The Analyst

Assay by nuclear magnetic resonance spectroscopy: quantification limits

Lee Griffiths* and Alan M. Irving

ZENECA Pharmaceuticals, Mereside, Alderley Park, Macclesfield, Cheshire, UK SK10 4TG

The use of an internal standard in NMR spectroscopy to determine the strength (assay) of a material incurs several errors. The sources and magnitude of these errors were examined so that they may be minimised. Minimisation of these errors will result in strength determinations of these materials with 95% confidence limits of ±1%.

Keywords: *NMR spectroscopy; signal-to-noise ratio; quantification*

NMR spectroscopy is able to give resonances the areas of which are directly proportional to the number of relevant nuclei.[1] With Fourier transformation and chemical shift resolution, this furnishes molecular stoichiometries. However, higher levels of quantitative accuracy are available and quantification of the population of one compound relative to another is feasible.

If it is assumed that all of the impurities in a sample are NMR (^{1}H) observable and if all the impurity resonances are then individually identified and integrated, the strength (assay) of the sample can be determined to high levels of accuracy. For example, for a single impurity of 1%, an NMR determination accurate to ±3% relative (a conservative but common estimate) would yield an overall error of 0.03%. The problem with this approach is the large amount of work involved in identifying all the minor components and the validity of the assumption that all impurities are NMR (^{1}H) observable.

Alternatively, the assay can be determined by measuring the integrals of the sample of interest against those of a standard weighed into the same solution[2,3] (referred to here as a 'strength' determination). Since the percentage errors are then a ratio of the whole sample and not just of the impurities, the errors are larger. Hence, if each integral carries a 3% error, a true 100% strength would, in principle, be observed as a strength of between 96 and 104% (the variance for the sample in question and the standard are added).

NMR is used increasingly in the pharmaceutical industry to calibrate other quantitative techniques such as HPLC and, in some circumstances, is used as the basis for setting specifications and to monitor impurity levels.[4–7] An assay precision of ±4% would be of limited use.

High magnetic fields and repetitive pulsing concomitant with Fourier transform NMR have given rise to higher sensitivity than hitherto. Higher fields have also increased dispersion, decreasing the probability of spectral overlap and increasing the specificity of the technique. This paper describes an investigation into the reliability of quantitative determinations by ^{1}H NMR and an investigation into the individual sources of error. Minimisation of these sources of errror leads to a strength precision of ±1%, increasing the applicability of this methodology.

Experimental

The effect of noise on peak area was determined using a Bruker AM-500 NMR spectrometer operating at 500.14 MHz. A sample of chloroform and $Cr(acac)_2$ was prepared in $CDCl_3$, $Cr(acac)_2$ was added to reduce the relaxation time and to broaden the resonance so that sufficient points were measured across the peak without having to use large data sets. A 90° pulse was employed such that errors in tip angle did not contribute to variation in area and 200 spectra were acquired.

A Fortran program was written to perform the peak area calculation: a summation of total intensity above the straight line of best fit (not necessarily horizontal) to 40 baseline points either side of the peak. The peak measured was between 20 and 27 points wide at half-height and the total number of points summed was 461. The point-to-point rms noise was determined using 15% of the baseline points (77) 5% from the high frequency limit. The determination of noise was about the line of best fit (again not necessarily horizontal) to those points.

In order to create different values of the signal-to-noise ratio, the receiver gain was varied from 1 to 8. In each case, the standard deviations of the point-to-point noise and the peak area were ratioed to a constant area for all values of received gain.

Experiments on integral variation across the spectrum and purity studies were performed on a Bruker AMX-400 NMR spectrometer operating at 400.13 MHz. A 5 mm QNP (^{1}H, ^{13}C, ^{19}F, ^{31}P) probe, yielding a 90 proton pulse of 11.5 μs, was used. Samples of ethyl picolinate, 3-pyridinepropanol and 2,6-di-*tert*-butyl-4-methylphenol were obtained from Aldrich. The relaxation delay plus acquisition time was 60 s and weighings were performed on a Mettler AT250 balance.

Results and discussion

If overall errors in assay measurement of ±1% are sought, it therefore follows that all individual sources of error (each of which has given rise to an error of a significant fraction of 1%) need to be considered. We have separated our treatment into systematic and random errors. Whereas systematic errors have been discussed before,[8] the precision sought in this work warrants re-examination.

Systematic errors

Resonances away from the carrier frequency are incompletely excited and incompletely detected. Provided the excitation field (B_1) is sufficiently intense, a rectangular rf pulse of finite width produces (after Fourier transformation) a sine shaped excitation profile, *i.e.*

$$\text{intensity} = \sin(2\pi sa)/2\pi s \quad (1)$$

where s is the frequency offset from the carrier and a is the width of the rectangular pulse in seconds. As the pulse width increases, the intensity of resonances away from the carrier are attenuated more. The chemical shift range in ^{1}H NMR is limited, but at higher magnetic field and/or as higher levels of precision are sought, off-resonance effects become significant.

With an 11.5 μs pulse width (that used in this work), a resonance 5 ppm (δ) away from the carrier (usually chosen to be the middle of the spectrum with quadrature detection) at 400 MHz would suffer a 0.4% decrease in intensity. At higher fields, it is more difficult to maintain short 90° pulses and resonances

move further apart in frequency. If high quantitative accuracy is required then lower tip angles can be used, the peak areas can be corrected explicitly using eqn. (1), or, if only two resonances are of interest, the carrier can be placed mid-way between them.

The above analysis is simplistic at finite B_1 fields. An 11.5 μs 90° pulse is equivalent to a 21.7 kHz decoupling field. A resonance 5δ from the carrier experiences a $B_{0\text{effective}}$ of 2 kHz at 400 MHz, $B_{\text{effective}}$ is then 5.3° above the *xy* plane in the rotating frame of reference, giving rise to a phase error which can be corrected and a further 0.4% error in intensity.[9] The expedients proposed previously for finite pulse width are effective, apart from reducing the pulse width, which has no effect. The rf power could be increased but arcing within the probe should be avoided.

Pulse tip angles interact with pulse intervals and relaxation to affect quantification in NMR. The conditions for optimum signal-to-noise ratio were well documented 30 years ago by Ernst and Anderson.[10] The conditions which give rise to optimum signal-to-noise ratio and quantification were investigated more recently by Cookson and Smith.[11] They recommended large tip angles. The full characterisation of errors incurred by incomplete relaxation requires the additional measurement of the spin–lattice relaxation times (T_1), which is a lengthy process. If pulse intervals are chosen to be long (say 60 s), these effects can be ignored at the expense of slightly degraded signal-to-noise ratios. The effect of the signal-to-noise ratio on area measurement is relatively simple to characterise; the information can be extracted from the existing spectrum and this is covered later.

The analogue to digital converter (ADC) imposes a limit on the vertical resolution in the free induction decay (FID), the transformed spectrum and hence the peak area. The maximum number that the ADC can accommodate is $2^{n-1} - 1$, where *n* is the number of bits, allowing for one sign bit and all other bits set. This gives 32 767 for a 16 bit ADC and 2047 for a 12 bit ADC. Having to set the pre-amplifier (receiver) gain by factors of two potentially halves the vertical resolution. Floating point Fourier transforms (now commonplace) ensure no further degradation of the vertical resolution. Since the intensity in the FID encodes the total area in the spectrum, the vertical resolution is only preserved for a spectrum consisting of a single resonance. At least two resonances are required for strength determinations. If these are present at the optimum equal areas, there would be a further halving of vertical resolution to one part in 8192 (0.01%) for a 16 bit ADC and one part in 512 (0.2%) for a 12 bit ADC. In practice, there will be far more peaks in the spectrum, further reducing the vertical resolution. For strength determination, a 12 bit ADC is at the limit of acceptability whereas 16 bit ADCs give negligible errors.

An adequate number of data points needs to be summed in order to describe fully the area of a peak. The number of points across a peak is a function of the peak width and digital resolution (spectral width/number of data points, strictly a reciprocal of resolution). The absorption spectrum and the dispersion spectrum contain alternative information and zero-filling to double the number of points acquired, before Fourier transformation, ensures that all the data are displayed. Zero-filling to a ratio > 2 does not improve the information content *per se*, but does improve the shape of peaks and, in that it is an interpolation procedure, improves the accuracy to which the peak position can be determined.[11]

As fewer points are used to describe the peak, the error in the peak area increases and more points need to be acquired or smaller spectral widths employed. The quantitative aspects of this were treated by Lindon and Ferrige.[8] In their work, they also considered the effect of finite vertical resolution, but, as alluded to above, with 16 bit ADCs and 32 bit computers, this is no longer important unless two peaks are widely divergent in intensity.

The results in this work diverge from those of Lindon and Ferrige. In order to test the effect of finite digital resolution on area measurement, a Fortran program was used to calculate 2000 points across a Lorentzian peak to a resolution of peak width/50. The effect of finite digital resolution was then simulated by summing every *n*th point. For all values of $n \neq 1$, there are *n* alternative positions from which one may start the summation. By varying the starting point the measured area oscillates such that the maximum and minimum areas are determined for each value of *n*; *n* was then in turn varied between 1 and 50. The maximum and minimum area limits are given as a function of digital resolution/linewidth in Fig. 1.

The results demonstrate that until the digital resolution is > 0.4 of the peak width, the maximum error in the area measurement is 0.1%. It will be found that peak widths commonly encountered, *i.e.*, of samples not degassed or suffering exchange broadening, are in the range 0.8–1 Hz. For a spectral width of 7 kHz (adequate to avoid most baseline distortions at 400 MHz) and using 64K points total data, the digital resolution is 0.11 Hz and the maximum area error is less than 0.01% smaller than the errors described by Lindon and Ferrige. At normal peak widths and data-set sizes, we find that the digital resolution will be adequate.

In order to include 100% of the peak area, an integral would have to extend to infinity in either direction. In order that realistic integral widths can be selected, the integral of a Lorentzian peak shape as a ratio of the total area is given in Fig. 2. The calculation of total area was done to 100 times the peak width (at half-height) in each direction and to a resolution of peak width/50. If the integral extends to 24 times the peak width in both directions, 99% of the total area is enclosed. At 39 times the peak width the area enclosed is 99.5%. In order to create insignificant errors (0.1%, say), the integral would need to extend 76 times the peak width in both directions.

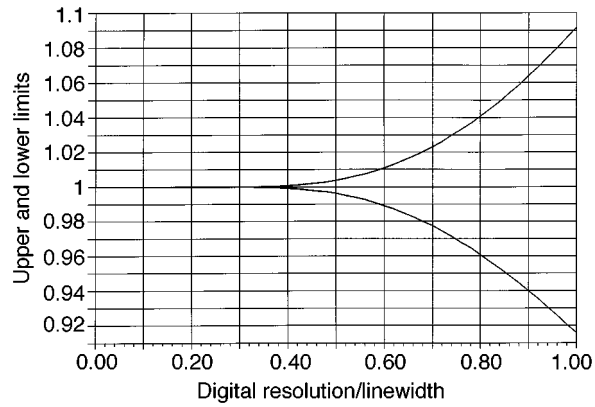

Fig. 1 Effect of digital resolution on area.

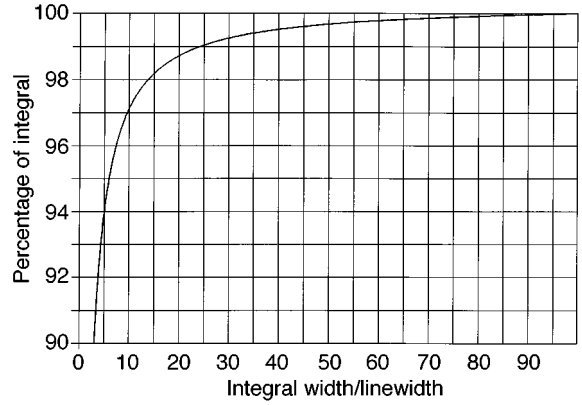

Fig. 2 Integral of Lorentzian peak.

Random errors

Finite noise in an NMR spectrum contributes to the random peak area determination. Whilst spectral repetition is a reliable way of determining the size of random error, it is time consuming. If the standard deviation of the area of each peak could be determined from a single spectrum, it would be a much more efficient way of determining the reliability of a quantitative result. It can be shown[12] that (provided the noise in adjacent points is not correlated), the standard deviation of the peak area is

$$\text{area}_{SD} = \sigma hn/\sqrt{n} = \sigma h\sqrt{n} \quad (2)$$

where σ is the point-to-point rms noise, *h* is the digital resolution (Hz per point) and *n* is the number of points integrated. The result is that minimising *h* and then minimising *nh* (the width over which the integral is determined) minimises the error in peak area. In minimising *nh*, consideration should be given to the issues raised in the previous section.

Increasing the total number of points acquired in the FID minimises *h* and maximises *n* in direct proportion, potentially decreasing area_{SD} in proportion to the square root of the FID size. This is an attractive proposition in that it potentially increases precision while the acquisition time increases to occupy time which is usually wasted waiting for relaxation. Unfortunately, this expedient does not decrease the error in peak area, since longer acquisition times lead to a compensating increase in point-to-point noise after Fourier transformation.

In order to test the validity of eqn. (2), 200 spectra of chloroform at four values of receiver gain (because some noise is introduced after the pre-amplifier, varying the receiver gain allows some variation in signal-to-noise ratio) were analysed to give standard deviations in peak area which are presented in Fig. 3 (40 baseline points). The slope of the plot was 4.6 times that predicted by eqn. (2). Either eqn. (2) is incorrect, or an additional factor is responsible for augmenting the standard deviation. Since the straight line fitted to the experimental points passes through the origin (within experimental error), the source of the enhanced slope scales with point-to-point noise.

A potential source of such an effect is the accuracy to which the straight baseline is fitted. The straight baseline fitted either side of the peak by our software (see Experimental) is equivalent to that performed during slope and bias corrections in commercial NMR integral software procedures. In order to test the effect of noise on baseline mis-set, the number of points used either side of the peak was reduced from 40 to 20 and then to 10. The resulting peak area standard deviations are also presented in Fig. 3. It can be seen that fewer baseline points increase the error of the determination.

In order to extract the function giving rise to the slopes in Fig. 3, the slopes were plotted arbitrarily against $1/\sqrt{m}$, where *m* is the number of baseline points fitted on either side of the peak. The resulting plot is given in Fig. 4.. The intercept of this plot should yield the equivalent of a slope in Fig. 3 for an infinite number of baseline points. This value is 1.5 times larger than that predicted by eqn. (2), a much closer fit than the isolated determination of 40 points. Indeed, within experimental error, it must be regarded as agreeing with eqn. (2).

Since eqn. (2) predicts the effect of point-to-point noise on the peak area, it would also be advantageous to be able to calculate the effect of point-to-point noise on the baseline mis-set. The area is that of a triangle of width *hn* and height a function of σ. Since the error can occur on both sides of the peak, the area is doubled. By analogy with eqn. (2), the area_{SD} as a result of baseline mis-set can be represented as

$$\text{area}_{SD} = \sigma hn/\sqrt{m} \quad (3)$$

The validity of this equation can be checked by reference to the slope of Fig. 4, *i.e.*, $\text{area}_{SD}\sqrt{m}/\sigma$. The slope is actually 0.84 times that predicted by eqn. (3) and is within experimental error.

The effect of point-to-point noise can therefore be summarised as

$$\text{area}_{SD} = \sigma hn\sqrt{1/n + 1/m} \quad (4)$$

Therefore, it follows that the least error is obtained by maximising the number of points used in the baseline determination.

Another likely source of error is the accuracy of phasing (mixing the real and imaginary spectra). In strength determinations, auto-phased peaks are rarely acceptable. It can be shown that as the phase error tends to 0°, a negative deviation in the signal of *x*% will give rise to an area error of 2*x*%. If optimum precision is required, the vertical expansion should be increased as much as possible and manual phasing performed.

The use of slope and bias corrections on integrals can lead to significant errors. A view often expressed is that, provided the baseline correction is done adequately, slope and bias corrections are not necessary. This is, of course, correct, but overlooks the fact that the integral is more sensitive to baseline imperfections than the baseline *per se*. Furthermore, most slope and bias correction software allows interactive adjustment of the linear baseline while still observing that sensitive indicator of baseline, the integral. Manual slope and bias corrections should therefore be regarded as a prerequisite for accurate strength determinations.

In the case of convoluted peaks the integral can be split, with or without subsequent baseline correction. In the former case, the peak area ratios will be in error but the sum of the areas will not. In the latter case, splitting and baseline correcting the integral are equivalent to subtracting a triangle of area whose apex is at the valley between the two resonances. This will reduce the measured area. Fig. 5 shows an exaggerated case in order to illustrate the effect, *i.e.*, two Lorentzian peaks 50 Hz

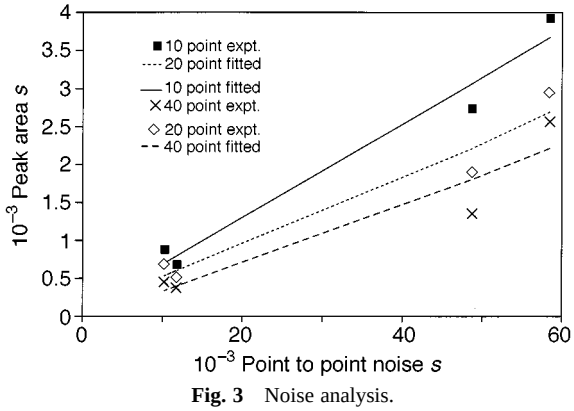

Fig. 3 Noise analysis.

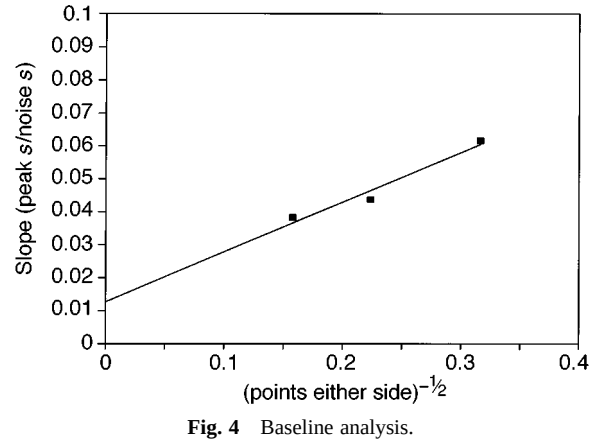

Fig. 4 Baseline analysis.

wide, separated by 200 Hz. The normal software provided with NMR instruments would effectively start and finish the triangular baseline at the start and finish of the integral. This would create significant negative area contributions in the region of the peak wings, resulting in further reductions in measured area. Our software, however, fits a baseline which starts from the valley between the peaks, but is tangential to the NMR spectrum and extrapolated until it meets the normal straight baseline in either direction, minimising the error in measured area. In Fig. 5, the upper integral is obtained without slope and bias corrections. The lower integral is a split integral, slope and bias corrected, but joined end to end so that the shortfall is clearly illustrated. When the peak separation/peak width is 50, the recovery is 99.4 and 99.7% at 100 times the peak width. Where one peak is less intense, the effect on the major peak is lessened. Where one peak is significantly broader, the effect on the other peak is lessened since its influence is a much closer approximation to the linear function subtracted.

Integrals should only be split between isolated resonances.

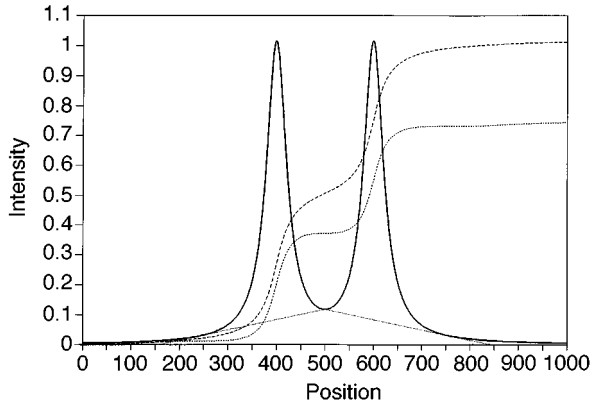

Fig. 5 Convolution of Lorentzian peaks.

Implications for assay determinations

In order to test the validity of the above error analysis, two samples with a wide spread of resonances were selected, *i.e.*, 3-pyridinepropanol and ethyl picolinate. These samples yielded five and six usable resonances, respectively (proton spectra in Figs. 6 and 7, respectively). In both cases the 90° pulse recycle time was 60 s, which is adequate for complete relaxation. The digital resolution in both cases was 0.11 Hz and the typical peak width was 0.8 Hz, giving a negligible ($< 0.01\%$) error from finite data points. Since a 16 bit ADC was used, the error due to 10 and 9 protons respectively, were typically one part in $32767/(2 \times 10) = 0.06\%$. An 11.5 μs excitation pulse was used, which with resonances at 3.5δ and 3.8δ away from the carrier, gave errors of 0.34 and 0.40%, respectively.

A total integral width of 116 Hz was chosen to be as wide as possible to optimise the baseline correction, but not to include the ^{13}C satellites. Always excluding ^{13}C satellites allows consistency, even in crowded regions of the spectrum, since resonances that do not have ^{13}C satellites (and would therefore be overestimated) are usually due to labile protons that are not used for quantitative purposes. One problem with integrating

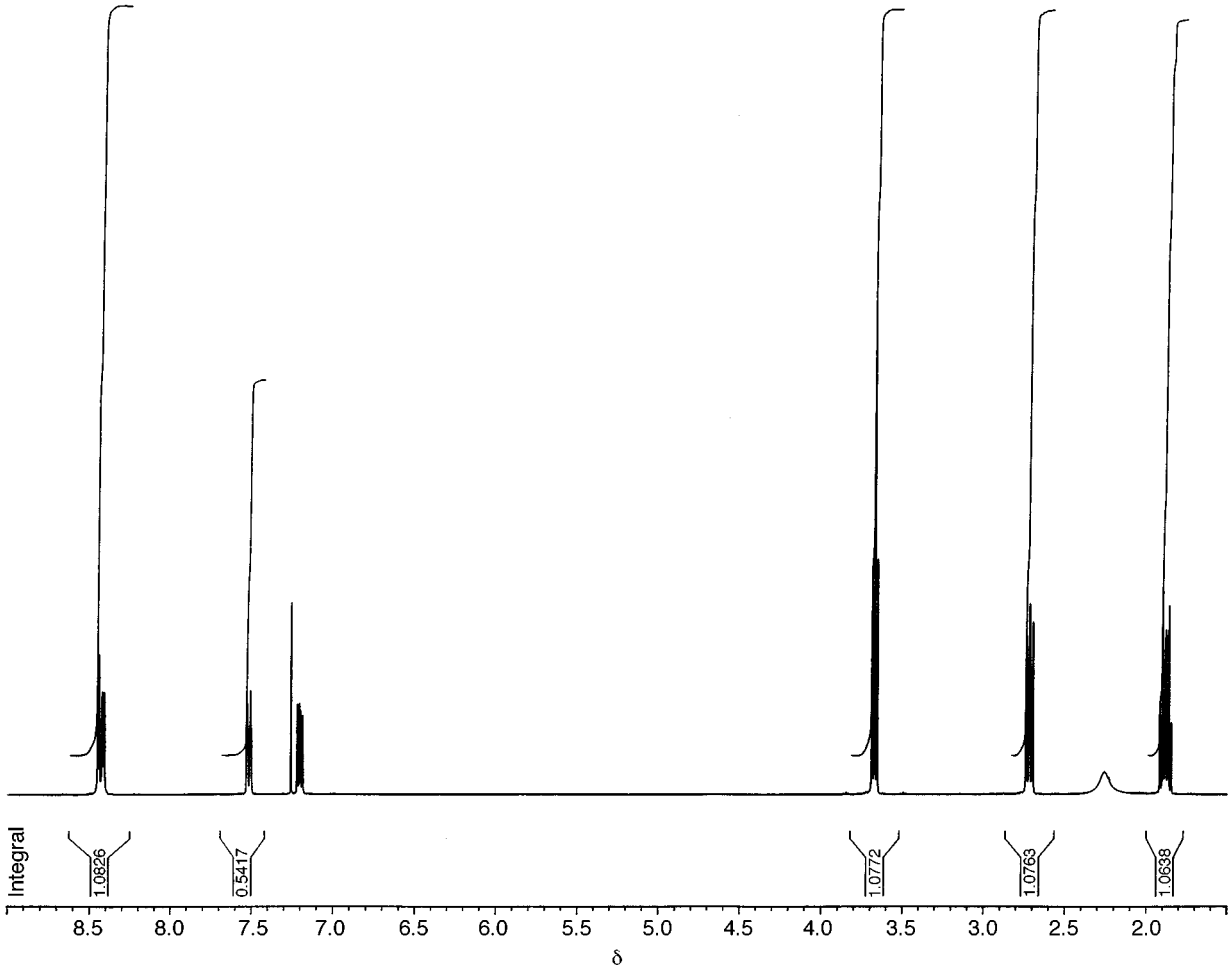

Fig. 6 400.13 MHz NMR spectrum of 3-pyridinepropanol.

inside the ^{13}C satellites, however, is that (disregarding the satellites) the whole resonance may not necessarily be integrated. However, for normal peak widths (say 0.8 Hz) 99.82% of the total area associated with the ^{12}C peak would be included, which is perfectly adequate.

The total number of points integrated for each peak was therefore 1084. Since the manufacturer's software was used in this instance to determine the integrals, separate control of integral and baseline points was not available. In order to extract values of the number of points employed for error prediction purposes, the baseline was arbitrarily allocated to the region of the spectrum outside that encompassing 99% of the peak area (24 times the peak width in each direction for a Lorentzian peak). With an average addition of 20 Hz for multiplet width, this left 269 points of baseline each side. For both compounds, the point-to-point noise of the least intense multiplet was the height of the ^{13}C satellites/2.5. If it is assumed that the rms noise is the peak-to-peak noise/5, and if it is assumed that the peak shape is Lorentzian (hence area = 1.57 × height × linewidth), then eqn. (5) yields an $area_{SD}$ of 0.36%. It is worth noting in passing that if commercial software allowed the whole width to be used to determine baseline, but only the part of the integral encompassing say 99% of the area to determine the area (rather than the two widths being the same), the above error would decrease from 0.36 to 0.20%.

These theoretical sources of error are given in Table 1 and combined as the sum of the percentage variance. The observed standard deviations (between integral ratios and molecular stoichiometry) for 3-pyridinepropanol and ethyl picolinate are 0.65 and 0.91%, respectively (Table 1). The former is closer to the theoretical 0.50% and consequently much more accurate than the 3% commonly assumed by NMR spectroscopists. The observed value for ethyl picolinate is much higher (*cf.*, theoretical value 0.54%) and warrants further investigation. The pyridinyl resonance at 7.46δ is close to the residual $CHCl_3$ resonance at 7.27δ while the methyl resonance at 1.43δ is close to the very broad water resonance. As indicated above, these neighbouring resonances are likely to be underestimated with slope and bias corrections. Both of these resonances in ethyl picolinate are indeed lower than the remainder and if they are arbitrarily removed from the analysis, the standard deviation of the integrals (of this admittedly reduced data set) reduces to 0.55%, exact agreement with our prediction. The agreement may be fortuitous since predictions based on the sum of variances are naive in assuming small linear errors.

Thus far, we have only considered errors within a molecule. As alluded to before, strength can be determined by the addition of a standard and comparison of the respective integrals of the standard and compound in question. The strength is then given by[2,3]

$$\text{strength\%}_{SA} = \text{strength\%}_{ST} \times \frac{I_{SA} H_{ST} M_{SA} A_{ST}}{I_{ST} H_{SA} M_{ST} A_{SA}} \qquad (5)$$

where I is the peak area, H is the stoichiometric amount of that proton, M is the molecular mass, A is the amount weighed into

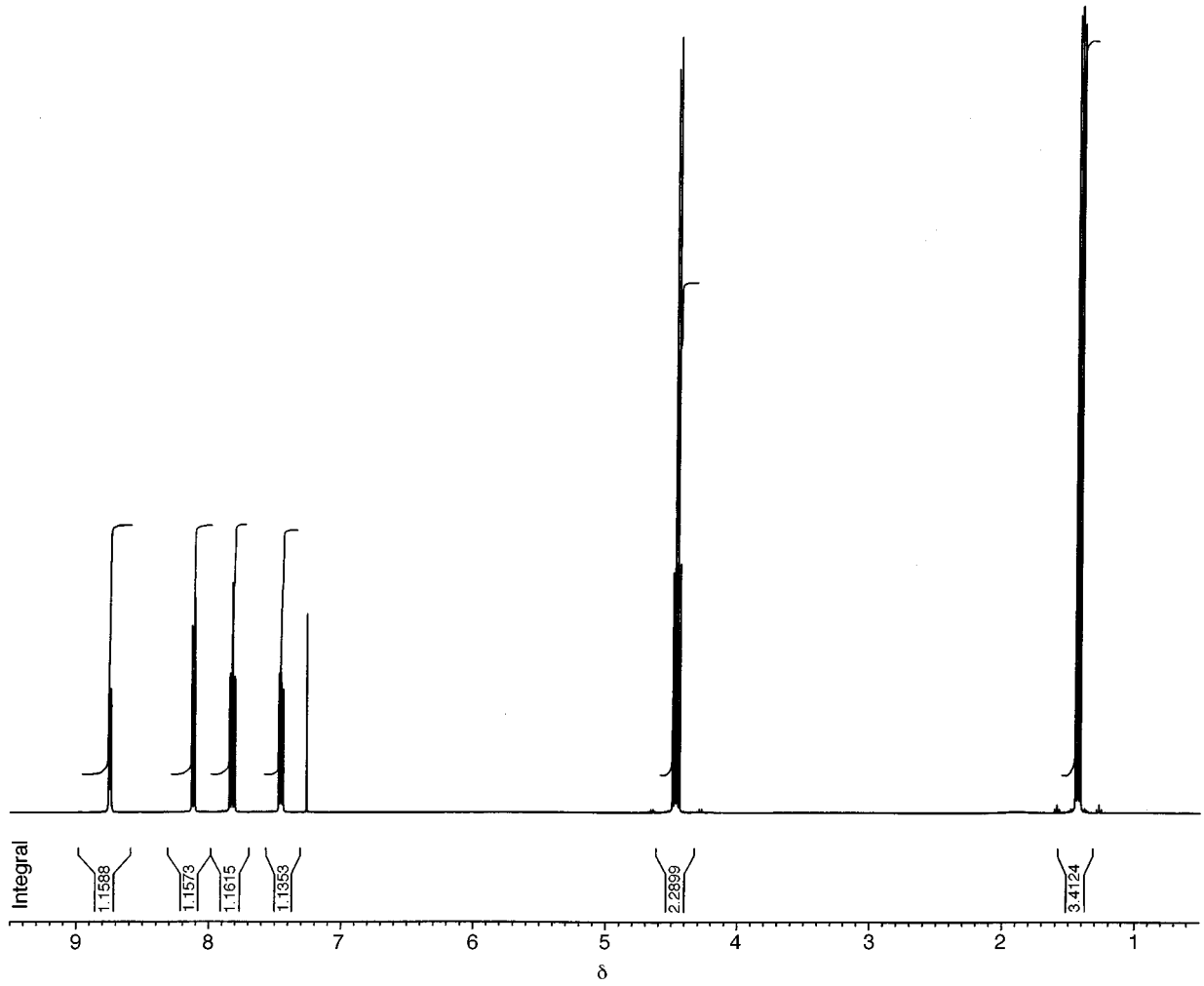

Fig. 7 400.13 MHz NMR spectrum of ethyl picolinate.

solution, SA denotes sample and ST denotes standard. Standards are chosen to be non-volatile and occupy normally vacant regions of the spectrum.

The use of eqn. (5) is not new, but the question so far as this work is concerned is how the sources of error enumerated above come together as the error in strength$_{SA}$%? The answer depends on the magnitude of the variables in eqn. (5) and their respective errors. Before a general case can be examined, typical values of the relevant variables need to be identified.

H is an integer with quantised error. If the assignment is correct, it will have zero error. Molecular mass can be determined to much greater accuracies than the quantitative ability of NMR and can also be regarded as having no error.

The choice of *A* is clearly a compromise: greater masses decrease the percentage weighing error and increase the signal-to-noise ratio, decreasing the error calculated by eqn. (4). However, greater masses increase the probability of incomplete dissolution which would cause errors that are far larger than all the others. In our laboratory, *A* is typically between 10 and 20 mg. The weighings are, however, into a vial (weighing approximately 7 g) in which the dissolution is performed. Repetition work with 1 and 10 g calibrated weights gave a standard deviation of 0.028 and 0.022 mg, respectively, giving a mean of 0.025 mg. Since the weighing of a sample is the difference of two weighings, this standard deviation should be increased by a factor of $\sqrt{2}$ to 0.035 mg, or 0.23%.

The experimental determination of the standard deviation of strength was performed on 2,6-di-*tert*-butyl-4-methylphenol (^{1}H NMR spectrum in Fig. 8). The signal-to-point-to-point noise ratio was typically between approximately 4096 and 2048, giving rise to a peak area standard deviation attributable to noise of 0.08%. The error in strength$_{ST}$% was low but not known exactly and therefore excluded. Otherwise the parameters were as before. The various contributions to error are summarised in Table 1.

In principle, the errors could be increased by a factor of $\sqrt{2}$ in order to account for the two compounds determined in a strength determination. This would yield a combined standard deviation of 0.64%. In practice, the assumption of the additivity of variance may be naive. In order to investigate the combined effect of these errors in a rigorous fashion, 10 000 simulations of various combinations of the parameters in eqn. (5) perturbed by the appropriate errors were performed using a statistical program using the SAS procedures.[13] The effect of dynamic range allowed for a total of 24 protons in the sample and 1 and 12 protons in the two standards. Additionally the following parameters were allowed to vary in magnitude (as opposed to error perturbation): the number of protons in sample between 1 and 3, the amount weighed between 10 and 20 mg, the molecular mass of sample and standard between 200 and 1000, the peak width between 0.6 and 1.5 Hz with an assumed log-normal probability distribution, the signal-to-point-to-point noise ratio between 500 and 2500 and strength$_{ST}$% between 99.1 and 99.6. The 95% confidence limits thus simulated were +0.91 and −0.56%, compared with the ±1.24% obtained by doubling the square root of the percentage sum of variance. The naive method produces a larger error.

The experimental determination of the standard deviation of strength was performed in duplicate, using all three usable resonances, with two different standards and three different operators (36 determinations in all). The standard deviation observed in strength%$_{SA}$ was 0.70%, giving an average strength of 99.5 and 95% confidence limits of ±1.4%, slightly larger than those predicted by the naive method and larger than by SAS procedures. The experimental results are only slightly higher than those obtained for variation across the individual spectra of 3-pyridinepropanol and ethyl picolinate.

The SAS underestimation of the width of the 95% confidence limits for the determination of strength$_{SA}$% for 2,6-*tert*-butyl-4-methylphenol could be due to additional factors not considered thus far: the choice of resonance within a given spectrum, choice of standard, operator, dissolution and sampling. Using an analysis of variance method, individual standard deviations were derived from the experimental data. The error due to choice of resonance (0.29%) measures the same random errors as the variation across the spectra of 3-pyridinepropanol and ethyl picolinate and is in agreement with the theoretical prediction in those cases (0.41%). The error due to choice of standard produced an error of 0.30% and this is likely to be a systematic error in the assigned purity of the standards. The choice of operator error (0%) should measure inconsistencies of approach; in our laboratory these are clearly low. The error produced by replication was 0.56% and not only measures various random errors, but clearly dominates the overall error. It is larger than that predicted theoretically by combining the effects of digital resolution, ADC resolution, noise and weighing (0.25%) and the shortfall must be explained by various random errors including dissolution and sampling.

Table 1 Standard deviations (σ) in NMR quantification

	σ (%)		
Error	3-Pyridine-propanol	Ethyl picolinate	2,6-Di-*tert*-butyl-4-methylphenol
Digital resolution	0.01	0.01	0.01
ADC resolution	0.06	0.06	0.04
Off-resonance effects	0.34	0.40	0.38
Noise	0.36	0.36	0.08
Weighing	—	—	0.23
Total (calculated: $\sqrt{\Sigma\text{variance}}$)	0.50	0.54	0.45
Total (observed)	0.65	0.91/0.55*	0.63†

* With reduced data set (see text). † A combination of two separate determinations (see text).

Scope for improvement

In order to improve quantification further, the integration software could be modified to separate baseline and peak area determinations, as suggested already. This would almost halve the error due to this factor, but the actual gain would also be directly proportional to the noise level. Tangential fitting of integral slope and bias correction would further reduce the error for crowded peaks.

Off-resonance effects could be halved by employing a 45° tip angle, *i.e.*, a decrease from 0.19 to 0.09% in the case of 2,6-di-*tert*-butyl-4-methylphenol. If the repetition time was left unchanged, the 1.4-fold decrease in signal-to-noise ratio would only increase the standard deviation attributable to noise from 0.08 to 0.11%, representing an overall gain in precision.

Finite signal-to-noise ratios can be a significant source of error. More material could be weighed and dissolved in larger volumes of solvent, but the cost of solvent and sample would have to be considered. In principle, more pulses could be acquired, but 64 pulses with a 60 s recycle time, giving a total time of 1 h, have already been used. Significant gains could only be realised at the expense of considerable amounts of instrument time. Higher field instruments would yield a higher signal-to-noise ratio, *e.g.*, a 600 MHz NMR spectrometer would yield approximately 1.7 times the signal-to-noise ratio of the 400 MHz spectrometer used in this work. This would decrease the error attributable to finite signal-to-noise ratio in 3-pyridinepropanol and ethyl picolinate from ±0.36 to ±0.21%. NMR at higher field would also allow more baseline points to be used between resonances and reduce errors in slope and bias correction in 'real' samples, *viz.*, where peak separation cannot

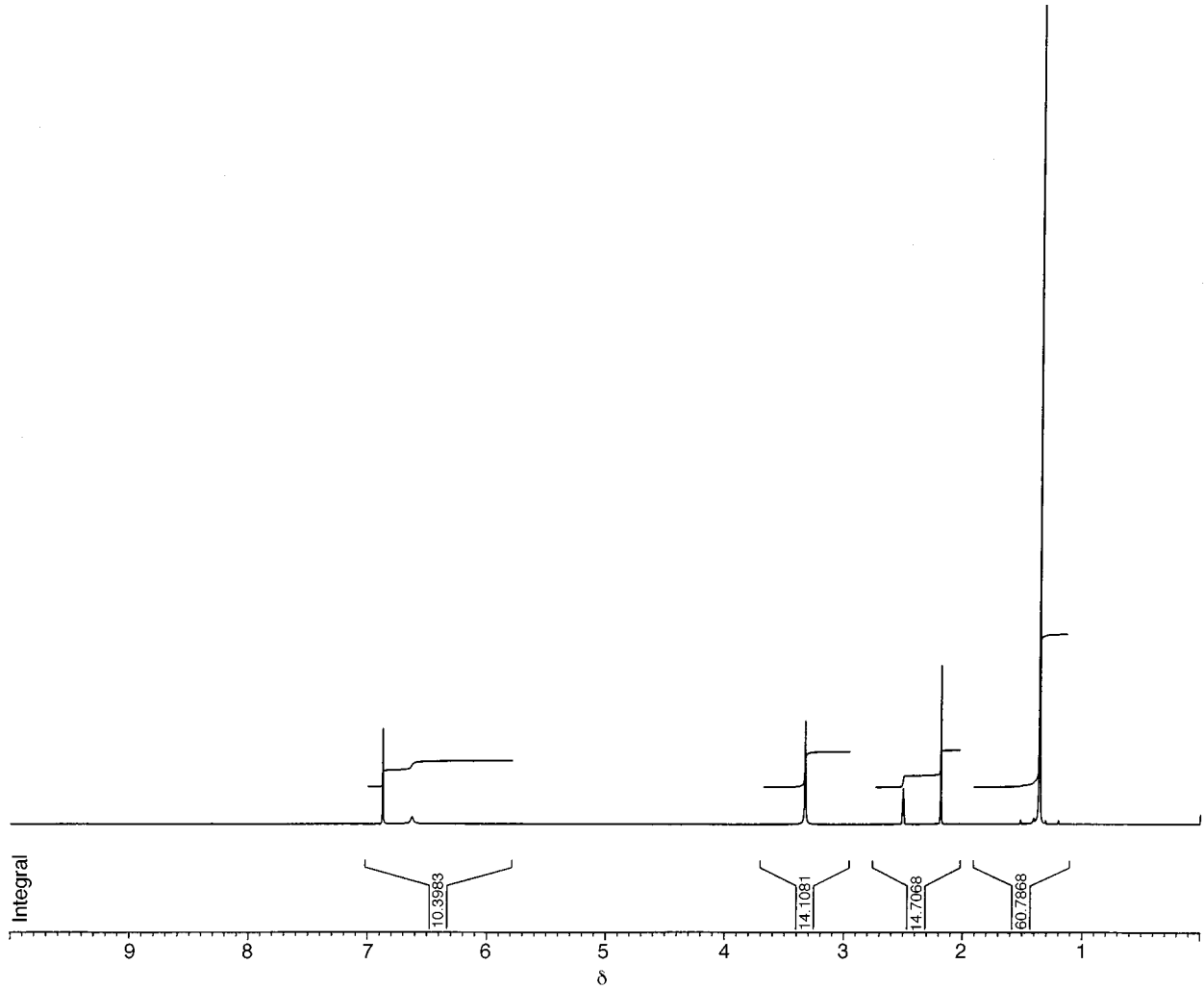

Fig. 8 400.13 MHz NMR spectrum of 2,6-di-*tert*-butyl-4-methylphenol.

be guaranteed as with the carefully chosen compounds in this work.

It is clear that if thought is given to what is done and care is taken in its execution, 95% confidence limits in assay determination using internal standards as low as ±1% are achievable at 400 MHz. At 600 MHz, and with improvements in software, the error could be as low as ±0.6%. It is also clear, however, that if care is not taken, or partially convoluted peaks are encountered, the errors will be much larger.

Conclusions

The commonly accepted quantification limits for NMR are overestimates. By employing the following procedure: (*a*) select a standard with as few resonances as possible; (*b*) weigh approximately 20 mg of sample and standard and ensure complete dissolution in approximately 1 cm^{-3} of solution; vary the weight ratio between sample and standard in proportion to molecular mass if divergent; (*c*) select the highest possible magnetic field for specificity and sensitivity; (*d*) employ an ADC $\geq$ 16 bits; (*e*) maximise the rf power to the practical limit; (*f*) employ a tip angle of 45° if the sensitivity is adequate; (*g*) set the carrier frequency mid-way between the sample and standard resonances; (*h*) employ a repetition time of $\geq$60 s and acquire as many pulses as practical; (*i*) zero-fill by a factor of two; (*j*) maximise vertical expansion during manual phase set; (*k*) integrate just inside ^{13}C satellites; (*l*) if possible, employ a phase routine which maximises points used to determine slope and bias correction but minimises those used to determine area; and (*m*) avoid integrating resonances closer than 100 times their respective linewidths; if unavoidable, use a routine which sets the slope and bias tangentially.

Assay can be determined to approximately ±1% by NMR, sufficiently precise to be used to calibrate other techniques. The factors which give rise to this error are well understood and can be predicted from the relevant experimental parameters and the signal-to-noise ratio attained.

Thanks are due to Gordon Thackray of ICI C&P for his help with some of the earlier statistical work, and Ian Jones and Rob Horton of ZENECA Pharmaceuticals for their help in the strength reproducibility work.

References

1 Attalia, A., and Bowman, R. C., *Anal. Chem. Symp. Ser., Anal. Spectrosc.*, 1984, **19**, 363.
2 Phillips, J. S., and Leary, J. L., *J. Chem. Educ.*, 1986, **63**, 545.
3 Peterson, J., *J. Chem. Educ.*, 1992, **69**, 843.
4 Griffiths, G., *Anal. Proc.*, 1984, **21**, 502.
5 Ismail, I. M., and Wilson, I. D., *Methodol. Surv. Biochem. Anal.*, 1986, **16**, 337.
6 Pieters, L. A. C., and Vlietinck, A. J., *J. Pharm. Biomed. Anal.*, 1986, **7**, 1405.
7 Lindgren, B., and Martin, J. R., *Pharmeuropa*, 1993, **5**, 51.

8 Lindon, J. C., and Ferrige, A. G., *Prog. Nucl. Magn. Reson. Spectrosc.*, 1980, **14**, 27.
9 Derome, A. E., *Modern NMR Techniques for Chemistry Research*, Pergamon Press, Oxford, 1987, p. 71.
10 Ernst, R. R., and Anderson, W. A., *Rev. Sci. Instrum.*, 1966, **37**, 93.
11 Cookson, D. J., and Smith, B. E., *Anal. Chem.*, 1982, **54**, 2593.
12 G. Thackray, personal communication.
13 *SAS Version 6.10*, SAS Institute, Cary, NC, 1996.

Paper 8/00625C
Received January 22, 1998
Accepted February 27, 1998