Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Journal of Pharmaceutical and Biomedical Analysis 43 (2007) 1094–1099

JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS

www.elsevier.com/locate/jpba

Direct quantification in bioanalytical LC–MS/MS using internal calibration via analyte/stable isotope ratio

Lars B. Nilsson *, Göran Eklund

Development DMPK & Bioanalysis Södertälje, Global Development DMPK & Bioanalysis, AstraZeneca R&D, SE-151 85 Södertälje, Sweden

Received 5 May 2006; received in revised form 14 September 2006; accepted 18 September 2006
Available online 2 November 2006

Abstract

The possibility to rationalize and simplify bioanalysis, without compromising the analytical quality, by omitting the calibration curves was studied. Using mass spectrometry (MS) and a stable isotope labeled internal standard it was possible to get equally good results by calculating the results directly from the analyte/internal standard area ratio and a predetermined response factor as by the traditional way, using a calibration curve run at the same occasion. To be able to use this simplified quantification method, that we call internal calibration, in its most simple form there are some prerequisites that must be considered: (1) The relative response should not be concentration dependent. (2) The relative response should be constant between batches/days. (3) The level of analyte in the internal standard should not be detectable. (4) There should be no influence from naturally occurring isotopes of the analyte on the internal standard peak area.

A bioanalytical LC–MS/MS method for a research compound was validated both with and without calibration curves and no significant differences were found regarding precision and accuracy. It was shown that all four prerequisites above were fulfilled. Validation data were very good for the whole concentration range, 0.010–30 μmol/L. Long-term data for QC samples showed excellent precision and accuracy.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Stable isotope labeled internal standard; Calibration curve; LC–MS/MS

1. Introduction

To improve precision and accuracy in chromatographic bioanalytical methods an internal standard is usually added to the samples prior to the sample work-up. A good internal standard, usually a slight chemical modification of the analyte, should have physico-chemical properties similar to the analyte; similar recovery, similar detector response and similar retention but still be chromatographically well resolved. Using MS, the analyte and internal standard peaks do not have to be chromatographically resolved as long as they have different masses. That introduces the possibility to use a stable isotope labeled analyte as internal standard. This compound, where a sufficient number ($n=3$ or more) of, e.g. ^{13}C or ^{15}N have been incorporated, will be virtually identical, chemically, physically and biologically, to the analyte and as close to an ideal internal standard as possible. This internal standard will have the same recovery, response and retention as the analyte. The stable isotope labeled internal standard will also compensate for matrix effects [1] and will have identical protein binding, a fact that have been utilized for the simultaneous determination of free and total concentrations in plasma [2]. Note also that deuterium labeling might introduce slight differences in the properties compared to the unlabeled compound [3] making deuterium labeled compounds less ideal as internal standards.

The ideal analytical system would be perfectly linear, the signal to concentration ratio would be constant regardless of concentration and then concentrations for unknown samples could be calculated by simply using a response factor. Unfortunately, the response from analytical systems are rarely constant and furthermore the response may vary from day to day due to, e.g. ageing and fouling of instruments, therefore a multi-point calibration curve is prepared and analyzed together with each batch of unknown samples. It has also been suggested that optimum precision and accuracy is obtained by using a minimum number of calibration points and perform multiple measurements on these [4]. Using analytical systems with known and proven linearity, e.g. LC-UV, this is a viable approach and it has been used at our laboratory for many years with excellent results.

* Corresponding author. Tel.: +46 8 552 27043; fax: +46 8 553 21570.
E-mail address: lars.b.nilsson@astrazeneca.com (L.B. Nilsson).

0731-7085/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2006.09.030

In bioanalysis, unknown samples with expected differences in concentration of a factor 1000 or more are not unusual, hence also the calibration curve has to cover this huge dynamic range. A common problem is that a small error at the higher concentrations might give rise to very high false intercepts, positive or negative, making the accuracy very low for the unknowns with low concentrations. In bioanalysis, the relative standard deviation is roughly the same for all data points; thus the absolute error, the variance, increases with increasing concentration (heteroscedastic data). Weighting is applied to transform data to a homoscedastic form, with constant variance, diminishing the influence of the high concentrations on the intercept when using calibration curves.

But calibration curves are not always linear and the mass spectrometer is a notoriously non-linear detector, mainly because the degree of ionization in the ion source decreases when the amount of ions increases. The quantification softwares of LC–MS systems contain several different algorithms for fitting the best line to non-linear calibration data. In LC–MS, the most used method for fitting a line to the points in the calibration curve is probably weighted ($1/X$ or $1/X^2$), quadratic non-linear regression. Still, the current US FDA guidelines for bioanalytical method validation [5] recommend: "The simplest model that adequately describes the concentration–response relationship should be used. Selection of weighting and the use of a complex regression equation should be justified". The application of different regression models to sets of bioanalytical data, trying to minimize the residual variance in the calibration curve, has resulted in suggestions of several sophisticated regression models [6–10]. In the present paper, instead of using statistical methods to transform data, a non-linear response is transformed to a linear calibration curve by using an internal standard that mimics the analyte.

In bioanalysis, according to our experience, roughly 20% of the total time for analysis is spent on preparation and analysis of calibration standards. To be able to omit these samples without sacrificing accuracy and precision would mean a significant step forward in productivity. When LC–MS/MS is used and a stable isotope labeled internal standard is available it is shown in this paper that the calibration samples can be omitted. The unknown concentrations in each sample can be calculated directly using internal calibration via the analyte/internal standard area ratio and a predetermined response factor. A bioanalytical method was validated using both the internal calibration method and traditional calibration curves. The results from the validations are compared and long-term data of the internal calibration method are presented.

2. Experimental

2.1. Model compounds

The analyte (*N*-[2-({(2*S*)-3-[(3*S*)-3-(4-chlorophenoxy) pyrrolidin-1-yl]-2-hydroxypropyl}oxy)-4-hydroxyphenyl]acetamide, hemi-fumarate salt) is a medium lipophilic compound, with an amine pK_a at 7.6 and a phenolic pK_a at 9.1. The internal standard was the analyte labeled with stable isotopes, two deuterium and three ^{13}C. Both compounds were synthesized at the Medicinal Chemistry Department, AstraZeneca R&D, Lund, Sweden. Stock solutions and dilutions were made in 0.025 M formic acid. The compounds were stable in solution for at least 3 months.

2.2. Procedures

Calibration samples and quality control (QC) samples were made by spiking of EDTA blank plasma. Calibration and QC samples were made from different weighings. The calibration samples ranged from 10 nmol/L to 30 μmol/L and calibration curves were prepared at six occasions. QC samples were prepared at four different concentrations, 0.025, 0.5, 5 and 25 μmol/L. To evaluate accuracy and precision, five QC samples at each of these four concentrations were analyzed at three different occasions. To evaluate the method in routine use, in total 112 QC samples collected at 14 occasions and during 2 months were compared.

All plasma samples were subjected to ultrafiltration prior to injection into the LC–MS/MS system. Using an eight-needle robot (Genesis RSP150, Tecan AG, Hombrechtikon, Switzerland) 120 μL plasma was transferred to a 96-well ultrafiltration plate with a collector plate (Multiscreen Ultracel PPB, Millipore Corp., Danvers, MA, USA) and 120 μL 500 nM internal standard in 0.05 M formic acid was added. The molecular weight cut-off of the ultrafiltration plate was 10 kDa. After mixing for 10 s, the plate was centrifuged at 2000 × *g* and 37 °C for 45 min. The collector plate with the ultrafiltrate, about 50 μL, was then placed in the cooled autosampler while awaiting injection.

2.3. LC–MS/MS

The chromatographic system consisted of a HTS PAL autosampler (CTC Analytics AG, Zwingen, Switzerland) with cooled cabinets, two LC-10AD pumps and an SCL-10A controller (Shimadzu Corp., Kyoto, Japan). The column was an ACE 3 C18, 2.1 mm × 50 mm (ACT, Aberdeen, Scotland). The injection volume was 20 μL and the pumps were run in a binary gradient mode at a flow rate of 0.35 mL/min. Mobile phase A was 5% acetonitrile in 0.025 M formic acid and mobile phase B was 95% acetonitrile in 0.025 M formic acid. The gradient went from 0 to 75% B between 0.70 and 2.70 min and then rapidly back to 0% B, the effective time between injections was 3.5 min. Detection was performed using a Micromass Quattro Micro mass spectrometer with MassLynx 3.5 software (Micromass Ltd., Manchester, UK). The instrument was operated in the positive electrospray ionization (ESI) mode and responses were measured using multiple reaction monitoring (m/z 421 → 254 for the analyte and m/z 426 → 254 for the IS).

2.4. Calculations

The internal calibration results were calculated using

$$C_A = \frac{\text{area}_A}{\text{area}_{IS}} \frac{C_{IS}}{\text{RF}} \quad (1)$$

where C is the concentration, A the analyte, IS the internal standard and RF is the response factor. This is a rearrangement of the equation for simple linear regression ($y = a + bx$) where the intercept a is zero and the slope b corresponds to the response factor (RF), $x = C_A$ and $y = \text{area}_A \times \text{area}_{IS}/C_{IS}$. Using calibration curves, the concentrations were calculated in the MassLynx software using linear regression with $1/x$ weighting.

The response factor was determined for every preparation of an internal standard working solution (remade every second or third month). When the new IS stock solution was made, also two stock solutions from different weighings of the analyte were made. Both the IS and the analyte stock solutions were diluted to 500 nM, two dilutions were made from each analyte stock solution. Autosampler vials, three vials for each combination, were filled with equal amounts of IS and analyte dilutions, giving in total 12 samples for analysis. Two injections were made for each sample and the response factor in Eq. (1) was the mean ratio for $\text{area}_A/\text{area}_{IS}$. Throughout the work reported here, corresponding to a time period of 10 weeks, a response factor of 1.0295 was used.

Analysis of variance (ANOVA) [11] was used to get a better understanding of within-run and between-run precision:

$$\text{within-run precision} = \frac{\sqrt{MS_W}}{\text{mean}_{all}} \times 100\% \qquad (2)$$

$$\text{between-run precision} = \frac{\sqrt{(MS_B - MS_W)/n}}{\text{mean}_{all}} \times 100\% \qquad (3)$$

where MS_W and MS_B are the mean squares within and between groups, respectively, and n is the number of determinations per group. Mean_{all} is the mean for all determinations at each concentration.

3. Results and discussion

3.1. Calibration curves and response factor

To be able to use internal calibration in the simplest form (Eq. (1)), the following requirements must be fulfilled: (a) a relative response that is not concentration dependent, (b) a relative response that is constant between batches, (c) an internal standard free from analyte and (d) no influence from naturally occurring isotopes in the analyte on the internal standard. These four requirements will be scrutinized below.

A stable isotope labeled internal standard can be expected to react exactly as the analyte as long as they are not separated. In a bioanalytical LC–MS/MS method this means that the internal standard will compensate for any irregularities in the analytical process after that the internal standard has been added to the sample and until the actual physical separation in the first quadrupole of the mass spectrometer. In ESI-MS, the degree of ionization in the ion source is strongly dependent of the amount of molecules [12] resulting in a non-linearity in the concentration/response ratio starting at rather low concentrations. If analyte and IS are eluted and ionized simultaneously, they will show the same degree of ionization and the linear range will be extended. This is shown in Fig. 1 where the absolute response decreases rapidly at higher analyte concentrations while the relative response is constant over the concentration range, i.e. the IS and the analyte show an identical decrease in response as the analyte concentration increases. Another proof of concentration independent relative response is linear calibration curves. During the validation of this method, calibration curves were run at six occasions and statistics for these are shown in Table 1. Despite the huge concentration ranges, all six calibration curves were perfectly linear, which is also reflected in the excellent correlation coefficients. Back calculated concentrations showed mean percentage bias values between 2.0 and 3.5% for all calibration curves. The bias was higher close to LOQ but otherwise evenly distributed over the whole concentration range. The coefficient of variation (CV) for the slope was very small, only 1.6%, indicating that the slope is constant over time when an ideal internal standard is used. The mean value for the slopes, 2.0566, or after correction for the IS concentration 1.0283, was also very close to the predetermined response factor, 1.0295. On the other hand, there was a huge variation in the intercept with a CV of 88%, reflecting the inherent weakness of calibration curves [13]; low concentrations are not always accurately determined, despite weighting. This weakness

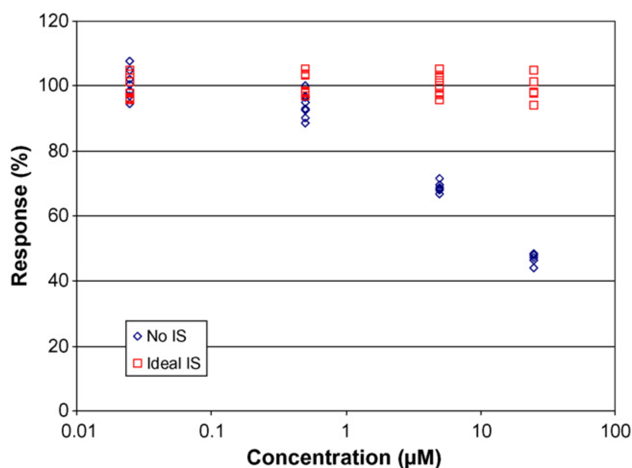

Fig. 1. Alteration of absolute (◇) and relative response (□) with analyte concentration. QC samples ($n = 8$) at four different concentrations were analyzed as described under Section 2. The absolute response is the ratio for the analyte peak area and the nominal injected concentration. The relative response is the ratio for analyte peak area/IS peak area and the nominal concentration.

Table 1
Calibration curve statistics

Range (μM)	n	Slope	Mean bias (%)[a]	Intercept	R^2
0.01–30	17	2.00290	2.3	0.00248	0.9986
0.01–30	10	2.03046	2.5	0.00054	0.9999
0.01–30	10	2.07566	2.7	0.00193	0.9996
0.01–20	6	2.06698	3.3	−0.00038	0.9994
0.01–20	6	2.09146	3.5	0.00441	1.0000
0.01–20	6	2.07227	2.0	0.00247	0.9998
Mean		2.0566		0.00191	0.9996
S.D.		0.0332		0.00167	
R.S.D. (%)		1.6		88	

[a] $100 \times |C_{calc} - C_{nominal}|/C_{nominal}$.

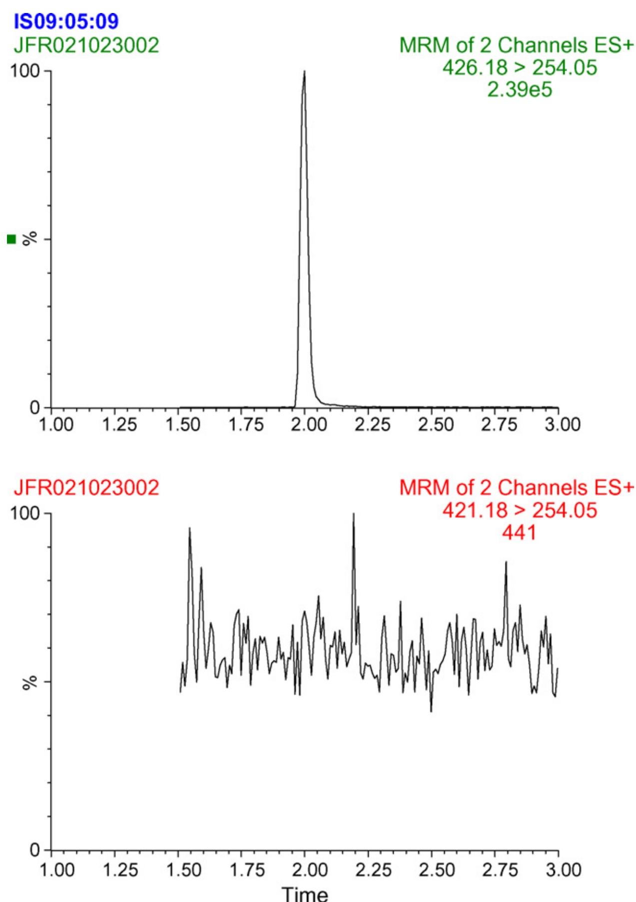

Fig. 2. MRM traces for IS (above) and analyte (below) after injection of a 5 μmol/L IS solution.

becomes especially evident when covering large concentration ranges.

Using internal calibration, an internal standard free from analyte is advantageous. In practice, internal standards are not always fully characterized and it is wise to check the percentage of unlabeled analyte in the internal standard. Injection of a 5 μmol/L solution of the internal standard showed no trace of the analyte (Fig. 2). As the limit of detection for the analyte was roughly 0.5 nmol/L, the IS contained less than 0.01% unlabeled analyte and was suitable for use in an internal calibration method.

Another source of non-linearity is the natural occurrence of stable isotopes, e.g. 1.1% of ^{13}C in carbon. For the present compound, an isotope pattern calculator tells that the natural occurrence of $M+5$ in the analyte would constitute about 0.05% of the compound. This implies that at high analyte concentrations the natural occurrence of $M+5$ would add to the area of the IS and give a non-constant relative response/concentration ratio. However, as this molecule is detected by a selected fragmentation transition, all the ^{13}C atoms need to be in the monitored part of the molecular ion and the natural occurrence of $M+5$ needs to be multiplied by the probability that all five ^{13}C positions are located in a particular part of the molecule. The part of the molecule that is cleaved off in the collision cell contains 8 carbon atoms out of the 21 carbon atoms in the intact molecule. The probability of having five naturally occurring ^{13}C atoms in this part of the molecule is $8/21 \times 7/20 \times 6/19 \times 5/18 \times 4/17$ which equals 0.0028. Therefore, the natural occurrence of $M+5$ in the detected part of the molecule is $0.0028 \times 0.05\%$, which equals 0.00014%. The injection of a 5 μmol/L analyte solution did not generate a visible IS peak, in agreement with the calculations above. At the upper limit of quantification (30 μmol/L), this isotope effect would add about 0.042 nmol/L to the added IS concentration of 500 nmol/L and this will not affect the linearity of the method. The importance of knowing the amount of unlabeled analyte in the internal standard and the natural occurrence of stable isotopes can be illustrated in a theoretical example. The calibration curves in Fig. 3 were simulated assuming an internal standard, labeled $M+4$ and containing 2% of the analyte, and an analyte, having 12 carbon atoms and 1 chlorine, containing 0.25% $M+4$. These calibration curves were compared against calibration curves for an analyte without isotope effects and with an internal standard free from analyte. The concentration range was set to 1–3000 arbitrary units. To get good accuracy at low levels (Fig. 3a) the concentration of the internal standard (5 arbitrary units) must be kept low, resulting in a severe non-linearity at high concentrations (Fig. 3b). Increasing the internal standard concentration to 200 arbitrary units will improve the linearity (Fig. 3c) but the accuracy at low levels will be lost (Fig. 3d).

Thus, all four stipulated requirements are fulfilled and internal calibration using the simple equation (1) can be used for direct calculations of the analyte concentrations in each sample.

3.2. Accuracy and precision

To show that internal calibration is a viable approach, an unbiased way of comparing the calculation methods is needed. An unbiased way is to compare the results for the QC samples as their nominal concentrations are known and as they are made from separate stock solutions independent of the calibration curve. Note that both ways of calculating the results use the same analyte peak area/internal standard peak area ratio and will thus show the same standard deviation. The difference is that internal calibration uses a constant intercept (zero) and slope while a new slope and intercept is determined each day of analysis when calibration curves are used. Table 2 shows the found mean concentrations when using calibration curves, the deviation from nominal concentrations and the within- and between-run precision values, evaluated using ANOVA. Table 3 shows the same figures of merit when using internal calibration. Both tables show excellent results and that the imprecision was almost exclusively due to within-run variations. After transforming all results to % of nominal and using ANOVA, also the difference between the calculation methods was evaluated. The within-group RSD was 5.2% while the between-group RSD was zero. It can be concluded the both calculation methods give equally good results.

A variation or trend, increase or decrease, in the response factor would also be revealed if QC samples are followed for a longer period. In Table 4, the statistics for the QC samples, two samples at each concentration on the first 14 occasions of analysis and during a period of 2 months, are given. The

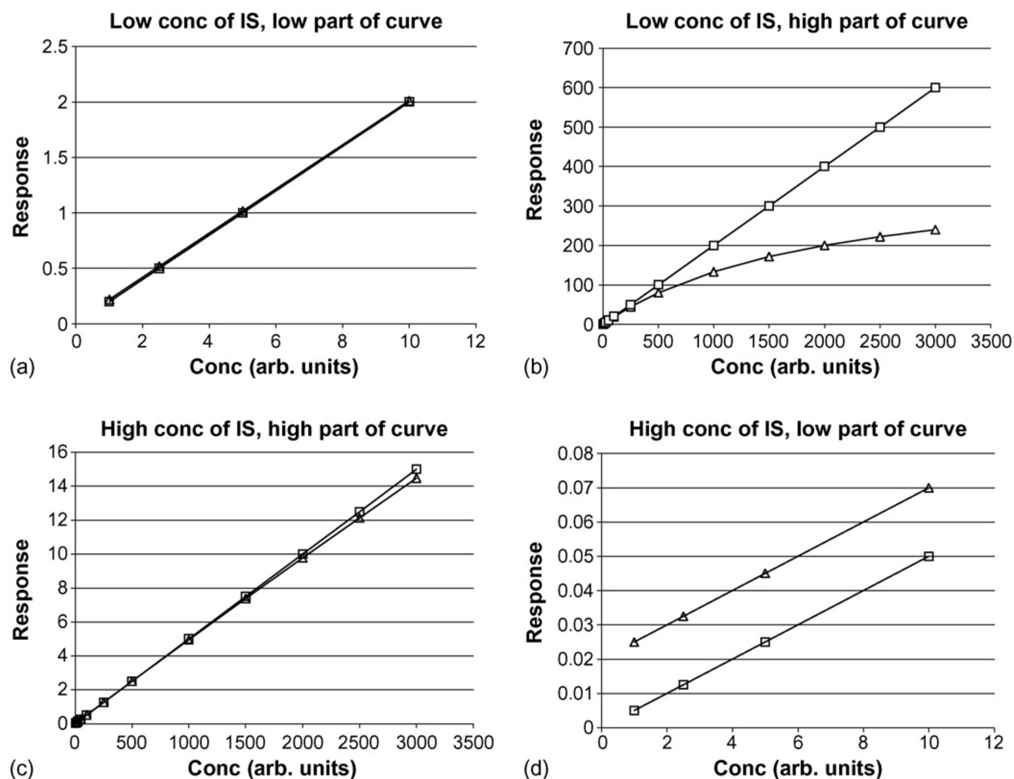

Fig. 3. Comparison of simulated calibration curves for an analyte without isotope effects and with an internal standard free from analyte (□) and an analyte with 0.25% $M+4$ and an internal standard containing 2% of the analyte (△).

Table 2
Accuracy and precision, using calibration curves (five QC samples at each concentration analyzed at three different occasions)

	Low QC (0.025 μM)	Medium QC (0.5 μM)	High QC (5 μM)	Extra high QC (25 μM)
Mean (μM)	0.0236	0.489	5.06	24.7
Deviation (%)	−5.8	−2.1	+1.1	−1.1
Within-run R.S.D. (%)	8.0	3.8	3.2	2.9
Between-run R.S.D. (%)	0	2.9	0	0

Table 3
Accuracy and precision, using internal calibration (five QC samples at each concentration analysed at three different occasions)

	Low QC (0.025 μM)	Medium QC (0.5 μM)	High QC (5 μM)	Extra high QC (25 μM)
Mean (μM)	0.0241	0.485	5.00	24.4
Deviation (%)	−3.6	−3.1	0	−2.3
Within-run R.S.D. (%)	7.7	3.8	3.2	2.9
Between-run R.S.D. (%)	0.8	1.3	0	1.6

Table 4
QC sample statistics ($n=28$ at each concentration)

	Low QC (0.025 μM)	Medium QC (0.5 μM)	High QC (5 μM)	Extra high QC (25 μM)
Mean (μM)	0.024	0.493	4.96	24.7
Deviation (%)	−2.0	−1.5	−0.7	−1.2
R.S.D. (%)	4.7	3.1	4.0	3.5
Lowest deviation (%)	−10.8	−6.4	−6.5	−6.4
Highest deviation (%)	+11.2	+4.4	+11.1	+6.9

Table 5
LOQ statistics

	Experiment 1 ($n = 5$)		Experiment 2 ($n = 10$)	
	CC	IC	CC	IC
Mean (nM)	8.6	9.6	10.7	10.9
Deviation (%)	−14.1	−4.4	7.4	8.5
R.S.D. (%)	9.0	7.8	9.4	9.1

CC: calculated using calibration curve; IC: calculated using internal calibration.

concentrations were calculated using internal calibration only. Eighty-four percent of the QC samples were within ±5% of the nominal concentration and all results were within the normally recommended ±15% acceptance limit [5]. Variations were low and almost the same over the whole concentration range. No trends were seen.

3.3. Limit of quantification

The accuracy and precision at the set lower limit of quantification (LOQ) were determined at two separate occasions. At the first occasion the first 5 out of 10 samples failed due to a technical problem, therefore a new set of 10 LOQ samples was analyzed. Results are shown in Table 5. The relative standard deviation was just below 10%. These two sets of LOQ samples were analyzed together with calibration curves 1 and 2, respectively (in Table 1). The difference in accuracy between the two calculation methods is mainly explained by the intercepts in the calibration curves and suggests an advantage for internal calibration over calibration curves when determining low concentrations.

3.4. Ion suppression and absolute recovery

Using ultrafiltration as sample work-up only proteins are removed and the ultrafiltrate still contains most of the salts and small molecules from the original sample and the risk for ion suppression is obvious. As most endogenous compounds are hydrophilic and eluting with the solvent front, a chromatographic capacity factor (k') of at least 3 is usually sufficient for avoiding severe ion suppression and an accompanying loss of sensitivity. For this method, the k' of the analyte was about 4.5. The signal intensity was compared for a sample dissolved in 0.025 M formic acid and in ultrafiltrate and the signal for the latter sample was about 10% lower. An experiment where the analyte was infused at a constant rate in the mobile phase stream and blank ultrafiltrate was injected showed that the signal was severely suppressed at $k' < 2$.

The absolute recovery for the analyte and for the internal standard, was found to be 76%. The compounds are protein bound to about 85% and the acidification prior to ultrafiltration is probably not sufficient to completely release the compounds.

Although a slight suppression and a somewhat low recovery were observed, no additional variation could be attributed to these observations as the ideal internal standard compensated well, as expected.

3.5. Routine use of internal calibration

After the validation period, this method using internal calibration was in use for almost 1 year and thousands of samples were analyzed. The total number of QC samples could be estimated to at least 500. A few QC samples, less than 1% of the total number of QC samples, have been outside the ±15% acceptance limit. No sample batches have been reanalyzed due to unacceptable QC sample results.

A constant response factor is absolutely essential. A small jump in response was noticed after an unplanned power failure. Differences in response factor have also been seen when changing from one MS instrument to another. Both changes are possibly due to differences/changes in MS calibration and parameter settings. Hence, a new response factor must be determined when transferring the method to another MS instrument. Using the same instrument and without major instrument breakdowns, the response factor seems to be very stable over long periods.

References

[1] B.K. Matuszewski, J. Chromatogr. B 830 (2006) 293–300.
[2] L.B. Nilsson, S. Schmidt, J. Pharm. Biomed. Anal. 24 (2001) 921–927.
[3] D. Wade, Chem. Biol. Interact. 117 (1999) 191–217.
[4] L. Renman, D. Jagner, Anal. Chim. Acta 357 (1997) 157–166.
[5] Guidance for Industry, Bioanalytical Method Validation, U.S. Food and Drug Administration, 2001 (available at http://www.fda.gov).
[6] J.L. Burrows, K.V. Watson, J. Pharm. Biomed. Anal. 12 (1994) 523–531.
[7] M. Mulholland, D.B. Hibbert, J. Chromatogr. A 762 (1997) 73–82.
[8] E.K. Kimanani, J. Pharm. Biomed. Anal. 16 (1998) 1117–1124.
[9] A.M. Almeida, M.M. Castel-Branco, A.C. Falcão, J. Chromatogr. B 774 (2002) 215–222.
[10] T. Singtoroj, J. Tarning, A. Annerberg, M. Ashton, Y. Bergqvist, N.J. White, N. Lindegårdh, N.P.J. Day, J. Pharm. Biomed. Anal. 41 (2005) 219–227.
[11] H. Rosing, W.Y. Man, E. Doyle, A. Bult, J.H. Beijnen, J. Liq. Chromatogr. Rel. Technol. 23 (2000) 329–354.
[12] P. Kebarle, L. Tang, Anal. Chem. 65 (2003) 972A–986A.
[13] G.K. Szabo, H.K. Browne, A. Ajami, E.G. Josephs, J. Clin. Pharmacol. 34 (1994) 242–249.