

CHEMOMETRY OF PLASMA DESORPTION MASS SPECTROMETRY

E. R. Hilf, W. Schlez, K. Kruse, A. Wullweber, U. Steenen,
St. Harsdorff, K. Koch, B. Nitzschmann, F. Kammer, W. Tuszynski*

AG Theoretische Physik III; *AG Molekular-und Biophysik; FB Physik, Universität Oldenburg,
2900 Oldenburg, P.O.B.2503,FRG; E-Mail: 091853 at Dolunil.

Abstract: PDMS, the 'Plasma' Desorption Mass Spectrometry, is a method to desorb even very large and fragile organic molecules up to some 10.000 a.m.u. by means of very fast, energetic (MeV/nucleon) heavy ions impinging a solid surface with sample molecules adsorbed on it. A brief sketch of the underlying physics of the process is given. Several programmes to analyze, calculate, and transfer PDMS spectra are described in detail.

INTRODUCTION

With PDMS, discovered by R. Macfarlane[1], one can measure mass spectra of large fragile organic molecules and their fragments by bombarding solid samples with fast heavy ions. In this and the following section we dwell on the peculiarities of PDMS as compared with other mass spectrometric methods, and give some intuition for its underlying physics before reporting the specific programmes developed to analyze experimental spectra.

The yield of desorbed molecules is largest if the energy deposition by the incoming ion is electronically resonant, that is if the ion has about the velocity of the electrons of the solid. The energy deposition time is short as compared to atomic vibration times in the solid and the solid sample stays comparatively cool.

The Physics of the process is complicated: The energy is deposited, transferred, and transmuted by the successive steps of Coulomb-scattering of the incoming ion with electrons of the solid, yielding a shower of free secondary electrons, which by coorganization thermalizes and by adiabatic expansion cools and disperses until the coupling of them to the solid atoms and molecules becomes dominant, the energy being transferred to them by exciton-phonon coupling. This mechanical energy then relaxes by forming an explosive shock blasting off the material of a microscopic crater of some 50 \AA^3 which disintegrates by desorbing molecules, atoms, or clusters.

Typically, the kinetic energy distribution of the resulting particles centers around some eV which may be described by a 'temperature' of some $10^4 K$. The internal molecular temperature is however low, say, room temperature. Thus molecules of up to some 50.000 amu have been detected unfragmented. The detection of pigments, such as chlorophyll, fatty acids, or oligosaccharides, has been well studied, references might be found in [4].

This fast, explosive, cold, energetic, efficient desorption complements the better known SIMS or FAB, where the incoming ion is as slow as the atoms in a solid and the resulting fragments often have more internal excitation. In short terms we ascribe the terms *entropic vs. thermal* desorption to the two alternative processes.

Experimentally the fast incoming ions of some *MeV* per *amu* are produced either by one of the large heavy ion accelerators, as e.g. operated at the GSI Darmstadt, or much cheaper but with a given large energy-, charge-, and mass-spread of the ions, by using the fission fragments of a radioactive Cf^{252} -source, sometimes called the *heavy ion accelerator of the poor man*—who however has to be cautious for the extreme toxicity of the source material.

The number of desorbed fragments per incoming ion is typically about 1000 to 10.000, but only about 10 of them are ionized and can thus be used for detection (in contrast e.g. to the laser desorption technique with its much higher abundance of desorbed material and its ions). By a careful optimization of the preparation of the sample, the choice of the matrix material, and the method of adsorbing the sample to be analyzed, one thus can often tune the yield to dramatically larger abundances.

The detection of the desorbed ions is achieved by applying a voltage of about some *10keV*, and by measuring the time of flight for a distance of about a meter. The arrival times are measured by detectors offering e.g. some 65.000 time bins of the width of, say, $1/2\text{nsec.}$, each able to detect one ion. At most 256 events can be measured for a given incoming ion. By accumulating the results for many events over a chosen measuring time of the order of seconds to hours, the final 'experimental channel spectrum' is gained, given by 65.000 channel contents.

These data sets are the starting point for the here presented analyzing programme packages. Some of the programs may, after adaption, be useful for quite different physical processes, if only their results end up in discrete spectra.

The optimum yield for many organic materials is often achieved by using a matrix of molecules with nitrogroups such as nitroglycerin, and by using as little material as necessary for a less than monomolecular adsorption layer.

The PDMS spectrometry technique has been developed mostly at nuclear Physics Laboratories as e.g. the Institut Physique Nucleaire (Y. Lebeyec), the Nuclear Accelerator Laboratory of the University of Uppsala (B. Sundquist), or the Institut für Kernphysik, T. H. Darmstadt (K. Wien). Commercial Machines (less sophisticated but more user-friendly) may be bought from *Applied Biosystems*, USA, or from *Selmi*, USSR.

The experimental set-up at Oldenburg has been started in 1987 and was supported for two years by a grant of the Marine Research Division of the German Science Ministry, with the aim to develop and test a PDMS spectrometer for application to marine organic molecules. Our PDMS spectrometer was developed by K. Wien at the Institut für Kernphysik, Technische Hochschule Darmstadt, and jointly constructed at Darmstadt and Oldenburg. The electronics, and especially the TDC, the time digital converter was drawn from the Institut Physique Nucleaire, of the C.N.R.S. at Orsay, by Y. Lebeyec.

PHYSICS OF PDMS

We briefly summarize the present knowledge about the mechanism of desorption of PDMS, to support the understanding of the subsequently described spectral evaluation programmes.

By measuring the primary energy deposition of a heavy ion, impinging a solid surface, R. Macfarlane[1] found in 1974 that besides the then wellknown maximum when the velocity of the incoming ion is in the range of the velocities of atoms in the solid (some keV), (socalled *Nuclear sputtering*) there is a much higher energy deposition maximum, if the incoming ion flies at the much larger velocity of the electrons to be met in the solid, some MeV, (Rydberg velocity). This was then termed *electronic sputtering*. Since Macfarlane assumed that a plasma might be formed, he named the process *Plasma Desorption Mass Spectrometry*.

The main features of the process may be inferred from some observations:

- there is a threshold for the desorption: only above a certain energy a substantial yield of desorbed molecules is observed. Thus we assume a *nonlinear* process.
- the resulting abundance is rather high, per incoming primary ion about 10^4 desorbed molecules, fragments, or even larger clusters. This indicates an effective and *collective* process,
- the desorbed ions are fast (some eV) but their internal temperature is essentially low (0.001eV) allowing for essentially unfragmented, large, fragile, organic molecules. This nonequilibration of the temperature of thermally coupled subsystem is typical for an energetic but short-time process, an *entropic* process.
- the desorption heavily depends on the electronic properties of the solid, being most ineffective for metals and class I superconductors, (which relax electronic energy by emitting photons instead of trapping excitons) demonstrating that the initially deposited *energy is transferred by electrons* in the solid.

The theory of the process[2] is thus described by an initial set-off of fast secondary electrons, which then thermalize in flight, coorganize and cool by expansion, finally coupling more and more to the lattice, forming excitons as a means to transport energy without charge by electrons, which then get finally trapped depositing their energy to the chemical bonds as mechanical energy. The amount for each bond may be estimated by time dependent quantummechanical perturbation theory, which gives

$$e^{V'(t)\delta t^2} \quad \text{with} \quad T \approx (\delta t/h)^{-1} \simeq 10^4 K$$

where the 'temperature' T is not an equilibrium property but a measure of the time duration of the electronic excitation passing.

We now follow the process from here on up to the extraction of the relevant information from the final outcoming spectra by describing the numerical programs developed by us.

MOLECULAR DYNAMICS

We simulate the relaxation of the initial mechanical stress by a Molecular Dynamics calculation, tracking Newton's equation for some 10^5 to 10^6 atoms bound to each other by atomic forces.

Numerous MD programme packages have been developed by other groups in the past for many purposes such as solid state physics or the desorption by nuclear sputtering. There, the atoms either oscillate only around their equilibrium locations or move near equilibrium below the sound velocity.

In contrast here we had to handle the following problems:

- a large number of atoms has to be tracked due to the three-dimensional geometry of the explosively relaxing area of some $100A^3$. We had to track up to 10^6 atoms. This needed extensive numerical calculations in the order of some 1000 CPU hours at a large Computer Centre.
- the boundary conditions have to be chosen in such a way that there is no artificial reflection of neither momentum nor energy. This can be achieved by a non-local boundary condition.
- simultaneously atoms of any velocity above and below sound velocity may occur. Thus the sound velocity has to be independent of the iteration step-time, which is the case for many numerical codes, but does not harm there because only velocities much lower sound velocity occur. We solved this problem by choosing a non-local iteration procedure.
- the calculation has to be extended over many time steps until the desorption process has terminated with its large global dislocations. We iterated different regions with individual time steps, the shorter, the more violent collisions are occurring.

Results are given in Ref.[3]. The desorption from a lattice is effective only for a certain range of atomic bond strengths: for too soft potentials one gets no effective shock, for an intermediate range, just where most of the interatomic bonds are found, a nice explosive shock cone is formed. For too high bond strength just the surface layer is blasted off.

After desorption initial clusters cool off by disintegrating or by evaporating an atom or molecule.

For chemists the short time reactions in these excited clusters of almost arbitrarily designable chemical composition are a new field of research, largely unexplored yet.

The ionization before (for preformed ions such as in salts) during (as in most cases) or after (seldom) desorption has not been quantitatively studied yet.

The fragmentation of molecules will be model-calculated in the next chapter. In the next but one chapter we will come back to the extraction of chemical information drawn from the finally resulting spectra.

A programme package *PDMS-MD* is set up at present, suitable to be used by other laboratories to gain training and insight by doing their own smaller scale test Molecular Dynamics calculations.

FRAGMENTATION OF MOLECULES

To calculate a fragmentation of *thermally* excited molecules, as occur e.g. in FAB, SIMS, EI, and many other mass spectrometry methods, is a well-known though complicated task: By a multi-normal-mode coupled-vibration-calculation one has to calculate the collective, energetically low-lying excitations which eventually lead to fragmentation.

For the *entropic* fragmentation as encountered here for PDMS, for times short to the vibration times a large amount of energy is offered to each chemical bond simultaneously. Thus one just has to calculate the number of locally excited states for each bond individually which directly lead to fragmentation (direct fragmentation channels).

Thus the chance for a chemical bond to be broken can be assumed to be more independent of the other ones. It should be proportional to the number of direct exit channels for that bond. The energy needed to break the bond, however, does not enter.

However, for most chemical bond types, the energy level density is related to the potential depth, because strong bindings have normally also short equilibrium distances.

There are, though, some special cases: these are covalent-strong bonds with, however, an unusual high number of exit channels. These different ways to break up at the same bond may be called *exit isomerism*. Examples are the glycosidic or the peptidic bond types. The ratio of their break chance to the one of a v.d.Waals type bond is only about 1/20, in contrast to an ordinary covalent bond where this ratio is almost zero.

Thus we have set up a programme package to calculate the fragment mass spectrum for a given molecule with known chemical structure[5].

For large organic molecules such as chlorophyll-a the tracking of all the numerous possible fragments due to the complicated netted chemical structure needs a large effort of numerical calculation. Due to a suitably chosen effective algorithm we need only about 500 Mflop for this molecule which still can be handled on a PC. A menue-surface and an embedding of DOS-tools makes it comfortable to work with. The molecular structure can be uniquely noted by working with a list of bonds. Some automatic checks are done to reduce the amount of possible mistakes in typing in large molecules. CRUNCHER has proven to be a successful tool in our scientific work to compare PDMS-Spectra with expectations from fragmenting bond breaking probabilities.

EXTRACTION OF MASS LINE INFORMATION FROM CHANNEL SPECTRUM

Before describing programmes to handle PDMS channel spectra, we give some information on their structure, to differentiate from experimental mass spectra gained by other methods.

The ions are accelerated by a voltage of, say, $U \approx 10\text{keV}$, daigned to be large as compared to the initial kinetic energies of some eV. The time of flight is about

$$t \approx \sqrt{L/(2eU)}\sqrt{m}.$$

The ions are counted by an electronic stop watch, with the start signal gained by counting the other of the two fission fragments. the first one being the primary ion causing the desorption.

The arrival times are counted by means of a grid of 65.000 time bins (channels).

In case an ion is counted, the following up (about 256) channels are blocked. Each channel is good only for one event per start signal. In total up to 256 ions per start signal can be registered.

Thus one deals with a discrete spectrum. In contrast to a continuum spectrum (as e.g. registered by GC) a finer time grid yields more electronic noise for the same number of to be counted ions.

The information-theoretical analysis for such discrete spectra has to deal with their often large background and their wide lines (the line width due to the initial velocity comes out to be $\approx \sqrt{(m)}$, the number of channels per mass unit is of course $\approx 1/\sqrt{(m)}$).

The analysis and programmes, presented here for PDMS, complement the work of D. Gingras [6] for continuous spectra with a low background.

The starting point is the experimental spectrum, an observation-time integrated flight-time channel spectrum of typically 268kByte.

Basically there are two alternatives to extract the relevant mass spectrum information:

- the *smoothing* view of a trained expert's eye. We use the one of our colleague W. Tuszynski who heads our experimental group.
- *Automatization* by suitably designed computer programmes.

The latter should simulate the expert's work by a quantitative algorithm. In addition, it should be sufficiently fast, efficient and automatizable.

The aim is to extract information for each observed fragment such as mass number, number of registered ions (without background), and, may be with less desire of the experimentalist, the width and asymmetry of the spectral mass line, for comparison to calculations of in flight cluster decay, sample surface roughness, etc.

We developed two methods:

1. Viewing the channel spectrum as a vector in the linear channel vector space we calculate for each channel the scalar product with an idealized mass line of a fixed line width (which is somewhat known here and taken to be $\approx \sqrt{(m)}$). We then pick the channel with the largest scalar product, to be the first found mass line and subtract its contribution to all channels. Then we repeat the whole procedure to find the next largest mass line. This method proved to be fairly save and sound, but it is rather CPU-time consuming and not suited for automatic scanning. We do have an operative programme package though.
2. The second method calculates a set of differently weighted sums over the channel spectrum for each mass line searched for. It will be called *Far-Moment Analysis* here. For details of the method see ref.[4].

The programme starts with the first channel and systematically scans the whole channel spectrum only once.

The summation-interval width for which the set of moments is calculated is an adjustable parameter. For the shipped programme package[7] it is automatically optimized by fitting to mass lines in the

area of $m = 100$, and then set to be slightly larger than the line width, growing $\approx \sqrt{(m)}$.

In principle, if all moments would be calculated, the exact original channel spectrum could be recovered. However, from the first, say, five moments, we calculate the desired quantities in the order of assumed decreasing interest to the expert: mass, number of ions, width, and asymmetry of mass line.

The background subtraction can be unanimously done for such moments, since they are sums over the fluctuating spectrum, in contrast to the ambiguous subtraction of an average to the original channel spectrum which would result in uninterpretable negative counts..

To stimulate the appetite of the reader we give some examples of applying PEAKS.

From the examples given in fig. 1, one may infer that the overall agreement to the expectation of an expert's eye may be viewed as to be sufficiently good.

The first example is a pigment, fluorescein[10]. The second example in fig. 1., Chloroform at $-150K^\circ$, shows experimentally [9] broad lines, probably due to a rather rough sample surface. Despite that, PEAKS seems to get the assumed information with some superfluous low lines though. Since PEAKS gives the total abundance of a line, a broad line may give more counts than a narrow large channel peak.

Finally we show the PDMS spectrum of an extract of a Wadden Sea sediment sample[10] as an example for a spectrum with a large background which tests the limits of automatic extraction.

This method of extraction is numerically the faster the larger the number of channels, as compared to other extraction methods, since we read-in each channel content and calculate its contribution to the moments only once.

SIMILARITY SEARCH IN A SET OF SPECTRA

SIMILARY is a programme package[11] to

- condition mass spectra to be compared by the usual algebra such as *add, subtract, and, or, not, scale, pick, scim*.. before one starts using comparing programme packages such as PEDAS[12].
- compare two spectra of a data base by a similarity measure[11].

For the similarity measure many different measure definitions have been proposed, including ones using mass-dependent weights (large masses are less abundant and are less well detected but often of larger interest). SIMILARY compares the result of different measures.

One wellknown measure is the euclidean norm of the scalar product of two spectra, as one is used to from ordinary space vector operations. The disadvantages are also known:

1. The length of a spectrum, its norm, does not have any physical meaning for this measure, in contrast to its analogy of space vectors, where it is just the measurable length of the vector.
2. The subtraction of an eventual background is logically not uniquely possible.
3. It is quadratically insensitive to the occurrence of small lines, although in comparing two different

spectra already a single even though small line occurring in one but not in the other spectrum may clearly say, these two samples are different, whereas a slightly different height of the largest line does not contain much information. From this viewpoint the smaller (even negative) the power of the channel content would be chosen the more suitable the measure would be for mass spectrometry.

As a compromise we use the first power and call this the *linear norm*. This norm has the advantage that the length of a spectrum is just the total number of counts, which just is used for calibration anyhow. Due to this linearity the eventual background can be subtracted, and this similarity measure is more sensitive as the euclidean one to an eventual small admixture of a substance. The only disadvantage of the linear norm is its more subtle nonlinear algebra[11].

The programme SIMILARITY is ready for shipping to other laboratories.

SYNTHESIS

SYNTHESIS[13] is a programme, designed to support the tedious search for chemical structure pieces from a given mass spectrum.

The analysis of it is given in ref.[14]. The usual procedure of looking for each set of three mass lines, the sum of two of the masses giving the third one, to define a *bond* between the first two fragments, is carried to the next steps of three and four bond objects. Because of the drastically increasing possibilities the user with his expertise is integrated in the process to cut off chemically senseless alternatives as early as possible. Clearly we use the neginformation (masses which do not occur), as an equally important information as the occurring lines.

EMS-NET

The EMS-NET system, the *European Plasma Desorption Mass Spectrometry Data Transfer*[15] has been designed to transfer and exchange mass spectra, as an example of experimental data sets, between external laboratories.

Data sets may contain mass spectra and all their necessary supplementary information.

The concept is that each laboratory keeps and takes care of its own spectra, and defines which of them should be accessible by other laboratories. EMS-NET then just provides the necessary data communication-software and search language as well as tools to update these automatically at all connected laboratories.

EMS-NET can be in principle installed for all operation systems, but at present works on VM/XA machines. Languages used are FORTRAN77, REXX, and the retrieval language SQL. The user-surface was integrated into ISPF. The user has however only to answer the questions in the masks on the screen shown to him.. For the transfer lines we use the academic research network, WIN in Germany, or bitnet elsewhere.

ACKNOWLEDGEMENTS

Computations were done at the Computer centres of the GSI Darmstadt, the Regionales Rechenzentrum Niedersachsen and of the University of Oldenburg.

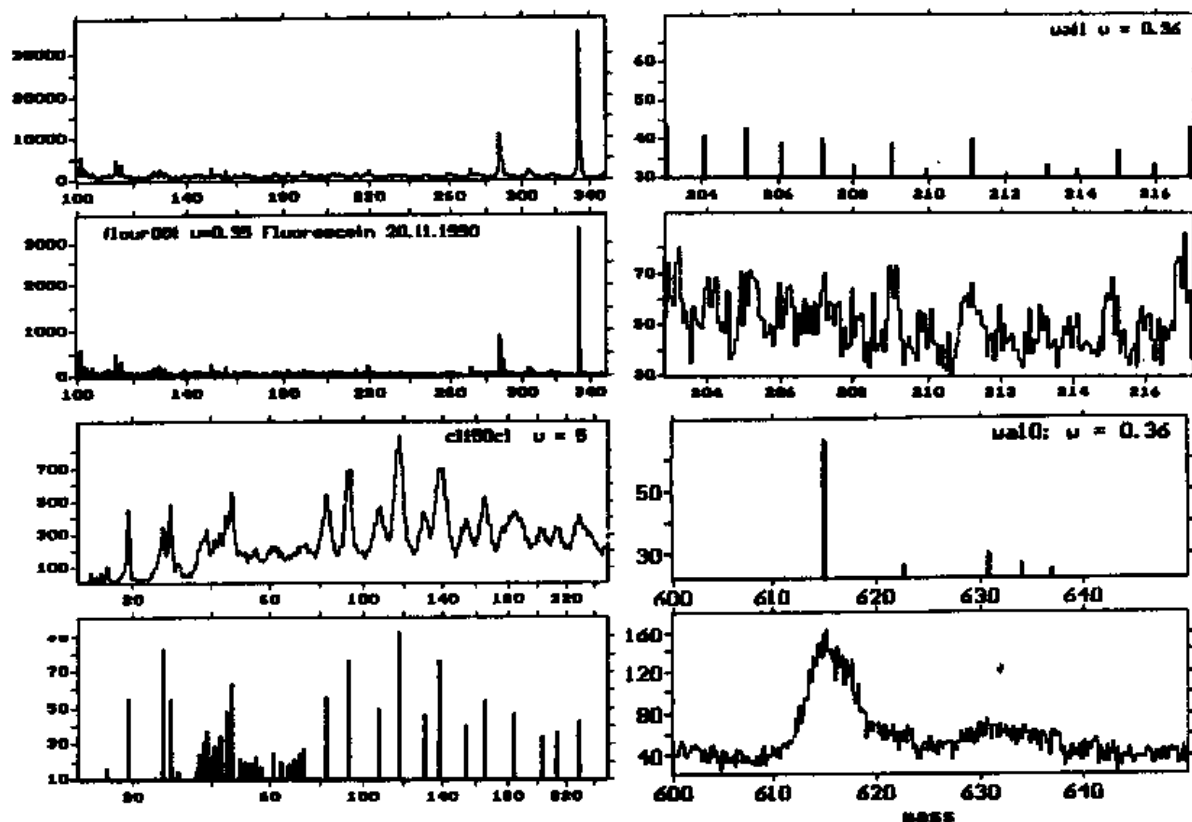


Figure 1 : PDMS spectra, and mass lines as extracted by PEAKS.

- 1 Fluorescein. Upper set: experimental PDMS channel spectrum[8]; lower set: mass lines gained by PEAKS, showing a generally good agreement
- 2 Chloroform. Upper set: experimental PDMS channel spectrum [9] with a large line width. Lower set: mass lines as of PEAKS,
- 3 Wadden Sea sediment extract[10]. Upper set: result of PEAKS for an integration width of 0.36 a.m.u., lower set: experimental channel spectrum, showing low statistics.
- 4 same sample as 3. The pheophytin peak with some rather low abundance side lines. Upper part: the PEAKS result, lower part is the experimental channel spectrum.

References

- [1] R. Macfarlane et al. (1974) *Biochem.Biophys.Research Comm.* 60:616
- [2] F. Kammer, (1991) Ph. D. Thesis, University Oldenburg
- [3] B. Nitzschmann, (1991) *Molecular Dynamics for PDMS*, Diploma Thesis, University Oldenburg
- [4] E. R. Hilf, W. Tuszynski; (eds.); (1990) *Mass Spectrometry of Non-Volatile Molecules for Marine Organic Chemistry*, publ.: WORLD SCIENTIFIC
- [5] K. Kruse, A. Dullweber, E. R. Hilf (1990) *CRUNCHER, a Programme Package to Calculate Fragment PDMS Spectra from Known Chemical Structure Molecules*; University Oldenburg
- [6] D. Gingras, (1989) in: *Computer-Einsatz in der Spektroskopie*, Int. Workshop, Gemen
- [7] W. Schlez, E. R. Hilf (1990) *PEAKS, a Programme Package to Extract Mass Line Informations from PDMS Channel Spectra*; Oldenburg,
- [8] W. Tuszynski, R. Angermann, F. Hillmann, et al. (1990) private communication prior to publication
- [9] K. Wien (1990) private communication prior to publication
- [10] W. Tuszynski, R. Angermann, F. Hillmann, K. Maier-Schwartz (1989) *Chlorophyll and Other Pigments in Photoactive and Buried Marine Microbial Mats*, in: ref.[4]
- [11] U. Steenen, E. R. Hilf (1990), *SIMILARITY, a Programme Package for Similarity Search in Spectral Bases*, Oldenburg
- [12] K. Varmuza (1989), *PEDAS*
- [13] P. Wagner, E. R. Hilf (1989), *SYNTHESIS, a Programme to Extract Chemical Structure Information Pieces from Mass Spectra*, Oldenburg
- [14] E. R. Hilf, St. Harsdorf, P. Wagner, W. Schlez (1989) *Chemometry of PDMS Spectra* in: ref.[4]
- [15] K. Koch, St. Harsdorff (1989) *EMS-NET, a Distributed Data Base for Experimental Data Transfer Between Different Laboratories with Automatized Access and Updating*, Oldenburg