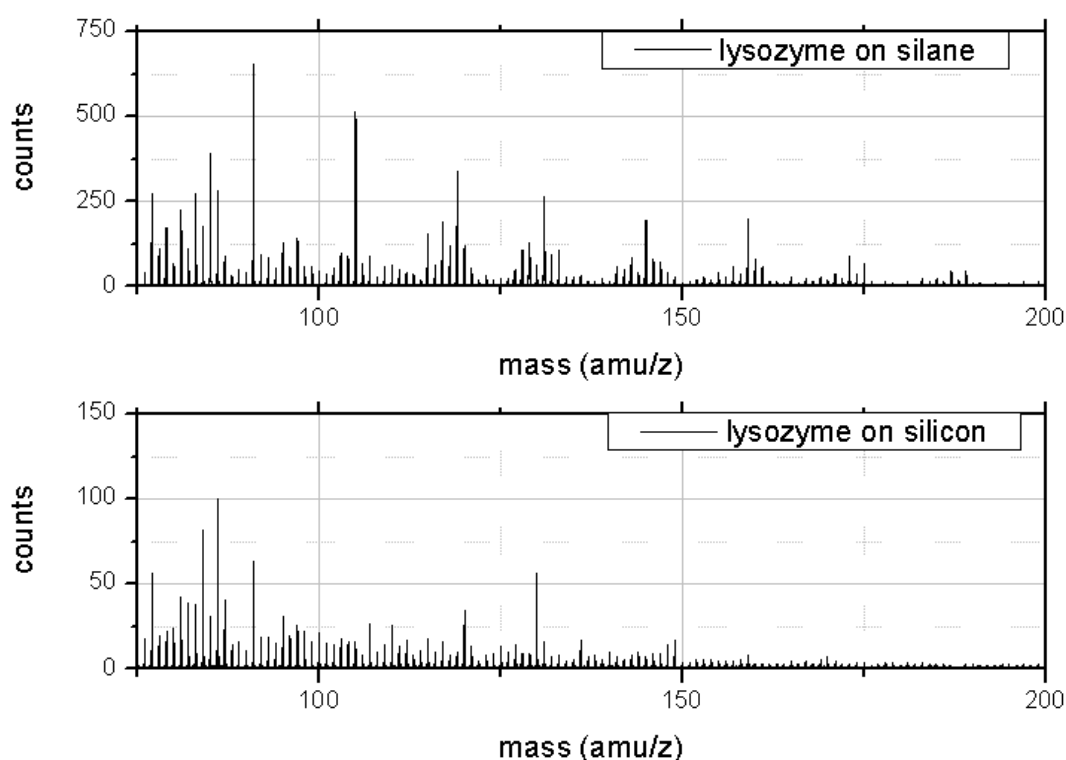


Figure 28: Comparison of mass spectra of cations from lysozyme films on silane or silicon

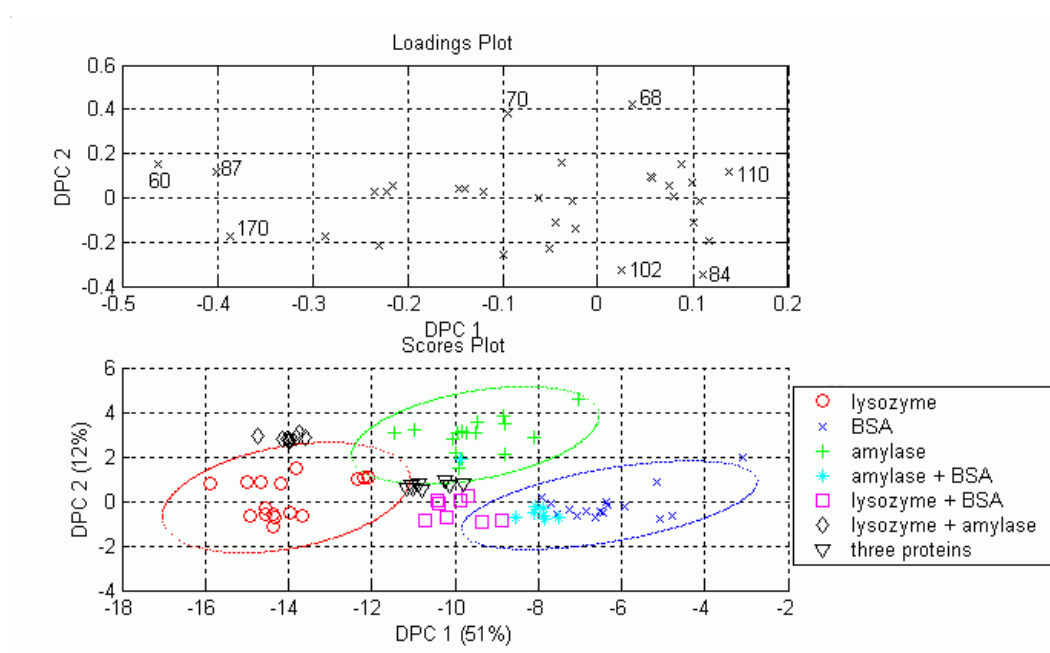


Figure 29: Results of DPCA on spectra of three proteins adsorbed to silicon from water

the other hand the projections of the mixture of lysozyme and amylase are very close to the lysozyme region and the projections of the mixture of amylase and BSA lie (with exception of one spectrum) completely in the probability ellipse of BSA. It can thus be presumed that in these two cases there is a preferential adsorption of lysozyme and BSA compared to amylase.

In the loadings plot the masses 60 amu/z, 87 amu/z and 170 amu/z show the strongest negative values on the first DPC. These masses are assigned to the amino acids serine, asparagine and tryptophane (see table 1). All of them are most abundant in lysozyme (see figure 26) causing lysozyme spectra to be projected to the left side of the scores plot. The mass 110 amu/z (histidine) has the highest positive loading on the first DPC. Since histidine is most abundant in BSA, this protein is projected to the right side of the scores plot. The positive side of the second discriminant principal component is dominated by proline (masses 68 amu/z and 70 amu/z) which is present in BSA and amylase at nearly equal proportions. Since the masses with strongest negative influence on the second DPC 84 amu/z (glutamine, glutaminic acid, lysine) and 102 amu/z (glutaminic acid) belong to amino acids most abundant in BSA, this protein is found in the lower part of the scores plot and amylase is found in the upper part.

Leave-one-out tests of the DPCA model show the following results: If spectra are assigned by their euclidean distances on the first two discriminant principal components, 47 out of 48 spectra (98%) are associated with the correct group. Using the probability ellipses for assignment, still 43 spectra (90%) are correctly recognised. This shows that the data are well represented by the DPCA model.

5.3.2 Proteins solved in 100 millimolar buffer solution

When the proteins are solved in 100 millimolar buffer solution, many sample sites exhibit a strong signal at 23 amu/z caused by sodium adsorbed from the buffer solution. Because of the matrix effect caused by sodium, spectra for analysis are acquired only at sites with small sodium concentrations but these are difficult to find. Four samples each are coated with lysozyme or BSA and four spectra of cations are acquired on each sample. The spectra are unit-mass binned and amino acid related masses are selected before doing DPCA.

In the scores plot (see figure 30) the two sample types are separated on the first discriminant principal component but the 95%-probability ellipses of BSA and lysozyme coated samples overlap to a large extent. The negative side of the first principal component is most strongly influenced by the masses 159 amu/z, 130 amu/z, 170 amu/z and 60 amu/z. These are linked to fragments of the amino acids tryptophane and serine (see table 1), which are more abundant in lysozyme than in BSA (see figure 26). On the other hand the masses 110 amu/z and 82 amu/z show the strongest positive loading. They are linked to histidine which is more abundant in BSA than in lysozyme. This explains why lysozyme can be found mostly in the negative range and BSA in the positive range of the first DPC in the scores plot.

In addition to the pure protein films, mass spectra acquired on two samples treated with a solution of lysozyme and BSA of equal mass concentrations in 100 millimolar buffer solution are projected into the scores plot. They are spread widely within the probability

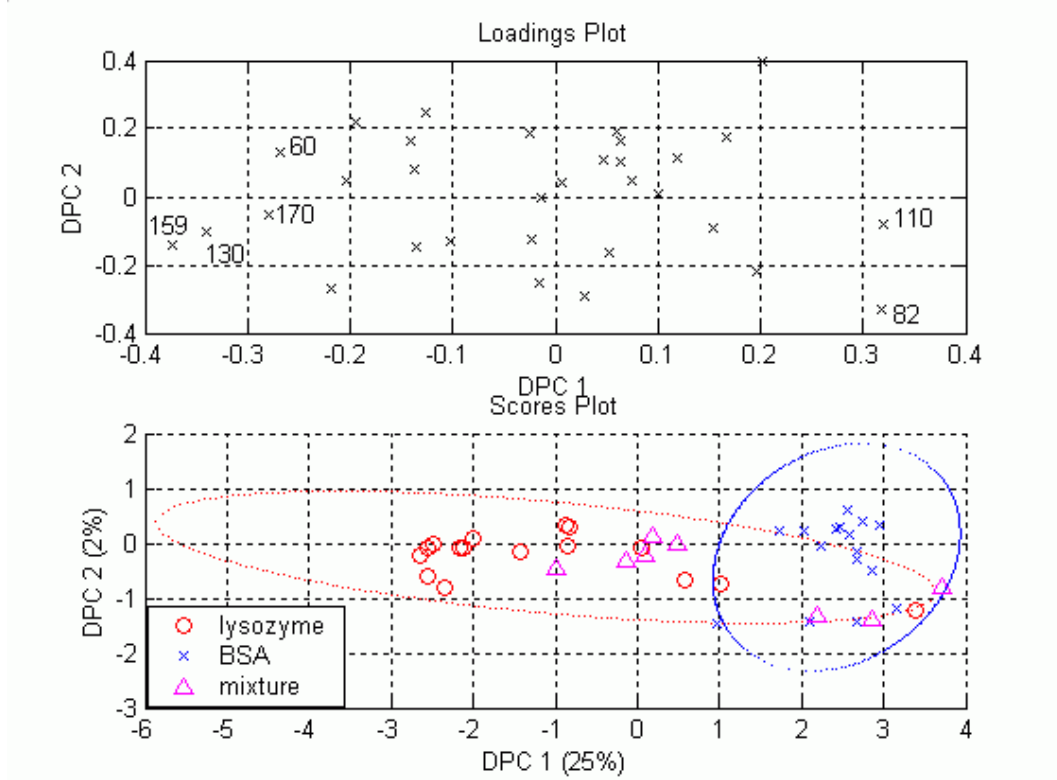


Figure 30: Results of DPCA of two proteins adsorbed on silicon from 100 mM buffer solution

ellipses of lysozyme and BSA.

The leave-one-out-test shows that in this case only 19 out of 32 spectra (59%) are assigned to the right protein using the probability ellipses. Due to the large within group spread, many of the spectra are projected into the overlap of the two ellipses preventing an unambiguous assignment. Using the euclidean distance on the first DPC, 28 spectra (88%) are correctly assigned. So the data are not very well represented by the two-dimensional DPCA model. Probably the high spread within the groups is caused by differing amounts of buffer salts remaining on the sample surfaces.

5.3.3 Proteins solved in 10 millimolar buffer solution

In the following experiments instead of a 100 millimolar buffer solution of sodium di-(hydrogen)phosphate (NaH_2PO_4) and disodium hydrogenphosphate (Na_2HPO_4), a ten millimolar buffer solution with pH 7.4 is used. Onto twelve silicon substrates lysozyme, amylase or BSA are adsorbed from this ten millimolar buffer solution. Protein concentrations are about 10^{-4} mols per litre. On each sample mass spectra of positively charged ions are acquired at four sites. The spectra are unit-mass binned and peaks resulting from amino acid fragments (see table 1) are selected before doing discriminant principal component analysis. The resulting scores and loadings plots are shown in figure 31. In the scores plot a very good separation of the different sample types is visible. This supports the assumption that in the preceding experiments with proteins adsorbed from buffer solution the mass spectra were strongly influenced by buffer salts. Now due to the

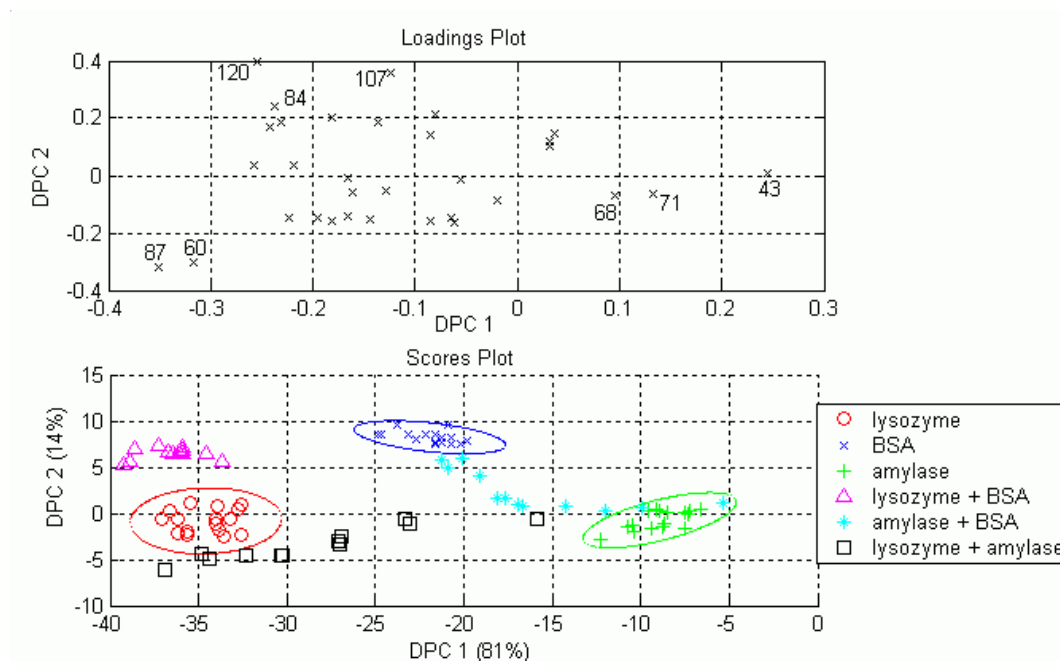


Figure 31: Results of DPCA of three proteins adsorbed on silicon from 10 mM buffer solution

smaller buffer concentration, a better removal of the buffer salts is possible. Hence the spread within the groups in the scores plot is largely reduced. Spectra acquired on nine samples treated with mixture solutions of two of the three proteins are projected into the scores plot. In the protein solutions used in this case equal molarities (approximately 10^{-5} mol/l) of two proteins are present. The projections of the mixtures of lysozyme and BSA can be found in a small region closer to the lysozyme probability ellipse than to the one of BSA indicating that more lysozyme than BSA adsorbed from the solution. With exception of one outlier the projections of mixtures of lysozyme and amylase are situated closer to lysozyme than to amylase. Thus it is assumed that mostly lysozyme adsorbed in this case. The projections of the three samples treated with mixtures of amylase and BSA are clearly differing from one another. Depending on the sample, the projections lie either close to the BSA region or close to the amylase region or between them. Hence the concentration ratio of the two proteins at the sample surface differs significantly between the three samples. Thus the co-adsorption process of amylase and BSA seems to be more sensitive to external influences like contaminations or temperature than the co-adsorption processes of the other mixtures.

The masses 87 amu/z and 60 amu/z show the strongest negative loadings on the first and second discriminant principal component. They are associated to the amino acids asparagine and serine (see table 1). These have highest abundance in lysozyme (see figure 26) which is thus projected in the lower left part of the scores plot. The positive side of the first DPC is most strongly influenced by the masses 43 amu/z (arginine), 71 amu/z (serine) and 68 amu/z (proline). The former two are most abundant in lysozyme and the latter one in BSA. But all of them are also highly abundant in amylase explaining its position on the right side of the scores plot. The amino acids phenylalanine (120 amu/z), tyrosine (107 amu/z) and glutamine, glutamic acid and lysine (84 amu/z) are all highly

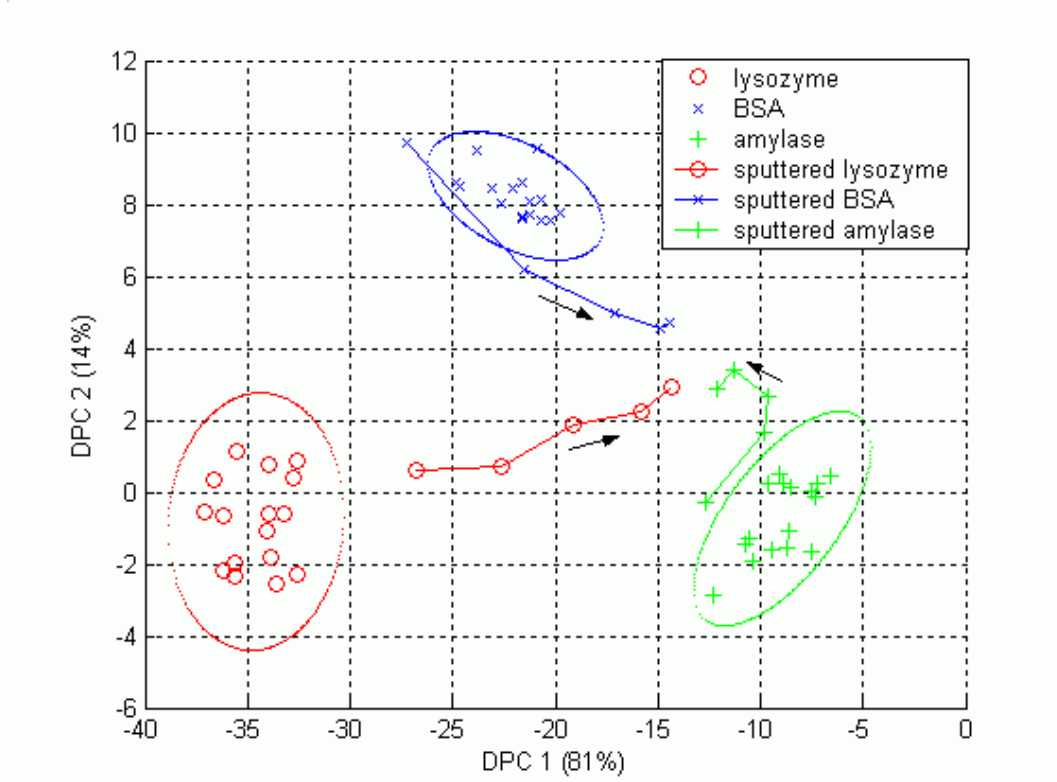


Figure 32: Projections of mass spectra of proteins adsorbed to silicon after sputtering into the scores plot. The arrows indicate the direction of increased sputter time.

abundant in BSA. With their strongly positive loadings on the second DPC they project the BSA spectra in the upper part of the scores plot.

The leave-one-out test confirms the good fitness of the DPCA model to represent the data. Using the euclidean distance on the first two DPC for assignment, all 48 spectra are correctly recognised. With the 95% probability ellipses, 46 spectra (96%) are assigned to the correct protein.

Similarly prepared samples of the three proteins on silicon substrates are sputtered with the unpulsed primary ion beam. Mass spectra of cations are acquired after 0, 5, 10, 15 and 20 seconds of sputtering. Using the primary ion current ($I \approx 10^{-9}$ A), the maximum sputter time ($t = 20$ s), the elementary charge ($e = 1.6 \cdot 10^{-19}$ C) and the scanning size ($A = 1.44 \cdot 10^{-2}$ cm²), the maximum primary ion dose D can be estimated as

$$D = \frac{It}{eA} \approx 9 \cdot 10^{12} \text{ cm}^{-2}. \quad (42)$$

These spectra are unit-mass binned and peak selection is performed before projecting them into the scores plot built with the preceding DPCA model. The projections are shown in figure 32. The arrows indicate the direction of increasing sputter time. The projections of samples of all three proteins quickly approach the same region in the scores plot. Thus even after short sputtering the protein layers are destroyed to a degree that makes it impossible to recognize the protein by DPCA. In the corresponding mass spectra, the intensities of all peaks with exception of the ones caused by the substrate are strongly reduced by sputtering.

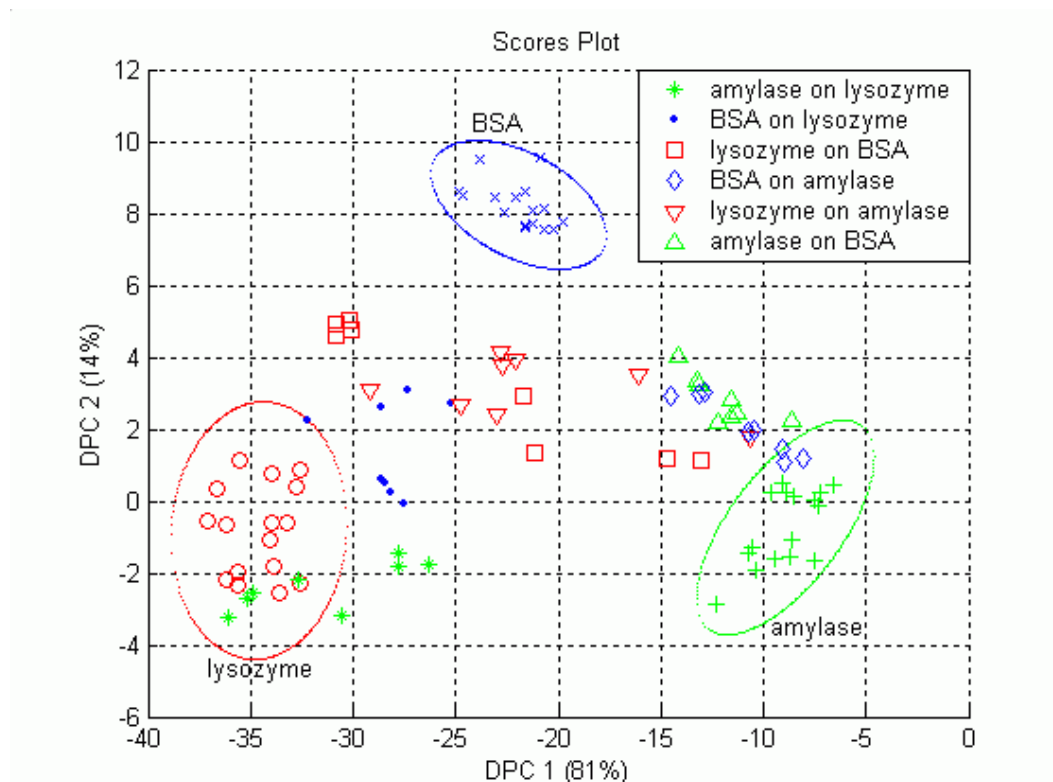


Figure 33: Projections of mass spectra of protein double layers adsorbed to silicon into the scores plot

Next it is tried to prepare double layers of two different proteins on silicon substrates. Therefore, after adsorption of the first protein, the samples are rinsed and given into a solution of the second protein for two hours. Protein molarities are approximately 10^{-4} mols per litre buffer solution. As before, a ten millimolar phosphate buffer solution is used. Two samples are prepared for each of the six possible combinations of the three proteins and on each sample mass spectra of cations are acquired at four different sites. The mass spectra are unit-mass binned and amino acid related peaks are selected. Then they are projected into the scores plot of the DPCA model created with the mass spectra of single component protein layers on silicon. The projections and the positions of the single component spectra on the first two discriminant principal components are shown in figure 33.

The projections can be interpreted as follows:

Amylase on lysozyme: These spectra are projected very closely to or into the lysozyme probability ellipse. Hence very little or no amylase adsorbs to the lysozyme layer or it is only loosely bound and removed in the rinsing step.

BSA on lysozyme: Again the projections are situated close to the lysozyme region suggesting that BSA adsorbs only in little amounts on the lysozyme layer.

Lysozyme on BSA: These projections show a large spread but they are all at approximately equal distances from the probability ellipses of lysozyme and BSA. Thus the two proteins are detected at equal amounts. This can be caused by mixing of the

two proteins in the second adsorption step or by a not tightly packed lysozyme layer. The latter would make the detection of BSA via holes possible. A non homogeneous lysozyme layer can also explain the large spread of the data points.

BSA on amylase: Since these spectra are projected closely to the amylase region, it is assumed that only little amounts of BSA adsorb onto the amylase film.

Lysozyme on amylase: Within the projections of these spectra the largest spread is found. Their positions range nearly from the lysozyme probability ellipse to the amylase ellipse. Thus these samples present a very inhomogeneous surface composition and there are large local differences in the amount of lysozyme adsorbed.

Amylase on BSA: These projections are situated close to the amylase region. Hence amylase adsorbs to the BSA layer and forms a relatively compact layer masking BSA from detection by ToF-SIMS.

The results for the adsorption of amylase on lysozyme are confirmed by amylase activity measurements. The group of Dr. Christian Hannig, Universitätsklinikum Freiburg, determined the immobilized activity of amylase adsorbed directly to silicon as 0.028 units per square centimetre. The activity measured for amylase adsorbed to lysozyme coated silicon is only ten percent as high (0.0028 U/cm²). This is compared to the amount of amylase estimated by the DPCA projections of the mass spectra of samples treated first with lysozyme and then with amylase solution. Figure 33 shows that lysozyme and amylase coated samples are separated by the first discriminant principal component. Thus the amylase fraction of a sample's surface composition f_{amy} is estimated by the distance of its projection on the first axis to the centres of the amylase (d_{amy}) and lysozyme (d_{lys}) probability ellipses:

$$f_{amy} = \frac{d_{lys}}{d_{lys} + d_{amy}}. \quad (43)$$

Hence it is assumed that a mass spectrum projected into the centre of the lysozyme ellipse ($d_{lys} = 0$) does not contain any amylase while a spectrum projected into the centre of the amylase ellipse ($d_{amy} = 0$) is caused by a pure amylase film. Calculating the surface fraction of amylase for all eight spectra of amylase adsorbed onto lysozyme a mean value of $f_{amy} = (16 \pm 5)\%$ is obtained. The error is calculated by the standard deviation of the values for different spectra. Bearing in mind the difficulties in determining the activity of immobilized enzymes caused by possible denaturation or obstructed substrate diffusion and the problems of quantification for secondary ion mass spectra due to differing ionization probabilities, the amylase activity fraction of ten percent lies remarkably close to the surface fraction estimated by ToF-SIMS and DPCA.

Unfortunately it was not possible to measure the activity of surface immobilized lysozyme. Probably it is deactivated by the adsorption or the subsequent drying. To determine the amount of BSA adsorbed to the samples, fluorescence microscopy is used. The obtained results are described in section 5.4.

The results of samples coated consecutively with two proteins can be correlated with the ones obtained from binary protein mixtures. From a mixture of lysozyme and amylase, little amounts of amylase adsorb to the silicon surface. Equally, only very small quantities

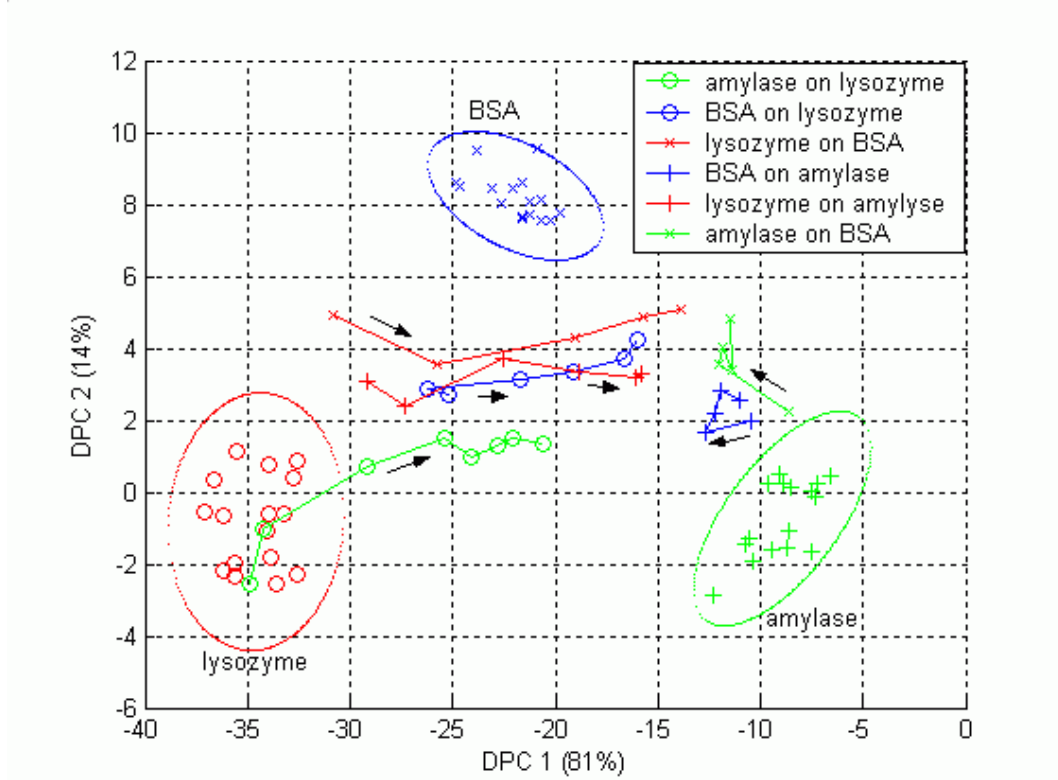


Figure 34: Projections of mass spectra of sputtered protein double layers adsorbed to silicon into the scores plot. The arrows indicate the direction of increased sputter time.

of amylase adsorb to a surface coated with lysozyme. Thus in both cases lysozyme hinders the adsorption of amylase. A similar but weaker effect can be observed for BSA and lysozyme. More lysozyme than BSA is adsorbed from the mixture, and BSA does not adsorb very well on a lysozyme coated surface. Hence lysozyme hinders the adsorption of BSA in both experiments. Due to the high spread in the data of amylase BSA mixtures, these cannot be compared to the data from the double layer experiments.

On the samples coated consecutively with two proteins, mass spectra of cations are also acquired after sputter times of zero to 35 seconds. The mass spectra are unit-mass binned and amino acid related peaks are selected. Then they are projected into the scores plot of the DPCA model created with the mass spectra of single component protein layers on silicon (see figure 34). The arrows indicate the direction of increasing sputter time.

For all samples the projections develop directly to the same region under sputtering. This is the same region, the projections of single protein layers developed to in figure 32. Thus sputtering destroys rapidly the whole protein layer making a discrimination of proteins after sputtering impossible. In particular the assumed double layer structure of the protein films cannot be detected. The experiments were performed twice for each sample type giving essentially the same results. For clarity figure 34 shows only the results of one series of experiments.

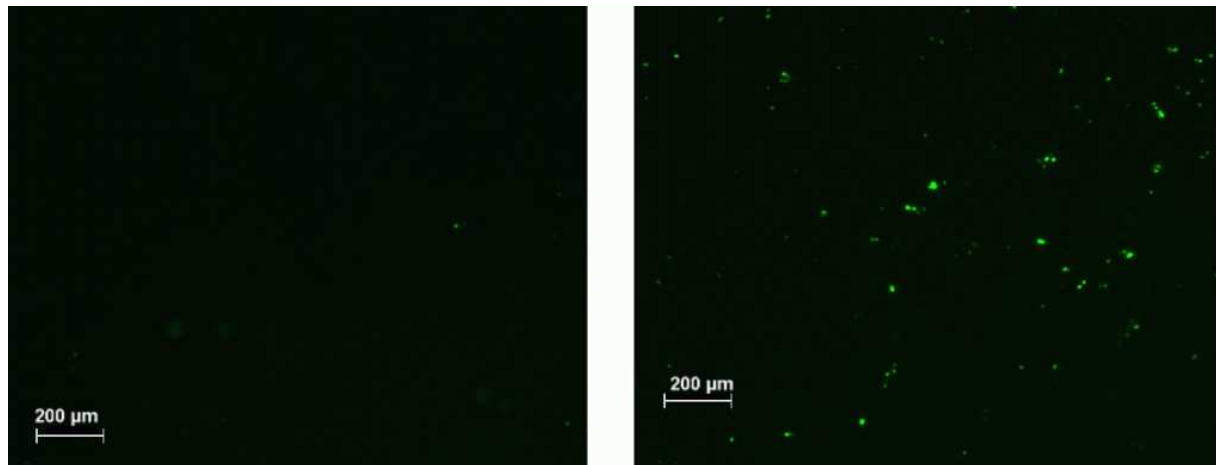


Figure 35: Fluorescence microscopy images of pure BSA fluorescein conjugate (left) and a mixture of BSA and lysozyme (right) adsorbed on silicon

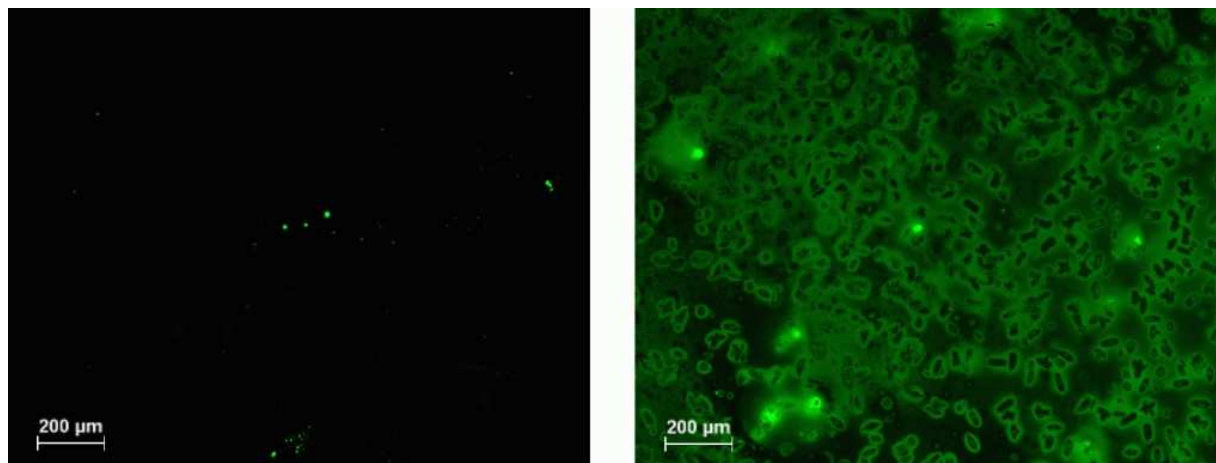


Figure 36: Fluorescence microscopy images of BSA fluorescein conjugate adsorbed to silane (left) and to silicon without rinsing of the sample after adsorption (right)

5.4 Fluorescence microscopy of protein films on silicon

Figures 35 and 36 show typical fluorescence microscopy images obtained from different samples. If the protein is adsorbed to silicon and the samples are rinsed after adsorption only few fluorescence spots are visible (see figure 35). The sample treated with a mixture of BSA fluorescein conjugate and lysozyme shows more fluorescence spots than the sample treated only with BSA fluorescein conjugate. Possibly most of the protein is removed in the rinsing step. This would also explain the strong substrate signal in the mass spectra acquired on similarly prepared samples. The images confirm the theory that BSA adsorbs to the substrate from a protein mixture as well as from a single protein solution. There might be different explanations for the stronger fluorescence shown by the mixture treated samples. It could be caused by possibly non-identical rinsing times removing more or less of the protein layer. The mixture of the two proteins might form a denser layer more difficult to remove by rinsing. Or the presence of lysozyme might enhance the fluorescence yield of fluorescein.

On the sample where the BSA fluorescein conjugate was adsorbed to silane, the fluorescence is only visible at a few spots on the sample surface, too. As the previous ones this sample was rinsed after protein adsorption. But since the proteins should bind covalently to the glutardialdehyde (see figure 16), it should not be possible to remove them by rinsing. There are the following possible explanations for the weak fluorescence: The fluorophore could lose its fluorescence properties by conformational changes of BSA upon adsorption. Or the protein does not bind covalently in the first place and can thus be removed by rinsing.

If the samples are not rinsed after protein adsorption to silicon, a strong fluorescence is visible. This supports the idea that most of the protein is removed from the silicon substrate in the rinsing step. The structures visible in the fluorescence image are probably crystallised buffer salts.

To find out whether they are covered with protein or not, the sample surfaces are investigated with scanning force microscopy in contact mode. On all of the rinsed samples globular structures with height differences of approximately four nanometres are visible covering most of the surface. The structures vary in their lateral size. In addition, larger structures with some tens of nanometres height and up to some hundreds of nanometres in diameter are found on the surface. Probably the small structure is a protein film, while the larger ones are protein agglomerates. As examples, images obtained on the two samples treated with a solution of BSA fluorescein conjugate and lysozyme in buffer solution are shown in figures 37 and 38. In the centre of figure 37 a smaller region was scanned for ten minutes with higher frequency and higher force than usual. Afterwards the surface level lies about two nanometres deeper in this scratched area because the protein film has been removed.

On the sample that was not rinsed after protein adsorption larger structures are visible (see figure 39). Their lateral size is some micrometres and their height is some hundreds of nanometres. These are probably crystallized buffer salts. The protein layer should be found underneath the buffer crystals and is not detectable by scanning force measurements.

From the scanning force microscopic images it can be concluded that most of the

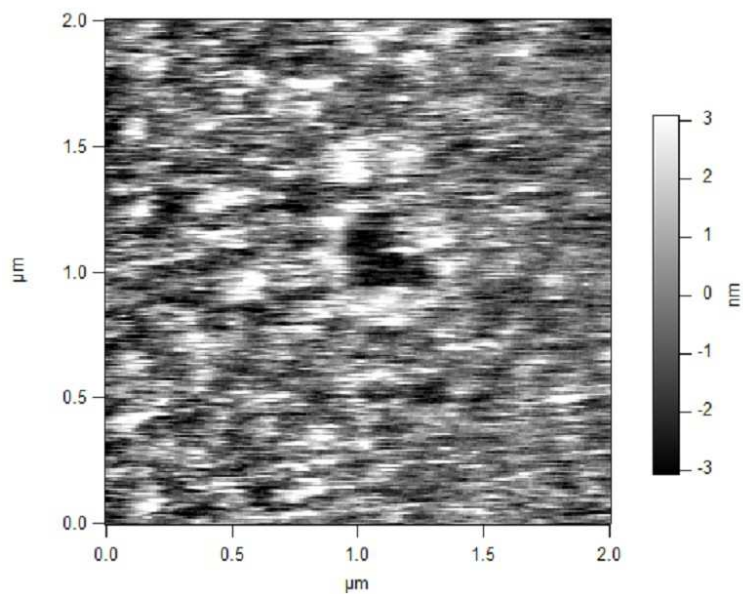


Figure 37: Scanning force microscope image of a silicon sample treated with a solution of BSA fluorescein conjugate and lysozyme rinsed after protein adsorption. The lower region in the centre of the image was scratched with the cantilever tip

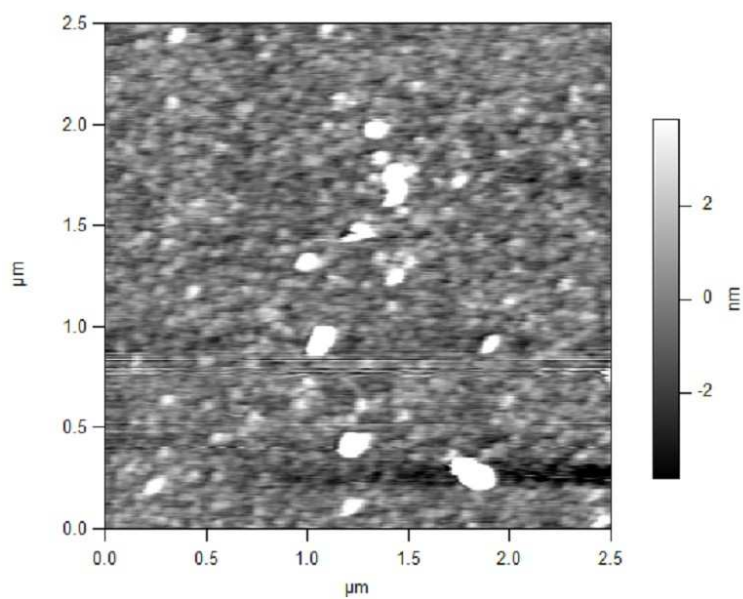


Figure 38: Scanning force microscope image of a silicon sample treated with a solution of BSA fluorescein conjugate and lysozyme rinsed after protein adsorption

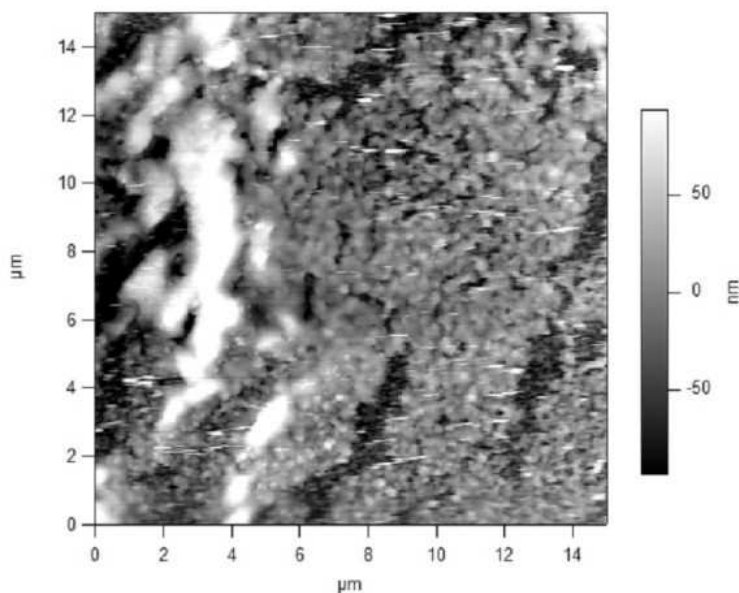


Figure 39: Scanning force microscope image of a silicon sample treated with a solution of BSA fluorescein conjugate not rinsed after protein adsorption

samples' surfaces is covered with a protein layer even after rinsing. Thus the fluorescence patterns on the samples do not reflect the protein distribution. Presumably the fluorescence yield of fluorescein is reduced due to interactions which can occur when the protein adsorbs to the surface and changes its conformation.

5.5 ToF-SIMS of protein films on dental implant materials

Single protein films are prepared on substrates of FAT and FAW. Of each protein three samples per substrate type are made. Mass spectra of cations are acquired by ToF-SIMS at four sites on each sample. Regardless of the substrate or protein used the spectra have a similar appearance. As an example a mass spectrum acquired on BSA coated FAT is shown in figure 40. A strong signal at 40 amu/z (calcium) and weaker ones at 23 amu/z (sodium) and 88 amu/z (strontium) are caused by the substrate. Additionally at nearly every mass up to 200 amu/z peaks caused by protein fragments or organic contaminants are visible. The visibility of substrate caused peaks shows that once again the proteins are not closely packed on the surface. The mass spectra are unit-mass binned and amino acid related peaks are selected. Of the peaks listed in table 1, the ones at 44 amu/z, 86 amu/z, 87 amu/z, 88 amu/z and 69 amu/z overlap with peaks caused by the substrate (^{44}Ca , ^{86}Sr , ^{87}Sr and ^{88}Sr) or by the primary ions (^{69}Ga). Hence these masses are not taken into account for multivariate analysis.

Performing DPCA on the data acquired on FAT substrates leads to the results shown in figure 41. In the scores plot the 95%-probability ellipses of the three proteins are well separated. In addition to the data acquired on FAT, spectra acquired on FAW substrates are projected into the scores plot using the loadings calculated with the FAT data. The projections lie very close to or within the probability ellipses of the corresponding protein. This shows that the intensity patterns of the selected peaks are only weakly influenced

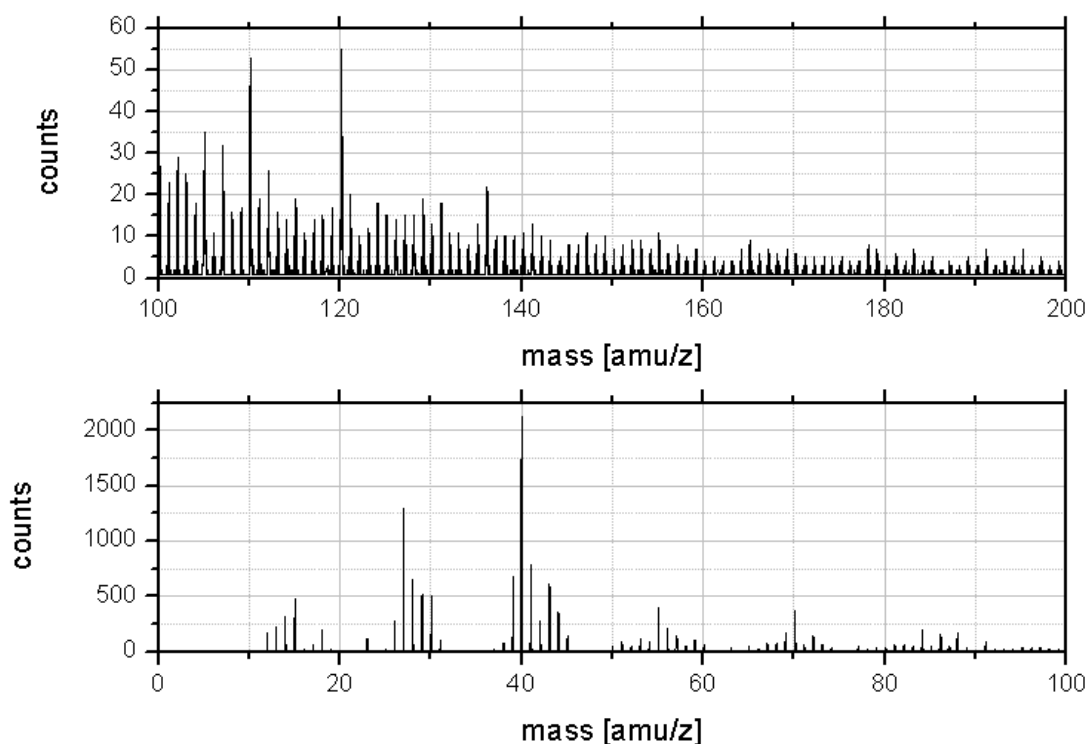


Figure 40: Mass spectrum of cations acquired on a FAT substrate coated with BSA

by the substrate.

In the loadings plot the masses 130 amu/z, 170 amu/z, 159 amu/z, 60 amu/z and 61 amu/z have the strongest negative values on the first discriminant principal component. They are caused by fragments of tryptophane, serine and methionine (see table 1). The former two are most abundant in lysozyme while methionine is most abundant in amylase but can be found to nearly equal amounts in lysozyme as well (see figure 26). Together these masses cause the lysozyme spectra to be projected to the left of the scores plot. The highest positive loadings on the first DPC are achieved by the masses 84 amu/z (glutamine, glutamic acid and lysine), 82 amu/z (histidine), 120 amu/z (phenylalanine) and 110 amu/z (histidine). All the corresponding amino acids but phenylalanine are most abundant in BSA projecting its spectra to the right of the scores plot. Finally, fragments of proline (68 amu/z), methionine (61 amu/z) and tyrosine (107 amu/z) have the strongest negative loadings on the second DPC. Since they are highly abundant in amylase, spectra of this protein are projected in the lower part of the scores plot.

A leave-one-out test confirms the quality of the DPCA model. Using the euclidean distance on the first two DPC to the centres of gravity of the protein groups for assignment, all 36 spectra are correctly assigned. With the probability ellipses, still 94% of the spectra are assigned to the right protein.

By performing DPCA on the spectra of proteins adsorbed to FAW similar results are obtained (see figure 42). Again the probability ellipses of the three proteins are well separated in the scores plot and the projections of spectra acquired on protein films on FAT lie mostly within the probability ellipse of the right protein.

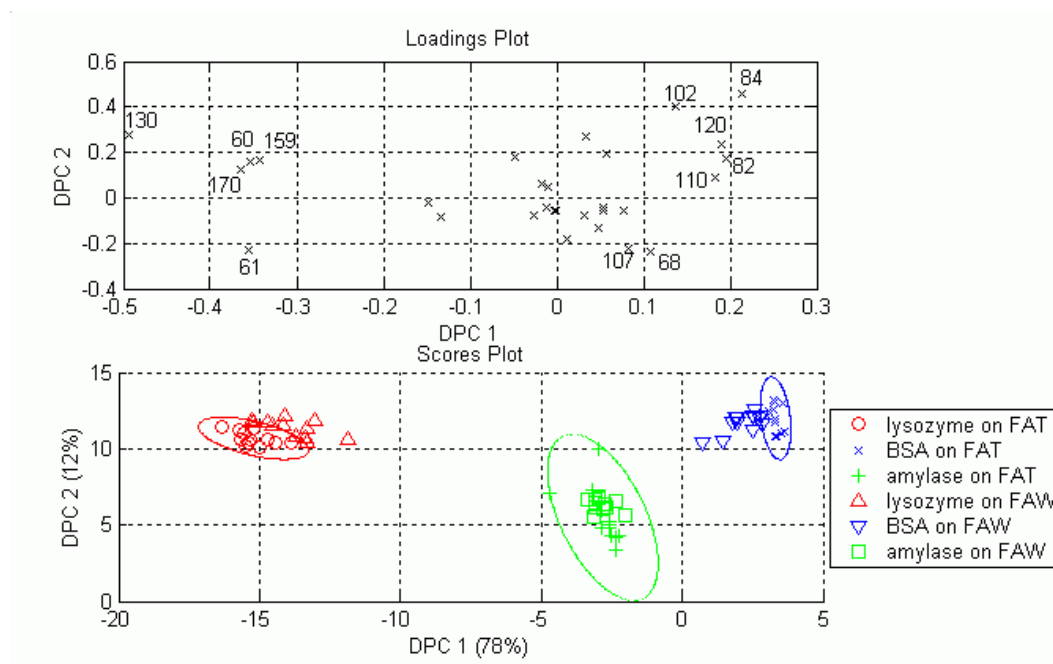


Figure 41: DPCA results of proteins adsorbed to FAT from 10 mM buffer solution

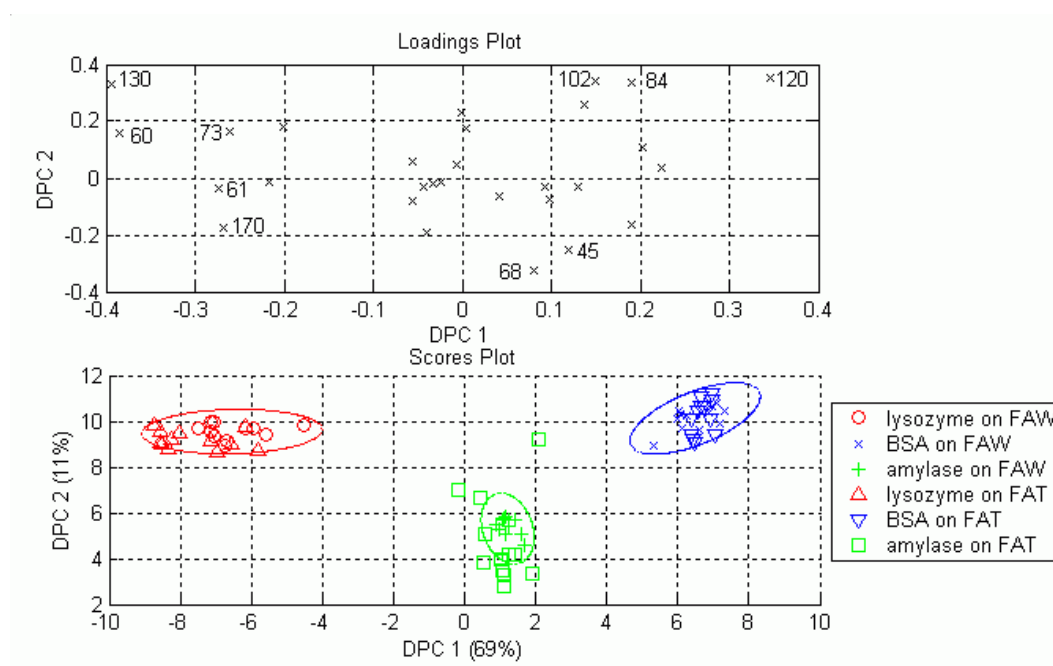


Figure 42: DPCA results of proteins adsorbed to FAW from 10 mM buffer solution

The loadings plot shows that the positioning of the proteins in the scores plot is mainly determined by the same amino acid fragments as before. Only two new fragments of significant loadings occur. On the negative side of the first discriminant principal component arginine (73 amu/z), which is most abundant in lysozyme, supports its positioning on the left side of the scores plot. Cysteine (45 amu/z) contributes with its negative loading on the second DPC in the positioning of amylase because it is highly abundant in this protein.

In the leave-one-out test the DPCA model for FAW substrates performs slightly worse than the one for FAT substrates. Using the euclidean distance, still all 36 spectra are correctly assigned but with the probability ellipses the percentage of spectra assigned to the right protein is reduced to 89%.

The similarities between the DPCA results of protein layers on the two substrate types show that very similar single component protein layers are formed on FAT and FAW substrates.

5.5.1 Adsorption from binary protein mixtures

Substrates of FAT and FAW are also treated with solutions of binary mixtures of two of the three proteins. Equal molarities of about $5 \cdot 10^{-5}$ mol/l of the proteins are solved in 10 mM buffer solution at pH 7.4. After protein adsorption time-of-flight mass spectra of cations are acquired at four sites on each sample. Of every possible binary mixture three samples are prepared on FAT and two samples on FAW. The spectra are unit-mass binned and amino acid related peaks are selected. They are projected into the scores plots of the DPCA models built with single protein spectra on the respective substrates. The resulting projections can be seen in figure 43 for FAT substrates and in figure 44 for FAW substrates.

The spectra of protein films on FAT substrates show a relatively large spread between different samples treated with the same type of mixture. The mass spectra of a mixture of lysozyme and amylase are projected in a region from the lysozyme probability ellipse to the middle between the probability ellipses of lysozyme and amylase. Thus it is assumed that mostly lysozyme adsorbs from this mixture but the proportions of lysozyme and amylase are different from experiment to experiment. Similarly the projections of the mass spectra of the amylase BSA mixture lie in a region from the BSA ellipse to the middle between the BSA and amylase ellipses. This indicates that BSA forms the major component of the adsorbed films but their exact composition differs between the experiments. In contrast the mass spectra of a mixture of lysozyme and BSA show less spread. They are found closer to BSA than to lysozyme. Hence in this case BSA seems to be the major component of the adsorbed film.

When the protein films are adsorbed to FAW substrates there are smaller differences between the projections of mass spectra of different samples treated with the same type of mixture. Now, with one exception, the projections of a mixture of lysozyme and amylase are situated very close to the lysozyme probability ellipse. Thus probably only lysozyme has adsorbed to the substrate. The mass spectra of BSA amylase mixtures are projected close to the BSA ellipse indicating that this protein is the major component of the adsorbed film. The projection of mixtures of lysozyme and BSA are closer to the

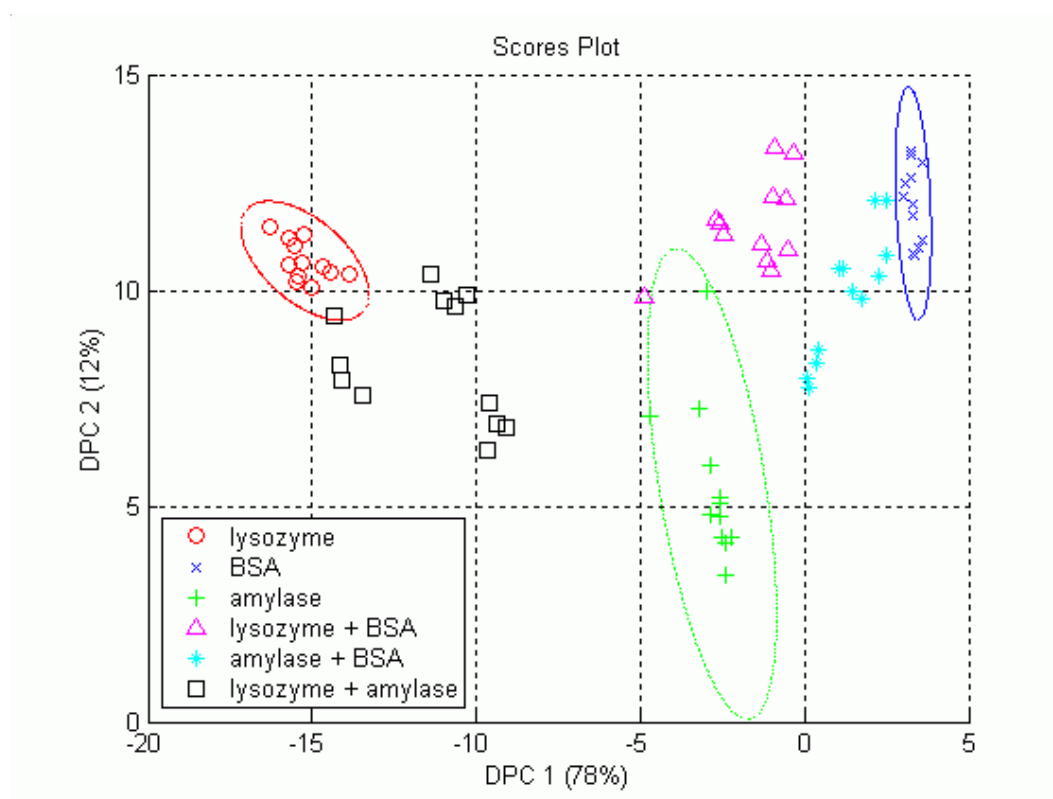


Figure 43: Projections of mass spectra of binary protein mixtures on FAT

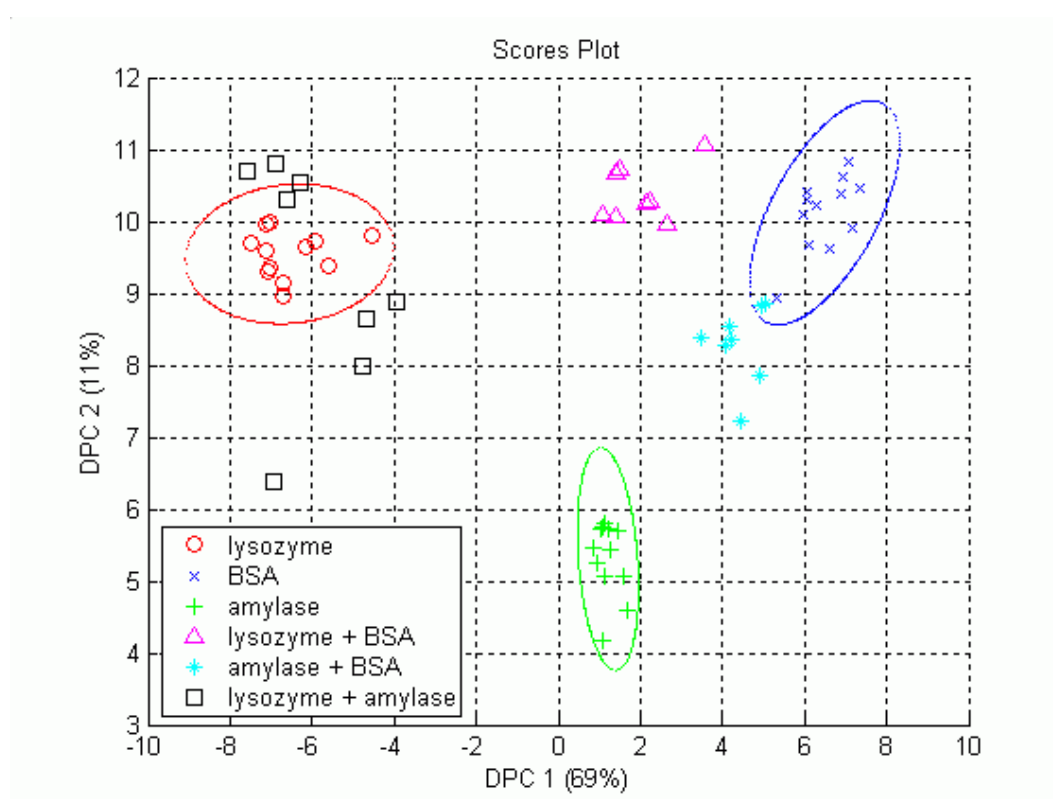


Figure 44: Projections of mass spectra of binary protein mixtures on FAW

proteins in solution	substrate		
	silicon	FAT	FAW
lysozyme and BSA	more lysozyme	more BSA	more BSA
amylase and BSA	spread from amylase to BSA	spread from equality to much more BSA	spread from more to much more BSA
lysozyme and amylase	spread from more to much more lysozyme	spread from equality to more lysozyme	much more lysozyme

Table 2: DPCA projection deduced compositions of protein films adsorbed from binary mixtures

BSA ellipse than to the one of lysozyme. Thus more BSA than lysozyme is detected on the samples. So the composition deduced from the DPCA projections of the protein films created with binary mixtures is similar on the two dental implant materials but it differs from the one on silicon substrates. This is not surprising, as an effect of the substrate on protein adsorption should be expected. These results are summarized in table 2. They suggest that BSA has the highest affinity to adsorb to the dental implant materials followed by lysozyme while amylase has the lowest affinity. On silicon substrates lysozyme shows the highest affinity while the relation between the ones of BSA and amylase is varying.

5.5.2 Sputtering of single protein layers

On two other sets of samples coated with single protein films, mass spectra are acquired after sputtering the surface with the unpulsed primary ion beam. The samples are prepared like the ones used to build the DPCA model. On one sample of lysozyme, BSA and amylase adsorbed to FAT and FAW substrates, mass spectra are acquired after sputtering for 0, 5, 10, 15, 20, 25 and 30 seconds. This corresponds to maximum primary ion doses of about 10^{13}cm^{-2} (see equation 42). The spectra are unit-mass binned, amino acid related peaks are selected, and the spectra are projected in the scores plots using the DPCA models created before. The projections are shown in figure 45 for FAT substrates and in figure 46 for FAW substrates. Under sputtering the projections move from the probability ellipses of the respective protein into the region between the three ellipses. After short sputtering times a recognition of the different proteins by their DPCA projections is not possible any more. At this stage it can be seen in the mass spectra that most of the organic material has been removed from the sample surface. The first one or two spectra corresponding to sputtering times of five to ten seconds of samples coated with lysozyme or BSA show a development heading to the amylase probability ellipse. Probably the different amino acids are not equally destroyed by sputtering. Thus amino acids related to lysozyme and BSA can be destroyed before the ones related to amylase. This makes the spectra more similar to amylase spectra and explains the development of sputtered lysozyme and BSA in the scores plot.

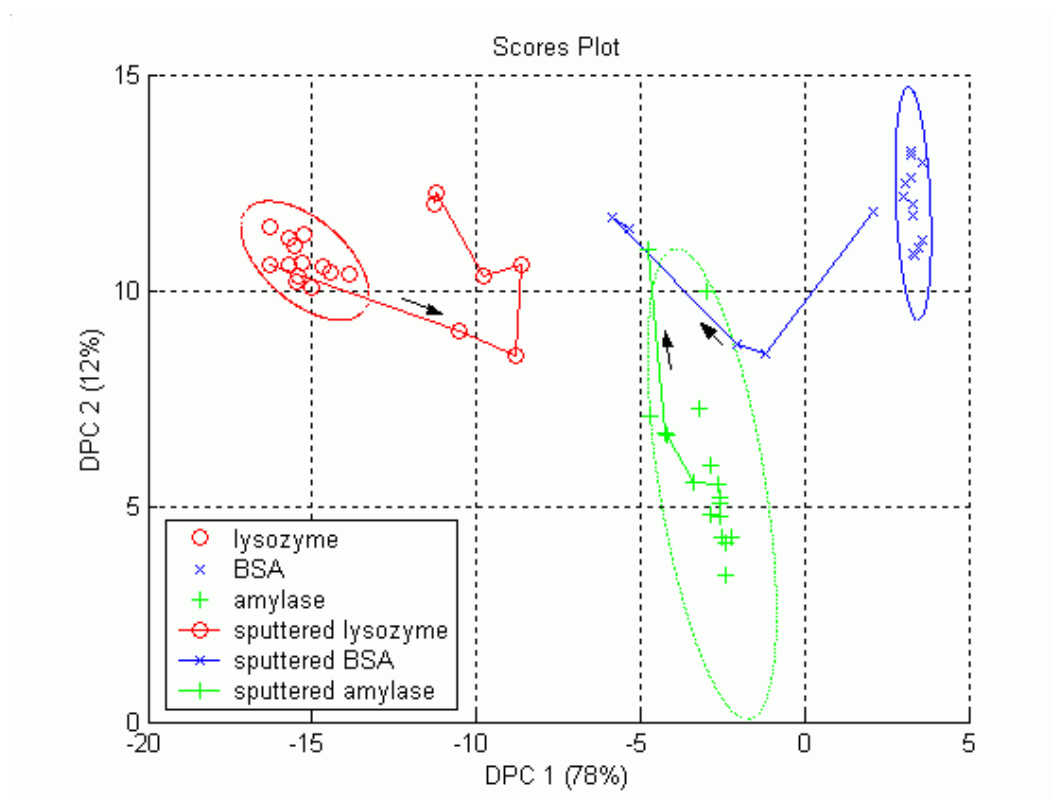


Figure 45: Projections of sputtered protein films on FAT

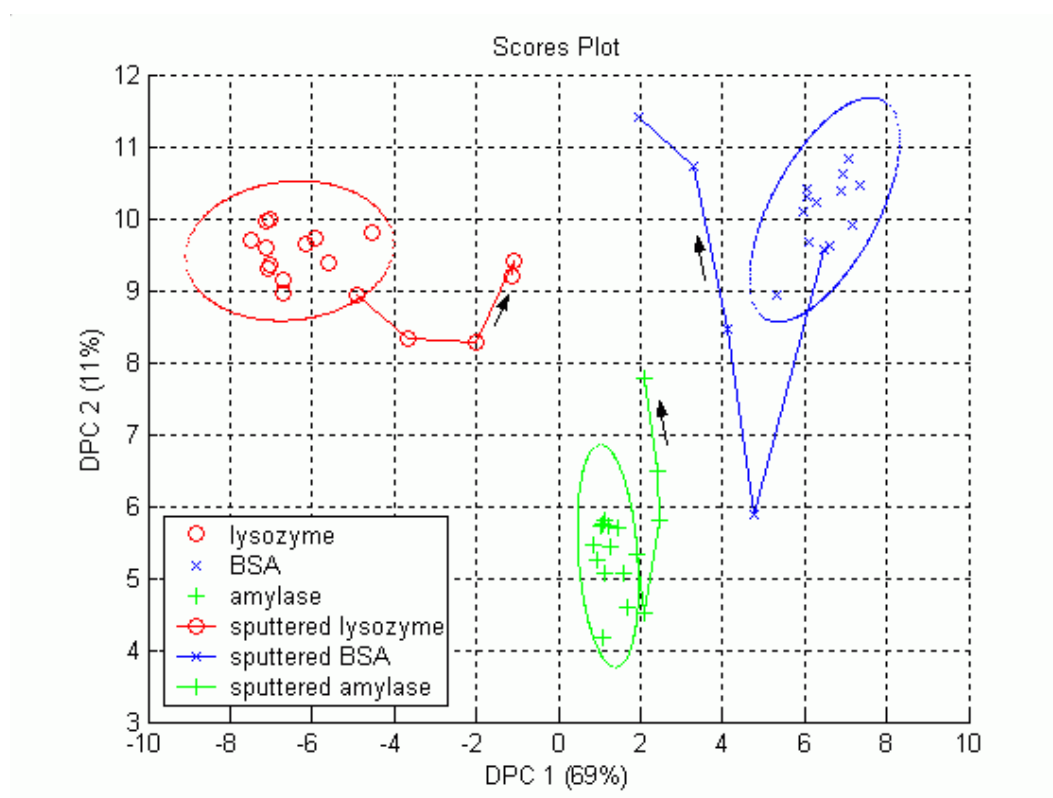


Figure 46: Projections of sputtered protein films on FAW

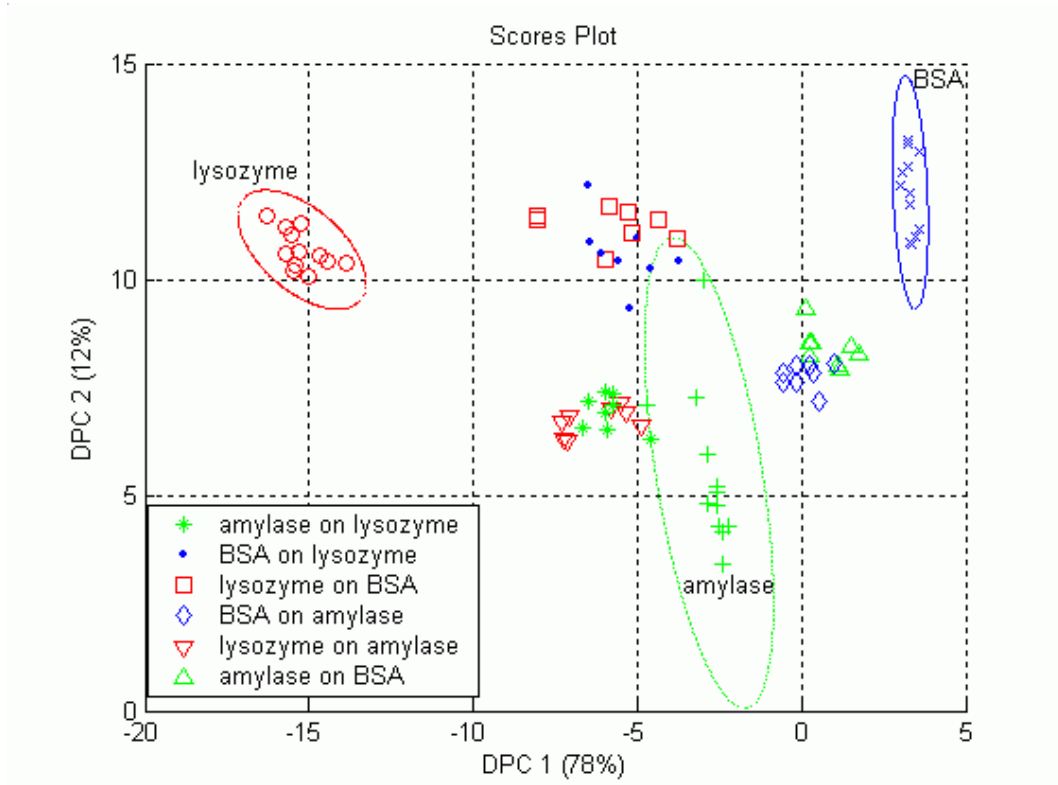


Figure 47: Projections of protein double layers adsorbed to FAT substrates

5.5.3 Consecutive adsorption of two proteins

As on the silicon substrates, the consecutive adsorption of two proteins is studied on the dental implant materials, too. Therefore the substrates are rinsed with buffer solution after the first adsorption step. Then they are put into a solution of the second protein. Proteins are solved at molarities of about $5 \cdot 10^{-5}$ mols per litre buffer solution. The latter is a ten millimolar phosphate buffer as before. Two samples are prepared for each of the six possible combinations of two proteins on the two substrate types. Mass spectra of cations are acquired at four sites on each sample. At one site on each sample mass spectra are also acquired after sputtering the surface with the unpulsed primary ion beam for five to thirty seconds. This corresponds to maximum primary ion doses of about 10^{13}cm^{-2} (see equation 42). The mass spectra are unit-mass binned and amino acid related peaks are selected. Then they are projected into the scores plots of the DPCA models built with mass spectra of single protein films.

The projections of the mass spectra of proteins adsorbed to FAT substrates are shown in figure 47 and reveal the following:

Amylose on lysozyme: The projections lie close to the amylose probability ellipse. Hence amylose is detected in high amounts indicating that it adsorbs well on the lysozyme layer.

Lysozyme on amylose: These projections are also close to the amylose ellipse. Thus only little amounts of lysozyme are detected indicating that it does not adsorb well on the amylose coated substrate.

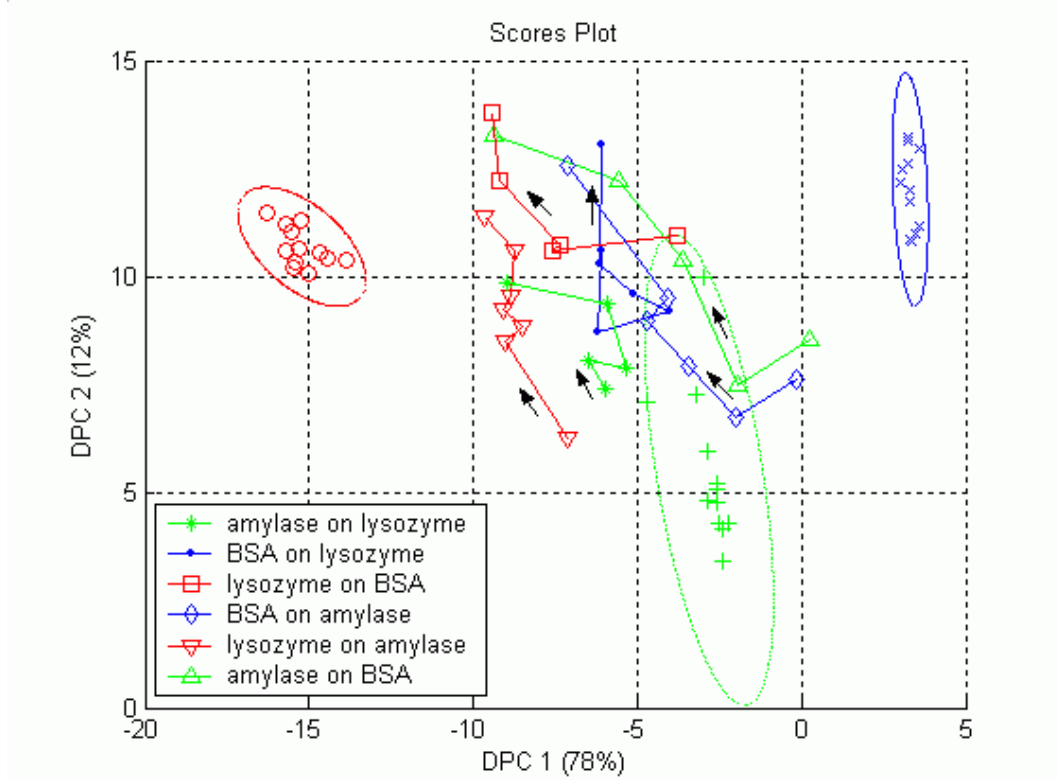


Figure 48: Projections of protein double layers adsorbed to FAT after sputtering. The arrows indicate the direction of increasing sputter time.

BSA and lysozyme: The spectra of both types of samples are projected in the middle between the BSA and lysozyme ellipses. As both proteins are detected, either the secondly adsorbed protein forms a porous layer making detection of the lower layer possible, or the two proteins form one mixed layer in the second adsorption step.

BSA and amylase: Here the same observations and conclusions as for BSA and lysozyme are made.

Comparing these conclusions to the ones drawn from experiments of simultaneous adsorption of two proteins, there seems to be a contradiction in the interaction of lysozyme and amylase. On the one hand amylase adsorbs well on lysozyme and lysozyme does not adsorb well on amylase. On the other hand more lysozyme than amylase adsorbs from a mixture of the two proteins. This raises the question why lysozyme is not covered by amylase as in the two step experiment. A possible answer can be found in the rinsing step between the two adsorption steps. It might induce conformational changes in the adsorbed lysozyme that favour the adsorption of amylase on it.

Figure 48 shows the effect of sputtering on the projections. The arrows indicate the direction of increasing sputter time in steps of five seconds between two consecutive spectra. For clarity only the projections of one series of mass spectra are shown. The projections of the other series develop mainly in the same way. As expected, the projections develop into a region between the three probability ellipses of single protein spectra. The development is not represented by straight lines because of unequal destruction of different amino acids by sputtering and because of noise, which gets more important with

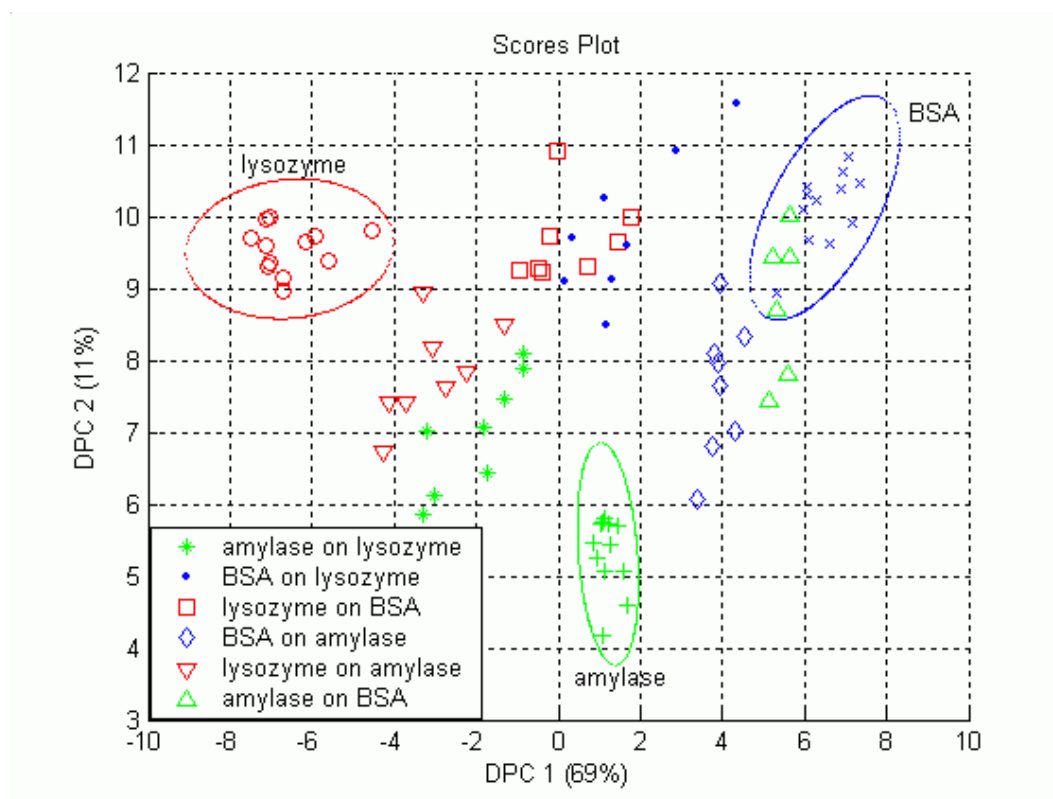


Figure 49: Projections of protein double layers adsorbed to FAW substrates

increasing sputter times because of the vanishing of the amino acid related peaks. The projections of spectra of samples coated with BSA and amylase behave in a very similar way. This indicates that, regardless of the order of the adsorptions, one mixed layer of the two proteins develops. If there were two layers one would expect a differing behaviour depending on the order of adsorptions as it can be seen for lysozyme and BSA. The effect that the projections of layers of amylase and BSA move closer to the amylase ellipse after the first five seconds of sputtering is also seen for pure BSA layers (see figure 45). Thus it does not show the detection of an increasing amount of amylase.

In figure 49 it can be seen that many of the projections of mass spectra of protein double layers adsorbed to FAW substrates are situated differently than the ones of FAT substrates (see figure 47). This shows an influence of the substrate onto the interactions between different proteins upon adsorption. In detail the following observations are made for FAW substrates:

Amylase and lysozyme: Contrary to the spectra on FAT substrates, the projections are situated more or less in the middle between the probability ellipses of amylase and lysozyme. Hence both proteins are detected at equal amounts. This can be explained either by the formation of a porous layer of the secondly adsorbed protein or by mixing of the two proteins in the second adsorption step.

BSA and lysozyme: Like the spectra on FAT substrates, the spectra are mostly projected into the middle between the ellipses of BSA and lysozyme. This leads to the same conclusions as before.

proteins	silicon	substrate	
		FAT	FAW
lysozyme on BSA	medium	medium	medium
lysozyme on amylase	spread from very little to high	little	spread from medium to high
BSA on lysozyme	little	medium	medium
BSA on amylase	spread from very little to little	medium	spread from medium to high
amylase on lysozyme	very little	high	medium
amylase on BSA	high	medium	spread from very little to little

Table 3: Results of DPCA on mass spectra of protein double layers. This table summarizes the amount detected of the secondly adsorbed protein

Amylase on BSA: These projections lie close to or within the BSA ellipse. Thus the amount of amylase detected is very small. It can be concluded that in contrast to the case of FAT substrates, amylase does not adsorb well to BSA coated FAW substrates.

BSA on amylase: Compared to the spectra on FAT substrates, these are projected in a larger region closer to the BSA ellipse than to the amylase ellipse. Thus BSA is detected in larger quantities than before. Hence BSA adsorbs in higher amounts to an amylase coated FAW substrate than to a corresponding FAT substrate.

The projections of the spectra of all sample types show a higher spread on FAW substrates than on FAT substrates.

Again the results of the consecutive adsorption of two proteins can be compared to the ones of the simultaneous adsorption. In contrast to the experiments on FAT substrates, the two types of experiments on FAW substrates lead partly to the same conclusions. From a solution of BSA and amylase mainly BSA adsorbs to the substrate. Equally only little amounts of amylase adsorb on a BSA coated sample. Thus in both experiments BSA hinders the adsorption of amylase. Since the projections of consecutive adsorptions of lysozyme and BSA or amylase and lysozyme are situated in the middle between the probability ellipses of their components, these proteins do not significantly hinder or favour the adsorption of one another. Hence there are no conclusions to be compared to the one-step adsorption experiments.

Table 3 summarizes the results of the consecutive adsorption experiments of two proteins for all three substrate types.

Figure 50 shows the development of the projections of mass spectra of protein double layers on FAW substrates under sputtering. The arrows indicate the direction of increasing sputter time in steps of five seconds between two consecutive spectra. For clarity only the projections of one series of mass spectra are shown. The projections of the other series develop mainly in the same way.

As before the projections of the spectra of all samples develop into the region between the three protein probability ellipses. Thus the protein layers are rapidly destroyed and

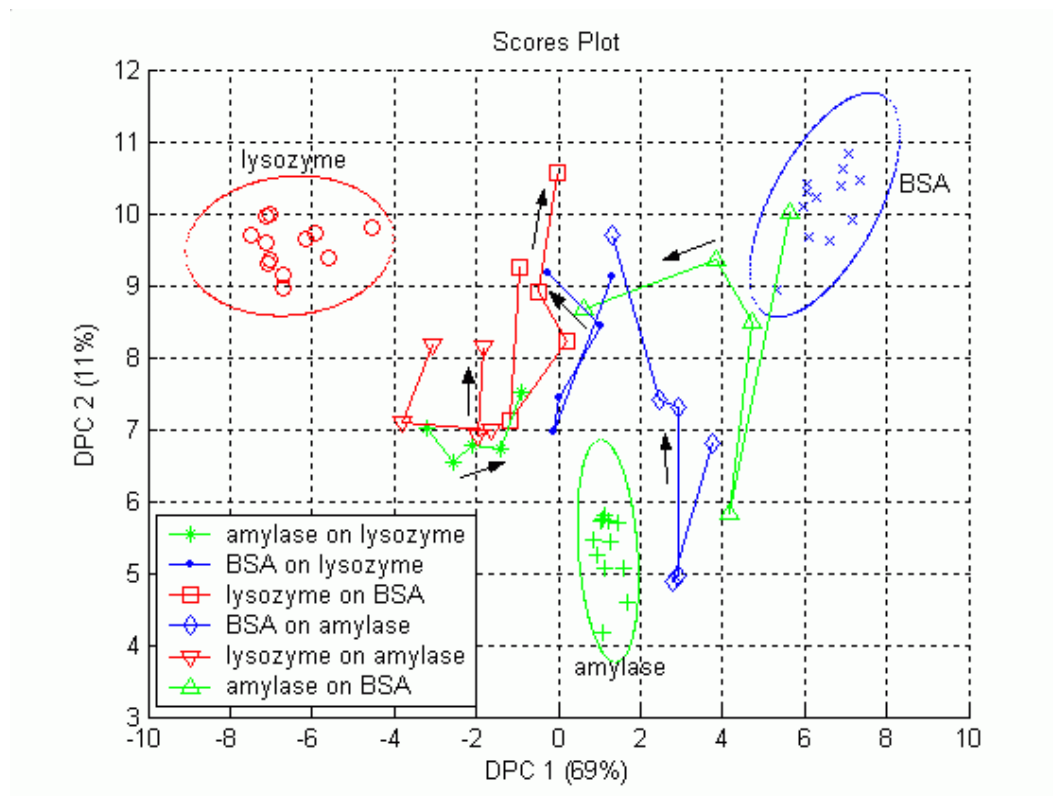


Figure 50: Projections of protein double layers adsorbed to FAW after sputtering. The arrows indicate the direction of increasing sputter time.

cannot be recognized by their mass spectra any more. After the last sputtering period only the substrate generated peaks show a noteworthy intensity in the mass spectra. The projections of most samples do not develop directly to their final position but show a strong shift to smaller values on the second discriminant principal component during the first five to ten seconds of sputtering. As this behaviour is seen for most samples regardless of their composition, it is not the effect of a rising proportion of amylose, which can be found at lower scores on the second DPC. Instead it is probably caused by the not equal destruction of different amino acids by the primary ion beam. A preferential destruction of amino acids with highly positive loadings on the second DPC explains the development of the projections.

6 Summary and Outlook

In this work the two experimental dental implant materials FAT and FAW, made of fluoroapatite particles embedded in polymer matrices, and films of the proteins lysozyme, amylase and bovine serum albumin (BSA), adsorbed to the two dental implant materials, were investigated with time-of-flight secondary ion mass spectrometry (ToF-SIMS) and the multivariate data analysis technique discriminant principal component analysis (DPCA).

The elemental surface composition of the dental implant materials was estimated from the intensity of the peaks caused by the ions of different elements in ToF-SIMS mass spectra acquired on samples of FAT and FAW. These compositions were compared to the surface composition of bovine tooth enamel obtained in the same way. It was found that the main surface components of all three materials are sodium, phosphorus, potassium, calcium, oxygen and fluorine. Additionally the two implant materials contain high quantities of aluminium and strontium. Due to the limitations of the ToF-SIMS technique for quantitative analysis, exact surface compositions cannot be measured by this method.

Protein adsorption experiments were at first performed on silanised silicon substrates, because silicon substrates offer a well-defined flat surface and the adsorption of proteins to silanised samples had already been studied with other methods within the work group. ToF-SIMS mass spectra of cations were acquired on samples coated with lysozyme or BSA. These spectra were dominated by a signal caused by silicon probably from polymerised silane on the sample surface. Additionally many peaks associated to amino acid fragments could be found in the mass spectra. MATLAB functions were written to select amino acid related peaks from the mass spectra, to subtract a background spectrum and to do DPCA on the mass spectra. Nevertheless it was not possible to distinguish between different proteins by their scores resulting from DPCA because of a large spread within the data. This spread was probably caused by varying amounts of polymerised silane on the sample surfaces.

In the next step proteins were adsorbed to unsilanised silicon substrates to prevent the influence of polymerised silane on the mass spectra. In this case, after selection of amino acid related peaks and DPCA, discrimination of samples coated with amylase, lysozyme or BSA by their mass spectra of cations was possible. The concentration of the buffer solution was varied to obtain the best discrimination of the proteins at a concentration of ten millimols per litre of the buffer salts sodium di(hydrogen)phosphate and disodium hydrogenphosphate. The results of DPCA were evaluated with leave-one-out (LOO) tests. These tests showed that the data were very well represented by the DPCA model, because 46 out of 48 spectra could be assigned to the right protein.

With the DPCA model developed with single component protein films, the adsorption behaviour in the cases of simultaneous or consecutive adsorption of two proteins to the same sample was studied. Therefore mass spectra obtained on the new samples were projected into the existing scores plot. Mutual influences between the different proteins upon adsorption were observed. It was concluded that the presence of lysozyme hindered the adsorption of amylase and BSA in the two studied cases. These results were partly confirmed by enzymatic activity measurements of amylase.

Finally, amylase, lysozyme and BSA were adsorbed to the dental implant materials.

Again the three proteins could be distinguished by the scores of their mass spectra of cations after selection of amino acid related peaks and DPCA. The peak selection had to be slightly modified, because some of the peaks were influenced by substrate generated peaks of the same nominal mass. A correct assignment of at least 89% of the spectra in LOO tests showed a good representation of the data by the DPCA models. Furthermore, a great similarity of the DPCA results obtained on the two different substrate types was observed. This shows the formation of similar single component protein films on the two dental implant materials.

Samples coated with two proteins adsorbed simultaneously or consecutively were also studied. Mass spectra acquired on these samples were projected into the scores plots of the corresponding DPCA model built with spectra from single component protein films. On both substrate types the experiments of simultaneous adsorption of two proteins led to the same results: BSA shows the strongest affinity to adsorb from the mixed solutions followed by lysozyme, while amylase has the smallest affinity. The consecutive adsorption of lysozyme and BSA on both substrates led to the detection of equal amounts of both proteins regardless of the order of adsorption. Thus either one mixed layer of the two proteins or a porous layer of the second protein had developed. On FAT there seemed to be a contradiction in the behaviour of lysozyme and amylase between the two types of experiments: Lysozyme hindered the simultaneous adsorption of amylase while it favoured the consecutive adsorption of amylase. This could be explained by conformational changes of lysozyme in the rinsing step between the consecutive adsorptions. By contrast both types of experiments on FAW suggest that the presence of lysozyme hinders the adsorption of amylase. Hence in this case the differences between the two substrate types clearly influence the co-adsorption of amylase and lysozyme.

Lysozyme has bacteriolytical effects in the acquired enamel pellicle while amylase favours the adsorption of bacteria and thus the formation of plaque. Thus the mutual influence of these two proteins upon adsorption is of special interest for the design of dental implant materials. This influence, as well as the ones between the other protein combinations, has been shown to be dependent on the substrate used. It should be further investigated why a given substrate favours the adsorption of one protein or another. Possible reasons are the electrical charge and the hydrophilicity of the substrate and the protein.

The co-adsorption results can be verified using other observation techniques. Nicole Lawrence, a member of our work group, is examining the adsorption of proteins on differently prepared silicon substrates with dynamic contact angle measurements and ellipsometry in the context of her Ph.D. thesis.

The solutions of maximum three proteins used in this work are by far less complex than human saliva. In future works more proteins should be used to approach the *in vivo* situation. Furthermore, the parameters pH and temperature are not constant in the oral cavity. So their influence should also be examined.

The influence of the buffer solution merits further investigation, too. It has been seen that the results strongly depend on the concentration of the buffer salts. The use of other than phosphate buffers could lead to new results.

So far no use was made of the possibility to image the sample surface by ToF-SIMS. It could be tried to bring different proteins on the sample at separated sites and to examine the resulting surface distribution.

It was not possible to create depth profiles of the two layer protein films because the unpulsed gallium ion beam used for sputtering the upper protein layer also destroyed the lower protein layer at the same time. This is due to the high penetration depth of the gallium ions. A possible solution to this problem is the use of a cluster ion source instead of the gallium ion source. If for example a buckminsterfullerene (C_{60}) is accelerated to the same total energy as a gallium ion, the energy per atom and thus the penetration depth is much smaller. Boussofiane-Baudin and others estimated that a buckminsterfullerene projectile at 20 kilo-electron volts deposits its energy within the first 30 Ångströms of an organic surface [4]. Hence most of the damage is limited to the removed layer and the lower layer remains intact for analysis. An additional effect of the lower penetration depth is a strongly increased secondary ion yield. Weibel and others have shown that using a buckminsterfullerene source improves the secondary ion yield on various organic surfaces by a factor of at least 30 compared to a gallium source [34]. For high mass fragments the effect is even stronger. Thus sample damage could be minimised by reducing the primary ion dose while maintaining strong secondary ion intensities due to the high yield. Other cluster ion sources as gold clusters (Au_n^+ with $n = 2 - 5$) or sulfur pentafluoride (SF_5^+) show comparable secondary ion yields as buckminsterfullerenes but the sample damage is smallest for the latter [34]. Another approach to make depth profiling possible can be the use of a cryo stage. By cooling down the samples the protein layers should be more stable against in depth destruction by the sputter beam.

The protein layers are certainly denatured due to dehydration upon exposure to the ultra high vacuum conditions in the ToF-SIMS apparatus. To preserve the protein layers in their physiological state for analysis, one should try to prevent denaturation for example by use of a cryo stage or by fixation with glutardialdehyde.

All these possibilities of further investigations show that this work has by no means exhaustively dealt with the investigation of protein films on dental implant materials by ToF-SIMS. Instead it has established a method of investigation, which can be used for future studies, that can finally lead to many new insights into the subject.

A Matlab functions

This section contains the source code of the MATLAB functions written for data analysis.

einlesen.m

```
function out=einlesen(prefix,start,stop,suffix,m)
%out=einlesen(prefix,start,stop,suffix,m)
%reads the files 'c:\tof- sims\ascii\prefixstartsuffix.asc'
%to 'c:\tof- sims\ascii\prefixstopsuffix.asc' containing
%unit- mass mass spectra exported from WinCadence and
%forms a (mx(stop- start))- data matrix with the
%first m masses in (stop- start) spectra

longprefix=['c:\tof- sims\ascii\',prefix];

k=start;
filename=[longprefix,int2str(k),suffix,'.asc'];

%reading the first WinCadence export:
x=load(filename);

%extracting the counts at the first m masses:
out=voll(x,m);

%loop for the other files:
for k=start+1:stop
    filename=[longprefix,int2str(k),suffix,'.asc'];
    x=load(filename);
    %new data is appended to the existing data matrix
    out=matrix(out,x);
end
```

voll.m

```
function filled = voll(in,m);
%filled = voll(in,m)
%extracts counts from a WinCadence output unit mass
%mass spectrum
%in: (?x3)-matrix: 1. column: channel numbers;
%    2. column: atomic mass units; 3. column: counts;
%    masses with zero counts are not exported
%m: highest mass extracted
%filled: (mx1)-vector: counts for m masses

c=0;
go=1;
k=1;
```

```

%delete first row if it contains mass 0:
if in(1,2)==0
    in=in(2:m+1,:);
end

%if the mass spectrum is complete,
%no masses with 0 counts must be added:
if in(m,2)==m
    go=0;
    filled(:,1)=in(1:m,3);
end

%add masses with 0 counts:
%loop for different masses:
while go==1

    %check if mass and row number match
    %and add the missing mass if they do not:
    if in(k-c,2)~=k
        filled(k,1)=0;
        %c accounts for the shift induced by adding masses:
        c=c+1;
    else
        filled(k,1)=in(k-c,3);
    end

    %stop when highest mass is reached:
    if k==m
        go=0;
    end
    k=k+1;
end

```

matrix.m

```

function X = matrix(X,in)
%X = matrix(X,in)
%adds a new measurement to a data matrix
%X: (mxn)-matrix: values of m variables in n measurements
%in: WinCadence output: unit mass mass-spectrum

[m,n]=size(X);

%extracting the counts from WinCadence output:
in=voll(in,m);

```

%appending the new data:

`X(:,n+1)=in(:,1);`

silanweg.m

function output=silanweg(input,background);

%output=silanweg(input,background);

%subtracts a background spectrum from mass spectra;

%the background is scaled to have the same value as

%the measurement at the 28th variable

%input: (m×n)-matrix: values of m variables in n spectra

%background: (m×1)-vector: background values of m variables

`[m,n]=size(input);`

`[mm,nn]=size(background);`

if mm~=m

`fprintf(1,'Error: Measurements and background must have
the same number of rows!')`

return

end

%loop to treat different measurements:

for k=1:n

`factor=input(28,k)/background(28,1);`

`output(:,k)=input(:,k)-factor*background(:,1);`

end

auswahl.m

function out=auswahl(in)

%out=auswahl(in)

%selects amino acid related peaks from unit mass

%mass spectra

`out(1,:)=in(30,:); %Gly`

`out(2,:)=in(43,:); %Arg`

`out(3,:)=in(44,:); %Ala`

`out(4,:)=in(45,:); %Cys`

`out(5,:)=in(60,:); %Ser`

`out(6,:)=in(61,:); %Met`

`out(7,:)=in(68,:); %Pro`

`out(8,:)=in(69,:); %Thr`

`out(9,:)=in(70,:); %Asn, Pro`

`out(10,:)=in(71,:); %Ser`

`out(11,:)=in(72,:); %Val`

`out(12,:)=in(73,:); %Arg`

`out(13,:)=in(74,:); %Thr`

`out(14,:)=in(81,:); %His`

```

out(15,:)=in(82,:); %His
out(16,:)=in(83,:); %Val
out(17,:)=in(84,:); %Gln, Glu, Lys
out(18,:)=in(86,:); %Ile, Leu
out(19,:)=in(87,:); %Asn
out(20,:)=in(88,:); %Asn, Asp
out(21,:)=in(98,:); %Asn
out(22,:)=in(100,:); %Arg
out(23,:)=in(101,:); %Arg
out(24,:)=in(102,:); %Glu
out(25,:)=in(107,:); %Tyr
out(26,:)=in(110,:); %His
out(27,:)=in(112,:); %Arg
out(28,:)=in(120,:); %Phe
out(29,:)=in(127,:); %Arg
out(30,:)=in(130,:); %Trp
out(31,:)=in(131,:); %Phe
out(32,:)=in(136,:); %Tyr
out(33,:)=in(159,:); %Trp
out(34,:)=in(170,:); %Trp

```

pca.m

```

function [PC, SCORE, V, mn]=pca(X);
%[PC, SCORE, V, mn]=pca(X)
%performs principal component analysis
%X: (mxn)-matrix: values of m variables for n measurements
%PC: (mxm)-matrix: loadings of m PC as m columns
%SCORE: (mxn)-matrix: scores of m PC for n measurements
%V: (m)-vector: variance captured by the m PC
%mn: (m)-vector: mean values of the m variables
%for details on the PCA algorithm see J. Shlens,
%A tutorial on principal component analysis, 10/12/2005

```

```

[m,n]=size(X);

```

```

%scaling by total intensity:

```

```

for k=1:n
    X(:,k)=X(:,k)/sum(X(:,k));
end

```

```

%mean-centering:

```

```

mn=mean(X,2);
X=X-repmat(mn,1,n);

```

```

%auxiliary matrix for SVD:

```

```

Y=X'/sqrt(n-1);

```

%singular value decomposition:

```
[u,S,PC]=svd(Y);
```

%variances -> m-vector V:

```
S=diag(S);
```

```
V=S.*S;
```

%projection -> (mxn)-matrix SCORE:

```
SCORE=PC'*X;
```

%saving loadings and scores:

```
save ploading.dat PC -ascii;
```

```
save pscore.dat SCORE -ascii;
```

dpca.m

function [loading,score,variance,varianceintra]=dpca(x,anz,rep)

%[loading,score,relvariance,varianceintra]=dpca(x,anz,rep)

%performs discriminant principal component analysis

%x: (mxn)-matrix: values of m variables for n measurements

%anz: number of groups

%rep: (1xanz)-vector: number of measurements per group

%loading: (mxm)-matrix: loadings of m DPC as m columns

%score: (mxn)-matrix: scores of m DPC for n measurements

%variance: (1xm)-vector: variance captured by the m DPC

%varianceintra: (mx1)-vector: pooled intra group variances of m variables

%for details on the PCA algorithm see J. Shlens,

%A tutorial on principal component analysis, 10/12/2005

%for details on DPCA see P.W. Yendle, H.J.M. Macfie,

%Discriminant principal components analysis,

%J. of Chemometrics 3, 589-600, 1989

```
[m,n]=size(x);
```

```
[sizerepx,sizerepy]=size(rep);
```

```
if sizerepy==1
```

```
    rep=repmat(rep,1,anz);
```

```
end
```

%scaling by total intensity:

```
for k=1:n
```

```
    x(:,k)=x(:,k)/sum(x(:,k));
```

```
end
```

%pooled intra group variances of the variables:

%-> varianceintra: (mx1)

```
varianceintra=zeros(m,1);
```

```

for k=1:m
    start=1;
    stop=0;
    for o=1:anz
        stop=stop+rep(o);
        varianceintra(k)=varianceintra(k)+var(x(k,start:stop));
        start=start+rep(o);
    end
end

```

```

%scaling by intra group standard deviations:
%-> scaled data matrix s
for k=1:m
    s(k,:)=x(k,:)/sqrt(varianceintra(k));
end

```

```

%scaled group means:
%-> (mxanz)-matrix sbar
for k=1:m
    start=1;
    stop=0;
    for o=1:anz
        stop=stop+rep(o);
        sbar(k,o)=mean(s(k,start:stop));
        start=start+rep(o);
    end
end

```

```

%(mxn)-matrix of scaled group means:
groupmeans=repmat(sbar(:,1),1,rep(1));
for o=2:anz
    groupmeans=[groupmeans repmat(sbar(:,o),1,rep(o))];
end

```

```

%PCA with SVD -> (mxm)-matrix of loadings loading:

```

```

%mean-centering:
mn=mean(groupmeans,2);
groupmeans=groupmeans-repmat(mn,1,n);
helpmatrix=groupmeans'/sqrt(n-1);
[u,singularvalues,loading]=svd(helpmatrix);

```

```

%projection -> (mxn)-matrix of scores score
score=loading'*s;

```

```

%variances of the DPC:
%-> m-vector variance

```

```

for k=1:m
    variance(k)=var(score(k,:));
end

```

```

%saving loadings and scores:
save dloading.dat loading -ascii;
save dscore.dat score -ascii;

```

plotpc2d.m

```

function plotpc2d(PC,SCORE,relvariance,anz,rep,omax)
%plotpc2d(PC,SCORE,relvariance,anz,rep,omax)
%creates 2D scores plots and loadings plots
%PC: (mxm)-matrix: loadings of m PC as m columns
%SCORE: (mxn)-matrix: scores of m PC for n measurements
%relvariance: (m)-vector: relative variance captured by the m PC
%anz: number of groups
%rep: (anz)-vector: number of measurements per group
%omax: highest dimension plotted

if omax<2
    fprintf(1,'Error: Maximum principal component must be grater 1!')
    return
end

%scaling variances if necessary:
if sum(relvariance)~=100
    relvariance=relvariance/sum(relvariance)*100;
end

[sizerepx,sizerepy]=size(rep);
if sizerepy==1
    rep=repmat(rep,1,anz);
end

%loop for different plots:
for o=2:omax

figure

%loadings plot:
subplot(2,1,1)
plot(PC(:,o-1),PC(:,o),'kx');
grid on
title('Loadings Plot');
xlabel(['PC ' int2str(o-1)]);
ylabel(['PC ' int2str(o)]);

```

```

subplot(2,1,2)

%scores plots:
hold on
min=1;
max=0;

%loop for different groups:
for k=1:anz
    max=max+rep(k);
    switch k
        case 1
            plot(SCORE(o-1,min:max),SCORE(o,min:max),'ro');
        case 2
            plot(SCORE(o-1,min:max),SCORE(o,min:max),'bx');
        case 3
            plot(SCORE(o-1,min:max),SCORE(o,min:max),'g+');
        case 4
            plot(SCORE(o-1,min:max),SCORE(o,min:max),'c*');
        case 5
            plot(SCORE(o-1,min:max),SCORE(o,min:max),'ms');
        case 6
            plot(SCORE(o-1,min:max),SCORE(o,min:max),'yd');
        otherwise
            plot(SCORE(o-1,min:max),SCORE(o,min:max),'k^');
            fprintf(1,'Cannot plot more than 6 groups!');
    end
    min=min+rep(k);
end
grid on
title('Scores Plot');
xlabel(['PC ' int2str(o-1) ' ( ' int2str(relvariance(o-1)) '%) ']);
ylabel(['PC ' int2str(o) ' ( ' int2str(relvariance(o)) '%) ']);

end

```

plotdpc2dmitellipse2.m

```

function plotdpc2dmitellipse2(PC,SCORE,relvariance,anz,rep,omax)
%plotdpc2dmitellipse2(PC,SCORE,relvariance,anz,rep,omax)
%creates 2D scores plots with probability ellipses and loadings plots
%PC: (mxm)-matrix: loadings of m DPC as m columns
%SCORE: (mxn)-matrix: scores of m DPC for n measurements
%relvariance: (m)-vector: relative variance captured by the m DPC
%anz: number of groups
%rep: (anz)-vector: number of measurements per group

```

%omax: highest dimension plotted

```
if omax<2
    fprintf(1,'Error: Maximum principal component must be greater one!')
    return
end
```

%scaling variances:

```
if sum(relvariance)~=100
    relvariance=relvariance/sum(relvariance)*100;
end
```

```
[sizerepx,sizerepy]=size(rep);
```

```
if sizerepy==1
    rep=repmat(rep,1,anz);
end
```

%loop to create more than one plot:

```
for o=2:omax
```

```
figure
```

%loadings plot:

```
subplot(2,1,1)
plot(PC(:,o-1),PC(:,o),'kx');
grid on
title('Loadings Plot');
xlabel(['DPC ' int2str(o-1)]);
ylabel(['DPC ' int2str(o)]);
```

%scores plot with different styles for different groups:

```
subplot(2,1,2)
hold on
min=1;
max=0;
```

%loop for different groups:

```
for k=1:anz
```

```
    %critical value of F-distribution for 2 variables,
    %rep(k) measurements and 95% confidence limit:
    F=Fdistribution(rep(k));
```

%calculate critical TSquare value for the ellipses:

```
    TSquare=2*(rep(k)-1)/(rep(k)-2)*F;
```

```
    max=max+rep(k);
```

```

switch k
    case 1
        plot(SCORE(o-1,min:max),SCORE(o,min:max),'ro');
        ellipse(SCORE(o-1:o,min:max),Tsquare,'r');
    case 2
        plot(SCORE(o-1,min:max),SCORE(o,min:max),'bx');
        ellipse(SCORE(o-1:o,min:max),Tsquare,'b');
    case 3
        plot(SCORE(o-1,min:max),SCORE(o,min:max),'g+');
        ellipse(SCORE(o-1:o,min:max),Tsquare,'g');
    case 4
        plot(SCORE(o-1,min:max),SCORE(o,min:max),'c*');
        ellipse(SCORE(o-1:o,min:max),Tsquare,'c');
    case 5
        plot(SCORE(o-1,min:max),SCORE(o,min:max),'ms');
        ellipse(SCORE(o-1:o,min:max),Tsquare,'m');
    case 6
        plot(SCORE(o-1,min:max),SCORE(o,min:max),'yd');
        ellipse(SCORE(o-1:o,min:max),Tsquare,'y');
    otherwise
        plot(SCORE(o-1,min:max),SCORE(o,min:max),'k^');
        fprintf(1,'Cannot plot more than six groups!');
end
min=min+rep(k);
end
grid on
title('Scores Plot');
xlabel(['DPC ' int2str(o-1) ' ( ' int2str(relvariance(o-1)) '%) ']);
ylabel(['DPC ' int2str(o) ' ( ' int2str(relvariance(o)) '%) ']);

end

```

ellipse.m

```

function ellipse(X,Tsquare,color)
%ellipse(X,Tsquare,color)
%plots a probability ellipse around data points
%X: (2xn)-matrix containing n measurements of 2 variables
%Tsquare: critical value of the TSquare distribution
%color: colour (r,b,g,c,m,y or k)
%for details on the algorithm see J.E. Jackson, A user's guide
%to principal components, John Wiley and sons inc., New York, 1991

[m,n]=size(X);
if m~=2
    fprintf(1,' \n Error: Data matrix must be 2 x n! \n');
return

```

end

%Centre of the ellipse:

mitte=mean(X,2);

%Mean-centered auxiliary matrix:

Y=X-repmat(mitte,1,n);

Y=Y'/sqrt(n-1);

%singular value decomposition to calculate

%principal components and standard deviations:

[u,S,PC]=svd(Y);

%standard deviations:

S=diag(S);

%rescaling of principal components

PC(:,1)=S(1)*PC(:,1);

PC(:,2)=S(2)*PC(:,2);

kmax=sqrt(Tsquare);

hold on

%calculation and plotting of the ellipse;

%number of points can be controlled

%by the stepwidth of the loop:

for k=0:0.04:kmax

g=k;

h=sqrt(Tsquare-k^2);

x=mitte+g*PC(:,1)+h*PC(:,2);

plot(x(1),x(2),color)

x=mitte+g*PC(:,1)-h*PC(:,2);

plot(x(1),x(2),color)

x=mitte-g*PC(:,1)+h*PC(:,2);

plot(x(1),x(2),color)

x=mitte-g*PC(:,1)-h*PC(:,2);

plot(x(1),x(2),color)

end

g=sqrt(Tsquare);

h=0;

x=mitte-g*PC(:,1);

plot(x(1),x(2),color)

x=mitte+g*PC(:,1);

plot(x(1),x(2),color)

Fdistribution.m

```

function ausgabe=Fdistribution(n)
%ausgabe=Fdistribution(n)
%finds the upper limit of the 95% confidence interval
%of an F distribution for 2 variables and n measurements;
%values are from the table in J.E. Jackson, A user's guide
%to principal components, John Wiley and sons, New York,1991;
%intermedial values linearly approximated

```

```

%tabulated values:

```

```

F= [1.9950000e+002
    1.9000000e+001
    9.5500000e+000
    6.9400000e+000
    5.7900000e+000
    5.1400000e+000
    4.7400000e+000
    4.4600000e+000
    4.2600000e+000
    4.1000000e+000
    3.9800000e+000
    3.8900000e+000
    3.8100000e+000
    3.7400000e+000
    3.6800000e+000
    3.6300000e+000
    3.5900000e+000
    3.5500000e+000
    3.5200000e+000
    3.4900000e+000
    3.4700000e+000
    3.4400000e+000
    3.4200000e+000
    3.4000000e+000
    3.3900000e+000
    3.3700000e+000
    3.3500000e+000
    3.3400000e+000
    3.3300000e+000
    3.3200000e+000
    3.2300000e+000
    3.1500000e+000
    3.0700000e+000
    3.0000000e+000];

```

```

%to find the right value the number of measurements has
%to be reduced by the number of variables:

```

```

n=n-2;

if n <1
    disp('Error: Number of measurements must be greater 2!')

elseif n<31

    %the first 30 values correspond to 1 to 30 measurements:
    ausgabe=F(n);

elseif n<40

    %linear interpolation:
    ausgabe=3.32-(n-30)*0.009;

elseif n<60
    ausgabe=3.23-(n-40)*0.004;

elseif n<121
    ausgabe=3.15-(n-60)*8/6000;

elseif n>120
    disp('Warning: Using approximation for
(number of measurements) >> 120')
    ausgabe=3;

end

```

loo.m

```

function loo(data,anz,rep)
%loo(data,anz,rep)
%performs DPCA, leaves each measurement out once
%and projects it into the scores plot
%data: (m×n) matrix: values of m variables in n measurements
%anz: number of groups
%rep: (anz)-vector: number of measurements per group

[m,n]=size(data);

[sizerepx,sizerepy]=size(rep);
if sizerepy==1
    rep=repmat(rep,1,anz);
end

```

```

%loop to leave out one measurement after another:
column=0;
for k=1:anz

    for j=1:rep(k)

        %reduceddata: matrix without 'column'th measurement
        column=column+1;
        if column==1
            reduceddata=data(:,2:n);
        elseif column==n
            reduceddata=data(:,1:n-1);
        else
            reduceddata=[data(:,1:column-1),data(:,column+1:n)];
        end

        %reducing number of measurements in concerned group:
        reducedrep=rep;
        reducedrep(k)=reducedrep(k)-1;

        %DPCA:
        [dloading,dscore,dvariance,dintvar]=dpca(reduceddata,anz,reducedrep);

        %scores plot:
        plotdpc2dmitellipse(dscore,dvariance,anz,reducedrep,2);

        %projection of the left out measurement:
        dpcaprojektion(dloading,dintvar,data(:,column),1,'ko');

        %halting program before the next measurement is left out:
        fprintf(1,'Press any key to continue! \n')
        pause

    end
end

```

loo3.m

```

function assignment=loo3(data,anz,rep,dimmax)
%assignment=loo(data,anz,rep,dimmax)
%performs DPCA, leaves each measurement out once
%and calculates its scores to assign it to the closest group.
%Closeness is defined by the euclidean distances to the
%center of gravity of a groups data points.

```

```

%data: (mxn) matrix: values of m variables in n measurements
%anz: number of groups
%rep: (anz)-vector: number of measurements per group
%dimmax: highest considered DPC
%assignment: (nx(anz+1))-matrix: first column: number of the
%           closest group to a measurement; 2.,3.,... column:
%           relative distance to the 1.,2.,... group

[m,n]=size(data);

[sizerepx,sizerepy]=size(rep);
if sizerepy==1
    rep=repmat(rep,1,anz);
end

%loop to leave out one measurement after another:
column=0;
for k=1:anz

    for j=1:rep(k)

        %reduceddata: matrix without 'column'th measurement:
        column=column+1;
        if column==1
            reduceddata=data(:,2:n);
        elseif column==n
            reduceddata=data(:,1:n-1);
        else
            reduceddata=[data(:,1:column-1),data(:,column+1:n)];
        end

        %reducing number of measurements in concerned group:
        reducedrep=rep;
        reducedrep(k)=reducedrep(k)-1;

        %DPCA:
        [dloading,dscore,dvariance,dintvar]=dpca(reduceddata,anz,reducedrep);

        %projection of left out measurement:
        score=dpcaprojektion(dloading,dintvar,data(:,column),1,'no');

        %loop to calculate distances to all groupmeans:
        min=1;
        max=0;
        for i=1:anz
            %calculate center of gravity:
            max=max+reducedrep(i);

```

```

        dpcmean=mean(dscore(1:dimmax,min:max),2);
        %calculate euclidean distance:
        distance(i)=0;
        for h=1:dimmax
            distance(i)=distance(i)+(score(h)-dpcmean(h))^2;
        end
        distance(i)=sqrt(distance(i));
        min=min+reducedrep(i);
    end

    %assign measurement to closest group:
    [minimum,minimumindex]=min(distance);
    assignment(column,1)=minimumindex;

    %calculate relative distances:
    for i=1:anz
        assignment(column,1+i)=distance(i)/minimum;
    end
end
end
end

```

pcaprojektion.m

```

function score=pcaprojektion>Loading,mitte,data,index,stil)
%pcaprojektion>Loading,mitte,data,index,stil)
%projects a data matrix using PC loadings into a scores plot
%loading: (mxm)-matrix: loadings of m PC as m columns
%mitte: m-vector: mean values of m variables for the original data
%data: (mxn)-matrix: values of m variables for n new measurements
%index: number of first displayed PC in the scores plot
%stil: style of the projected data points
%score: (mxn)-matrix: projected scores of m PC for n new measurements

[m,n]=size(data);

%scaling of measurements by total intensity:
for k=1:n
    data(:,k)=data(:,k)/sum(data(:,k));
end

%mean centering:
centered=data-repmat(mitte,1,n);

%projection:
score=loading'*centered;

%plotting can be prevented with style 'no':

```

```

if stil=='no'
    return
end

```

```

%plotting:
hold on
plot(score(index,:),score(index+1,:),stil);

```

dpcaprojektion.m

```

function score=dpcaprojektion>Loading, varianceintra, data, index, stil)
%dpcaprojektion>Loading, varianceintra, data, index, stil)
%projects a data matrix using DPC loadings into a scores plot
%loading: (mxm)-matrix: loadings of m DPC as m columns
%varianceintra: (mx1)-vector: pooled intra group variances of m variables
%data: (mxn)-matrix: values of m variables for n measurements
%index: number of first displayed DPC in the scores plot
%stil: style of the projected data points
%score: (mxn)-matrix: projected scores of m DPC for n measurements
%for details on DPCA see P.W. Yendle, H.J.M. Macfie,
%Discriminant principal components analysis,
%J. of Chemometrics 3, 589-600, 1989


```

```

[m,n]=size(data);

```

```

%scaling of measurements by total intensity:

```

```

for k=1:n
    data(:,k)=data(:,k)/sum(data(:,k));
end

```

```

%scaling by pooled intra group variance:

```

```

for k=1:m
    scaled(k,:)=data(k,:)/sqrt(varianceintra(k));
end

```

```

%projecting:

```

```

score=loading'*scaled;

```

```

%plotting can be prevented with style 'no':

```

```

if stil=='no'
    return
end

```

```

%plotting:
hold on
plot(score(index,:),score(index+1,:),stil);

```

zuordnen.m

```

function zuordnung=zuordnen(score,loading,intvar,anz,rep,data,dimmax)
%zuordnung=zuordnen(score,loading,intvar,anz,rep,data,dimmax)
%assigns new measurements to a group of measurements
%by the euclidean distance of the projection into the
%DPC space calculated with the original measurements to the
%centers of gravity of the original groups
%anz: number of groups
%rep: (1xanz)-vector: number of measurements per group
%loading: (mxm)-matrix: loadings of m DPC as m columns
%score: (mxn)-matrix: scores of m DPC for n measurements
%intvar: (mx1)-vector: pooled intra group variances of m variables
%dimmax: dimensionality of the DPC space
%zuordnung: (nx(anz+1))-matrix: first column: number of the
%           closest group to a measurement; 2.,3.,... column:
%           relative distance to the 1.,2.,... group

[m,n]=size(data);
[sizerepx,sizerepy]=size(rep);
if sizerepy==1
    rep=repmat(rep,1,anz);
end

%DPCA projection of the new data:
newscore=dpcaprojektion(loading,intvar,data,1,'no');

%loop for different new measurements
for j=1:n

    %select one measurement:
    newscorej=newscore(:,j);

    min=1;
    max=0;

    %loop for different groups:
    for i=1:anz
        max=max+rep(i);

        %center of gravity of the group:
        groupmean=mean(score(1:dimmax,min:max),2);

        %calculate euclidean distance:
        distance(i)=0;
        for h=1:dimmax
            distance(i)=distance(i)+(newscorej(h)-groupmean(h))^2;
        end
    end

```

```
        distance(i)=sqrt(distance(i));
        min=min+rep(i);
    end

    %assign new measurement to closest group:
    [minimum,minimumindex]=min(distance);
    zuordnung(j,1)=minimumindex;

    %calculate relative distances to other groups:
    for i=1:anz
        zuordnung(j,1+i)=distance(i)/minimum;
    end
end
```

References

- [1] Research collaboratory for structured bioinformatics protein data bank. <http://www.rcsb.org/>.
- [2] A.M. Belu, D.J. Graham, and D.G. Castner. Time-of-flight secondary ion mass spectrometry: techniques and applications for characterization of biomaterial surfaces. *Biomaterials*, 24:3635–3653, 2003.
- [3] G. Binnig, C.F. Quate, and Ch. Gerber. Atomic force microscope. *Physical Review Letters*, 56(9):930–933, 1986.
- [4] K. Boussofiane-Baudin, K. Bolbach, A. Brunelle, S. Della-Negra, P. Hakensson, and Y. Le Beyec. Secondary ion emission under cluster impact at low energies (5-60 keV); influence of the number of atoms in the projectile. *Nuclear Instruments and Methods in Physics Research B*, 88:160–163, 1994.
- [5] Wolfgang Demtröder. *Atome, Moleküle und Festkörper*, volume 3 of *Experimentalphysik*. Springer, Berlin, third edition, 2005.
- [6] G. L. Friedli. *Interaction of deamidated soluble wheat protein (SWP) with other food proteins and metals*. Phd thesis, University of Surrey, 1996.
- [7] Hans-Joachim Galla. *Spektroskopische Methoden in der Biochemie*. Georg Thieme, Stuttgart, 1988.
- [8] Paul Geladi and Bruce R. Kowalski. Partial Least-Squares Regression: A Tutorial. *Analytica Chimica Acta*, 185:1–17, 1986.
- [9] Hubert Gnaser. *Oberflächenmassenspektrometrie*, 2007.
- [10] D.J. Graham, M.S. Wagner, and D.G. Castner. Information from complexity: Challenges of TOF-SIMS data interpretation. *Applied Surface Science*, 252:6860–6868, 2006.
- [11] Michael Guilhaus. Principles and Instrumentation in Time-of-flight Mass Spectrometry. *Journal of Mass Spectrometry*, 30:1519–1532, 1995.
- [12] C. Hannig, M. Hannig, and T. Attin. Enzymes in the acquired enamel pellicle. *European Journal of Oral Science*, 113:2–13, 2005.
- [13] Christian Hannig, Thomas Attin, Matthias Hannig, Elvira Henze, Kirsten Brinkmann, and Ronald Zech. Immobilisation and activity of human alpha-amylase in the acquired enamel pellicle. *Archives of Oral Biology*, 49:469–475, 2004.
- [14] Christian Hannig, Jörg Hoch, Klaus Becker, Matthias Hannig, and Thomas Attin. Lysozyme activity in the initially formed in situ pellicle. *Archives of Oral Biology*, 50:821–828, 2005.

- [15] M. Hannig and A. Joiner. The structure, function and properties of the acquired pellicle. In R. M. Duckworth, editor, *The Teeth and Their Environment*, volume 19 of *Monographs in Oral Science*, pages 29–64. Karger, Basel, 2006.
- [16] Joachim Hartung, Bärbel Epelt, and Karl-Heinz Klösener. *Statistik*. Oldenbourg, München, 1982.
- [17] J.E. Jackson. *A user's guide to principal components*. John Wiley and sons, inc., New York, 1991.
- [18] I.T. Jolliffe. *Principal component analysis*. Springer, New York, second edition, 2002.
- [19] David Kriesel. Ein kleiner Überblick über Neuronale Netze, 27. Aug. 2006.
- [20] J-B Lhoest, M.S. Wagner, C.D. Tidwell, and D.G. Castner. Characterization of adsorbed protein films by time of flight secondary ion mass spectrometry. *Journal of Biomedical Materials Research*, 57:432–440, 2001.
- [21] Albert Light. *Proteins - Structure and Function*. Prentice-Hall, Inc., Englewood Cliffs, 1974.
- [22] D.S. Mantus, B. D. Ratner, B.A. Carlson, and J.F. Moulder. Static Secondary Ion Mass Spectrometry of Adsorbed Proteins. *Analytical Chemistry*, 65:1431–1438, 1993.
- [23] R. Michel and D.G. Castner. Advances in time-of-flight secondary ion mass spectrometry analysis of protein films. *Surface and Interface Analysis*, 38:1386–1392, 2006.
- [24] Matthias Mondon. *Untersuchungen zur Proteinadsorption auf medizinisch relevanten Oberflächen mit Rasterkraftspektroskopie und dynamischer Kontaktwinkelanalyse*. Phd thesis, Universität Kaiserslautern, 2002.
- [25] Jochen Ohnsorge and Reimer Holm. *Rasterelektronenmikroskopie: Eine Einführung für Mediziner und Biologen*. Georg Thieme Publishers, Stuttgart, second edition, 1978.
- [26] O.D. Sanni, M.S. Wagner, D. Briggs, D.G. Castner, and J.C. Vickerman. Classification of adsorbed protein static ToF-SIMS spectra by principal component analysis and neural networks. *Surface and Interface Analysis*, 33:715–728, 2002.
- [27] F. Schmitt. *Cantilevermodifizierungen zur Untersuchung der Biofilmbildung auf Implantatmaterialien und der Bestimmung der Adhäsion von Botrytis Cinerea Konidien auf Modelloberflächen mit Hilfe der Rasterkraftmikroskopie*. diploma thesis, TU Kaiserslautern, 2006.
- [28] A. Schnieders, D. Rading, and M. Fartmann. Short course sims part 1 - fundamentals. In *SIMS Europe 2006*, Münster, Germany, 2006.
- [29] J. Shlens. A Tutorial on Principal Component Analysis, 10. Dec. 2005.

- [30] Lindsay I. Smith. A Tutorial on Principal Components Analysis, 26. Feb. 2002.
- [31] M.S. Wagner and D.G. Castner. Characterization of Adsorbed Protein Films by Time-of-Flight Secondary Ion Mass Spectrometry with Principal Component Analysis. *Langmuir*, 17:4649–4660, 2001.
- [32] M.S. Wagner, M. Shen, T.A. Horbett, and D.G. Castner. Quantitative time-of-flight secondary ion mass spectrometry for the characterization of multicomponent adsorbed protein films. *Applied Surface Science*, 203-204:704–709, 2003.
- [33] M.S. Wagner, B.J. Tyler, and D.G. Castner. Interpretation of Static Time-of-Flight Secondary Ion Mass Spectra of Adsorbed Protein Films by Multivariate Pattern Recognition. *Analytical Chemistry*, 74:1824–1835, 2002.
- [34] Daniel Weibel, Steve Wong, Nicholas Lockyer, Paul Blenkinsopp, Rowland Hill, and John C. Vickerman. A C60 Primary Ion Beam System for Time of Flight Secondary Ion Mass Spectrometry: Its Development and Secondary Ion Yield Characteristics. *Analytical Chemistry*, 75:1754–1764, 2003.
- [35] N. Xia and D.G. Castner. Preserving the structure of adsorbed protein films for time-of-flight secondary ion mass spectrometry analysis. *Journal of Biomedical Materials Research*, 67(A):179–190, 2003.
- [36] Y. Yao, E. A. Berg, C. E. Costello, R. F. Troxler, and F. G. Oppenheim. Identification of Protein Components in Human Acquired Enamel Pellicle and Whole Saliva Using Novel Proteomics Approaches. *The Journal of Biological Chemistry*, 278(7):5300–5308, 2003.
- [37] P.W. Yendle and H.J.H. Macfie. Discriminant Principal Components Analysis. *Journal of Chemometrics*, 3:589–600, 1989.

Acknowledgements

It would have been impossible for me to create this work without of the help of several people. I would especially like to thank the following:

Prof. Dr. Christiane Ziegler for giving me the possibility to do my diploma thesis on an interesting subject in her work group as well as for the freedom and helpful advice she gave to me for dealing with it.

apl. Prof. Dr. habil. Hubert Gnaser for an introduction to the ToF-SIMS technique and for fruitful discussions about the use of it.

Dr. Wolfgang Bock for solving several technical problems with the TRIFT apparatus.

Manfred Strack for help with all the other technical problems of everyday laboratory life.

Nicole Lawrence for all kinds of help throughout my first months in the scientific world, for many discussions about the mysteries of protein physics as well as for an introduction to scanning force microscopy.

Adam Orendorz for being helpful on various subjects and for sharing office and laboratory with me.

Dr. Stefan Trellenkamp and Dr. Bert Lagel, Nano+Bio Center Kaiserslautern, for taking scanning electron microscopy images of my samples and for help on their interpretation.

The whole group “Interfaces, Nano Materials and Biophysics” for their warm welcome and the pleasant atmosphere.

My parents for their moral and financial support throughout my studies.