USING MASS SPECTROMETRY FOR DRUG METABOLISM STUDIES

Absorption AD Distribution ME Excretion

Pharmacokinetics (PK)

Dose NCE (Drug) PO/IV

Plasma — A
Drug Levels — & PK Parameters LC-MS/MS

Brain — D
Drug Levels — LC-MS/MS or MS IMAGE — MALDI-MS/MS

Bile/Urine — ME
Metabolite ID & Drug Levels — LC-MS/MS

EDITED BY
Walter A. Korfmaecher

CRC PRESS
Chapter 4

Matrix Effects: Causes and Solutions

Hong Mei

4.1 Introduction

Due to the inherent selectivity and sensitivity of tandem mass spectrometry combined with the separation power of a liquid chromatographic system, liquid chromatography–tandem mass spectrometry (LC–MS/MS) has become the method of choice for quantitative analysis for both drug discovery and drug development studies in most pharmaceutical companies. When LC–MS/MS was first introduced, it was generally assumed that the high specificity and selectivity of LC–MS would eliminate extensive sample preparation and reduce time required for the chromatographic analysis, thus making LC–MS/MS a much better method than the classical HPLC/UV methods.1 The fast turnaround time of bioanalytical analysis with the use of LC–MS/MS has greatly accelerated pharmaceutical research. However, in more recent years, matrix ionization suppression issues in LC–MS/MS assays have been reported2–4 and these matrix effects have become one of the most important causes for failures and errors in bioanalysis.5

The existence of different matrix components in study samples as compared with calibration samples can cause many fold errors in accuracy, which can invalidate the analytical results and the calculation of pharmacokinetic parameters based on these data. Matrix ion suppression (the most common matrix effect) not only affects quantitative analysis in pharmacokinetic studies, but can also hamper qualitative analysis in metabolite identification studies.
For example, the severe ion suppression caused by nonvolatile salts in bile and urine can make even major metabolites undetectable. With the increasing number of LC–MS/MS assays that are applied to more complex matrices, such as cell cultures, plasma, bile, urine, feces, and tissue samples, the issue of matrix effects has gained more and more attention. Extensive studies have been conducted to obtain a better understanding of the mechanism of electrospray ionization and the factors that contribute to ionization suppression. At the same time, different approaches for the evaluation of matrix effects have been introduced, and various strategies for overcoming matrix effects have been proposed.

### 4.1.1 What are matrix effects

In the FDA guidelines for bioanalytical validation, matrix effects are defined as “interference from matrix components that are unrelated to the analyte.” This broad definition includes both ion enhancement and ion suppression; these effects can be caused by ionization competition of co-eluting components, “cross-talk” from metabolites or internal standards, signal enhancement caused by in-source fragmentation of metabolites, and low or variable analyte recovery due to strong binding of analytes to biological matrices. For bioanalytical LC–MS/MS assays, matrix effects usually refer to signal reduction or enhancement caused by co-eluting components that are not related to the analytes. Matrix effects can cause significant errors in precision and accuracy, thereby invalidating the assessment of pharmacokinetic results based on these LC–MS/MS assays. Compared to ion enhancement, ion suppression is more problematic in that it will reduce the sensitivity of the assay. If one cannot control the variability caused by matrix effects, both ion enhancement and ion suppression can be challenging, because both will result in poor reproducibility of results. When matrix effects cause differential suppression or enhancement between calibration samples and study samples, the accuracy of the assay results will be significantly affected.

Based on our limited understanding of LC–MS/MS matrix effects, the following common perception of matrix effects have been widely accepted: (1) atmospheric pressure chemical ionization (APCI) is less sensitive than electrospray ionization (ESI) in regard to matrix effects, and (2) extensive sample preparation may be required due to the need to separate the analyte from co-eluting endogenous matrix components. Recently, studies have shown that APCI can also exhibit severe matrix effects and that exogenous material can also be a major cause of ionization suppression. Therefore, a thorough understanding of the mechanisms of matrix effects can help one to avoid the problem of matrix effects.

The aim of this chapter is to discuss the possible mechanisms of LC–MS/MS matrix effects, to provide the guidelines for evaluating matrix effects and to propose strategies for overcoming matrix effects. By using this information, researchers should be able to develop faster and more reliable LC–MS–MS assays that are devoid of matrix effect problems.
4.2 Recent Literature Review

4.2.1 Mechanistic studies of matrix effects

The mechanism of matrix induced ion suppression or ion enhancement is still not fully understood, this is in part due to the fact that the mechanism of electrospray ionization has proven to be very difficult to establish. However, extensive investigations have been conducted to gain a better understanding of the electrospray ionization process and the causes of matrix induced ion suppression. Some investigations have focused on each individual step involved in the production of ions from solution phase to gas phase, while others have tried to identify the source of the interfering matrices.

4.2.1.1 Ionization process for analyte and matrix components in ESI

There are a number of papers that have described details of how ions are first generated in the solution phase and then converted to gas phase ions in the electrospray ionization (ESI) process.\textsuperscript{3,8–12} Basically, there are four critical steps in ESI that are important to mass response: (1) excess charge generation in the Taylor cone and ESI droplets; (2) uneven fission of parent droplets to very small, highly charged offspring droplets that readily transform to gas phase ions; (3) formation and transformation of gas phase ions and (4) separation of neutrals from charged ions. In ESI, the liquid is an electrolyte solution that is continuously flowing into a high voltage capillary tip where primary droplets are formed in the Taylor cone that is generated at the capillary tip. Due to the charge separation that is caused by the voltage gradient, these droplets have excess charge that exists on the surface of the droplet while solvated paired ions or neutrals are present in the inner part of the droplet (inner phase). The concentration of excess charge is determined by the flow rate and applied voltage, and its production rate is equal to the maximum rate of production of vapor phase ions. With applied heating and desolvation gas, continuous solvent evaporation at constant charge leads to droplet shrinkage and uneven fission to form offspring droplets from the surface phase of parent droplets, thus with significantly higher mass-to-charge ratio on the offspring droplets. The inner phase of large parent droplets contains ion pairs that will be less possible to be detected by the mass spectrometer. The repeated evaporation and uneven droplet fission leads ultimately to gas phase ions by one of two model mechanisms: evaporation from droplet surfaces during the fission process (the charged residue model),\textsuperscript{13} or formation of final droplets containing only one ion at the end of fission process (ion evaporation model).\textsuperscript{14} Gas phase ions would undergo gas phase ion reactions in the atmospheric ion sampling regions. Finally, the ultimate gas phase ions in the sampling region will be sampled through an orifice, into the differentially pumped regions of the mass spectrometer, while the neutrals, solids and liquids will be blocked by various means, e.g., an interface metal plate (skimmer) and curtain gas. Any matrix components that interfere with
any of the above-mentioned process could affect the ionization efficiency of an analyte. Many studies have already proven that for ESI, the intensity of ion signal is dependent on the chemical nature of the analyte, as well as many other factors, including the presence and concentration of electrolytes in the liquid, \(^{15}\) volatility of the solvent, \(^{11,16}\) surface activity of the droplet, \(^{9,12,17,18}\) presence of nonvolatile components, \(^{15,19}\) flow rate of electrosprayed solution, \(^{20}\) concentrations of other ionizable species, \(^{9,12}\) and competition of gas phase ion-transfer reaction between analytes and other ionized ions. \(^{11,16,22,23}\)

### 4.2.1.2 Property of analyte and ESI response

In addition to instrumental parameters, the most important factor that determines ESI responses is the physical and chemical nature of the analyte and co-eluting components. There are two models, the ion-evaporation model and the equilibrium-partitioning model, that have been developed to predict the ESI response of an analyte based on the properties and concentrations of the analyte and co-eluting components.

#### 4.2.1.2.1 Evaporation rate and ion-evaporation model

Based on the ion evaporation model (IEM) of gas phase ion formation described by Irbarne & Thomson, Tang and Kebarle proposed an equation to predict ESI response of analyte A in the presence of electrolytes (E) and other components (M) using the evaporation rates and the concentrations. \(^{14,21,24}\) The typical equations are listed below.

Two components: \[ I_A = fp \frac{k_a[A]}{k_a[A] + k_e[E]} I. \tag{4.1} \]

Three components: \[ I_A = fp \frac{k_a[A]}{k_a[A] + k_m[M] + k_e[E]} I. \tag{4.2} \]

As described by Tang and Kebarle, the product \(fp\) is a factor that was assumed to be independent of the chemical nature of the ions, \(f\) is the fraction of charges on the droplets that are converted to gas phase ions (desolvation efficiency) and \(p\) is the ion-sampling efficiency of the system. The bracketed ions are the concentrations of the analyte (A), a matrix component (M) and electrolyte species (E) and the \(k\)'s are the rate constants of ion evaporation. \(^{21}\) The rate constant for each ion can be calculated experimentally from the free energy of activation (\(\Delta G\)). \(^{3}\) The value of \(\Delta G\) depends on the number of the charges (\(N\)) and the radius of the droplet (\(R\)), and the distance of ion charges from the surface of the droplet (\(D\)). \(D\) reflects the extent of solvation. Strongly solvated ions, such as \(\text{Li}^+\), hold on strongly to a large number of solvent molecules and have larger \(D\), thus need more energy to evaporate. Generally, the ion
evaporation rate constant $k$ increases with $N$, decreases with $R$ and decreases with $D$.\textsuperscript{3,11}

Basically, this model predicts that the ions with higher charge density and lower desolvation will have higher ESI responses. This was the first mathematical model that established the dependence of ion intensity on its concentration. It also successfully modeled the ion suppression effect of NH$_4$Cl on a series of analytes.\textsuperscript{11} Furthermore, it also explained the saturation of calibration curve by suggesting that competition occurs when the sum of electrolyte species ($I_E$) and analyte ion ($I_A$) exceeds the total fixed available current ($I$). It demonstrated that the ion intensity, $I_A$, depends only on the ratio of $k_a/k_b$, but not on the individual $k_a$, as illustrated by Kebarle and Peschke in the following relationship developed using Equation 4.1:\textsuperscript{11}

\[
\frac{I_a}{I_b} = \frac{k_a[A]}{k_b[B]} \quad \text{or} \quad \frac{I_a}{I_b} = \frac{k_a}{k_b} \quad \text{when} \quad [A] = [B].
\] (4.3)

However, this model was built using data generated by a variety of metal cations that have different ion evaporation rates but no surface activity. Furthermore, this model of ion evaporation failed to predict the ESI responses for more complex organic molecules where surface activity plays an important role. The factor of surface activity had to be accounted for in the explanation of the analyte response at the high concentration range; however, surface activity was not included in the mathematical equation.\textsuperscript{21} Due to this reason, this model can only predict the response within a narrow range of analyte concentration. Furthermore, due to the exponential relationship between $k$ and $\Delta G$, a small experimental deviation in $\Delta G$ will cause a significant difference in $k$. As a result, the calculated theoretically generated rate constant ($k$) based on an experimentally obtained $\Delta G$ for the same ion exhibited a large range of values.\textsuperscript{11}

### 4.2.1.2.2 Surface activity and partitioning-equilibrium model

With the consideration of the importance of surface activity, Enke developed the partitioning-equilibrium model to predict the ESI response of an analyte with a single charge in the presence of matrix components based on the surface activity of the analyte and co-eluting components as well as the competition for the limited number of excess charge sites on the surface of the initial droplet without invoking the effect of ion evaporation.\textsuperscript{9} Excess charge on the surface of the Taylor cone and the droplets is generated by the intense electric field at the ESI capillary tip. The concentration of excess charge $[Q]$ is equal to the circuit current ($I$) divided by the product of the Faraday constant ($F$) and flow rate ($G$). In other words, $[Q]$ is determined by the applied voltage and flow rate. Therefore, the rate of production of surface excess charge is a constant at a fixed experimental condition. Thus, $[Q]$ is also the upper limit for the concentration of observable ions generated by the electrospray process and
equal to the sum of the concentrations of all the charged species on the surface, e.g., \([A^+]_s\) and \([E^+]_s\). Therefore, \([Q]\) can be described by the following equation:
\[
[Q] = IF/\Gamma = [A^+]_s + [E^+]_s.
\] (4.4)

As described previously, ESI droplets can be divided into two parts: the charged surface phase and the neutral interior phase. In all proposed mechanisms of gas ion formation, gas phase ions that are freed from the liquid phase are those charged ions at the droplet surface, even though ions are free to partition between the surface and interior phases. Therefore, ions that are better able to partition into and stay inside the surface phase would expect to have higher ESI responses than ions trapped in the interior phase. Surface affinity or surface activity is closely related to the nonpolarity of a molecule. Usually, higher hydrophobicity will lead to higher surface activity. An equilibrium-partitioning coefficient \((K)\) was used in this model and defined for each analyte as the ratio of its concentration on the droplet surface phase to that in the interior phase. For the analyte ion \(A^+\) and the necessary electrolyte ion \(E^+\), the equilibrium partition reactions and their partition coefficient can be described as follows:

\[
(A^+X^-)_i \leftrightarrow (A^+)_s + (X^-)_i,
\]

\[
K_A = [A^+]_s[X^-]_i/[A^+X^-]_i,
\] (4.5)

when \([A^+X^-] \gg [A^+]_s\), and \(C_A = [A^+X^-] + [A^+]_s\),

\[
K_A = [A^+]_s[X^-]_i/C_A.
\] (4.6)

\[
(E^+X^-)_i \leftrightarrow (E^+)_s + (X^-)_i,
\]

\[
K_E = [E^+]_s[X^-]_i/[E^+X^-]_i,
\] (4.7)

when \([E^+X^-] \gg [E^+]_s\), and \(C_E = [E^+X^-] + [E^+]_s\),

\[
K_E = [E^+]_s[X^-]_i/C_E.
\] (4.8)

Where \(X^-\) denotes the counter ions, \(C_A\) and \(C_E\) are the total analyte concentration and the total electrolyte concentration, respectively. The two components \(A^+\) and \(E^+\) are both competing for the supply of a fixed number of surface charges. Therefore, this equation and equilibrium constant for this competition can be expressed as

\[
(A^+X^-)_i + (E^+)_s \leftrightarrow (A^+)_s + (E^+X^-)_i,
\]

\[
K_A/K_E = [A^+]_s[E^+X^-]_i/[A^+X^-]_i[E^+]_s,
\] (4.9)

when \([A^+X^-] \gg [A^+]_s\), and \([E^+X^-] \gg [E^+]_s\),

\[
K_A/K_E = [A^+]_sC_E/[E^+]_sC_A.
\] (4.10)
\([A^+]_s\), can be expressed as follows, if we combine Equations 4.10 and 4.4, as demonstrated by Enke.

\[
[A^+]_s = \frac{C_A K_A}{C_A K_A + C_E K_E} [Q]
\]

(4.11)

If we agree to the assumption that the mass response of a certain ion is proportional to the concentration of that ion in the surface phase of droplet, then the ESI response of analyte \(A\) (\(R_A\)) can be expressed as follows:

\[
R_A = p f \frac{C_A K_A}{C_A K_A + C_E K_E} [Q].
\]

(4.12)

Following the convention of Kebarle, \(p\) and \(f\) are the efficiency of and sampling efficiency of the system, respectively. As pointed out by Enke, this equation has exactly the same form as Equation 4.1, except the values of \(k\) in Equation 4.1 are the evaporation rate constants while the values of \(K\) in Equation 4.12 are the equilibrium-partitioning coefficients.\(^9\) Equilibrium-partitioning coefficients (the values of \(K\)) reflect the basicity, charge density and nonpolarity of charged molecules. The basicity guarantees that molecules carry protons, while their charge density and nonpolarity determine how likely they are to stay on the droplet surface.\(^25\)

High surface activity also has a sequential enriching effect on ESI response through uneven fission. For the uneven fission process, it was believed that the offspring droplet was generated from the surface phase of its parent, and thus attained the significantly enhanced mass-to-charge ratio on the offspring droplets.\(^25\) As illustrated in Figure 4.1, with this uneven fission process, the concentration of surface active ions can be much higher in the ultimate offspring droplets, while the concentration of a non surface-active compound will be reduced. In order to theoretically model this effect, Cech and Enke extended the partitioning-equilibrium process from initial ESI droplets to the offspring droplets and created the charge overlap model.\(^12\) The modeling results demonstrated that the effect of uneven fissioning of mass and charge compounded the effect of partitioning within an ESI droplet and make the issue of droplet surface affinity even more important in determining ESI response.\(^12\)

Low solvation energy usually correlates to high surface activity. Therefore, compounds with high surface activity will have high evaporation rate constants. As a result, both the ion evaporation model (IEM) and the charged residue model (CRM) predict the similar dependence of the ion intensities observed in ESI. But this does not mean that the role of ion evaporation can be ignored. With the partitioning-equilibrium model, ions that have no surface activity, such as alkali ions, \(\text{Li}^+, \text{Na}^+, \text{K}^+, \text{Cs}^+\), are expected to exhibit approximately the same ion intensities as solutions containing the alkali salts (\(\text{M}^+\text{X}^-\)) at the same concentration. However, increasing values of \(k\) were
observed from Li\(^+\) to Cs\(^+\),\(^{26}\) therefore, both ion evaporation and surface activity play important roles in ESI.\(^{10}\) For compounds that have no surface activity, ion evaporation plays a major role. Most new pharmaceutical candidates have a hydrophobic region on the molecules,\(^{27}\) and the equilibrium-partitioning model is more appropriate for estimating their ESI responses. Since the response of ESI is highly dependent on the hydrophobicity of analytes, one could predict the MS response based on the retention time on reversed-phase HPLC.\(^{28}\)

Unlike the ion-evaporation model, which can only predict the MS response within certain range using the same \(k_a/k_e\), this model successfully predicted the MS response in a wide range of concentrations (10\(^{-9}\) to 10\(^{-3}\) M) with the same value of \(K_A/K_E\).\(^{15}\) Furthermore, the effect of \(C_A\), the \(K_A/K_E\) ratio, \(C_E\) and [\(Q\)] on the \([A^+]_s\) can be simulated for a better understanding of the contribution of each component.\(^{15}\) As illustrated by Enke’s group, the analyte surface concentration \([A^+]_s\) is a quadratic function of \(C_A\), \(C_E\), \(K_A/K_E\) and [\(Q\)].\(^{15}\) As shown in Figure 4.2, when \([A^+]_s\) is plotted against \(C_A\), two portions of the whole curve, a linear portion at lower \(C_A\) and a saturated portion at higher \(C_A\), are generally observed. At the low \(C_A\) region, \([A^+]_s\) is proportional to \(C_A\) because there is plenty of extra charge for analyte ions (\(C_A \ll [Q]\)), regardless of the value of \(K_A/K_E\). With \([A^+]_s\) approaching [\(Q\)], in another expression, when \(C_A \geq [Q] + C_E/(K_A/K_E)\), saturation occurs. The start of this turning point and the curve shape at the saturated region is controlled by the value of \(K_A/K_E\). Analytes with higher \(K_A/K_E\) values (e.g., analytes with higher surface activity)
had a wider linear range and a sharper response slope. On the other hand, analytes with lower $K_A/K_E$ values (e.g., analytes with lower surface activity) had a narrower linear response range, with their response curve showing more curvature and gradually reaching $[Q]$.\(^{15}\)

This model also predicts that the analyte response will decrease as electrolyte concentration ($C_E$) increases. However, this is contradicted by the observed data where the analyte response increases to a maximum as $C_E$ increases to $10^{-4}$ M and decreases with further increases of $C_E$.\(^{15}\) It was explained that an increase of $C_E$ increases the conductivity of the solution and thus increases the spray current $[I]$ and the excess charge $[Q]$. With the increased of $[Q]$, more analyte ions, but not electrolyte ions, can be ionized at the surface phase and transferred to gas phase due to its higher $K_A/K_E$ ratio. However, further increase of $C_E$ causes a loss in ion transfer efficiency ($p$) or desolvation efficiency ($f$), thus reducing the analyte response with a further increase of $[Q]$.\(^{15}\) Overall, this model simplified the effect of salt and therefore can only be used to predict the analyte responses at salt concentrations less than $10^{-5}$ M.

With the information provided by this model, we now learn that high surface-active ionic contaminants are undesirable, not only because their high-intensity peaks may interfere with the analyte mass spectrum, but also because they will suppress the spectrum of the analyte by competing for the limited excess charge on the droplets. It is worth pointing out that the relative ion yields represented by the relative values of the coefficients $K_A, K_B, K_E$ etc.,
depend on a complex sequence of events which require consideration not only of bulk to surface equilibriums and ion evaporation rate, but also droplet evolution schemes and other experimental settings. For example, the surface partitioning coefficient of a low surface-active analyte can be higher when smaller initial droplets are formed with less fission steps required before gas phase ion formation begins, such as in nanospray techniques.

4.2.1.2.3 \( pK_a \) and solvent pH

Since ESI requires protonation or deprotonation, it was believed, initially, that \( pK_a \) and pH of solution would play an important role in ESI response. However, protonated ions of basic analytes can be observed when the pH is higher than the analyte \( pK_a \). One possible explanation for this phenomenon is that the fixed amount of surface excess charge is dictated by the solution flow, and the applied voltage, not the solution pH. In other words, the analyte or other components that stay on the surface phase of the ESI droplet can be protonated even if their \( pK_a \) values are below the solution pH. Besides, uneven fission process will enrich the signal of the surface-active components but not those ion pairs that stayed inside the droplet. Furthermore, gas phase ion reactions can also generate charged ions from neutrals. Therefore, \( pK_a \) and solvent pH are not as important as the surface activity in producing an ESI response.

4.2.1.3 Possible mechanisms for ion suppression in ESI

The understanding of how the ESI response is controlled by instrument settings, properties of the analyte and co-eluting components have brought us closer to understanding the mechanisms of ion suppression. The following section will focus on the possible mechanisms involved in both the solution phase and the gas phase, as shown schematically in Figure 4.3.

4.2.1.3.1 Competing for limited surface excess charge

Even though there are different mechanisms (IEM and CRM) for gas phase ion formation, both theories agree that initial gas ions are generated from the surface phase of ESI droplets. Both the ion-evaporation model (Equation 4.1) and the partitioning-equilibrium model (Equation 4.12) are used to predict that the ESI response in the presence of other components is based on competition for a fixed amount of ESI current (\( I \)) or excess charge (\( Q \)) which is controlled by the applied voltage and the flow rate. Since \( I \) is proportionally correlated to \( Q \) as expressed by Equation 4.4, the competition for \( I \) and for \( Q \) are equivalent. Furthermore, since the excess charges all reside on the surface of the droplets, competition for the limited charge or competition for the limited surface space are both possible. Based on Equation 4.12, it is clear that when total ion concentration in the droplet exceeds \([Q]\), there will be a competition among the ions for the excess surface charge. Matrix induced ion suppression can be
explained in part as the competition for the limited excess surface charge. The matrix components with higher $K$ are more surface active, therefore they would be expected to out-compete the low $K$ analytes for the limited excess charge or limited space on the initial droplet surface. The uneven fission process will produce an even more profound ion suppression effect when surface-active matrices are present. The high surface-active matrices will out-compete the low surface-active analyte in each uneven fission process and occupy the droplet surface in each subsequent offspring droplet. In this case, analytes would be preferentially left in the interior neutral phase of each preceding droplet and therefore became undetectable.

Surfactants are molecules with both polar and hydrophobic regions and known to prefer the air–liquid interface. Due to their high affinity to the droplet surface, surfactants are expected to have high ESI responses. Many experiments have shown that surfactants significantly suppress the ESI response of other analytes.$^{12,14}$ Therefore, it is not hard to understand that surfactants, such as Tween 80, that are used as dosing excipients to improve the solubility of drug candidates, could cause significant ESI ion suppression for co-eluting analytes in LC–MS/MS assays.$^{34-36}$

Polymers that are used as co-solvents to improve the solubility of hydrophobic compounds usually have both hydrophilic and hydrophobic
parts in the same molecule, therefore these polymers also have high surface activity. The attainment of improved solubility using PEG400 is through the bridging effect of this polymer between the hydrophobic analyte and water. The backbone of polyethelene (—CH₂CH₂—)_n of PEG400 will associate with the nonpolar part of hydrophobic compounds via hydrophobic interaction while the terminal hydroxyl group (—OH) will hydrogen bond with water. At the same time, this hydrophobic backbone of polyethelene also provides for sufficient surface activity of PEG400. Therefore, if PEG400 is contained in the plasma sample and cannot be separated from analytes, matrix effects will often be observed. Several reports have described severe matrix effects from plasma samples obtained from laboratory animals dosed with formulations containing PEG400.³⁴–³⁷

Lipophilic components such as long-chain (C12–C16) fatty acids, glycerophosphocholine lipids, phosphatidylethanolamine, phingomyelins, and triacylglycerols in plasma and tissue sample all have high surface activity, therefore these components can be part of the cause of ion suppression effects. It was demonstrated that lyso-phosphatidylcholine (C16:0, C18:0, C18:2) present in serum contributed to the matrix effects observed in an assay for verapamil.³⁸ In our laboratory, we have observed that hydrophobic matrix effects are more often observed in tissue samples, especially in brain samples; part of the reason for this effect might be that brain tissue contains more lipid components that are surface active than those found in plasma samples.

4.2.1.3.2 Incomplete evaporation

It is a well-known fact that the presence of nonvolatile salts such as phosphate and sulfate in the mobile phase is deleterious for ion sources of LC–MS/MS systems due to the deposition of solid material onto surfaces of the source. Nonvolatile components in biological sample can also cause significant ion suppression for early-eluting compounds. Ions that are generated in droplets can only be detected after they are emitted into the gas phase, therefore evaporation is a critical process for the ultimate gas phase ion generation. The efficiency of gas phase ion generation depends on the evaporation efficiency or desolvation efficiency (f), size and charge of the initial ESI droplets. Nonvolatile material in the biological sample can change the volatility, viscosity, and conductivity of the sprayed solution, causing incomplete evaporation and weak Taylor cone emission, hindering the process of uneven fission, and decreasing the efficiency of gas phase ion generation; this effect results in a reduction of the number of analytes that are converted to the gas phase and then detected by the mass spectrometer system.¹⁵ In order to test these hypotheses, King and colleagues designed a set of experiments comparing the amount of analyte and nonvolatile components depositing on the interface plate with or without nonvolatile material present in the sprayed solution.¹⁹ If the nonvolatile components cause incomplete evaporation, then both the analyte and the nonvolatile components would stay in the solution phase and be sprayed onto the interface plate. The tested nonvolatile samples were
ammonium sulfate and extracted plasma samples prepared by protein precipitation, liquid–liquid extraction and solid phase extraction, respectively. It was shown that the samples prepared by protein precipitation contained the most nonvolatile material. Furthermore, it was demonstrated that the amount of nonvolatile matrix components was correlated to the extent of ion suppression. With more nonvolatile material present, more analytes were deposited on the interface plate than were transferred to the gas phase.19

Instrument interface designs with inefficiency in the heating and desolvation processes are more prone to matrix effects due to the cause of incomplete evaporation when analytes are eluted with nonvolatile components. For example, it has been shown that one design of the APCI probe made by Micromass had issues in terms of low sensitivity39 and was more prone to matrix effects than other vendors’ designs for APCI probes.40,41 It was also demonstrated that decreasing the interface chamber pressure by attaching a roughing pump would improve the APCI response of this one design;39 these findings indicated that desolvation or ion transmission characteristics of the unmodified Micromass APCI interface were not optimized. This might be one of the reasons that this APCI probe design was more prone to matrix effects than other vendors’ APCI probes. A new Micromass APCI probe (IonsSabre®, Micromass, UK) has been introduced and its design includes an increased heating capacity and efficiency by using an optimized ceramic heater with gradient heating distribution, so that the efficiency of desolvation is improved therefore the ionization efficiency and sensitivity of this new design should be better than the previous design. Currently, the most recent generation of APCI probes of tandem mass spectrometers by different manufactures are all made with enhanced heating capacity and efficiency, some with even improved gas dynamics. The advanced Turbo V® source for the Sciex API 4000 MS/MS system is equipped with dual ceramic heaters and improved gas dynamics which maximize the desolvation efficiency, thus providing greater efficiency in ionization and increased sensitivity and reduced peak tailing caused by cross-contamination at the same time. In a recent report, it was demonstrated that fewer matrix effects were observed with the same set of samples prepared by protein precipitation using the Sciex API 4000 as compared with the Sciex API 3000.42

It has been demonstrated that polar matrices cause matrix effects mostly due to incomplete evaporation as opposed to neutral evaporation.19 If this type of ion suppression was caused by the competition of charge from matrix components, then analytes could exist as neutrals in the gas phase in ESI, but would be ionized by APCI. King and colleagues built a combined ESI–APCI source, expecting to see an improved signal when using the corona discharge for neutrals in gas phase.19 However, no improvement was observed under these conditions for rat plasma samples prepared by protein precipitation; these data indicated that the amount of neutral analytes existing in the gas phase was negligible. Therefore, the nondetectable analytes in this sample set must have existed as either liquid or solids in the ESI source.
4.2.1.3.3 Ion pairing

Ion suppression caused by strong acids, like trifluoroacetic acid (TFA), has been a problem for ESI applications in proteomics. This type of ion suppression was originally believed to be primarily due to the high surface tension and the high conductivity of the solution which resulted in unstable spray effects. However, further studies have shown that the ion suppression caused by TFA is also due to its strong ion-pairing effect with basic analytes. The strong ion-pairing of TFA with basic analytes keeps analytes in the interior neutral phase of ESI droplets and prevents the analytes from partitioning to the surface phase. Based on this mechanism, a solution of TFA fix was successfully proposed and tested. The TFA fix is a method to reduce the ion suppression caused by TFA using a post-column addition of a solution containing a high concentration of propionic acid in 2-propanol at the flow rate half that of the mobile phase flow rate. As proposed by Apffel and colleagues, when TFA ion-pairs with a basic analyte, both TFA and the analyte are restricted in the interior neutral phase and therefore cannot be released to the gas phase, even though TFA is a volatile acid. When a weak and less volatile acid, such as propionic acid, is added at high concentration, its mass effect will compete with TFA for ion pairing with the basic analyte and in turn releases TFA into the surface phase of droplets that go into the gas phase. At the same time the weaker association between weak acid and basic analyte will make more analytes partition into the surface phase as droplets that are released to the gas phase. This ion-pairing mechanism was further corroborated by the observation of ion enhancement of “almost neutral” compounds, such as diphenylthiourea, in the presence of 0.2% TFA. As a strong acid, TFA cannot pair with the analyte in this situation; however, it does improve the protonation of the analyte.

4.2.1.3.4 Competition for protons in gas phase

ESI is a soft ionization technique that involves transferring the ions from solution phase to gas phase. The detected gas ions are initially generated from solution phase, and only those very stable singly charged alkali ions will stay in the gas phase without any chemical reactions. Since the gas phase environment is different than the solution phase, many ions can be modified in the gas phase after they are initially generated in the solution phase. These gas phase chemical reactions, such as charge neutralization, charge stripping and charge transfer, can also have a significant effect on ESI response. The proton transfer reaction is a fundamental chemical reaction that has been investigated in both solution and gas phases. Several studies have demonstrated that gas phase proton transfer reactions occur in ESI.

The order of basicity in the solution phase can differ from that in the gas phase, thus a proton transfer to a strong gas phase base can enhance the formation of some ions while suppressing the formation of others. A special study was conducted to study the impact of gas phase proton affinities on the
ESI responses. This study demonstrated that ions present in solution can be altered in the gas phase by the presence of molecules that are stronger gas phase ions. If matrix components are stronger gas phase ions, they will out-compete gas phase analyte ions for protons and suppress the response of analytes. It has been suggested that matrix effects caused by PEG400 are due to gas-phase proton competition.

4.2.1.4 Matrix sensitivity: ESI or APCI?

In order to understand what ionization mode is more subjective to matrix effects, we first need to understand the differences in the two ionization modes. Even though there are several models that have been proposed to explain how gas phase ions are produced via charged droplets in electrospray, it was generally agreed that for ESI, most ions are generated in solution phase followed by transferring to the gas phase; while for APCI, molecules (not ions) are first vaporized into gas phase followed by being ionized by the corona discharge process. In other words, the ionization efficiency can be affected by matrix components in both solution phase and gas phase for ESI, while only in gas phase for APCI. The difference in ionization process in these two different modes has been utilized to determine whether the dominant impact of matrices is in solution phase or gas phase. It was found that ESI is more susceptible to plasma matrix effects than APCI, because these matrices usually contain large amounts of nonvolatile components and high concentrations of electrolytes: all these have been proven to have a dominant effect in the reduction of ionization efficiency in the solution phase, and thus reduce the amount of analytes being transferred to the gas phase. Compared to APCI, there are more steps involved in ESI that are susceptible to matrix effects. Matrix components can affect ionization efficiency at any of the steps of ion formation starting from the initial ESI droplet formation to the final gas phase ion reaction, therefore ESI is believed to be more susceptible to matrix effects than APCI. However, ionization in APCI could also be affected to a significant degree if the matrices were dominated by a large amount of other ionizable species that can compete with analyte ions for gas phase protons. Generally speaking, for the matrix components that will affect solution phase ion formation, such as the droplet evaporation process, charged droplet formation and uneven fission, ESI will be more sensitive than APCI. For matrix components that affect gas phase ion reactions, both ESI and APCI will be affected to a significant degree. However, there are also exceptions: when there are defects in the interface design, such as insufficient heating and desolvation capacity for the APCI interface, then APCI can be more susceptible to matrix effects.

4.2.1.5 Differences in interface design

It is generally believed that ESI is more subject to matrix effects than APCI. However, it was found that different instrument interface designs could also
affect the tolerance of matrix effects with different ionization modes. A comparative study using three different triple quadrupole mass spectrometers was performed to evaluate their sensitivity toward matrix effects; the three tandem MS systems were the following: PE Sciex (Concord, Ontario, Canada) API 3000, Micromass (UK) Quattro Ultima® and Thermo-Finnigan (USA) TSQ 7000 API-2.41 Identical sets of samples containing the same amount of several test compounds were prepared using protein precipitation. The matrices for these test samples were either HPLC grade water or rat plasma A or rat plasma B and they were analyzed using these three different tandem mass spectrometers. Identical mobile phases, gradient (one fast and one slow on each system), flow rate, column and HPLC systems were used to reduce the assay variation to only be the various ion sources. Assays were performed using both APCI and ESI on all the samples with all three mass spectrometers. Table 4.1 summarizes the observations with the fast gradient where both the analyte and the internal standard (ISTD) eluted at 1.9 min. Significant ion suppression of both the analyte and ISTD in rat plasma B was observed when using the Micromass Quattro Ultima system with both ionization modes and using the Finnigan TSQ system with ESI mode, while no ion suppression was observed using Sciex API 3000 with either APCI or ESI mode (Table 4.1). Table 4.2 summarizes the same comparison with the slow gradient where the analyte eluted at 4.0 min and the ISTD eluted at 3.9 min. Consistent analytical results were obtained using the Sciex API 3000 interfaced with both ionization modes and using the Finnigan TSQ instrument interfaced with the ESI source. However, weak ion suppression was observed in rat plasma A using the Finnigan TSQ with the APCI mode. It is interesting to note that for the same set of samples under the same slow gradient, the mass responses of the two compounds in plasma A and B showed more variation using the Micromass Quattro Ultima interfaced with the APCI mode, while mass responses obtained by ESI were quite consistent (Table 4.2). These compelling data suggest that the matrix effects in HPLC–MS/MS assays are not only ionization mode (APCI, ESI) dependent, but can also vary between different vendors’ source designs. It was found that under the same ionization mode, different instruments showed different sensitivity to the same matrix (Table 4.1), and that for the Micromass Quattro Ultima, the APCI mode was even more sensitive to the observed matrix effects than the ESI mode.41

4.2.1.6 Nature of matrices: hydrophilic versus hydrophobic

Identifying the nature of matrices will provide useful information for overcoming the matrix effects. Most drug candidates are small molecules with log $P$ ranging from 1 to 5 and are retained on a reversed-phase HPLC column.27 By manipulating pH, the composition of the mobile phases and the gradient46 or selecting a mini-bore column,47 one can easily separate the drug candidates from those polar and nonvolatile matrices that are typically eluted at an earlier retention time on a reversed phase HPLC system. Thus, hydrophilic or polar matrices are not an assay problem for relatively
Table 4.1  Relative mass responses (%) in two different batches of rat plasma obtained using different API sources with the following fast HPLC gradient (from 10%B to 100%B in 1 min, held for 1.9 min, and back to 10%B in 0.1 min) and a shallow gradient (from 40%B to 100%B in 4 min, held for 1.5 min, and back to 40%B in 0.1 min)

<table>
<thead>
<tr>
<th>Technique</th>
<th>Matrices</th>
<th>Micromass Quattro Ultima</th>
<th>Sciex API 3000</th>
<th>Finnigan TSQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CMPD_8</td>
<td>ISTD</td>
<td>CMPD_8/ISTD</td>
</tr>
<tr>
<td>APCI</td>
<td>Rat plasma (A)</td>
<td>99</td>
<td>94</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Rat plasma (B)</td>
<td>53</td>
<td>50</td>
<td>107</td>
</tr>
<tr>
<td>ESI</td>
<td>Rat plasma (A)</td>
<td>87</td>
<td>81</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>Rat plasma (B)</td>
<td>87</td>
<td>72</td>
<td>122</td>
</tr>
</tbody>
</table>

Mobile phases A and B: 10 mM ammonium acetate and 0.005% acetic acid (v/v) in water/methanol (80/20, v/v) and 10 mM ammonium acetate and 0.005% acetic acid (v/v) in water/methanol (10/990, v/v). Flow rate: 0.8 mL/min. Column: Metachem Basic, 5 μ, 4.6 x 50 mm. (Source: Mei et al. Rapid Commun. Mass Spectrom. 17(1), 97, 2003. With permission.) APCI, atmospheric pressure chemical ionization; ESI, electrospray ionization. Values in bold type indicate matrix effects.
Table 4.2  Relative mass responses (%) in two different batches of rat plasma obtained using different API sources with the following slow HPLC gradient: %B to 100% B in 4 min, held for 1.5 min, and back to 40% B in 0.1 min.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Matrices</th>
<th>Micromass Quattro Ultima</th>
<th>Sciex API 3000</th>
<th>Finnigan TSQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>APCI</td>
<td>Rat Plasma (A)</td>
<td>206 163 127</td>
<td>111 110 101</td>
<td>83 81 103</td>
</tr>
<tr>
<td>APCI</td>
<td>Rat Plasma (B)</td>
<td>82 69 119</td>
<td>106 106 100</td>
<td>96 93 103</td>
</tr>
<tr>
<td>ESI</td>
<td>Rat Plasma (A)</td>
<td>114 113 101</td>
<td>115 110 105</td>
<td>113 102 111</td>
</tr>
<tr>
<td>ESI</td>
<td>Rat Plasma (B)</td>
<td>109 109 100</td>
<td>108 98 110</td>
<td>97 91 107</td>
</tr>
</tbody>
</table>

Mobile phases A and B: 10 mM ammonium acetate and 0.005% acetic acid (v/v) in water/methanol (80/20, v/v) and 10 mM ammonium acetate and 0.005% acetic acid (v/v) in water/methanol (10/990, v/v). Flow rate: 0.8 mL/min. Column: Metachem Basic, 5 μ, 4.6 × 50 mm. (Source: Mei et al. Rapid Commun. Mass Spectrom. 17(1), 97, 2003. With permission.) Values in bold type indicate matrix effects.
hydrophobic new chemical entities (NCEs). Using the post-column infusion technique, it has been found that the majority of the problem matrices in plasma samples are those polar components that elute at earlier retention times on a reversed-phase HPLC system.\textsuperscript{5,47} Hydrophobic matrices that usually exist in smaller amounts are often revealed as a narrow dip in the later retention time of a reversed-phase chromatogram when using the post-column infusion technique. For example, fatty acids and phosphatidylcholine with long carbon chains (C16–C18) are endogenous hydrophobic matrices, which have a high potential to co-elute with pharmaceutical compounds.\textsuperscript{38} While these types of matrix issues are relatively challenging, they are still manageable by carefully separating these components from analytes of interest using either various sample preparation techniques or HPLC adjustments. The most difficult matrix effect problems are those caused by hydrophobic components existing in relatively large amounts and with retention times that overlap the analytes.\textsuperscript{48–50} In this situation, the chromatographic system does not provide sufficient separation; in some cases, these matrices can also overload the column and carry over to next injection, causing huge assay variations from injection to injection.\textsuperscript{48} One good example of such situation is the matrix effect caused by a polymeric material that existed in one set of samples;\textsuperscript{41} it was demonstrated that the polymeric material was rather hydrophobic and eluted over a wide range that overlapped with the analytes, resulting in significant ion suppression for compounds that eluted in that part of the chromatogram.

### 4.2.1.7 Source of matrix effect

As opposed to the visible UV interferences, LC–MS/MS matrix effects are often described as unknown and nonvisible.\textsuperscript{51} This characterization mystified LC–MS/MS matrix effects and deterred our efforts in searching for their source. If matrix effects can be described as exogenous, known, or constant as opposed to endogenous, unknown, or variable, then they will become manageable obstacles. The separation of large amounts of hydrophobic matrices imposes special challenges to bioanalytical assays in the drug discovery environment where fast turn-around time is required regardless of whether the assay is easy or difficult. Therefore, identifying the source of matrices becomes crucial, especially for those difficult hydrophobic matrices. Intentionally avoiding matrices that exist in relatively large quantities is much easier than blindly finding a way to separate them from the analytes. Some exogenous materials such as plasticizers and anticoagulants can be present in relatively large amounts compared to the analyte of interest and some of the plasticizers are very strong gas phase ions. In addition, some anticoagulants, such as Li-heparin, are strong ionizing agents, which have a high potential to cause matrix effects. Thus, the following study was designed to identify the possible exogenous sources of matrix effects. Compounds with a significant hydrophobicity range which would elute at various retention times were employed as markers for ion suppression evaluation. Rat plasma obtained
from one source but stored in different plastic tubes was used for study samples. Aqueous solutions served as control samples to be assayed at the same time as study samples in order to identify the source of the matrix ion suppression as either the plasma or tube used to store the plasma. A typical chromatogram for these eight markers is presented in Figure 4.4 with retention times ranging from 1 to 7 min, representing majority of drug discovery compounds in terms of log $P$. The matrix effects observed with plasma or water in these test tubes are summarized in Table 4.3. Overall, there were 22 observations of matrix effects across most regions of the chromatographic gradient. Sixteen of these involved polar components that were restricted to the early-eluting Compound 1 and 2, and 12 of these examples involved exogenous components, which affected both early-eluting compounds and late-eluting compounds. The full mass scan data suggested that the matrix responsible for

![Figure 4.4](image-url)
the ionization suppression for late-eluting compounds in the Li-heparin/ Microtainer tube was some type of plasticizer (or release agent) used in this brand of tube (Figure 4.5). Typically, if the source is unknown, such severe matrix effects can result in a significant sample preparation effort to separate these matrices from the analytes of interest. With the cause identified, these matrix effects now can be simply avoided by using the proper brand of tubes for processing and storing both plasma samples and spiked plasma standards.41

Another study was designed to study the impact of different anticoagulants on LC–MS matrix effects using the same type of strategy. The markers were added to water and rat plasma containing different types and increasing amount of anticoagulants. No significant matrix effect was observed for all the test compounds with up to 29% of Na-heparin and Na2-EDTA in serum. However, an enhanced mass signal of CMPD 1 with increasing concentrations of Li-heparin in serum was observed, as shown in Figure 4.6. As shown in Table 4.4, the normalized mass responses of eight test compounds in serum are

<table>
<thead>
<tr>
<th>Test tube</th>
<th>CPMD1</th>
<th>CPMD2</th>
<th>CPMD3</th>
<th>CPMD4</th>
<th>CPMD5</th>
<th>CPMD6</th>
<th>CPMD7</th>
<th>CPMD8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fisher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Plasma</td>
<td>12</td>
<td>160</td>
<td>111</td>
<td>101</td>
<td>91</td>
<td>103</td>
<td>98</td>
<td>95</td>
</tr>
<tr>
<td>Li-heparin/Sarstedt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>109</td>
<td>97</td>
<td>103</td>
<td>103</td>
<td>108</td>
<td>81</td>
<td>92</td>
<td>104</td>
</tr>
<tr>
<td>Plasma</td>
<td>9</td>
<td>39</td>
<td>44</td>
<td>67</td>
<td>101</td>
<td>84</td>
<td>86</td>
<td>100</td>
</tr>
<tr>
<td>K3-EDTA/Sarstedt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>16</td>
<td>155</td>
<td>110</td>
<td>107</td>
<td>86</td>
<td>82</td>
<td>83</td>
<td>95</td>
</tr>
<tr>
<td>Plasma</td>
<td>11</td>
<td>127</td>
<td>100</td>
<td>97</td>
<td>93</td>
<td>101</td>
<td>90</td>
<td>100</td>
</tr>
<tr>
<td>Na-heparin/Vacutainer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>130</td>
<td>159</td>
<td>117</td>
<td>92</td>
<td>96</td>
<td>91</td>
<td>88</td>
<td>97</td>
</tr>
<tr>
<td>Plasma</td>
<td>17</td>
<td>160</td>
<td>107</td>
<td>99</td>
<td>96</td>
<td>94</td>
<td>94</td>
<td>97</td>
</tr>
<tr>
<td>Corning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>130</td>
<td>111</td>
<td>109</td>
<td>98</td>
<td>101</td>
<td>98</td>
<td>90</td>
<td>98</td>
</tr>
<tr>
<td>Plasma</td>
<td>14</td>
<td>164</td>
<td>110</td>
<td>97</td>
<td>96</td>
<td>99</td>
<td>105</td>
<td>98</td>
</tr>
<tr>
<td>Li-heparin/Microtainer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>113</td>
<td>149</td>
<td>110</td>
<td>95</td>
<td>84</td>
<td>60</td>
<td>47</td>
<td>59</td>
</tr>
<tr>
<td>Plasma</td>
<td>13</td>
<td>146</td>
<td>98</td>
<td>85</td>
<td>82</td>
<td>65</td>
<td>46</td>
<td>57</td>
</tr>
<tr>
<td>Nunc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>104</td>
<td>108</td>
<td>102</td>
<td>96</td>
<td>97</td>
<td>97</td>
<td>91</td>
<td>92</td>
</tr>
<tr>
<td>Plasma</td>
<td>11</td>
<td>162</td>
<td>105</td>
<td>95</td>
<td>90</td>
<td>102</td>
<td>92</td>
<td>91</td>
</tr>
<tr>
<td>USA/Scientific Plastics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>100</td>
<td>131</td>
<td>121</td>
<td>111</td>
<td>104</td>
<td>74</td>
<td>87</td>
<td>107</td>
</tr>
<tr>
<td>Plasma</td>
<td>18</td>
<td>145</td>
<td>106</td>
<td>94</td>
<td>95</td>
<td>74</td>
<td>79</td>
<td>97</td>
</tr>
</tbody>
</table>

Caused by endogenous materials.
Caused by exogenous materials.
Caused by both.
Figure 4.5  Comparison of LC/MS chromatograms and mass spectra of pure water or plasma in Li-heparin/Microtainer tubes and blank plasma obtained from a contract research organization (CRO) (Source: Mei et al. Rapid Commun. Mass Spectrom. 17(1), 97, 2003. With permission.)
listed with increasing concentrations of Li-heparin, the anticoagulant, Li-heparin, could also affect the response of CMPD 2 at the higher concentration of Li-heparin. Li$^{+}$ and some transition metal ions have been used as cationizing agents for characterizing many chemicals, such as polymers and lipids with mass spectrometric detection.\textsuperscript{52–55} It was demonstrated that the ionization efficiency of glycerophosphocholine lipids,\textsuperscript{52} phosphatidylethanolamine,\textsuperscript{54} triacylglycerols,\textsuperscript{53} and polyglycols\textsuperscript{55} were enhanced in the presence of Li$^{+}$ ions by the formation of lithiated adducts, which can be further fragmented through low-energy collision-induced dissociation.\textsuperscript{55} On the other hand, the potential effect of ion enhancement from Li-heparin treated plasma has not been reported by bioanalytical mass spectrometrists. Our own data show that Li-heparin can produce ion enhancement for certain hydrophilic compounds. It is also possible to observe matrix effects for hydrophobic compounds when

\begin{table}[h]
\centering
\begin{tabular}{lcccccccc}
\hline
\textbf{Li-heparin\% (v/v)} & \textbf{CPMD1} & \textbf{CPMD2} & \textbf{CPMD3} & \textbf{CPMD4} & \textbf{CPMD5} & \textbf{CPMD6} & \textbf{CPMD7} & \textbf{CPMD8} \\
\hline
0 & 100 & 100 & 100 & 100 & 100 & 100 & 100 & 100 \\
2 & 139 & 100 & 92 & 96 & 95 & 91 & 96 & 99 \\
5 & 139 & 120 & 95 & 97 & 100 & 89 & 124 & 104 \\
9 & 142 & 113 & 95 & 97 & 100 & 89 & 124 & 104 \\
17 & 163 & 100 & 97 & 96 & 104 & 106 & 121 & 100 \\
29 & 197 & 134 & 105 & 97 & 105 & 106 & 109 & 97 \\
\textbf{Control with} & \textbf{98} & \textbf{106} & \textbf{94} & \textbf{89} & \textbf{100} & \textbf{104} & \textbf{98} & \textbf{92} \\
\textbf{29\% Na-heparin} & & & & & & & & \\
\hline
\end{tabular}
\caption{Relative mass responses (%) of eight compounds in serum with increasing percentage of Li-heparin. (Source: Mei et al. Rapid Commun. Mass Spectrom. 17(1), 97, 2003. With permission.)}
\end{table}

\begin{figure}
\centering
\includegraphics[width=0.7\textwidth]{figure4.6}
\end{figure}

Shaded area indicates ionization enhancement caused by Li-heparin.
Li$^+$ is present in relatively larger amounts that can overload a column. Therefore, Li-heparin should be avoided as the anticoagulant for plasma samples when the samples are to be assayed by HPLC–MS/MS.

Other exogenous materials that have potential matrix effects are dosing excipients. Multiple studies have demonstrated that excipients such as PEG400, propylene glycol, Tween 80 and even hydroxypropyl beta cyclodextrin (HPBCD) used for either intravenous or oral formulation, can cause significant matrix effects in both ESI and APCI modes.\textsuperscript{34,35} Another unique aspect of this type of matrix effect is that it is variable with time, since these dosing excipients will also undergo an absorption, distribution, and elimination process in animals which will be reflected in the plasma samples collected at different time points.\textsuperscript{34,35,56} Even though techniques such as appropriate sample purification or employing negative ionization mode can diminish the problem, avoiding such excipients is a safer alternative.\textsuperscript{34} If one cannot avoid these dosing excipients, then it is important to evaluate their effect, if any, on the LC–MS/MS system that is used for assaying samples that include these excipients.

### 4.2.2 Evaluation of matrix effect

As discussed above, LC–MS/MS matrices can be challenging, in part, due to their character of being unknown or unseen, as opposed to LC/UV assays where interferences can be seen. As a result, the separation of analytes from those unknown matrices in LC–MS/MS assays becomes more difficult. Thus, evaluation of matrix effects should be the first step in solving the problem. Several approaches have been developed to evaluate the matrix effects using different experimental techniques and each has its own advantages and disadvantages.

#### 4.2.2.1 Post-column infusion

In order to directly observe the location of ionization suppression in an LC–MS/MS assay, Bonfiglio and colleagues\textsuperscript{51} developed a post-column infusion scheme that has been widely adopted by many laboratories. In this scheme, as presented in Figure 4.7, blank sample extracts are injected on the HPLC column under conditions chosen for the assay while a constant amount of analyte is infused into the HPLC stream before it enters the mass spectrometer. Ion suppression caused by matrices is shown as the variation of MS response of the infused analyte, as compared to the response from the injection of blank mobile phase. This approach was successfully employed to detect potential matrix inconsistencies between assay samples and standard samples in drug discovery studies,\textsuperscript{5} to demonstrate that a minibore column coupled with a fast gradient is very efficient for separating endogenous polar matrices from analytes,\textsuperscript{47} and to study matrix effects caused by dosing excipients.\textsuperscript{35} It has been recommended that one should run the same test two or three times to ensure that late eluting matrix components will not interfere with subsequent injections.\textsuperscript{5} Even though this method has the advantage of showing
the region in the chromatogram of ion suppression, it only provides a semiquantitative picture of matrix effects that is not easy to be tabulated or graphed for comparison when many different matrices with a large number of compounds are studied. Also, this approach is not suitable for more than ten samples, since it is hard to be automated. Furthermore, as pointed out by Weng’s group, this method is not so efficient in that after any change in the extraction or chromatographic elution procedures, the infusion experiment must be repeated. Finally, since the post-column infusion is usually performed at relatively high concentrations, it can contaminate the source, generating a high background signal and reducing sensitivity; when this happens, instrument cleaning must be conducted to solve the problem.57

A more efficient and practical approach was also proposed by Weng’s group. Following the identification of the matrix region using the post-column infusion method, one should conduct a full mass scan LC–MS experiment to identify the matrix ions in the region of matrix effects. These matrix ions, instead of a post-column infusion of a high concentration of analytes, can be used as matrix markers for developing better sample extraction or chromatographic separation.57

4.2.2.2 Direct comparison

When there are multiple compounds or multiple matrices that need to be evaluated, the most efficient and straightforward method to detect matrix effects and obtain the extent of matrix effects is to run a set of samples containing the same amount of analytes and internal standards in (1) matrix-free solvent, (2) blank matrix used to prepare calibration standards, and (3) blank matrices obtained from different sources or pre-dose blank sample plasma (plasma obtained from animals before dosing with an analyte). Figure 4.8(A) schematically presents this strategic procedure. Matrix effects can be determined if the difference of the MS responses in different matrices are greater than 25%. If the MS response in pre-dose blank sample plasma is within 25% of the MS response in standard plasma, then this method can be
used for quantitation in drug discovery studies. For example, this method was efficiently utilized to compare the matrix effects among three different tandem mass spectrometers and to investigate the endogenous and exogenous sources of matrix effects (vide supra).41

Another use for this method is to simultaneously compare the effectiveness of different internal standards in correcting matrix effects and to compare the efficiency of different sample cleaning procedures in removing matrices from plasma obtained from multiple lots and different species, thus speeding up method development.58 Large amounts of data can be easily graphed to facilitate the decision making process. This method can also be modified with

Figure 4.8  (A) Schematic diagram of evaluation of matrix effects using direct comparison—pre-spiking approach. (B) Schematic diagram of evaluation of matrix effects using direct comparison—post-spiking approach.
the post-spiking method to separate the recovery factor from total matrix effects as illustrated in Figure 4.8(B). The detailed pre-spiking and post-spiking procedures are discussed in the following section.

4.2.2.3 Pre-spiking and post-spiking comparison

In order to separate the recovery loss from matrix suppression, one can use this pre-spiking and post-spiking approach, where pre-spiking refers to adding standards and internal standards before sample preparation and post-spiking refers to adding standards and internal standards after sample preparation.\(^59,60\) As demonstrated in Figure 4.9, response I was produced by the neat analyte solution, free of matrix effects and binding loss. Response II was obtained from pre-spiking procedure and reflected the loss from both analyte recovery and matrix effects. Response III, which was obtained from the post-spiking procedure only reflected the loss from matrix effects. Therefore, a matrix effect can be calculated as response III/response I, recovery equals response II/response III, and process efficiency equals response II/response I.\(^60\) This technique is especially helpful when complicated sample preparation procedures, such as solid phase extraction and liquid–liquid extraction, are used, since these procedures, unlike protein precipitation, are more subject to analyte loss.

![Figure 4.9](image)

**Figure 4.9** Schematic diagram of evaluation of matrix effect using both pre-spiking and post-spiking approaches. Recovery = (response II/response III) × 100%. Matrix effect = (response III/response I) × 100%.
### 4.2.2.4 Standard addition to incurred samples

For clinical samples, the evaluation of matrix effects should be performed using the appropriate biological matrix obtained from different sources. However, the most meaningful evaluation of matrix effects for discovery studies is to use samples obtained from the tested animals. Therefore, the pre-dose plasma sample is the most appropriate sample for matrix effect evaluation for pharmacokinetic studies. However, there are times when even this approach will not ensure that matrix effects will not be an issue; for example, pre-dose samples cannot provide any information if the dosing excipients are the cause of the matrix effects. Also, for drug disposition studies where the analyte level in individual tissues needs to be determined, it is impractical to obtain pre-dose samples from the same animal. In these situations, one can consider using the so-called standard addition method to evaluate matrix effects. In this method, the unknown tissue homogenate (X) and X with an added known amount of analyte (\(X + A\)) are analyzed along with calibration standards; it is important that all the preparation and processing procedures are all the same for X, \(X + A\), and calibration standards. Using the calibration standards, one can obtain the observed concentration of X and \(X + A\). The expected \(X + A\) can be obtained by adding the observed X and added A. The matrix effects can then be evaluated by comparing the observed \(X + A\) with the expected \(X + A\). The added known amount of analyte should be at least 20 times the limit of quantitation (LOQ) and about equal to the amount of the unknown, as shown in Table 4.5, so that one can differentiate the matrix effects from the other variations. The matrix effect then can be evaluated as the difference of measured \(X + A\) and expected \(X + A\), and if the difference is greater than 25% it can be considered to be due to matrix effects. Usually the volume of tissue homogenate is sufficient, so that different values of \(X + A\) can be prepared, with A covering a range of three orders of magnitudes, such as \(X + 25\), \(X + 250\), and \(X + 2500\). Values of A that give a better evaluation can then be selected. Where sample volume is sparse, two assays can be taken for the evaluation. The unknown with calibration standards can be assayed first to obtain the observed unknown. The unknown sample with

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Observed X (ng/g)</th>
<th>Added A (ng/g)</th>
<th>Expected (X + A) (ng/g)</th>
<th>Observed (X + A) (ng/g)</th>
<th>Diff%*</th>
<th>Matrix effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2500</td>
<td>2500</td>
<td>5000</td>
<td>3000</td>
<td>-40</td>
<td>Suppression</td>
</tr>
<tr>
<td>B</td>
<td>1000</td>
<td>1000</td>
<td>2000</td>
<td>2000</td>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>C</td>
<td>500</td>
<td>500</td>
<td>1000</td>
<td>1500</td>
<td>50</td>
<td>Enhancement</td>
</tr>
<tr>
<td>D</td>
<td>100</td>
<td>100</td>
<td>200</td>
<td>220</td>
<td>10</td>
<td>No</td>
</tr>
<tr>
<td>E</td>
<td>50</td>
<td>100</td>
<td>150</td>
<td>100</td>
<td>-33</td>
<td>Suppression</td>
</tr>
<tr>
<td>F</td>
<td>5</td>
<td>100</td>
<td>105</td>
<td>120</td>
<td>14</td>
<td>No</td>
</tr>
</tbody>
</table>

*Diff% = [(Observed – Expected)/Expected] × 100.
added analyte can be assayed during a second run. Using this procedure, it is possible to not only evaluate the extent of the matrix effects, but also obtain the correct analytical results at the same time (see Section 4.3.10).

4.3 Current strategies for overcoming matrix effects

4.3.1 Introducing the minimum amount of sample

The extent of the matrix effect is dependent on the amount of matrices in the injected sample. Specifically, the observed ion suppression is proportional to the amount of matrices in the sample that enters the ion source at the same time as the analyte. As shown in Figure 4.10, when the volume of the injected sample that is free of matrix interference is increased, an almost linear response was observed; however, with the increase in volume of a protein-precipitated plasma sample, the response did not increase, instead the response showed a saturation tendency. In other words, increasing the injection volume of processed biological samples does not necessarily guarantee an increase of analyte response. The increasing amount of interfering matrices might compete with the analytes for ionization, and limit the maximum response for the analyte. In one report, it was noted that at the sample injection volume of 5 µL, ion enhancement was observed while at an injection volume of 20 µL, ion suppression was observed. The same phenomenon has been observed in our laboratory for some compounds; one example of this is shown in Figure 4.10. One possible mechanism proposed by Enke’s group might explain this phenomenon; at a low injection volume, the electrolyte content of the matrices increased the ion conductivity and thus increased the amount of excess charge, resulting in ion enhancement. However, with the increased sample volume, a further increase of electrolyte would cause a loss in the ion transfer efficiency.

![Figure 4.10](image_url) Effect of injection volume of protein precipitated plasma on the mass responses of caffeine.
(p) or desolvation efficiency (f) and the effect could be greater than the enhancement caused by the increased excess charge, resulting in a reduced or saturated analyte response. Therefore, we recommend introducing minimum amount of sample in order to avoid ion suppression problems. With the recent improvement in the sensitivity in the new generation of tandem mass spectrometers made by different vendors, introducing more diluted or lower volume of sample is an achievable approach, and this should be the first step to reduce the chance for having matrix ionization suppression problems.

4.3.2 Minimizing the build-up of nonvolatile material in the ionization source

As many studies have shown, ion suppression can be caused by nonvolatile components. The accumulation of nonvolatile material on the ion source interface plate will increase the electrical resistance and therefore will reduce the ion flow (I) at the applied voltage of same value and decrease the ESI signal intensity. At the worst, salt deposits on the metal surfaces can even result in a complete loss of ion transmission. Therefore, minimizing the build-up of nonvolatile materials will help to maintain the instrument’s sensitivity. Approaches such as employing a divert valve, which only delivers the portion that contains the analytes into the MS while diverting the unwanted eluate to waste, or using a splitting device, which splits an appropriate amount of eluate into mass spectrometer and sends the rest to waste, or injecting the minimum volume of sample and cleaning the interface plate on a regular basis can all be very effective techniques for maintaining the sensitivity of the instrument.

4.3.3 Avoiding exogenous matrices

Knowing the details about the collection of bioanalytical samples can be crucial for obtaining reliable bioanalytical results, as we have stated, exogenous material such as plasticizers, Li-heparin or dosing excipients can be the cause of significant matrix effects. Therefore, it is strongly recommended that detailed information about sample collection should be obtained in order to avoid potential exogenous matrix effects. The same brand of tubes should be used for processing and storing both spiked plasma standards and unknown plasma samples. The plastic tubes or 96-well plates should be pre-tested to determine their potential for matrix effects before adopting a brand for routine use in the laboratory. Li-heparin should be avoided for the samples to be quantified using an LC–MS-MS assay. If PEG400 has to be used as the dosing excipient, additional steps need be carried out to ensure that it does not cause any interference. Detailed methods for dealing with matrix effects, such as using good chromatographic separation, solid phase extraction, or liquid–liquid extraction have been described and compared in the studies by Tong et al. and Shou et al.
4.3.4 Application of internal standards

The rationale for using an internal standard (IS) to correct for matrix effects is based on the assumption that it will experience the same extent of ion suppression or ion enhancement as the analyte. Therefore, under this rationale, even though the MS response of an analyte can be enhanced or suppressed, and the response ratio (response of analyte divided by the response of the IS) will stay the same. In many situations, especially when using an isotope-labeled analyte for the IS, this assumption is true. However, an isotope-labeled IS is usually unavailable for assays in the drug discovery stage. Often internal standards that we use are analogs of the analytes, but they can differ from the analytes in terms of their log $P$ and $pK_a$. Furthermore, recent studies have shown that the IS can sometimes interfere with the signal of analyte or vice versa via cross-talk or ionization competition. Even co-eluting compounds as the IS sometimes cannot correct for the matrix effect if they have a different $pK_a$ to the analyte. As shown in Table 4.6, phenylpropanolamine (PPA) and pseudoephedrine (PSE) were spiked into a rat plasma sample and they eluted at same time on the HPLC system when the extracted sample was injected, but only PSE experienced ion suppression. Therefore, the use of an IS cannot always guarantee the correction of matrix effects. Careful studies need to be carried out to select a good IS with a matched $pK_a$ and log $P$ and appropriate concentration level, which sometimes is impractical for drug discovery studies.

4.3.5 Preparing standards or quality control samples using the pre-dose samples

By using blank plasma obtained from the same animal that was dosed by the test compounds or blank plasma from the same batch of animals to prepare calibration standards, the matrix difference between the calibration standards and the samples can almost be eliminated, except for the matrix effects caused by dosing excipients. If there is not enough pre-dose blank plasma for making the calibration curve, quality control (QC) samples from the pre-dose plasma can be prepared. Matrix effects can be corrected with the information provided by QC samples at different levels. For example, if consistent matrix effects were observed for QC samples at different levels, then a constant correction factor could be used to correct the matrix effects.

<table>
<thead>
<tr>
<th>Drug</th>
<th>Water</th>
<th>Plasma</th>
<th>Plasma w/PEG400</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPA</td>
<td>100</td>
<td>92</td>
<td>101</td>
</tr>
<tr>
<td>PSE</td>
<td>100</td>
<td>83</td>
<td>58</td>
</tr>
</tbody>
</table>

Shaded area indicates ionization suppression.
4.3.6 Diluting samples with blank standard plasma

When sensitivity is not an issue, the differences in the matrix between standards and samples can be reduced by diluting samples with blank standard plasma (plasma used to prepare calibration standards). Our laboratory has been using this method to analyze samples with high concentrations; this technique is well suited for rising dose pharmacokinetic studies. This approach was also found effective for overcoming the matrix effects that can be caused by PEG400 being used in the dosing formulation.37

4.3.7 Post-column addition of signal-enhancing agents

Ionization suppression caused by the high concentration of electrolytes can be reduced by the addition of so-called signal-enhancing agents. A high concentration of electrolytes can change surface tension and modify the volatility of the sprayed solution; this can lead to the formation of larger final droplets containing higher percentages of water and electrolytes. Furthermore, some electrolytes can form strong ion pairs with analytes, preventing analytes from partitioning to the droplet surface. Therefore, ionization suppression can be reduced by adding some modifying agents that either change the surface tension or volatility of the sprayed solution in favor of formation of analyte gas ions or reducing the ion pairing of analytes.

If only lowering surface tension is needed, then post-column addition (PCA) of surface tension lowering agents, such as methanol, 2-propanol or acetonitrile can be very effective. The ratio of mobile phase flow rate to the flow rate of the PCA agents can be easily optimized. With these conventional surface tension lowering agents, the signal can be enhanced 3–16-fold depending on analytes and methods of sample preparation. However, these agents also have a higher volatility than water and most electrolytes and, therefore, they will evaporate earlier and leave more electrolytes in the final droplets, thus further reducing the ionization suppression caused by ion competitions requires additional properties of these agents, such as an optimized volatility and \( pK_a \) value.

The most famous solution for improving the signal of a strong base in a TFA-containing mobile phase is to employ a mixture containing both propionic acid and 2-propanol. Propionic acid is a weak acid with an optimum volatility. Propionic acid has a volatility lower than TFA which leads to the earlier evaporation of TFA from the droplets, thus replacing strong analyte/TFA pairs with weak analyte/propionic acid pairs that release the analyte ion to the gas phase. If the weak acid is too volatile, such as acetic acid, then TFA will still strongly ion pair with the analyte causing ionization suppression. Unlike other weak acids, such as valeric acid, propionic acid still has sufficient volatility to prevent it from causing rapid desolvation of the droplets.

For improving the signal of acids in negative ESI, 2-(2-methoxyethoxy)-ethanol (2-MEE) is a good choice. One or two of the following additives: formic acid, ammonium formate, acetic acid, and ammonium acetate, is
usually added to the mobile phases for improving the retention time and peak shape. The ion suppression in this condition is caused by high surface tension and ion competition between the analyte and co-eluting electrolytes, such as \((\text{AcO}^-)\), in the mobile phase. The ionization suppression usually gets worse with the increase of these additives. Conventional surface tension lowering agents, such as methanol or acetonitrile, can only improve the part of the problem that is caused by high surface tension, while 2-MEE not only can improve the signal by lowering the surface tension, but can also improve the signal by reducing the \(\text{AcO}^-\) concentration in the final droplet. This is because 2-MEE is a surface tension lowering agent with optimal volatility properties; it has a boiling point of 193°C, higher than that of water (100°C) and acetic acid (118°C). Therefore, water and AcO\(^-\) will be evaporated earlier than 2-MEE in the ESI spray formation. In this way, the final droplet will be smaller and contain more 2-MEE and a lower percentage of water and \(\text{AcO}^-\).

Even though the post-column addition approach is quite effective in reducing the ionization suppression problem, it is mostly applicable for metabolite identification type studies where usually a small number of samples are to be analyzed. For quantification of a large number of samples, one can try to add these modifiers to the mobile phases as described in the following section.

### 4.3.8 Modifying the mobile phase

The ideal mobile phases for ESI are those with an optimum surface tension that facilitates the generation of a stable spray. For positive mode, the most popular mobile phases are methanol/water or acetonitrile/water or methanol/acetonitrile/water with a weak acid (such as acetic or formic) added to the solution at the concentration range from 0.005% to 0.05%, or with a weak acid buffer (such as formic acid–ammonium formate or acetic acid–ammonium acetate). The low pH of such mobile phases can facilitate the protonation of analytes with basic functional groups. The neutral salts (ammonium formate or acetic acid–ammonium acetate) are useful for facilitating the ionization of polar or neutral analytes through adduct formation. However, some strong bases will have very short HPLC retention times or exhibit peak tailing and may need further additives for chromatographic improvement reasons.

TFA is a commonly used additive in HPLC for reducing the peak tailing of basic compounds on silica-based columns. However, TFA is also notorious for its ionization suppression of the ESI signal for basic compounds. As stated above, post-column addition of propionic acid is an effective way to reduce the ion suppression caused by TFA. Based on the mechanism of reducing ionization suppression by adding propionic acid post-column, one can reasonably assume that adding propionic acid directly to a mobile phase should also work. Propionic acid is a weaker acid than TFA; when they co-exist in mobile phase, it is a very weak competitor of TFA; for ion pairing
with the analyte, therefore it should not affect the chromatographic peak shape of the analyte. However, when the driving force for evaporation of TFA is introduced in the electrospray process, the mass action will make propionic acid a good competitor of TFA for ion pairing with the analyte. It has been shown that by adding propionic acid (0.1 to 0.5%) propionic acid directly to mobile phases containing either 0.025% or 0.05% TFA, ESI sensitivity improved by 2–5 fold for basic compounds such as sildenafil, fluconazole, nicotine, midazolam, and isoniazid. For negative ionization mode LC–ESI–MS assays, it is necessary to use a solvent that creates stable anions. The mixture of halogenated solvents with methanol is a very good system for analysis of oligonucleotides in the negative mode, due to the fact that halogenated solvents can form stable anions through electrochemical reduction processes. Examples of these mixtures that have been reported are hexafluoropropionic acid with methanol and 2,2,2-trifluoroethanol with methanol.

4.3.9 Separation of matrices and analytes by sample preparation

The most common means of obtaining maximum sensitivity and signal reproducibility is through comprehensive sample clean-up and purification, even though sometimes it can be very time consuming. The commonly used procedures are protein precipitation (PPT), liquid–liquid phase extraction (LLE) and solid phase extraction (SPE).

For drug discovery studies, where a large number of compounds need to be assayed with a relatively small number of samples per study, it is not practical to develop a unique sample preparation procedure for each compound using SPE or LLE. Thus, PPT has become the main methodology for plasma sample preparation due to its simplicity and universality for almost all small molecules. PPT can be easily automated or semi-automated using a robotic liquid handler and 96-well plates, and it has been implemented in many pharmaceutical companies for drug discovery bioanalytical applications. Studies were conducted to optimize the PPT procedure based on effectiveness of protein removal and matrix effects in LC–MS. It was found that the most efficient protein precipitants for protein removal were zinc sulfate, acetonitrile, and trichloroacetic acid. These precipitants all have excellent protein precipitation reproducibility. However, using either acids or zinc sulfate as precipitants may cause potential degradation of analytes or hydrolysis of some conjugates such as glucuronides and sulfates, while organic precipitants usually will not cause degradations and they are usually compatible with mobile phases. Furthermore, acetonitrile precipitation generally has the lowest ionization suppression potential. Due to all these reasons, acetonitrile has become the most common precipitant for LC–MS assays for small molecules. Even though the PPT approach only removes the protein and leaves behind other matrices from the sample, HPLC usually can provide a satisfactory separation of the analyte and most of the problematic matrices. It has been shown that through a wise selection of the HPLC column in terms of size and
separation mode, HPLC can be very efficient and effective in removing the polar matrices.\textsuperscript{47}

For drug development applications where high quality results are required for a small number of compounds with a large number of samples per study, developing a specific sample preparation method is more common; in this case, SPE and LLE are widely utilized techniques. SPE is probably the most popular technique owing to its ease of automation\textsuperscript{76} and to the availability of a wide variety of commercial sorbent materials.\textsuperscript{77,78} The advantages of sample preparation by SPE include the removal of nonvolatile salts and the attainment of a relatively clean extract with reduced amounts of potentially interfering matrix components. The selectivity of SPE is usually achieved by choosing or mixing appropriate sorbents and designing an effective washing and elution scheme. All these can be achieved at the expense of time and experience. Several papers have details on how to select sorbents and how to design elution schemes to achieve separation for various analytes with different properties.\textsuperscript{76–85}

Automated SPE can be achieved by either using an on-line column switching format or an off-line 96-well format. Generally speaking, the parallel format (e.g., 96-well format) has a much higher throughput than the serial format and is more suitable for large studies; however, with the serial format, it may be easier to achieve complete automation without human intervention. The fastest parallel processing system can achieve speeds of up to 400 samples per hour.\textsuperscript{82}

There are also various on-line extraction techniques using direct injection of plasma (for more on this topic, see Chapter 5); some examples include restricted access media (RAM),\textsuperscript{86} turbulent flow chromatography,\textsuperscript{87–89} molecularly imprinted polymer extraction,\textsuperscript{90} and on-line solid phase extraction. RAM columns are columns made of a hydrophilic external surface and a hydrophobic internal surface in silica particles with controlled pore sizes. The separation of small molecules from biological matrices is achieved by the combination of size-exclusion and partition chromatograph. A limitation associated with this approach is the relatively long run times (5–15 min) and the potential sample instability in biological fluids. Its effectiveness in removing the protein and other matrices is similar to using PTT coupled with HPLC.\textsuperscript{91}

Turbulent flow chromatography is another direct-injection sample preparation technique. In this method, the separation of small molecules from large molecules and polar matrices is obtained by nonlaminar flow of the mobile phase through use of large particles (50\,\mu m) for the stationary phase. Further separation of analytes from other components can be achieved with column switching to an analytical column. With the demand for higher throughput, generic turbulent flow chromatography was developed for routinely removing protein and polar matrices from biological samples;\textsuperscript{88,89} if the washing protocol, including both acidic and basic wash,\textsuperscript{85} is followed even more matrix components can be removed from the sample. By adding on-line dilution of the eluate, optimal usage of switching valves and dual extraction
column procedures, the problems of peak fronting or tailing, variable recovery, and high carryover can be dramatically reduced.89

Molecularly imprinted polymer (MIP) extraction is a very selective cleaning procedure for analytes. The custom made MIP will only retain the components with the same steric and chemical properties as the analytes, therefore the matrix components can be completely and efficiently removed. The key step of this approach is the preparation of steric and chemical molecular imprints by polymerization of functional and cross-linking monomer in the presence of a templating ligand, or imprint species.90 Even though there are reports of successful applications in bioanalysis,92,93 issues of template bleeding, peak broadening and tailing due to strong nonspecific adsorption to the polymer still need to be solved. Therefore, the wide application of this technique depends on its commercial availability and a better understanding of this technique by application scientists.

One of the more popular on-line solid phase extraction apparatus is the Prospekt® system. Its popularity is due to its incorporation of single-time use solid phase extraction cartridges that elute directly into the HPLC system via three switching valves, which not only provide cleaner samples but also eliminate carry-over issues.76,94 The separation power of this system can be enhanced by coupling columns with different separation modes, resulting in the so-called two dimensional or multi-dimensional LC.48,50 The 96-well format solid phase extraction system is perhaps the most suitable method for processing a large number of samples.82,95 The recent development of a 96-well format with a small bed volume of the membrane features reduced back pressures, eliminated bed channeling, increased sample capacity, and improved repeatability and reproducibility; this system also provides mixed separation modes and a reduced eluate volume and has made SPE even more attractive for many users.96–98

The routine strategy for SPE is to retain the analyte and wash out the interfering compounds. Recently, a reversed approach was proposed to effectively remove basic matrices by retaining the matrices using strong cation exchange (SCX) while washing out analytes.99 This is most applicable for bioassays of multiple analytes, since in this situation the highly specific extraction is unlikely to be successful for all analytes. In this process, the plasma sample was first basified and deproteinated by using acetonitrile containing ammonium hydroxide followed by the separation using SCX. It was found that recovery for most compounds that have \( pK_a < 8 \) was satisfactory (>74%).99

Liquid–liquid extraction is believed to be a highly selective sample preparation method that provides extracts that show the least amount of ion suppression, thereby allowing for reproducible and accurate LC–MS/MS analysis. LLE gained popularity when semi-automated100–103 and automated LLE104 with 96-well format became commercially available. It has been reported that the extraction efficiency in these small tubes can be improved by using small inert particles with an average diameter of 1 mm to increase the extraction surface.104,105
4.3.10 Standard addition method

If finding an effective sample preparation method with appropriate internal standards is too time-consuming, the standard addition method to correct for matrix effects can be tried. This method requires at least two LC–MS/MS assays—the first with the unknown sample (X) and the second with the unknown sample spiked with a known amount of standard (X + A). If the mass response is linear between unknown sample (RX) and spiked unknown sample (RX + A), the following relationship can be established as shown in Figure 4.11, from which the concentration of analyte in unknown sample (X) can be calculated.

\[
\frac{X}{RX} = \frac{(X + A) - X}{RX + A - RX} \Rightarrow X = \frac{A(R_X)}{R_{X+A} - R_X}.
\]  \hspace{1cm} (4.13)

This method is extremely suitable for analyzing a small number of tissue samples, where sample volume is sufficient for preparation of multiple combinations of X + A. In this situation, if linear response range could be established with all the X+As, one might not need to prepare a calibration curve. As described in the section on matrix evaluation using standard addition, if one can prepare multiple combinations of X + A, then only one run will be sufficient. The most appropriate A can be chosen; in other words, the X + A that still has an MS response in the linear range, but has a sufficient difference from the response of X to allow for the calculation to be accurate. This method was successfully used to quantify toxins in scallops where the degree of signal suppression varied from scallop to scallop.\textsuperscript{106}

Our laboratory has used this method to analyze brain tissue samples, where hydrophobic matrix effects are usually more severe than in plasma samples and can vary from animal to animal. With only limited sample preparation, the

![Figure 4.11](#)  \hspace{1cm} Relationship between X and X + A for the method of standard addition.
matrix effects caused by different matrices in different animal brains can be easily corrected. As shown in Table 4.7, without the standard addition method, 56 and 60 ng/g as the brain levels for each rat would be reported; however, with the standard addition method, it is clearly shown that there was about a 50% signal enhancement for the samples compared to the standards; therefore, the correct concentration levels should be 25 and 26 ng/g, respectively, as calculated using Equation 4.13.

### 4.3.11 Using a nano-splitting device

It is well known that reducing the ESI flow rate to nanoliters per minute leads to increased desolvation, ionization and ion-transfer efficiency. Experiments have clearly demonstrated that the ESI signal can be dramatically enhanced with nanoflow ESI conditions for those low surface activity compounds, such as oligosaccharides and glycosides. It has been reported that nanoelectrospray is more tolerant of samples that contain nonvolatile salts because of its ability to generate smaller, more highly charged droplets. The advantage of nano-ESI in overcoming the matrix effects can be rationalized based on the fundamentals of the ESI process. The slower flow rate will reduce the size of the initial ESI droplets that in turn require fewer uneven fission processes and less solvent evaporation prior to ion release into the gas phase. The uneven fission process is known to be a cause of many ESI matrix effects. The uneven fission will enhance the surface-active components, such as surface-active matrices, to compete with less surface-active components, such as polar analytes, for the limited surface charge. Fewer uneven fission processes will minimize the competition between surface-active matrices and polar analytes, and thereby lead to a higher signal for polar analytes. In order to prove the effect of flow rate on the ESI responses of low surface-active compounds,
Schmidt and colleagues designed an elegant experiment with a mixture of compounds that have differences in surface activity, a low surface activity compound (turanose) and a high surface activity compound (n-octyl-glycogyranoside). Because both turanose and n-octyl-glycogyranoside are uncharged in solution, the difference of ion signal intensity at different flow rate should only reflect the contribution of the difference in surface activity. The suppression of turanose by n-octyl-glycogyranoside can be completely eliminated at a flow rate of a few nL/min while the suppression increases to 5-fold at flow rates greater than 50 nL/min.

While using a few nL/min might not be practical for routine LC–MS quantification, reducing the flow rate to 100–200 nL/min using a nano-splitting device is still a feasible approach. The same laboratory has tried to apply this system for both metabolite identification and bioanalytical quantitation. The advantage of this nanosplitting device as compared to using a capillary LC column is that samples still can be run with faster LC flow rates on conventional HPLC columns. Therefore, there is no need to be concerned about overloading the column with matrix or rapid deterioration of the column, as would be the case when using a capillary column. Furthermore, the chromatographic quality is not affected by this splitting device. With this splitting device, ionization suppression was dramatically reduced, so that more metabolites that had been suppressed in the faster flow rate (200 µL/min) were able to be detected. When it was applied to bioanalytical quantitation, one could still produce a calibration curve with a linear dynamic range similar to the standard interface, with an improved LOQ and chromatographic performance (for more on this topic, see Chapter 12).

4.3.12 Switching instruments and ionization modes

As shown in a recent study, ionization of the analyte is a very complex phenomenon; it can be affected by many factors, including by co-eluting components and instrument settings. In an earlier section, we have also shown that matrix effects can be different with different ionization modes or a different brand of instruments. Matrix effects observed in one brand of instrument might not be seen in another brand. Therefore, if the source and cause of a matrix effect is unknown and another brand of instrument is available, it may be easier to switch to the second vendor’s system. In this case, attempts should be made to evaluate the matrix effects in the other instrument using both the ESI and the APCI modes, and then use the second instrument to analyze the samples if a satisfactory evaluation is obtained.

4.4 Conclusions

Matrix effects are one of the most important causes for failures and errors in bioanalytical LC–MS/MS assays. With increasing applications of LC–MS/MS
for pharmaceutical research, the issue of matrix effects has received more and more attention. We have discussed possible causes of matrix effects based on the mechanisms of electrospray ionization, summarized different approaches for evaluation of matrix effects and proposed the strategies for overcoming matrix effects in this chapter.

Matrix effects can be caused by polar nonvolatile components as well as hydrophobic volatile components via different mechanisms. Polar nonvolatile components could cause ion enhancement at low concentrations when excess charge is increased by the increase of conductivity of spayed solution at optimum electrolyte levels. Polar nonvolatile components can also cause ion suppression at high concentrations when they change the physical property of sprayed solution that reduces the desolvation efficiency ($f$) or the ion transfer efficiency ($p$). Nonpolar hydrophobic components can cause matrix effects by competition for the limited surface excess charge in the ESI process. Both polar and nonpolar matrix components can cause ion suppression by strong ion-pairing with the analytes or by competing for gas phase protons. Ion enhancing agents that exist in biological samples, such as Li$^+$, can also induce matrix effects. Endogenous components as well as exogenous components, such as plasticizers, Li-heparin, and dosing excipients, can cause significant matrix effects.

Matrix effect evaluation is the first step for solving the issues of matrix effects. Post-column infusion, direct comparison of mass response in different matrices using pre-spiking or post-spiking approaches are the most common procedures. The purpose of the evaluation is to locate the range, to know the relative hydrophobicity, to identify the source of matrix effects, to select appropriate internal standards, and to compare the effectiveness of separation procedures.

Different strategies need to be used for solving matrix effects with different causes and sources. For the matrix effects caused by polar nonvolatile components, efficient separation from relatively hydrophobic analytes can be achieved using mini-bore reversed phase columns or turbulent flow chromatography. For polar nonvolatile analytes that are hard to separate from polar matrices, either the nano-spray technique with a concentric nano-splitting device or post-column addition of a signal-enhancing agent could be useful. For a small amount of endogenous hydrophobic matrices, modifications can be made to the chromatographic separation conditions, or other sample preparation procedures, such as off-line or on-line solid phase extraction or liquid–liquid extraction should be tried. For matrix effects caused by the ion pairing of TFA, a sufficient amount of propionic acid/2-propanol should be added to the mobile phase or these can be added to the post-column eluate. For extensive matrix effects caused by a relatively large amount of hydrophobic matrices in tissue samples, the method of standard addition should be tried. The following procedures should be adopted whenever possible:

1. Employ an appropriate internal standard.
2. Introduce the minimum amount of sample into the assay system.
3. Minimize the build-up of contaminants on the MS interface.
4. Use pre-dose samples to prepare the calibration standards or QCs.
5. Avoid exogenous matrices.
6. When all else fails, try a different ionization mode or a different brand of instrument.

With a better understanding of the mechanisms of matrix effects, application scientists should be able to work with manufacturers to design better MS interfaces that not only increase the analyte sensitivity but also minimize or eliminate the problem of matrix effects.

References


62. Liang, H.R., Foltz, R.L., Meng, M., and Bennett, P., Ionization enhancement in atmospheric pressure chemical ionization and suppression in electrospray ionization between target drugs and stable-isotope-labeled internal standards in


93. Koeber, R. et al. Evaluation of a multidimensional solid-phase extraction platform for highly selective on-line cleanup and high-throughput LC–MS analysis of


