MALDI Ionization: The Role of In-Plume Processes

R. Knochenmuss*\(^1\), R. Zenobi*\(^2\)

1) Novartis Pharma AG
Postfach
4002 Basel, Switzerland
tel: ++41 61 324 4541, fax: ++41 61 324 8868
rknochenmuss@gmx.net

2) Laboratory for Organic Chemistry
Swiss Federal Institute of Technology
ETH Hönggerberg
8093 Zürich, Switzerland
tel: ++41 1 632 4376, fax: ++41 1 632 1292
zenobi@org.chem.ethz.ch

* corresponding authors
Table of Contents

Introduction

Types of Secondary Reactions

Proton Transfer

Cation Transfer

Electron Transfer

Electron Capture

Signal Intensities and Suppression Effects

Factors Influencing Relative Ion Intensities

Concentration Dependence of Ion Signals

Matrix Suppression Effect

Matrix Suppression and Matrix Ion Interconversion Reactions

Analyte Suppression Effect

Role of Clusters

Ionization Potentials of Clusters

Proton Transfer Reactions of Clusters

Recent Cluster Models

Multiply Charged Ions

IR vs. UV MALDI

Prospects for Improved/Quantitative MALDI
**Introduction**

The factors determining the observed ion distribution in MALDI continue to be of considerable interest, in the hope that the method can be placed on a rational and predictive foundation. There have been a number of qualitative mechanistic proposals which include a wide variety of rather different processes. Knochenmuss et al. have recently proposed that secondary reactions in the MALDI plume may in many cases be the dominant determinant of the final, detected mass spectrum. This quantitative, thermodynamically-based proposal was built on earlier qualitative indications that analyte ions are formed either predominantly or in part via secondary reactions with matrix or metal ions. The thermodynamic approach also extends earlier studies of systematic influences on MALDI spectra.

One motivation for separating primary and secondary events in MALDI is one of time scale. The laser pulse typically lasts 3-5 nanoseconds (N₂ or Nd:YAG lasers) but the time required for expansion to collision-free densities is much longer, many microseconds. With the possible exception of preformed ions that are liberated later (e.g. by cluster evaporation, vide infra), the primary ions will be generated during the laser pulse or within the excited state lifetime of the matrix (also only a few nanoseconds). In the
expanding plume, however, reactions between ions and neutrals will continue as long as there are collisions. If the number and energies of these collisions is high enough, any thermodynamically favorable processes can proceed to equilibrium. Note that the time scales of primary and secondary reactions may or may not be similar, the key consideration is the sequence of the reactions.

To give a more concrete example, primary matrix ionization in a commonly used UV-MALDI matrix (2,5 dihydroxybenzoic acid, DHB) has been shown to have a time scale for the dominant process of about 2 nanoseconds.\textsuperscript{33} Since plume velocities are about 500-1000 m/s, the forward expansion is 1-2 micrometers. Energy deficit studies suggest that some analyte ions may be formed later, at considerable distances (tens of µm), and long times (10s to 100s of nanoseconds) after the laser pulse.\textsuperscript{4,20,25,34} Similar results were obtained with pulsed extraction TOF experiments.\textsuperscript{21,35} Simulations of the plume show that in this time range the plume density is still up to 10% of the pre-desorption solid.\textsuperscript{36} Plume temperatures of about 500 K have also been observed.\textsuperscript{37,38} The mean free path of ions and molecules is quite short, only a few molecular diameters, and the collision rate high.

It should be clearly noted that this review does not consider (or rule out) any particular primary ionization events or processes. It should therefore not be considered a complete picture of MALDI ionization. Complementary
literature, including other contributions to this issue, should be consulted for a more complete picture. However, regardless of the means by which initial ions are created or liberated, they will always be subject to the secondary reactions considered here. In such ion-molecule reactions, matrix is the dominant neutral partner, simply because it is nearly always present in substantial excess: the ion to neutral ratio in MALDI has been reported as $10^{-4} - 10^{-7}$.\textsuperscript{39,40} Other reactions may take place, such as ion-ion recombination, but they will be less frequent.

**Types of Secondary Reaction**

**Proton Transfer**

Proton transfer is probably the most important secondary reaction in MALDI. Many analyte classes, especially proteins and peptides, are detected predominantly in protonated form in MALDI.

Neutral analytes such as peptides and proteins have rather high proton affinities (PAs), typically at least 900 kJ/mol and often much greater.\textsuperscript{41-45} This is significantly greater than PAs of typical matrices, which are between 850 and 900 kJ/mol.\textsuperscript{15,30,46,47} It is therefore expected that plume proton transfer reactions of the primary protonated matrix, MH\textsuperscript{+}, with analyte are efficient (this does not exclude the additional presence of "preformed" AH\textsuperscript{+} in the plume):
\[ \text{MH}^+ + \text{A} \rightarrow \text{M} + \text{AH}^+ \]

Strong evidence for thermodynamic equilibrium of matrix/analyte plume proton transfer reactions has recently been presented by Kinsel, \textit{et al.}\textsuperscript{48} Mixtures of conventional MALDI matrices with various amino acids having increasing gas-phase basicities (GB’s) were studied. Solid mixtures of matrix and amino acid were ground to a fine powder, and a thick sample affixed to the metal substrate using adhesive. MALDI was performed at 337 nm and ToF spectra were acquired under continuous extraction conditions.

The data were evaluated analogously to the kinetic method for the determination of gas-phase thermodynamic properties.\textsuperscript{49} If equilibrium is achieved and no kinetic barrier exists, a plot of the natural log of the acceptor to donor ratio versus the GB of the acceptor is linear. (The amino acids were chosen for minimal intramolecular hydrogen bonding that could lead to substantial entropic barriers to protonation.) Figure 1 shows results obtained for mixtures of the MALDI matrix α-cyano-4-hydroxycinnamic acid with 5 amino acids. The kinetic method plot is clearly linear, validating the equilibrium picture of secondary reactions in the plume.

\textbf{FIGURE 1}

This approach yields additional information about the plume. For example, the x-intercept should correspond to the GB of the proton donor. In Figure 1
the intercept is at 845.4 kJ/mol. If this is correct, it is substantially lower
than the 858 kJ/mol GB of the deprotonated 2,5-dihydroxybenzoic acid
radical, consistent with the observation that α-cyano-4-hydroxycinnamic
acid causes more analyte fragmentation than DHB, and the concept that
analyte internal energy is strongly dependent on the exothermicity of the
proton transfer reaction.

In addition, the slope of the plot yields the effective temperature of the
proton transfer reaction. From Figure 1 an effective temperature of 1733 K
is obtained. This value is substantially higher than other estimates, which is likely due to the atypically high proportion of analyte (matrix : analyze molar ratio=1:4) used in these experiments. The Kinsel group has
recently found that as the matrix : analyte ratio is increased the effective
temperatures drop substantially. At matrix-to-analyte ratios of 15:1
effective temperatures in the 600 – 800 K range are obtained, more in line
with other experiments.

As briefly noted above, one indication of matrix-analyte proton transfer
was recently found in the correlation of analyte internal energy with the
exothermicity of the reaction. Other potential contributors to analyte
activation (e.g. matrix sublimation temperature) were ruled out. This result
is evidence for extensive plume reactions (and against preformed
protonated analyte), but does not demonstrate equilibrium.
Similar indications for the role of plume proton transfer reactions are known for peptide mixtures. Analyte basicity has often been found to correlate well with the relative strength of protonated signals, although there are indications that other analyte characteristics also contribute. Analyte acidity or basicity can be modified by adding protonation (or deprotonation) "tags" to analytes. These functional groups may lead to preformed ions in the prepared sample as well as enhanced reactivity in the plume.

MALDI is frequently performed with acidic matrices and positive ion spectra are measured. If the analyte is expected to have a low proton affinity, it can be preferable to measure negative ions, and to use a more basic matrix. Extensive lists of basic matrices have been compiled. In these cases primary deprotonated matrix can efficiently abstract protons from the neutral analytes, and/or will not donate protons to deprotonated analyte.

Sometimes analytes may give poor MALDI signal because of their solution-phase behavior. Oligonucleotides or phosphorylated peptides tend to be deprotonated in dilute solution and then form salts as the sample crystallizes. Such salts are then unlikely to protonate in the gas phase. By means of appropriate additives, they can be at least partially embedded in the solid sample in neutral (i.e. non-salt) form. Neutrals of this type
are found to be much easier to protonate in the gas phase, enhancing sensitivity. There is also a clear inverse correlation between matrix proton affinity and protonated oligonucleotide signal, as expected from the secondary ionization model.⁵⁴

Proton transfer can play a role in the generation of the more exotic ions that sometimes are observed in MALDI. Adduct ions of complex composition sometimes appear, with a net charge of 1. The low charge state is often achieved by loss of protons. For example, ions may contain multiple divalent metals (e.g. Ca²⁺) yet have a net charge of +1 due to proton loss.⁶⁶ Certain metals seem to favor this pathway, particularly as complex adducts with deprotonated matrix.³,⁶⁶-⁶⁸ The energetics of proton transfer to neutral matrix in this reduction process have been shown to be favorable.¹

Matrix is apparently not the only source of protons. Wong et al.⁶⁹ used matrices possessing no (solution-phase) exchangeable protons as well as deuterated solvents and deuterated matrices. Non-matrix proton sources examined were exchangeable solvent protons, analyte molecules, and traces of water taken up by hygroscopic solvents. These authors concluded that fast proton exchange between excited matrix and the surrounding solvent is unimportant and therefore not a necessary pathway for analyte protonation. This is in agreement with two earlier studies where derivatization of matrix hydroxyl groups was employed.²⁴,⁷⁰ Protonated
analyte could always be observed, although the efficiency of analyte protonation greatly varied depending on the exact experimental conditions used.

Related to proton transfer is the possibility of hydrogen atom transfer. Sometimes signals are found at M+1, M+2 or even M+3, where M is the expected molecular weight. An example is the observation by Calba et al.\textsuperscript{71,72} of H atom transfer in MALDI ionization of photochromic systems, producing an intense signal of (M + 3H)\(^+\) in the mass spectrum. From experiments with deuterated matrix, it was shown that the H atoms originated from matrix molecules, but this did not reveal at what stage of the MALDI process they are transfered. However, extensive "scrambling" due to ion-molecule reactions in the plume is probably involved. This known to occur in a wide variety of positive ions.\textsuperscript{73} Hydrogen atoms that are not exchangable in solution or in neutral molecules can be extremely labile in ions.

**Cation Transfer**

Adducts of analyte with various cations can either be a nuisance when other signals are already strong (e.g. protonated species), or very useful when they are weak. Particularly for the analysis of synthetic polymers, cations are often intentionally added to enhance signals.
The thermodynamics of possible plume cation transfer reactions is slowly becoming known, as is their role in MALDI. Cation affinities are much lower than proton affinities. For matrices, Na\(^+\) affinities are in the range of 150-170 kJ/mol. Potassium affinities are even lower, on the order of 50 kJ/mol. For most amino acids Na\(^+\) affinities are >150 kJ/mol, and for dipeptides they are >160 kJ/mol. Nucleobases have higher Na\(^+\) affinities of 164-190 kJ/mol, as do carbohydrates (>160 kJ/mol).

The first question regarding cation adducts in MALDI is why they give the dominant signal in certain cases, although proton affinities are always higher. Again the answer lies in secondary plume reactions. Cationization is typically important for analytes with a low proton affinity, less than that of the matrix. Analyte must therefore compete with neutral matrix for available protons in plume reactions. Polystyrene, for example, probably has a proton affinity near that of benzene (750 kJ/mol), so it will clearly not be able to compete with typical matrices having proton affinities of >850 kJ/mol. Cationization thus becomes the ionization route of last resort.

Among the more popular matrices used for cationizing analytes is dithranol. It has one of the lowest Na\(^+\) affinities of those measured, 150 kJ/mol. Transfer of the cation to most analytes of the categories noted above is exothermic by 15-40 kJ/mol. The low sodium affinity of the matrix is could
well be the reason it has empirically become so popular for this purpose. Should dithranol-Na\(^+\) complexes form, they will readily transfer the sodium ion to analyte. Equally important if preformed analyte-Na\(^+\) complexes exist, the matrix will not abstract the alkali ion. The barrier for Na\(^+\) transfer has been found to be typically below 10 kJ/mol\(^\text{74}\) on the same order of magnitude as the thermal energy available in the warm plume. Cation transfer is therefore not expected to be kinetically limited.

**Electron Transfer**

Molecular radical cations or anions are observed for a number of matrices. Chemical intuition suggests that these could be quite reactive in the plume. The simplest type of matrix-analyte ion-molecule reaction, electron transfer, has been demonstrated for certain combinations of matrix and analyte\(^\text{27,77}\). As illustrated in Fig. 2, if the matrix ionization potential (IP) is greater than the IP of the analyte, reaction can be efficient, otherwise the analyte is not ionized:

\[
M^+ + A \rightarrow M + A^+ \quad \Delta H = \text{IP}(A) - \text{IP}(M)
\]

**FIGURE 2**

Such reactions are not widely observed in UV-MALDI since the IPs of matrices are rather low, usually not greater than those of the analytes\(^\text{2,78}\).
However, they have been utilized to enhance the signal of low ionization potential analytes. Primary matrix cations are believed to abstract electrons from analytes in secondary reactions, leading to analyte radical cations.

The reverse reaction, electron transfer from neutral matrix to analyte ions, is very important in determining the observed MALDI spectrum. Any doubly charged (localized) ions are reduced to the singly charged state, but further reduction is not possible. A simple example is that of a divalent metal ion, e.g. Cu$^{2+}$. The gas-phase second IP of Cu is 20.292 eV, compared to, for example, a matrix IP of 8.05 eV (2,5 DHB). The reaction:

$$\text{Cu}^{2+} + \text{M} \rightarrow \text{Cu}^{+} + \text{M}^+$$

$$\Delta G = \text{IP(M)} - \text{IP(Cu}^{+}) = -12.24 \text{ eV}$$

where M is matrix, is exothermic by about 12 eV. If any Cu$^{2+}$ is somehow present in the plume, it will not survive plume collisions, but reduction will not continue because the first IP of copper is only 7.726 eV.

The same holds for divalent ion-neutral analyte complexes. The $(\text{ACu})^+ \rightarrow (\text{ACu})^{2+} + e^-$ reaction energy (second IP of ACu) will be less than that of the free Cu case, but not enough to change the overall picture. The second IP
would have to drop below 8 eV, which, due to the Coulomb energy involved, is very unlikely. For example, \((\text{DHB-Na})^+ \rightarrow (\text{DHB-Na})^{2+} + e^-\) requires 11.2 eV, and the IPs of singly protonated peptides are above 10.5 eV.\(^{80}\)

Rashidezadeh et al.\(^ {68}\) studied reduction of higher oxidation state cationization agents for the detection of cationized polystyrene by MALDI-MS. Their results are consistent with the secondary plume reaction model in all cases but one. Working with dithranol and all-trans-retionic acid as matrices and Ag trifluoroacetate, Cu(acac)\(_2\), Cr(acac)\(_3\), and Pd(acac)\(_2\) as adducts for cationization, they noted an inability of Cr to cationize polystyrene. In the low mass range, \((\text{Cr(acac})_2)^+\) and \((\text{Cr(acac})_2\text{H})^+\) ions were observed, i.e. Cr occurred in the +2 and +3 oxidation states, as opposed to other metals that were mostly observed in the form of matrix clusters with metals in the +1 oxidation state. The first, second, and third ionization potentials are on the order of 7, 18, and 33 eV for all these metals.\(^ {79}\) Thus, given the range of matrix ionization potentials, collisions between higher oxidation state metal ions or metal complexes with matrix should result in +1 oxidation state species, as observed for Pd and Cu. The reason for slow neutralization of Cr ions has been proposed to be slow exchange of matrix into the first coordination sphere, limiting the electron transfer rate.\(^ {3}\) This represents one of the few currently suspected cases where kinetics rather than thermodynamics may be the prime determinant of the ions observed.
Hunsucker et al.\textsuperscript{61} studied mono- and trimetallic complexes of Ir, Ru, and Rh with polypyridyl ligands. They observed what appears to be reduction during MALDI of the metal centers, which are di- or trivalent in solution. Reduction was only partial, and the remaining positive charge was compensated by retained PF\textsubscript{6}\textsuperscript{-} ligands. Again questions arise of kinetically limited electron transfer in strongly bound complexes of large ligands. Unfortunately, their data neither reveals the reduction mechanism nor the nature of the reductant. A possible interpretation is reduction by collisions with matrix in the MALDI plume.

**Electron Capture**

One important aspect of a recently proposed mechanism invokes charge-dependent capture rates for free electrons to explain the prevalence of +1 ions in MALDI.\textsuperscript{18} Multiply charged ions and capture rates will be discussed below, only the general concept of electron capture will be considered here.

For free electrons to play a significant role in MALDI they must be present in substantial quantities. A large negative polarity signal can indeed sometimes be observed at very short time-of-flight. This is probably due to electrons. However, they are not necessarily always abundant.
It has recently been shown in an ICR instrument that electrons are only abundant when the MALDI sample is thin or does not cover the metallic sample holder. A thick sample which completely covers the metal yields 50 times less electrons but a similar amount of negative matrix ions. It was concluded that the vast majority of electrons that appear in UV-MALDI are due to photoelectric emission from the "dirty" metal substrate. The presence of matrix or contaminants in small amounts on the metal enhances the emission via band bending and the associated reduction in work function.

Electrons thus appear to be an artifact of substrate irradiation, and not intrinsic to MALDI. They can have an effect on the mass spectrum because they lead to an excess of negative charge in the plume. The positive ion spectrum of a thin sample is about a factor of 10 weaker than that of a thick sample, apparently due to partial neutralization by electrons. Not only is the signal strength affected, the excess negative charge leads to lower charge states, as recently shown for copper complexes.

The negative ion mass spectrum of a thin sample is slightly stronger than that of a thick sample, but about a factor of two. This is expected, since the electron affinities of molecules similar to typical matrices are more favorable than -1 eV, leaving M⁻ as the primary negative charge carrier (and subsequent products, see below). These effects may be less pronounced in TOF-MALDI experiments in which electrons are not
constrained by the magnetic field, and may be quickly extracted by large electric fields.

Electron capture mass spectra of several MALDI matrices have been recently measured and lend support to the above concepts. The matrices studied all have peaks in the capture cross section at low electron energy. This is advantageous for MALDI since most proposed electron emission processes would generate electrons with no more than a few electron volts of energy. Many matrices undergo efficient post-capture fragmentation, leading to a variety of products. Loss of protons and carboxyl groups are frequent decay routes. Particularly important is the generation of (M-H)_-, in some cases with 100% efficiency. This ion is of obvious importance for possible proton transfer reactions leading to deprotonated analyte ions.

**Signal Intensities and Suppression Effects**

Secondary ion-molecule reactions not only explain which ions are observed in the final mass spectrum, they also help to understand relative ion intensities. This is particularly important if UV-MALDI is to be developed into a quantitative method. In addition to relative intensities, absolute yields also need to be predictable, which will require better understanding of the primary ionization events as well.
Factors influencing relative ion intensities

The thermodynamic model of ref. 1 has high predictive value if some conditions are met:
- good mixing of matrix and analyte in the sample
- sufficient collisions in the plume
- thermally accessible activation energies at typical plume temperatures

The first condition is readily understood: without a homogenous sample preparation the results of any MALDI experiment are unpredictable in many respects. This is often one of the most difficult aspects of MALDI to control.

The second condition was shown in refs. 1 and 85 to be directly correlated with the laser fluence. Near the ion production threshold, the plume may be insufficiently dense to attain thermodynamic equilibrium. Fluences comfortably above threshold (i.e. a "normal" MALDI experiment) appear to give a dense plume and predictable results. This is illustrated in Figure 4.

**FIGURE 4**

Should relevant secondary reactions have high activation energies, an kinetically limited ion distribution could arise. At present very few cases are known or suspected (see the discussion of chromium complexes above). Proton transfer reactions generally have quite low barriers if $\Delta G$ is more negative than about -25 kJ/mol.\textsuperscript{86} The PA differences between matrix and
most biomolecules are usually greater than this. Cation transfer reactions have been less well studied, but for sodium ions the barrier to transfer appears to be typically 5-10 kJ/mol, and so thermally quite accessible in the plume.

Concentration Dependence of Ion Signals

For purposes of quantitation, concentration calibration curves must be feasible. Ideally a linear relationship between analyte concentration in the prepared sample and the MALDI ion signal should be observed. Even when homogenous samples can be prepared, this is very often not the case. As shown in refs. 87,88 in many cases there are dramatically non-linear concentration effects. These can also be understood in terms of secondary ion-molecule reactions, as discussed in the next section.

Matrix Suppression Effect

The most dramatic effect involving signal intensities in MALDI is the matrix suppression effect (MSE).\textsuperscript{87-90} If enough analyte is present in a MALDI sample, the matrix ions can be completely suppressed, as shown in Figure 5. An "ideal" mass spectrum results in which only analyte is observed. In those cases for which thermodynamic data is known, it has always been found that suppression is fully consistent with the thermodynamic model.\textsuperscript{1,85} Matrix suppression has been observed many times, for many combinations of matrix and analyte. It is clearly not an effect of a few special analytes on a few matrices.
When secondary reactions of primary matrix ions with analyte are thermodynamically favorable, they can proceed to completion (i.e. equilibrium) and are generally reagent limited. When analyte is in low concentration, it is limiting and matrix ions are left over. When analyte is abundant, matrix ions are limiting and analyte ions are the sole products. As a consequence, above a certain concentration, analyte signal cannot be increased by adding more to the sample.

The typical concentration curve for a suppressing analyte/matrix pair typically has a linear region at quite low analyte concentration, a plateau at medium concentrations and a sudden transition to suppression at high concentrations.\textsuperscript{88} The plateau region is not well understood and may have to do with lateral diffusion ranges within the plume. An example is shown in Figure 6.

In the opposite polarity to that in which MSE is observed, analyte signals are generally weaker or even very weak. There is normally a reasonably linear concentration curve with no major structure. These observations are consistent with either i) preformed ions as the major ion source (i.e. little
plume chemistry), or ii) secondary ionization pathways with poor driving force.

The first possibility is straightforward, the second deserves some comment. Consider the case where matrix to analyte proton transfer is strong, so that the MSE will be observed in positive mode:

$$MH^+ + A \rightarrow M + AH^+$$

The corresponding reaction in negative mode, proton transfer from neutral analyte to deprotonated matrix is not the inverse of the positive mode reaction:

$$(M-H)^- + A \rightarrow M + (A-H)^-$$

It may occur even if the first reaction is very favorable. At the same time, if the first reaction is favorable this one is probably much less so, because the acidity or basicity of a neutral molecule is strongly correlated with the acidity or basicity of the corresponding protonated or deprotonated species.

There is a pronounced molecular weight effect on the MSE.\textsuperscript{87,88,90} Larger molecules yield suppression at a higher concentration in the sample.\textsuperscript{87,88} This is consistent with a model in which suppression appears when the analytes are separated by roughly 2 layers of matrix.\textsuperscript{88} The MSE concentration is dependent on the physical size of the analyte as well as its molecular mass.
Matrix Suppression and Matrix Ion Interconversion Reactions

A remarkable aspect of the MSE is that all matrix ions vanish, regardless of analyte ion type. For example a protonating analyte can suppress matrix radical cations and matrix alkali adducts as well as protontated matrix.\textsuperscript{87,88} This has been quantitatively interpreted in terms of secondary plume reactions between matrix species.\textsuperscript{1}

At least for the matrix 2,5 DHB, it has been shown that all these matrix ions can be interconverted by reactions with neutral matrix.\textsuperscript{1} These are sufficiently close to isoenergetic that they should be facile under plume conditions. It is then clear why all ions are suppressed in the MSE. When a highly favorable matrix-analyte reaction depletes one matrix species, all others will be efficiently depleted through this channel as well.

The Analyte Suppression Effect

Analogous to the MSE, the thermodynamic model of the plume predicts that one analyte should be able to suppress another. This analyte suppression effect (ASE) was systematically sought and found for a test system,\textsuperscript{1} but may very well be a common phenomena. An example of the effect is shown in Fig. 7. It is well known, for example, that only a subset of peptides generated by a tryptic digest are observed in a MALDI spectrum. If the individual peptides are separated and analyzed separately, they are all observable. Peptide-peptide suppression in a complex mixture has been
found to be strongly correlated with gas-phase basicity.51

**FIGURE 7**

The ASE is fundamentally more complex than the MSE. It may appear due to competition for the same or different primary matrix ions, or via direct reaction between analytes. The concentration dependence may therefore become complex. The most straightforward ASE is probably direct competition for protonated matrix as suggested in early work.\textsuperscript{28,89} ASE was demonstrated and quantitatively treated for both similar (protonated) and dissimilar (protonated vs. sodiated) analytes.\textsuperscript{1} As for MSE, the latter must proceed via ion interconversion reactions.

**Role of Clusters**

There has been a recent surge of interest in the role of clusters in MALDI ionization. Again the question needs to be subdivided into parts pertaining to primary or secondary ionization. Only secondary ionization will be discussed here.

**Ionization Potentials of Clusters**

The ionization potentials of matrix clusters are reduced compared to the free molecules. In the case of 2,5 DHB, this has been directly measured by molecular beam methods. As seen in Fig. 8, the free molecule IP is at 8.054
eV,\textsuperscript{78} while that of clusters of up to 10 molecules drops to 7.82 eV.\textsuperscript{91} The trend with size has already reached a plateau at this size so there may be only a very slow trend downward at higher sizes.

**FIGURE 8**

The IPs are relevant for electron transfer reactions as noted above. The decreased IPs of clusters means that they have a decreased ability to ionize analyte. Free matrix ions are therefore more likely to be relevant to secondary electron transfer processes than clusters.

The IPs of matrix/analyte clusters can also be decreased with respect to the free species. This effect has been found to be more dramatic than the IP decreases of matrix clusters. For example the complex of DHB with 4 prolines was found to have an IP of 7.0 eV, or a reduction of over 1 eV from the lowest IP constituent (DHB, IP=8.054 eV\textsuperscript{78}).\textsuperscript{92} Although the IPs were not determined, very strong indications for similar decreases have been found for clusters of DHB with the tripeptide valine-proline-leucine,\textsuperscript{93} and the matrix sinapinic acid with proline, methionine and prolyl-methionine.\textsuperscript{5}

This is potentially very significant for charge transfer reactions in the plume. It brings the IP of a neutral peptide/matrix complex down to the range where it can be ionized by electron transfer to matrix radical cations.
The IP reduction in clusters may also be very important for primary mechanisms. Two-photon ionization with the usual 337 nm nitrogen or 355 nm Nd:YAG lasers becomes possible when IPs are below 7.4 or 7.0 eV, respectively. Primary ionization effects are, however, outside the scope of this review.

Proton Transfer Reactions of Clusters

Little information on the proton affinities of neutral matrix clusters is available. Due to delocalization, the proton affinity of already deprotonated clusters is less than that of the free, deprotonated matrix, as has been demonstrated for dimers. For ferulic acid and sinapinic acid the PA decreases going from monomer to dimer were close to 100 kJ/mol (monomers both 1400 kJ/mol). Similarly, the proton donating ability of a protonated cluster will be less than that of a single protonated matrix molecule.

The decreased reactivity of clusters vs. monomers means that single molecules will be the most active species (and are the most numerous), and limit the thermodynamics. Clusters should have no effect on the thermodynamically predicted outcome, although they may well affect the kinetics.
Recent Cluster Models

Matrix/analyte clusters take center stage in a couple of models about MALDI ion formation that are discussed more extensively in the contributions by Karas and coworkers and Tabet and coworkers in this issue. The argumentation in these models is based on signals of multiply charged analyte \((A + nH)^{n+}\) and of analyte multimers \((An + H)^+\), usually minor peaks in MALDI mass spectra. While it is unlikely that cluster models can capture all essential features of MALDI ionization, they may have some value for explaining the appearance of these signals in the mass spectra. Also, an interesting aspect of these proposals is an attempt to explain certain similarities between MALDI and electrospray mass spectra. Much work remains to be done to put cluster-based models on a firmer theoretical basis. In particular, thermodynamic and kinetic parameters of the species and reactions involved should be determined. They also will have to be put to the test in terms of their ability to make predictions. At any rate, these proposals have recently re-energized and enriched the discussion about MALDI ion formation.

It has been clear for some time that not only small clusters but sizeable chunks of material are liberated in the MALDI process. This has been shown both experimentally\(^ {35,94}\) as well as in simulations.\(^ {95-99}\) Furthermore, recent MS\(^ n\) experiments give direct evidence for matrix cluster ions as precursors in MALDI.\(^ {100}\) The number of clusters in the plume is small compared to the
number of free gas-phase molecules, (although the mole fraction contained in clusters is large) on the order of $10^{-3}$ to $10^{-4}$. However, the MALDI ion yield is small, too, on the order of $10^{-4}$. It is thus, in principle, conceivable that ion formation mechanisms based on clusters can account for all the observed ions.

The model of Karas and coworkers starts from possible similarities between electrospray ionization and MALDI and asks questions about the dominance of singly charged ions in MALDI mass spectra. These authors' model assumes that large biopolymers exist in the form of multiply protonated precursors in the acidic environment of the matrix, and that they are liberated within chunks or clusters of matrix upon laser irradiation of the MALDI sample. The clusters would then lose neutral matrix in the selvege region and decrease in size down to the embedded analyte ion. The model further assumes the generation of a substantial number of electrons by photoionization that serve to reduce the charge state of the embedded analyte ion. Most analyte should thus be reduced to charge state 0, but some analyte, called "lucky survivors," can survive as singly charged ions. While introducing some new concepts, this model is still unclear in several respects: (i) it must be shown that multiply protonated biopolymers exist within matrix clusters (rather than zwitterionic forms or ions whose charge is balanced by counterions close by). In this respect the recent studies of pH indicators in solid matrices are a step in the right direction. (ii) As noted
above, photoionization is not a viable mechanism for the production of excess electrons,\textsuperscript{78,91} a much more likely reducing agent would be neutral matrix or matrix anions.\textsuperscript{1} (iii) Electrons, even if they are present in the MALDI plume\textsuperscript{103} are presumably rapidly extracted from the plume in time-of-flight MALDI experiments performed in continuous extraction mode. (iv) Most importantly, the model is presented as a unified theory for MALDI ion production. This clearly cannot be the case in view of the many different matrix/analyte combinations that surely require a range of scenarios to be considered for MALDI ionization.

Tabet and coworkers have also recently presented their view of the role played by multiply charged precursor clusters in the formation of MALDI ions.\textsuperscript{104} Experimental investigations included the dependence of the abundance of nonspecific protonated insulin clusters on laser energy, acceleration voltage, and extraction delay time using sinapinic acid as a matrix.\textsuperscript{105} The concept presented in this work is that ions detected in MALDI TOF experiments result in part from fragmentation of larger clusters, where the laser and extraction conditions determine their internal energy and thus propensity to fragment. These authors found that in linear mode TOF mass spectrometry, the abundance of neutrals produced from fragmentation of (insulin)$_n$H$^+$ clusters increased with laser energy. The heaviest clusters were found to be less favored at higher laser energy, supporting the idea of more extensive fragmentation with increasing laser
energy. However, the total contribution of neutrals to the cluster ion signals was found to be small, below 10%, and differences in time of flight due to recoil energy in flight direction of fragmenting ions were always less than 10 ns. The abundance of cluster ions relative to the monomer was found to be independent of the acceleration field strength and of the extraction delay. Finally, the initial velocity of all cluster species was nearly identical. Overall, the findings point to a single origin of bovine insulin cluster ions, followed by fast fragmentation (desolvation) before any significant acceleration of the ions occurs.

These findings are in contradiction to those of Kinsel et al.,\textsuperscript{25} who studied the flight times of insulin clusters formed in MALDI TOF experiments using HCCA as matrix. They found systematic deviations from the expected Gaussian flight time distributions as a function of cluster size: tailing towards lower m/z for the protonated pentamer and towards higher m/z for the protonated dimer. The signals could be fitted with a two-component Gaussian profile. This led the authors to conclude that two different sources exist for the production of these cluster ions: a prompt source, for example due to direct liberation of preformed ions; and ions formed with some time delay, i.e. due to gas-phase ionization processes in the plume.

In another study, Livadaris et al.\textsuperscript{106} presented investigations of cluster ions of peptides in the 1000 to 1700 Da mass range, using a range of matrices and experimental conditions. The contribution of a broad signal that was
shown to be due to metastable decay in the ion source (i.e., not PSD) was found to be strongly dependent on the nature of the matrix, with low proton affinity matrices such as HABA or HCCA giving the strongest contribution. Their interpretation of the data discounts gas-phase ion-molecule reactions, but is instead based on the size and stability of precursor cluster ions, both of which are surmised to increase with decreasing matrix proton affinity. The authors argue that larger clusters can provide more efficient internal energy relaxation, and thus survive in the extraction region for some longer time before fragmenting. This model is interesting, but the concept that the cluster size is simply a function of analyte proton affinity seems unlikely. Crystal morphology or the interaction of the laser with the solid matrix may play a more important role. Furthermore, it is unclear whether metastable dissociation in the source results from cluster stabilization (longer survival) or rather from cluster activation; the latter would be compatible with a gas-phase thermodynamic picture.

The model of Ref. 104 discusses formation of multiply protonated analyte ions starting from matrix-analyte clusters. Assuming several protonated matrix molecules are initially present in a cluster, it is suggested that the protons are transferred to analyte in a stepwise fashion as the cluster shrinks in size, because the proton affinities of small matrix clusters or even free matrix molecules are expected to be smaller than that of a large cluster. However, as noted above, the thermodynamics of proton transfer from matrix to analyte will be most favorable for single matrix molecules, which
are also the most numerous species. As a result, matrix monomers can be used as the limiting case to evaluate this mechanism of multiple protonation. UV matrices have proton affinities in the range of 850 - 900 kJ/mol. Even in low charge states, measured peptide and small protein incremental proton affinities are below 850 kJ/mol. Proton transfer from matrix to analyte is not thermodynamically allowed. It thus seems extremely unlikely that larger matrix aggregates are a universal means for transferring protons to solvated multiply charged analytes as proposed in this model. This is not to say that multiply charged ions do not form, these are discussed next.

**Multiply Charged Ions**

The presence, or rather general lack of, doubly or higher charged ions in MALDI has been regarded by some authors as particularly significant. Within the thermodynamic picture of secondary processes, such ions are not excluded and their abundance can be qualitatively considered. The prediction is that the propensity for multiple charging increases with molecular size. As molecules become larger, the charges can be sufficiently separated to become more and more independent. The internal Coulomb energy of the system becomes sufficiently low that reduction via electron or proton transfer reactions with neutral matrix no longer occur, the ion remains multiply charged. Significant thermochemical data now exists for multiply charged species, allowing these processes to be considered
recently, the effects of chain length on proton transfer properties of peptides have been reexamined in detail. Consistent with earlier work, shorter peptides have markedly lower gas-phase acidities.

Highly charged small to medium sized ions can essentially always be reduced to the +1 state by reaction with neutral matrix. This includes both proton and electron transfer reactions. Thermochemical data show, for example, that multiply protonated peptides in the range of 1500 Da can be deprotonated by any one of several neutral matrices. There is no indication at present that the thermodynamic model does not hold for higher charge states in MALDI. Some recent studies on small molecules (porphyrins) claiming to show this are in fact excellent examples of the principles laid out in ref. 1 Specifically, presumably preformed doubly charged ions were only observed in the mass spectrum as +1 ions. This was interpreted as reduction by free electrons. The charge sites on these porphyrins are, however, just as close as those in gramicidin S, which was shown in ref. 1 to be readily reduced by matrix.

**IR vs. UV MALDI**

Some authors have noted that in those cases where IR and UV lasers can be used with the same matrix, the mass spectrum is quite similar. This is true over a wide range of laser pulse widths and wavelengths. A straightforward explanation for this is provided by the plume reaction
model of ref. 1. Even if the primary ions generated by different lasers are not the same, which may well be true, secondary plume reactions should convert them to the same thermodynamically most favorable ions. In this picture, it would be surprising if the final mass spectra were not similar. Secondary reactions should be especially extensive for IR MALDI, since much more sample is typically ablated per laser shot, resulting in a denser plume. In this context it is interesting to note that the secondary reaction model provides a straightforward explanation for the spot size effect. The analyte signal strength has been found to depend on the diameter of the irradiated sample spot, in both IR and UV MALDI. Larger spots give more signal, consistent with the expected denser plume. This has been quantitatively treated in a recent model. Clearly more information is needed to test the quantitative applicability of the model to the IR case.

**Prospects for Improved/Quantitative MALDI**

The secondary ion-molecule reaction model provides a useful framework for planning a MALDI experiment. It will become more useful as more thermodynamic data for matrices and analyte classes becomes available.

A planned analysis begins with the choice of ion to be detected. This, in turn, depends on the range of matrix properties available, and the nature of the analyte. If the analyte is likely to have a proton affinity significantly higher (lower) than that of the available matrices, protonation (deprotonation) should be the goal, due to generally better signals.
Otherwise less strongly bound cation adducts can be sought. If the analyte is unlikely to compete with neutral matrix for both protons and cations, it should be derivatized. In special cases the analyte may be expected to have a very low IP, in which case the matrix can be selected to yield radical analyte cations.

The sample should be prepared so that the analyte crystallizes in a reactive form. Analyte salts are generally undesirable if protonation is sought. They are also generally undesirable for further cationization, but may be easily dissociated to yield deprotonated analyte in some cases. Sample pretreatment for desalting may be necessary, the pH may be adjusted, or additives may help to precipitate the desired form.

If it is likely from the thermodynamics (highly favorable reactions leading to one analyte ion) that matrix suppression or analyte suppression effects could appear, the concentration ratios must be considered. Particularly for small molecules, suppression can be highly desirable to simplify the spectrum. At present, for complex mixtures there appears to be no straightforward method to ensure that all components are observed (analyte suppression effect). However, suppression effects in one polarity will probably not occur, or not in the same way, in the opposite polarity. This may expand the range of components detectable in a single sample. Otherwise a separation step must be used before MALDI.
For predictability, the analysis should generally be carried out with sufficient laser fluence for adequate plume collisions. The exception is when plume reactions are expected to deplete the desired ions, in which case one should operate as close to threshold as possible. The loss of resolution which can occur at higher laser fluences can be compensated by delayed extraction, which also allows secondary reactions to more easily reach their end point.

The excess energy released in secondary reactions can be sufficient to induce fragmentation. Depending on whether this is desirable or undesirable, matrix can be selected accordingly, to modulate the energy released in these reactions.

For positive ion mass spectrometry, photoelectrons released from the metal sample support may be deleterious. Either a thick sample, covering the metal, or a non-metallic support should be used. For negative ions, photoelectrons may add to the signal strength, so a thin sample is preferred.

Quantitation in MALDI is known to be difficult. Among the largest obstacles is inconsistency and inhomogeneity of sample preparation. Without this, thermodynamic considerations are of limited use. Sample inconsistency dictates the use of internal standards or standard additions. Secondary ion-molecule thermodynamics suggests that a standard of similar chemical class to the analyte should be used. If suppression is expected or possible, the
concentration range should be carefully considered, or a non-suppressing matrix selected. Mixtures can exhibit complex interferences, so quantitation should be performed on simple, few component samples whenever possible. Averaging over many laser shots and a large sample area\textsuperscript{116} are also clearly desirable.

**Acknowledgments**

The authors thank Prof. Gary Kinsel for providing Figure 1, and Prof. Patrick Limbach for providing Figure 2. This work was supported in part by Schweizerische Nationalfonds grant number 21-63558.
Literature Cited


(99) Zhigilei, L. V.; Garrison, B. J. *Appl. Phys.* 2000, 88, 1


Figure 1. Kinetic method plot derived from MALDI-ToF mass spectra of mixtures of the matrix α-cyano-4-hydroxycinnamic acid with the amino acids, G, A, V, I and F. The protonated amino acid ion signal, $XH^+$, was compared to the matrix ion signals. The plot versus the gas-phase basicity (GB) of the amino acids is linear, with a fitted intercept of 845 kJ/mol. Figure courtesy of G. Kinsel.
Figure 2. Part (a): MALDI-TOFMS spectrum of triphenylphosphine (m/z 262.3, IE = 7.8 eV) using 1,4 diphenyl-1,3-butadiene (m/z 206.27, IE = 8.5 eV) as a matrix. Part (b): MALDI-TOFMS spectrum of triphenylphosphine analyzed using perylene (m/z 252.3, IE = 6.96 eV) as a matrix. In (a) the low IP of triphenylphosphine thermodynamically allows electron transfer to matrix radical cations, whereas in (b) electron transfer is thermodynamically disfavored. Figure courtesy of P. Limbach.
Figure 3. Comparison of negative (a,c) and positive (b,d) mode MALDI mass spectra of DHB for thin (a,b) and thick (c,d) sample layers, recorded with an FT-ICR mass spectrometer. The results are consistent with photoelectron production from metal surfaces in the thin samples. Note the poor signal-to-noise ratio in (b) which results from neutralization of positive ions by photoelectrons. Effects of electrons are amplified by the presence of the magnetic field and the absence of electric fields in the drift region between the MALDI target and the trapped ion cell of the FT-ICR. Adapted from 81.
Figure 4. Fluence dependence of matrix and analyte signals for the tripeptide Gly-Gly-His in the matrix sinapinic acid (SA). Thermodynamic considerations predict matrix suppression, which is observed at higher fluences. At lower fluences, lack of suppression is attributed to insufficient plume density and collisions. Adapted from ref. 85.
Figure 5. Example of the matrix suppression effect in UV MALDI. The peptide substance P was mixed with the matrix 2,5 dihydroxybenzoic acid in two molar concentration ratios. At the higher analyte concentration, all matrix signals have been suppressed. Adapted from ref. 87.
Figure 6. Dependence of the matrix/analyte signal ratio (negative ions) on the concentration of analyte in the sample. The matrix was 3-amino quinoline (3AQ), and the analyte was MOPS (3-morpholino-propanesulfonic acid). Apparent is the non-linear behavior associated with appearance of the matrix suppression effect. The solid line is a guide for the eye, and intended only to show the approximate behavior. Adapted from ref. 88.
Figure 7. Example of the analyte suppression effect. Excess peptide substance P can suppress valinomycin in the positive mode MALDI mass spectrum. The matrix was 2.5 dihydroxybenzoic acid. Note that substance P is protonated, but valinomycin is cationized with sodium. Suppression of dissimilar ion types is believed to require intermediating ion-molecule reactions involving matrix species. Adapted from ref. 1.
Figure 8. Two-color, two-photon ionization thresholds for clusters of the MALDI matrix 2,5 dihydroxybenzoic acid. The exponential fit reaches an asymptotic value of 7.82 eV at large sizes, well above twice the nitrogen laser photon energy of 7.36 eV. Adapted from ref. 91.