

Making an HPLC Calibration work (Part 2)

Preparation of Calibration Standards

It is normal to make up a master standard, and then prepare a series of dilutions to cover the required calibration range. Care must be taken here. Any mistakes, and all subsequent results will be incorrect.

“There are times when the HPLC conditions can change during a sequence of samples. The longer the runtime and the more samples in the sequence, the greater the likelihood of this happening. Sometimes these changes affect the detector response and hence affect the validity of the calibration.”

Author Details:

Dr Stuart Jones
Tel: 01634 294001
Email: stuart@laserchrom.com

PREPARATION OF THE MASTER STANDARD.

- Use a volume and weight which can be measured really precisely without difficulty. Using too low a mass of sample in too small a volume builds in a minimum error level into the calibration which is completely unnecessary.
- Remember that glass pipettes are usually more accurate than pipettors, and allow for volumetric errors caused by temperature variations in the lab. Volumetric glassware is accurate at 20°C. Check the balance calibration with a standard weight every day to ensure that its calibration remains as it was on the day it was calibrated.
- Check the shelf life of the master standard material, and its level of hydration. It may be necessary to subdivide it when new, and then use it from small containers to prevent changes due to degradation or absorption of water. Be sure to know the level of hydration of the material because concentration levels will be incorrect if the wrong number of water molecules of hydration are included in the calculation.
- It is a good idea to test the shelf life of the master standard solution when made up. It may be longer than expected, and can save a lot of time later.
- Remember to use spotlessly clean glassware. It is so much nicer not having extra peaks in the master standard.

Preparation of the dilutions

- Use volumes which can be measured very precisely. Working with too small a volume introduces an unnecessary potential for error.
- Consider the effects of lab temperature on the dilution accuracy.

Choose the sample solvent wisely.

- Volatile solvents make it extremely hard to maintain concentration, especially whilst pipetting.
- Samples should dissolve reasonably easily. An ultrasonic bath or shaker can help, but beware of changing the temperature whilst dissolving the sample, because of the inherent volumetric errors which can occur later.
- A UV transparent sample solvent will reduce the size of the void volume peak, making integration of fast eluting peaks easier.
- It must be instantly miscible with the eluent, and in the case of gradient elution, with the initial composition of the eluent.
- It should be no stronger an eluent than the eluent itself, and in the case of gradient elution, with the initial composition of the eluent.

Bracketed Calibration

There are times when the HPLC conditions can change during a sequence of samples. The longer the runtime and the more samples in the sequence, the greater the likelihood of this happening. Sometimes these changes affect the detector response, and hence affect the validity of the calibration. We can monitor changes by running a QC standard periodically through the sequence, but this does not update the calibration. We can re-run the calibration standards periodically during the run, but this will either average with the previous calibration or replace it, and either way, it means that every few samples, the calibration changes, making it hard to compare results. We

could ignore the changes, but this means that the calibration accuracy becomes progressively worse during the sequence. The solution is to use bracketed calibration.

Essentially this means running the calibration standards at the beginning of the sequence and at the end, and makes the assumption that any changes occurred in a linear manner during the sequence. The data system then changes the calibration incrementally from the beginning to the end, and applies this to the results. If the assumption that the change was linear is correct, the data should then all be correctly quantified. Not all data systems have this function, and for long runs it is very useful.

THE USE OF ZERO IN A CALIBRATION CURVE

Why a calibration curve may not pass through zero

It is clear that if a blank is injected, a value of zero should be obtained. If it is not, this needs to be addressed as a matter of urgency. However it is not necessarily true that a calibration curve should pass through zero.

The first consideration is that when working close to the lower limit of detection, there will be a minimum sample concentration which gives rise to a detector response. This concentration and below will all give a detector response of zero. So in this context, it would be normal to expect the slope to intercept with the y axis at a point below zero.

It is not normal to get a positive intercept with the y axis. However if this occurs, it is important to be convinced of the validity of the data points that have been run. A positive intercept should at least raise a question mark. It is possible that the slope is correct but the calibration has a constant bias (ie all points are too high by a fixed amount) and hence all results will be low, by a constant amount. Or it could be that highest standard is correct but the slope of the curve is incorrect, thus giving the highest error at low concentrations.

The calibration is valid only between the minimum and maximum standard levels that have been run. It is not valid to extrapolate above or below these levels, and if samples present values outside of the calibration range, they must either be diluted to bring them within the range, or the calibration must be extended by running a higher or lower standard.

Note that a disadvantage of using Standard Addition Calibration is that there is no possibility to detect errors relating to zero intercept.

Using Zero as a point on the calibration curve

Following on from the above, if the curve passes through zero, it is acceptable to use zero as point on the curve. Be careful about forcing the line through zero. It may be better to allow the data system to compute the curve with zero as a point on the curve rather than as a fixed point. If the curve does not pass through zero, then zero must be ignored altogether. Forcing zero when the least squares curve clearly does not pass through will induce an error which is most apparent at low concentrations.

The importance of running a blank

There are three reasons for running a blank:

1. To prove that zero is zero!

- There may be instances where a trace of the sample of interest is present in the blank, such as when analysing low levels of sodium in deionised water.

- If running a gradient, there may be traces of the sample of interest on the column.
 - This test also identifies any contaminated solvent.
- So a blank is very important to confirm the zero point for the analysis

2. Checking for carryover. If the autosampler wash routine is not sufficiently rigorous, traces of one sample may appear in the following sample. This is most apparent when running samples of high concentration, so to test for carryover, run the highest concentration sample or standard, followed by a blank containing solvent only. If even a trace of carryover exists, the wash routine should be modified to eliminate it. Check that the autosampler washes the outside as well as then inside of the needle. And be careful not to solve this problem using a strong eluent as wash solvent. A variable volume autosampler may inject 10ul of sample and 190ul of wash solvent in a 200ul loop. So if a strong eluent is used as wash solvent, really broad peaks will result.

3. Identifying ghost peaks

Using isocratic elution, ghost peaks should not occur. However if a sample presents with an unusual impurity which has a retention time longer than the normal run time, it will appear as an unusually broad peak in a subsequent chromatogram. This will not be identified by running standards, but if ghost peaks arise, run a series of 2-3 blanks after a sample and see what comes out!

However with gradient elution, ghost peaks are commonly observed, and it is important to identify them before starting analysis. They usually arise from the aqueous part of the eluent, and their peak height is proportional to the length of time spent re-equilibrating the column after each run.

LIMIT OF DETECTION AND LIMIT OF QUANTITATION

Limit of Detection – LOD

The limit of detection is a measure of the lowest concentration which can be detected using a particular HPLC method with certainty that it is a peak. It is generally accepted that a signal:noise ratio of 3:1 is the limit of detection, so it will be necessary to identify the concentration which gives a peak at three times the height of the noise.

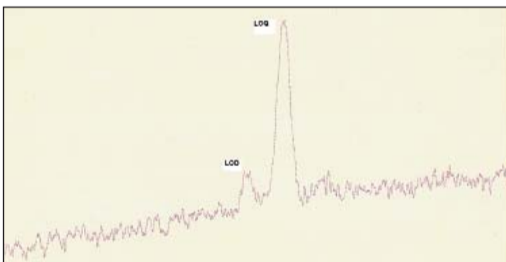


Figure 1.

There are problems associated with this. Noise is complex, arising from electronic noise, lamp fluctuations and spikes, pressure pulsations etc, and each of these occurs with a different frequency. The observed noise will be any or all of the above superimposed on each other. As a consequence, the noise is not constant, and it can be difficult to determine exactly what the noise level is. Noise also changes with flow rate, increases as the detector lamp ages, and changes when the peak width setting is altered. So although the limit of detection will be specified as a concentration level, the noise level from which it was derived should also be stated.

If the LOD is unacceptably high, there are various things that can be done to reduce the noise, such as a new lamp, the use of a pulse dampener, using screened signal cables if an analogue signal is used etc. If the LOD is still not low enough, a more sensitive detection method should be used, such as fluorescence instead of UV.

When a peak is detected which is below the limit of detection, it will be ignored by the data system if the LOD is specified in the method, and the report will be None Detected.

Limit of Quantitation – LOQ

The limit of quantitation is the lowest sample level which can be integrated and analysed with an acceptable accuracy. The smaller the peak, the harder it is for the tick marks to be placed in the correct position by the data system, and the greater the percentage error which results when the tick marks end up in the wrong places. A small error in the position of the baseline results in a large percentage error in peak area.

Limit of detection is also set as a function of the signal: noise ratio. Many labs accept 10:1, and live with the 20-30% error which can result. Others prefer 30:1 because of the inherently greater accuracy. As with the LOD, the LOQ is very sensitive to changes in the noise level, so for sensitive analyses, the noise level must be monitored regularly, and corrections made when it rises to an unacceptable level. This may mean changing the detector lamp more often than would otherwise be necessary, and maintaining the pumps regularly.

When the data system detects a peak which is below the LOQ but above the LOD, the report will be a Trace.

Integration by Height or Area

Sample concentration is proportional to peak area. So unless there is a good reason to do otherwise, we integrate peaks by area. However there are times when it is hard to measure a peak area accurately, especially using integration parameters to do it automatically. In this situation, we may find it preferable to integrate by height instead of area.

It is **valid** to integrate by height instead of area when we have **sharp, Gaussian (symmetrical) peaks**.

It is **preferable** to integrate by height instead of area when

- Noise levels are high and it is hard to place the start and end tick marks
- There are small peaks eluting close to the beginning or end of the peak, which may sometimes get included with the main peak.
- The sample is a natural product extract with many components
- Some peaks are not totally resolved

It is not necessary to integrate all peaks by height. This choice can be made for each peak separately. But once the preferred integration method has been selected, it must be used for standards and samples alike.

Because of the complexity of the sample matrices, many hospital labs integrate by height as the default. For most other labs, the default should be integration by area.

Note that integration by height is much more sensitive to changes in retention times. As peaks retain longer, they become broader and the height decreases. So if using a reference peak to correct for variations in retention time, be careful because it will find the peaks but the results may well be inaccurate. If retention times change, run the calibration again.

FACTORS WHICH CAN THREATEN THE INTEGRITY OF A CALIBRATION

1.Errors.

The four main sources of errors in HPLC are the weighing, dilution, injection and integration. Errors can be systematic (the same error occurs every time) or random (we get a different result each time).

Systematic Errors affect accuracy. Sources of error include balance or pipettor calibration, incorrect integration settings, using the wrong flow rate, or using the wrong calculation (eg when calculating the concentration of calcium in a standard, using the wrong number of water of hydration molecules). Systematic errors can be very hard to detect in-house.

Random Errors affect precision. Sources include poor technique (eg spilling sample on the balance pan so it is included in the weighing but not in the solution, not shaking a flask, or allowing solvent to drip out of a pipettor whilst transferring from one flask to another), or air in the autosampler syringe.

If we get a calibration that looks like this, we have errors all over the place:

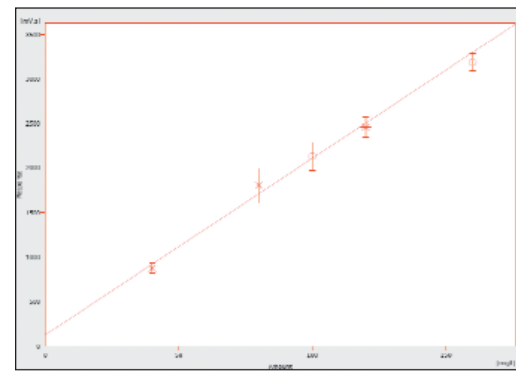


Figure 2.

It is clear that the dilutions have not been made up carefully, and then there are major variations in result between the two duplicate injections at each level. Two standards are not even included in the least squares fit of the line.

2.Changes in the Noise level

Noise arises from several sources, as discussed earlier. If the noise level increases, for example as a result of the lamp nearing the end of its life, then the integration parameters will need to change in order to integrate peaks correctly. An increase in noise, especially spiking noise from a worn lamp, can make it very hard to get the start and end tick marks in the correct places.

3.Retention Time Reproducibility

Retention time consistency is important for two reasons. First, we identify peaks by their retention times, so if these change, the peaks may fall out of their time windows, or worse, fall into another time window if the peaks are close together. And second, because peak areas can change with changing retention time. As an example, a 200mg/l standard which eluted at 32 mins gave a result of 200mg/l. But when the eluent strength was increased, and its retention time reduced to 27 minutes, the integration result showed 170mg/l. Using the old eluent, and the result returned to 200mg/l.

Peak identification is more reliable if a diode array detector is available, so that the peak identity can be confirmed by comparing the observed spectrum with that stored in the computer library. If using LC-MS, then the molecular ion will usually make peak identification straightforward.

If retention times change, the first place to look is the void volume peak. If its retention time has also changed, the problem is a flow rate issue. Either the user has entered an incorrect value, or there is a leak. Changes in flow rate are also accompanied by a corresponding change in system back pressure, and a change in the drop rate into the waste bottle.

If it is not a flow rate issue, then a change has occurred which has changed k' from the Resolution Equation. ie a change in temperature, %B in the eluent, pH, buffer concentration etc.

If the retention times are there or thereabouts, but the actual times change a little from one injection to the next, change the check valves. Or at least try cleaning them in an ultrasonic bath.

4.Temperature

An HPLC system measures sample concentration. But the volume of the solvent used to dissolve the sample changes with temperature. Hence the apparent sample concentration is also temperature dependent. When using volumetric glassware it is very important that the lab temperature is close to 20°C, or that a correction is made for temperature. The same applies to dilutions.

If this becomes a problem, it may help to make up standards and samples on a weight/weight basis instead of weight/volume. The procedure is:

- Weigh the flask
- Tare
- Weigh in the sample
- Record the weight - weight 1
- Make up to the line with solvent
- Record the weight again - weight 2

The dilution is weight1/weight 2. eg 1g in 108g, instead of 1g in 100ml.

Another source of temperature-related errors is the chilled autosampler tray. Samples placed in the tray take a while to equilibrate at 4°C, and during this time the volume in the vial decreases. Hence it is very important to allow time for this to reach equilibrium before making the first injection. How long does it take? Inject a sample over and over until reproducible retention times are achieved.

5.Purity of standards

It is important to use the purest standards available. If a 97% purity standard is used, the calibration will probably be out by 3%. Some labs correct for this by multiplying their standard levels by 0.97, but this too can lead to errors. If the supplier has run out of 97% material, they could easily supply a higher grade – 99.9% could still pass the QC test for 97% grade! The only way to be sure that the calibration is correct is to use the highest purity possible.

6.Peak Shape Changes and Peak Efficiency

Resolution is proportional to peak efficiency (N is one of the three parameters in the resolution equation.) So as peak efficiency decreases, so does resolution, and at some point this causes a problem.

If a void develops in the top of a column, the peak shape deteriorates, and this is accompanied by a gradual increase in back pressure. As peaks become broader, the area tends to increase, so at best the calibration should be run again. The problem comes if this develops during an overnight sequence, because the accuracy becomes progressively worse during the night.

The solution is to use system suitability criteria. The System Suitability Test (SST) is usually an optional extra with data handling packages but is very useful. It allows us to specify a minimum peak efficiency which is acceptable, and when N falls below this level, the results are no longer accepted. SST also allows us to specify a retention time range, so that if peaks fall out of their retention time range, the results will be rejected.

Poor peak shape can also result from an extra-column void. This has the same effect on the peak shape, but is not accompanied by rising back pressure, and does not tend to get worse during the night! It comes from using the wrong ferrule lock distance for a tubing connection, or using wide bore tubing from the autosampler to the column, or the column to the detector. The error will be immediately apparent the first time a sample is run, so it is worth watching the first sample run through to check all is well before saying goodnight.

7.Changes in detector settings or a new lamp

If a method is modified, be sure to check that setting such as wavelength or response time have not been changed.

A visit from a service engineer to change the lamp will also make a big difference to peak heights. The lamp energy will be much higher with a new lamp, resulting in much larger peaks. So the calibration will have to be run again. In some cases where a method has been developed using a detector with the old lamp, peaks may even go offscale with the new lamp. This should be checked, especially if the calibration linearity becomes worse with the new lamp.

8.Integration Settings

Some integration settings affect where the start and end tick marks are drawn, thus affecting the area of the peak. If these settings are modified, perhaps to solve another problem, the calibration will have to be run again using the new peak areas. This does not mean re-running the chromatograms, but simply reloading them into the calibration so that it takes account of the new peak areas.

Another problem area is the handling of unresolved peaks. Where the ratio of the sizes of the two peaks is 3:1 or less, we would always drop a perpendicular. Where the ratio of their sizes is 4:1 or greater, it would be normal to skim off the rider peak. Where the size ratio is between 3:1 and 4:1, care needs to be exercised, and the selection is made to ensure that peaks are always separated using the same method. It is vital to ensure that all samples are treated the same way, or the calibration fails. This is normally handled by setting the skim ratio. But in some data systems the default is to drop a perpendicular, and rider peaks have to be skimmed using timed events. Either way, it is vital that integration uses one method or the other.

The final warning from integration settings is concerning the use of timed events. This could be a wavelength change, or turning the integration on and off to avoid a baseline anomaly, or detection of negative peaks. Using timed events works very well when retention times are constant. But when the retention times vary, the result can be chaos! So the guide is to use timed events only when necessary.

9.Presence of unexpected impurities

The presence of an unexpected component in a sample can cause problems. It may co-elute with a sample of interest. Or it may at least elute close by

and make integration difficult.

This is unlikely to occur when analysing a product made in-house, such as a flavoured water, but can easily occur when analysing natural products. It is especially important to consider this in the hospital environment, where the lab has no control on the drugs a patient may be taking, or their choice of diet, or any metabolic disorders they may have. In this situation, the analyst must be aware of the possibility for errors resulting from the presence of such a peak.

CONFIDENCE IN RESULTS

It is very important to have confidence in the result. There is no reason not to! We know exactly what is in the standards, we can see if the integration is correct, and if we have a good calibration with good precision in the results, we have every reason to be confident. Where QC limits are wide, it is perhaps less important, but where important decisions must be made (such as to accuse an Olympic athlete of taking drugs, or to release a ship load of product around the world) it is very important to have confidence in the calibration and the results derived from it.

This is where method validation comes in. But essentially it boils down to:

Check the accuracy. This is very hard to do in-house, because systematic errors are very hard to detect. So it will probably mean sending out samples for external analysis.

Check the precision. This has to be done in-house. Optimise all random errors, then run samples repeatedly and calculate the standard deviation. Using statistical theory, assuming the results follow a normal distribution, 95% of results will be within +/- 2 standard deviations and 99.7% will be within +/- 3 standard deviations.

Stuart Jones runs a wide range of HPLC courses, both at Laserchrom, and at locations around the UK, including one covering Integration and Calibration in HPLC. The Stuart Jones HPLC Academy is the training division of Laserchrom HPLC Laboratories Ltd, and details of all the courses may be found at www.hplccourses.com. On-site courses are available.