MAGNETIC RESONANCE IN CHEMISTRY, VOL. 35, 505–519 (1997)

Investigation of the Advantages and Limitations of Forward Linear Prediction for Processing 2D Data Sets

William F. Reynolds,[1]* Margaret Yu,[1] Raul G. Enriquez[2] and Ismael Leon[3]
[1] Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 3H6, Canada
[2] Instituto de Quimica, Universidad Nacional Autonoma de Mexico, Circuito Exterior, Ciudad Universitaria, Coyoacan, Mexico, D.F., 04510, Mexico
[3] Centro de Investigaciones in Quimica, Universidad Autonoma del Estado de Morelos, Cuernavaca, Morelos, Mexico

Although the potential advantages of f_1 forward linear prediction for the processing of multi-dimensional NMR spectra are well established, this method is surprisingly little used for 2D spectra used for organic structure determination. A detailed investigation of the advantages and limitations of f_1 forward linear prediction for this purpose is reported. This is a reliable technique which is particularly useful for ^{1}H-detected ^{13}C–^{1}H shift correlation spectra, allowing up to 16-fold linear prediction of the ^{13}C axis of HSQC spectra. In general, the use of linear prediction allows one to obtain comparable 2D spectra in one quarter of the time or double the sensitivity in a comparable time relative to similar spectra without linear prediction. The one exception is the absolute value COSY spectrum, where linear prediction beyond a factor of two gives poor results. Linear prediction is generally superior to zero filling as a time-saving technique, although the difference between the two approaches disappears as the f_1 data point resolution approaches the natural linewidth. By contrast, f_2 forward linear prediction is not recommended.

Magn. Reson. Chem. **35**, 505–519 (1997) No. of Figures: 10 No. of Tables: 2 No of References: 29

Keywords: 2D NMR; linear prediction; zero filling

Received 12 February 1997; revised 24 March 1997; accepted 24 March 1997

INTRODUCTION

The potential utility of forward linear prediction for processing multidimensional NMR data sets has long been recognized.[1,2] This technique is particularly useful as a means of minimizing the acquired data matrix and consequently the total acquisition time for 3D and 4D NMR spectra of proteins.[3] By contrast, as pointed out in a recent paper on the use of forward linear prediction for 2D NOESY spectra,[4] this technique has been surprisingly little used for 2D spectra of smaller organic molecules such as natural products. For example, checking the experimental sections of papers on organic structure determination by 2D NMR in the 1995 and 1996 volumes of this journal revealed only eight papers (out of approximately 100) which explicitly mentioned the use of linear prediction,[4–11] three of which came from this research group.[5,6,8] We find this failure to make more extensive use of linear prediction in 2D NMR very surprising since we have consistently found it to be a useful and reliable method for improving the resolution and sensitivity of 2D spectra. This is particularly true for ^{1}H-detected ^{1}H–^{13}C shift correlation spectra[12] where one is faced with the problem of digitizing a wide ^{13}C spectral window. Reference 4, while providing quantitative information concerning the accuracy of NOESY cross-peak areas with linear prediction, offered only qualitative comments concerning the resolution and sensitivity advantages of this approach. Consequently, we decided to carry out a quantitative evaluation of the advantages and limitations of forward linear prediction for other common 2D experiments used in organic structure elucidation. In addition, we wished to compare the relative merits of linear prediction and the more commonly used method of zero filling. Since our goal was to encourage the more widespread use of linear prediction, we used the standard linear prediction software package provided by the spectrometer manufacturer in order to evaluate the advantages which would be routinely available to non-expert users.

* *Correspondence to*: W. F. Reynolds.
Contract grant sponsor: NSERCC.
Contract grant sponsor: DGAPA (UNAM).

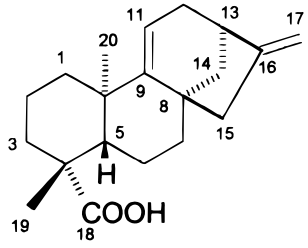
RESULTS AND DISCUSSION

Forward f_1 linear prediction of HSQC data sets

As a first basic test of the reliability of linear prediction, three HSQC[13] data sets were acquired for the aliphatic

 CCC 0749–1581/97/080505–15 $17.50

proton and carbon region of our standard test molecule, kauradienoic acid (**1**).[14] HSQC was chosen rather than HMQC[15] because the limiting $f_1(^{13}C)$ resolution is determined by the natural ^{13}C linewidth for HSQC but by the ^1H multiplet width for HMQC (since the ^1H multiplet structure appears along both f_1 and f_2 with the latter sequence[12,16]). The sample contained 6 mg of **1** dissolved in *ca.* 1 ml of $CDCl_3$. All three data sets used an f_2 spectral window of 1770 Hz, an f_1 spectral window of 6300 Hz and 256 f_2 data points, zero filled to 512. In the first experiment (see Table 1), we collected 16 transients per time increment and 256 time increments (total acquisition time 5 h 42 min) and linearly predicted the f_1 data to 1024, with zero filling to 2048. Experiment 2 had four transients per time increment and 1024 time increments (total acquisition time 5 h 43 min) with f_1 zero filling to 2048 and no linear prediction. Experiment 3 had four transients per time increment and 256 increments (total acquisition time 1 h 26 min), linearly predicted to 1024 with zero filling to 2048. All other acquisition and processing parameters were identical for the three experiments (see Experimental). All data were acquired in the hypercomplex mode to generate phase-sensitive spectra.



1

Our assumption was that, provided that the raw data had a sufficient signal-to-noise ratio to allow accurate prediction, the first experiment should show twice the signal-to-noise ratio of the second experiment in about the same overall time while the third should show a comparable signal-to-noise ratio to the second in one quarter of the time. The signal-to-noise ratio was measured for ^1H cross-sections through individual $^{13}CH_2$ peaks. The latter were chosen since they have the lowest signal-to-noise ratio of the different CH_n groups in **1**. Noise levels were measured from ^1H spectral regions with no peaks to avoid any contributions from t_1 ridges. These measurements showed that the average signal-to-noise ratio for CH_2 groups was 237:1 in the first experiment, 117:1 in the second and 116:1 in the third, i.e. almost exactly the ratios predicted above. Two typical CH_2 cross-sections from experiments 1 and 2 are displayed in Fig. 1, clearly illustrating the higher sensitivity for the experiment with linear prediction.

Contour plots for the three experiments, illustrated in Fig. 2, show an additional interesting and unexpected feature. It appears that the experiments using linear prediction, and particularly the first experiment, show weaker t_1 ridges from the CH_3 singlets. There are many sources of instability which may lead to t_1 ridges[17] with the actual origins being dependent on the individual spectrometer and its environment.[18] This observation,

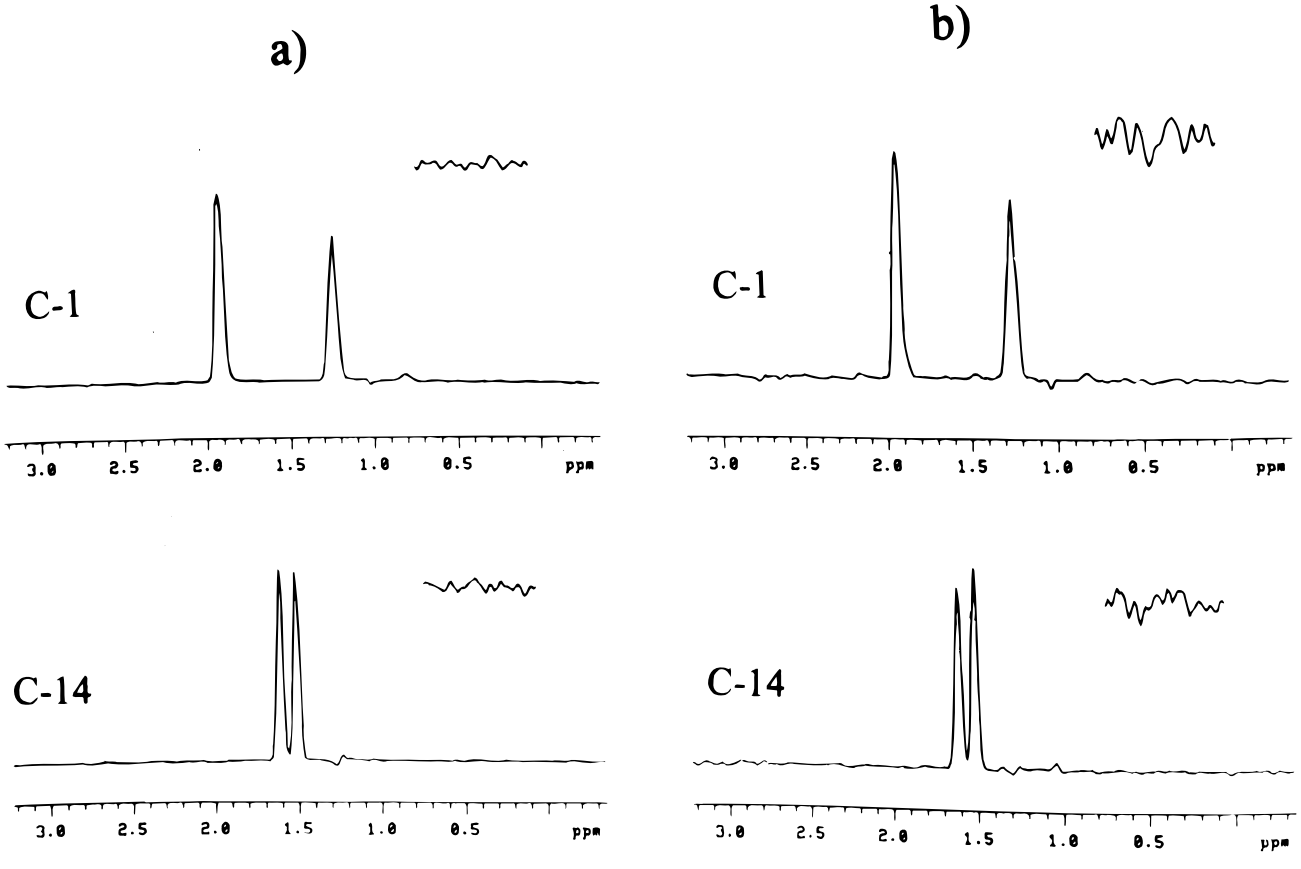


Figure 1. Typical cross-sectional spectra for CH_2 groups of **1** obtained in (a) experiment 1 and (b) experiment 2. Spectra for a given carbon are plotted at the same vertical scale with noise insets ×10.

Table 1. Acquisition and processing parameters for various 2D experiments involving kauradienoic acid (1) discussed in the text

Experiment[a]	Segment[b]	Sample[c] mg	NT[d]	SW2[e]	SW1[f]	NP[g]	TP2[h]	NI[i]	LP1[j]	TP1[k]	LP2[l]
1	HSQC	6	16	1770	6300	256	512	256	1024	2048	—
2	HSQC	6	4	1770	6300	256	512	1024	—	2048	—
3	HSQC	6	4	1770	6300	256	512	256	1024	2048	—
4	HSQC	6	16	3500	16 000	512	1024	256	1024	2048	—
5	HSQC	6	4	3500	16 000	512	1024	1024	—	2048	—
5a	HSQC	6	4	3500	16 000	512	2048	1024	—	2048	1024
6	HSQC	6	4	3500	16 000	512	1024	256	1024	2048	—
6a	HSQC	6	4	3500	16 000	512	1024	256	4096	8192	—
6b	HSQC	6	4	3500	16 000	512	1024	128	1024	2048	—
6c	HSQC	6	4	3500	16 000	512	1024	64	1024	2048	—
6d	HSQC	6	4	3500	16 000	512	1024	128	512	1024	—
6e	HSQC	6	4	3500	16 000	512	1024	256	—	2048	—
6f	HSQC	6	4	3500	16 000	512	1024	128	—	2048	—
6g	HSQC	6	4	3500	16 000	512	1024	64	—	2048	—
6h	HSQC	6	4	3500	16 000	512	1024	256	1024	8192	—
6i	HSQC	6	4	3500	16 000	512	2048	256	1024	2048	1024
6j	HSQC	6	4	3500	16 000	512	2048	256	1024	2048	—
7	HSQC	1.5	16	3500	16 000	512	1024	256	1024	2048	—
8	HSQC	1.5	4	3500	16 000	512	1024	1024	—	2048	—
9	HSQC	1.5	4	3500	16 000	512	1024	256	1024	2048	—
9a	HSQC	1.5	4	3500	16 000	512	1024	256	—	2048	—
9b	HSQC	1.5	4	3500	16 000	512	1024	828	—	2048	—
9c	HSQC	1.5	4	3500	16 000	512	1024	64	—	2048	—
9d	HSQC	1.5	4	3500	16 000	512	1024	256	4096	8192	—
9e	HSQC	1.5	4	3500	16 000	512	1024	256	1024	8192	—
10	HMQC	6	16	3500	16 000	512	1024	256	256	2048	—
11	HMQC	6	4	3500	16 000	512	1024	1024	—	2048	—
12	DQCOSY	1.5	8	1770	1770	2048	4096	256	2048	4096	—
12a	DQCOSY	1.5	8	1770	1770	2048	4096	256	—	4096	—
12b	DQCOSY	1.5	8	1770	1770	2048	4096	256	512	4096	—
13	DQCOSY	1.5	8	1770	1770	2048	4096	2048	—	4096	—
14	COSY	1.5	16	3500	3500	1024	2048	1024	1024	2048	—
14a	COSY	1.5	16	3500	3500	1024	2048	256	1024	2048	—
14b	COSY	1.5	16	3500	3500	1024	2048	512	1024	2048	—
14c	COSY	1.5	16	3500	3500	1024	2048	256	—	2048	—

[a] Experiments labelled only with a number correspond to the original experiment described in the text while those with a number followed by a letter, e.g. 5a, correspond to a reprocessed data set.
[b] Sequence used in experiment. The COSY experiment was run in the absolute value mode and all others in the phase-sensitive mode, using the hypercomplex method.
[c] Samples were either 6 or 1.5 mg of **1** dissolved in *ca*. 1 ml of $CDCl_3$.
[d] Number of transients collected per time increment.
[e] f_2 spectral window.
[f] f_1 spectral window.
[g] Number of f_2 data points.
[h] Total number of f_2 points after zero filling.
[i] Number of f_1 time increments used in processing data set.
[j] Extent of linear prediction along f_1 axis. Where no number appears, no linear prediction was carried out.
[k] Total number of f_1 points after zero filling.
[l] Extent of linear prediction along the f_2 axis. Where no number appears, no linear prediction was carried out.

which we have reproduced in repeated experiments, represents another advantage of linear prediction.

While the experiments described above clearly demonstrate the sensitivity and resolution advantages of forward f_1 linear prediction, we carried out a number of additional experiments designed to probe possible limitations. First, the three experiments described above were repeated (as experiments 4–6, respectively) using ^1H and ^{13}C spectral windows (3500 Hz and 16 000 Hz, respectively) which included all protons and protonated carbons for **1**. The number of ^1H data points was doubled to 512 (with zero filling to 1024), with all other parameters kept identical with those listed previously (see Table 1). The new experiments were run for two reasons. The first was to confirm that the conclusions described above still held when using wider spectral windows and the second was to allow us to test the usefulness of more extensive linear prediction. As noted previously,[12] there is no sensitivity or resolution advantage in linear prediction (or straight data acquisition) which provides data point resolution which is less than the f_1 linewidths. The use of a wider spectral window allowed us to test more extensive f_1 linear prediction without running into this limitation.

The observed average signal-to-noise ratios were 200:1, 108:1 and 122:1 for experiments 4, 5 and 6,

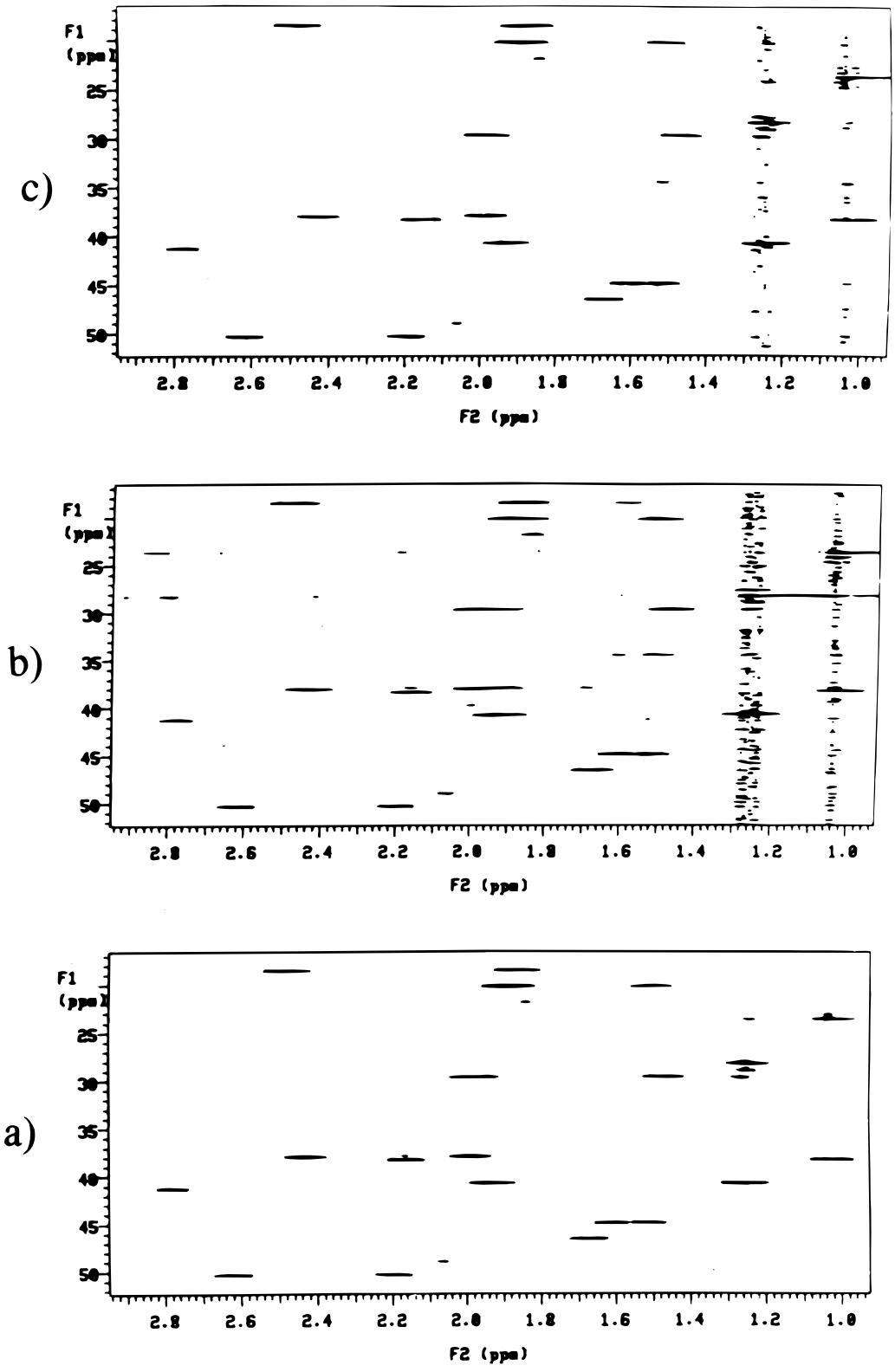


Figure 2. Contour plots for the full HSQC spectra of **1** obtained in the first set of experiments with (a) experiment 1, (b) experiment 2 and (c) experiment 3. All spectra are plotted with the same vertical scale and threshold. Note the increased t_1 ridges for experiments 2 and 3.

respectively. These are similar to those noted above, confirming that the advantages of linear prediction hold, even with wider spectral windows. Experiment 6 was then reprocessed with linear prediction from 256 to 4096 (with zero filling to 8192), i.e. 16-fold linear prediction. This experiment was chosen because the weaker signal-to-noise ratio compared with experiment 4 should provide a more rigorous test of extended linear prediction. To our surprise, the spectrum with 16-fold linear prediction (experiment 6a) had an average signal-

to-noise ratio for CH_2 cross-sections which was almost twice that for the spectrum with fourfold linear prediction (239:1 compared with 122:1). As observed previously,[12] this was mainly due to the increase in peak height associated with the linewidth narrowing in f_1. This line narrowing arises because extended linear prediction provides access to a larger fraction of the interferogram envelope, providing better data point resolution and allowing the use of less extreme apodization. However, there are limits to the gains which can be realized in this manner (see later). Typical CH_2 cross-sections are illustrated in Fig. 3 and contour plots for one region of the 2D spectra are shown in Fig. 4. Figure 4(b) is plotted at half the vertical scale of Fig. 4(a) to allow for the approximate twofold increase in signal intensities. The narrower lines are apparent in Fig. 4(b). Finally, in order to confirm the accuracy of extended linear prediction, ^{13}C chemical shifts were estimated from the mid-points of ^{13}C signals in the contour plot for experiment 6a. The estimated chemical shifts are listed in Table 2, along with the ^{13}C chemical shifts determined from a 1D ^{13}C spectrum for the same sample. The maximum difference is 0.03 ppm, demonstrating that the extensive linear prediction has not resulted in any significant errors in peak position.

Table 2. Comparison of ^{13}C chemical shifts determined from a 2D HSQC spectrum of 1 with 16-fold linear prediction and from a regular 1D ^{13}C spectrum

Carbon	δ_c [b]	δ_c [c]
1	40.74	40.75
2	20.12	20.15
3	38.25	38.28
6	18.47	18.46
7	29.64	29.66
11	111.48	111.49
12	37.92	37.92
13	41.24	41.24
14	44.93	44.94
15	50.30	50.31
17	105.49	105.47
19	28.22	28.24
20	23.60	23.61

[a] In $CDCl_3$ measured relative to internal tetramethylsilane.
[b] From 2D HSQC spectrum (experiment 6a).
[c] From 1D ^{13}C spectrum.

We next investigated the minimum number of increments necessary for reliable linear prediction. This was done by reprocessing experiment 6 using only part of the time increments. Three spectra (experiments 6, 6b and 6c, respectively) were compared (1) using all 256 increments and linearly predicting to 1024, (2) using the first 128 time increments and linearly predicting to 1024 and (3) using the first 64 time increments and linearly predicting to 1024. Assuming that linear prediction still works well with a smaller number of time increments, then similar quality spectra would be expected in each case, as was observed above when comparing experiments 5 and 6. In fact, this is not the case. Contour

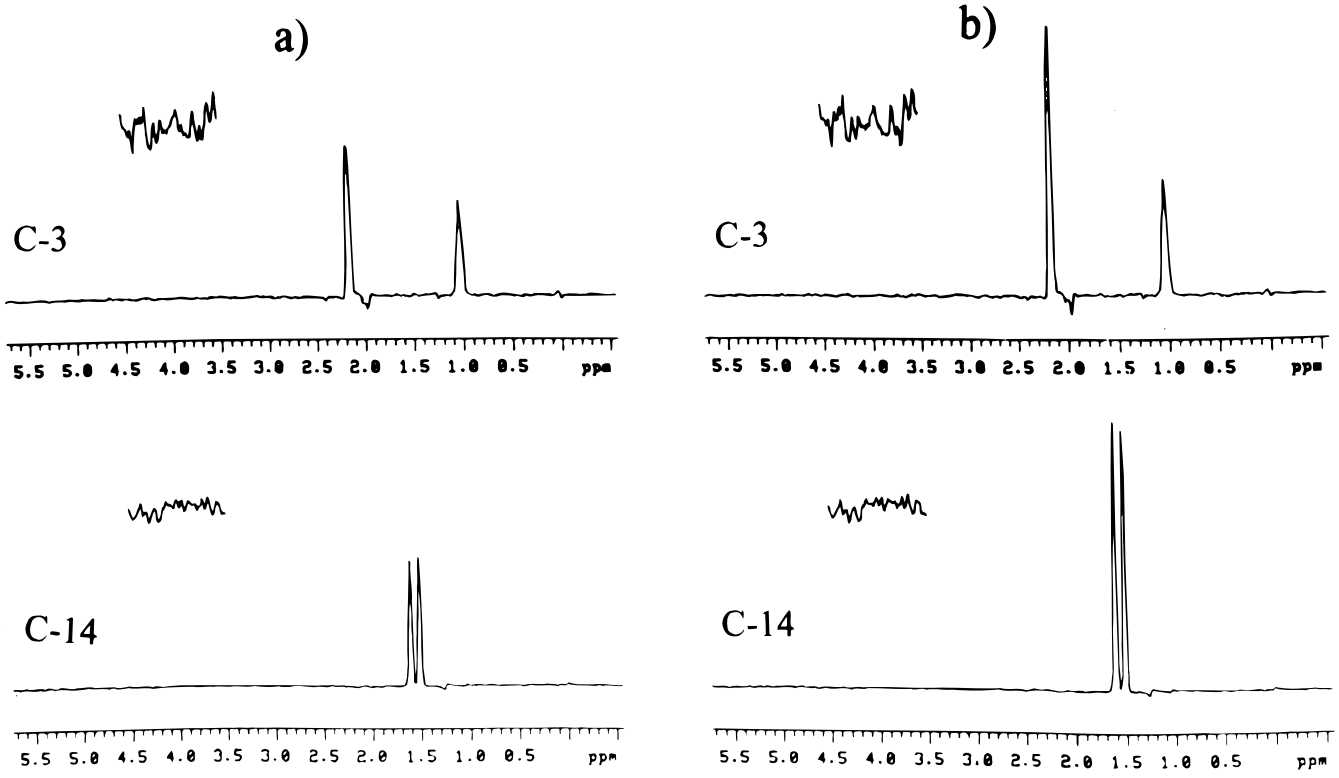


Figure 3. Cross-sectional spectra for two CH_2 groups of **1** obtained in the second set of experiments with (a) experiment 6 with linear prediction to 1024 and (b) experiment 6a with linear prediction to 4906. Spectra for the same carbon are plotted at the same vertical scale with noise insets ×20. The full spectra were obtained with a ^{13}C spectral window of 16 000 Hz and an ^{1}H spectral window of 3500 Hz.

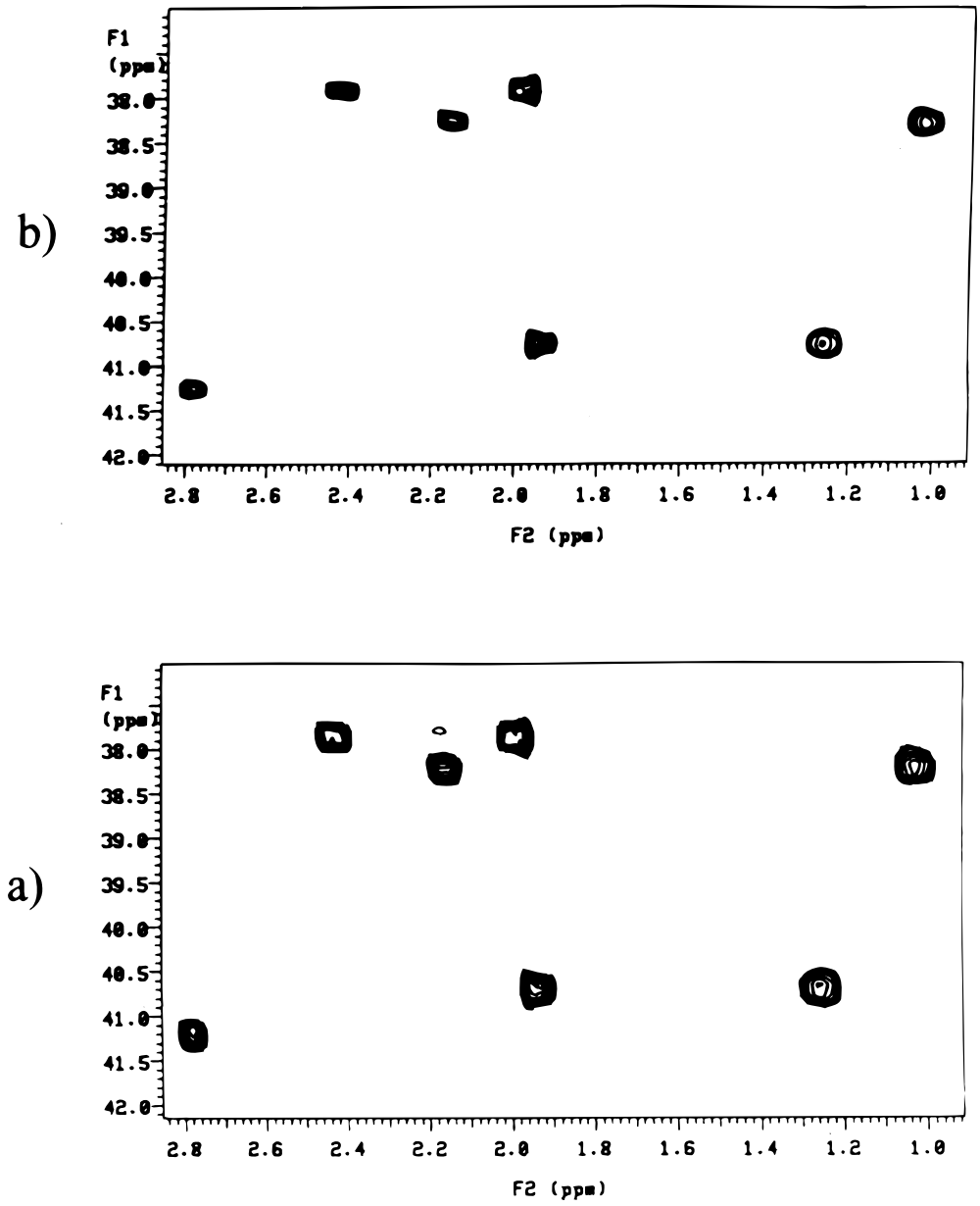


Figure 4. Contour plots for a region of the HSQC spectra of **1** obtained with (a) experiment 6 with linear prediction to 1024 and (b) experiment 6a with linear prediction to 4096; (b) is plotted at the same threshold as (a) but with half of the vertical scale to allow for the increased intensity shown in Fig. 3(b). ^{13}C peaks from bottom to top of each spectrum correspond to C-13, C-1, C-3 and C-12.

plots for the three spectra are illustrated in Fig. 5. They show a significant deterioration in quality from 4- to 8- to 16-fold linear prediction. Reprocessing the data from experiment 6 using the fist 128 time increments combined with linear prediction to 512 instead of 1024 gave a spectrum (experiment 6d) which showed no significant distortions but with a twofold increase in f_1 linewidth. Thus, for this particular series of experiments, it appears that the collection of 256 increments with linear prediction to 1024 (or further if necessary) provides the optimum combination of resolution, sensitivity and quality of spectra. Fewer increments lead to imperfect spectral prediction or loss of resolution while the use of a greater number of increments offers no advantages over linear prediction. In principle, the optimum number of increments should depend on the particular molecule–experiment combination, being determined by factors such as natural f_1 linewidth, f_1 spectral window and the number of peaks being detected. However, in practice, we have observed that 256 increments usually represents a good choice for HSQC, HMQC and HMBC experiments with wide ($>15\,000$ Hz) f_1 spectral windows when acquiring spectra for natural products and other organic compounds up to molecular mass 2000.

Finally, to check whether these conclusions would still be valid for more dilute solutions, a new sample containing 1.5 mg of **1** (a fourfold dilution) was prepared and three new data sets (experiments 7–9) were obtained with parameters identical with those in experiments 4–6. In this case, experiment 7 showed an average CH_2 cross-section signal-to-noise ratio of 59:1, experiment 8 27:1 and experiment 9 28:1. Since the signal-to-noise ratios are again essentially identical with

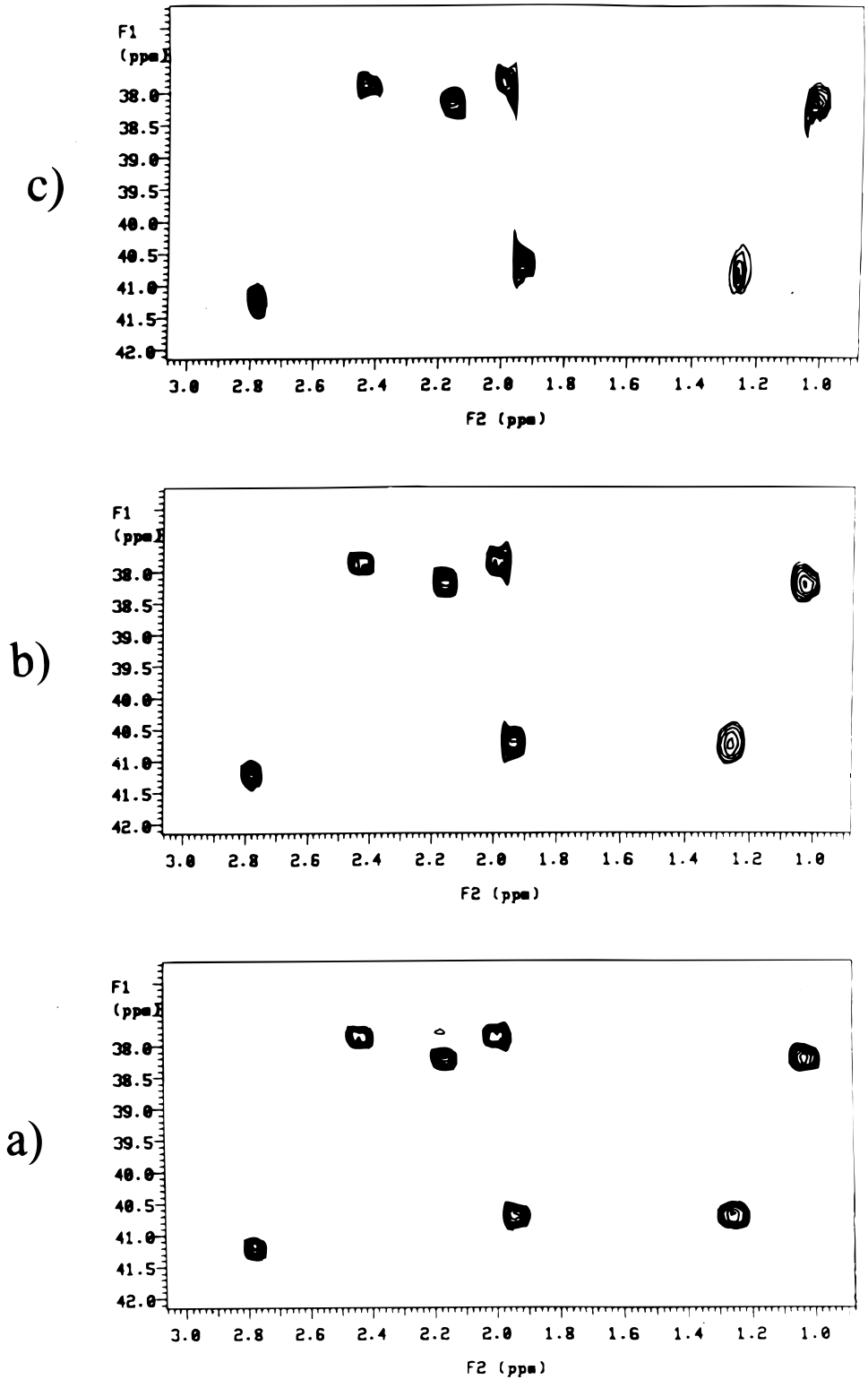


Figure 5. Contour plots for (a) experiment 6, (b) experiment 6b and (c) experiment 6c. All spectra are plotted at the same vertical scale and threshold.

the expected values, it appears that fourfold linear prediction is still reliable, even for a relatively dilute solution.

The results presented above demonstrate that linear prediction provides a very robust method for minimizing the time and maximizing the sensitivity of 2D HSQC spectra. It might be argued that our results are biased by the use of a test molecule with relatively low molecular mass (300) and that the shorter ^{1}H and ^{13}C relaxation rates of larger molecules might lead to less satisfactory results with linear prediction. However, based on earlier results, we are confident that this is not true. For example, Babcock *et al.*[4] reported increased resolution and sensitivity with linear prediction for

NOESY spectra of a nucleotide decamer. Similarly, we have successfully used linear prediction for HMQC, HMBC, COSY, TOCSY and ROESY spectra for a number of compounds with molecular masses in excess of 1000.[19,20] For example, Fig. 6 shows three expansions of crowded regions of a coupled HSQC spectrum obtained for a saponin currently under investigation in our laboratory.[20] It consists of a triterpene with seven saccharide units. The spectrum, obtained with 1024 f_2 data points and 256 time increments, with linear prediction to 2048, covers the sugar region (f_1 spectral window 6300 Hz, f_2 spectral window 1770 Hz). Even with such a large data set for a large natural product (molecular mass >1600), individual chemical shifts can

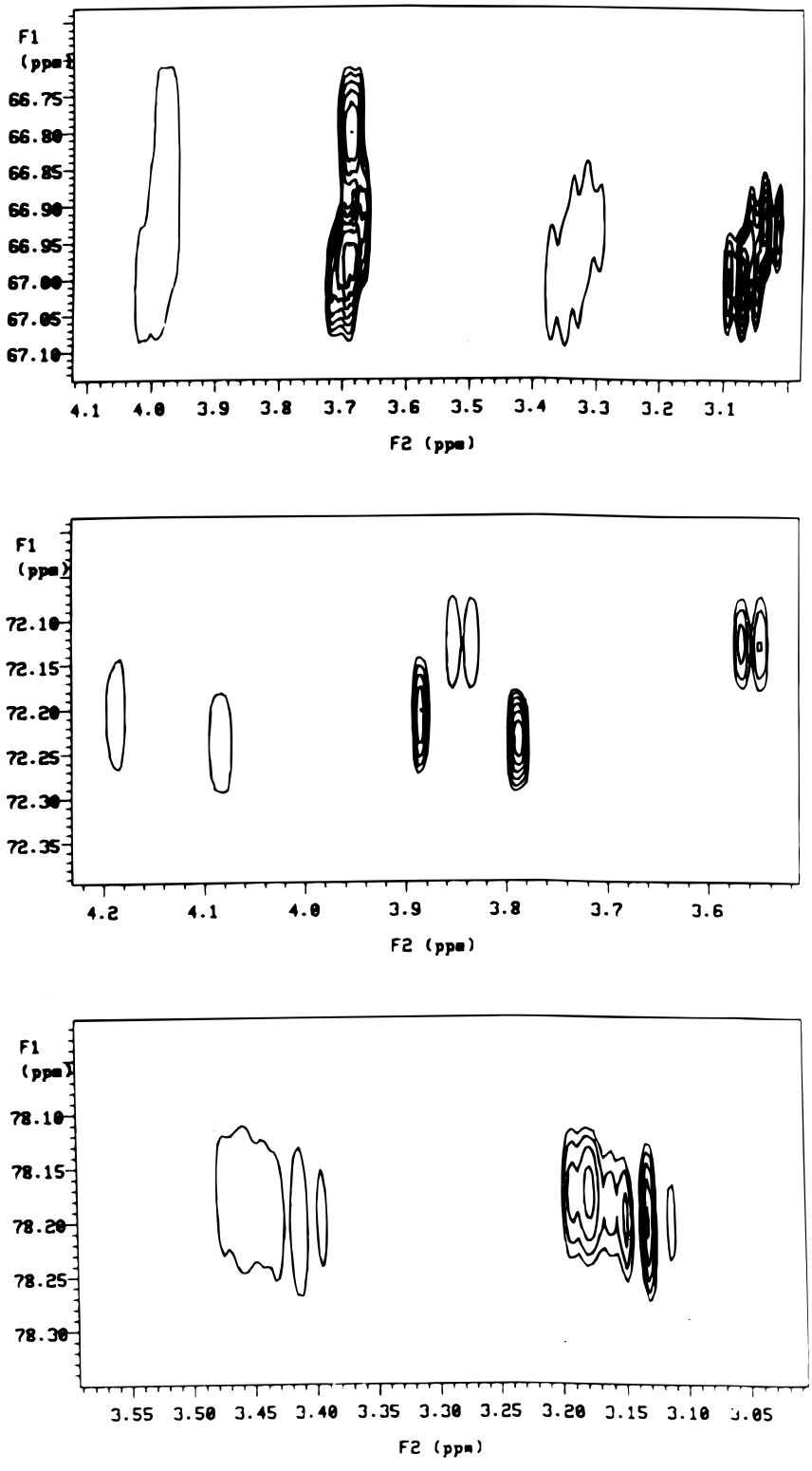


Figure 6. Expanded contours plots illustrating spectrally crowded regions for a coupled HSQC spectrum of the saponin described in the text. Each CH peak appears as an antiphase doublet with splitting = $^1J_{CH}$. Filled contours are positive and open contours are negative.

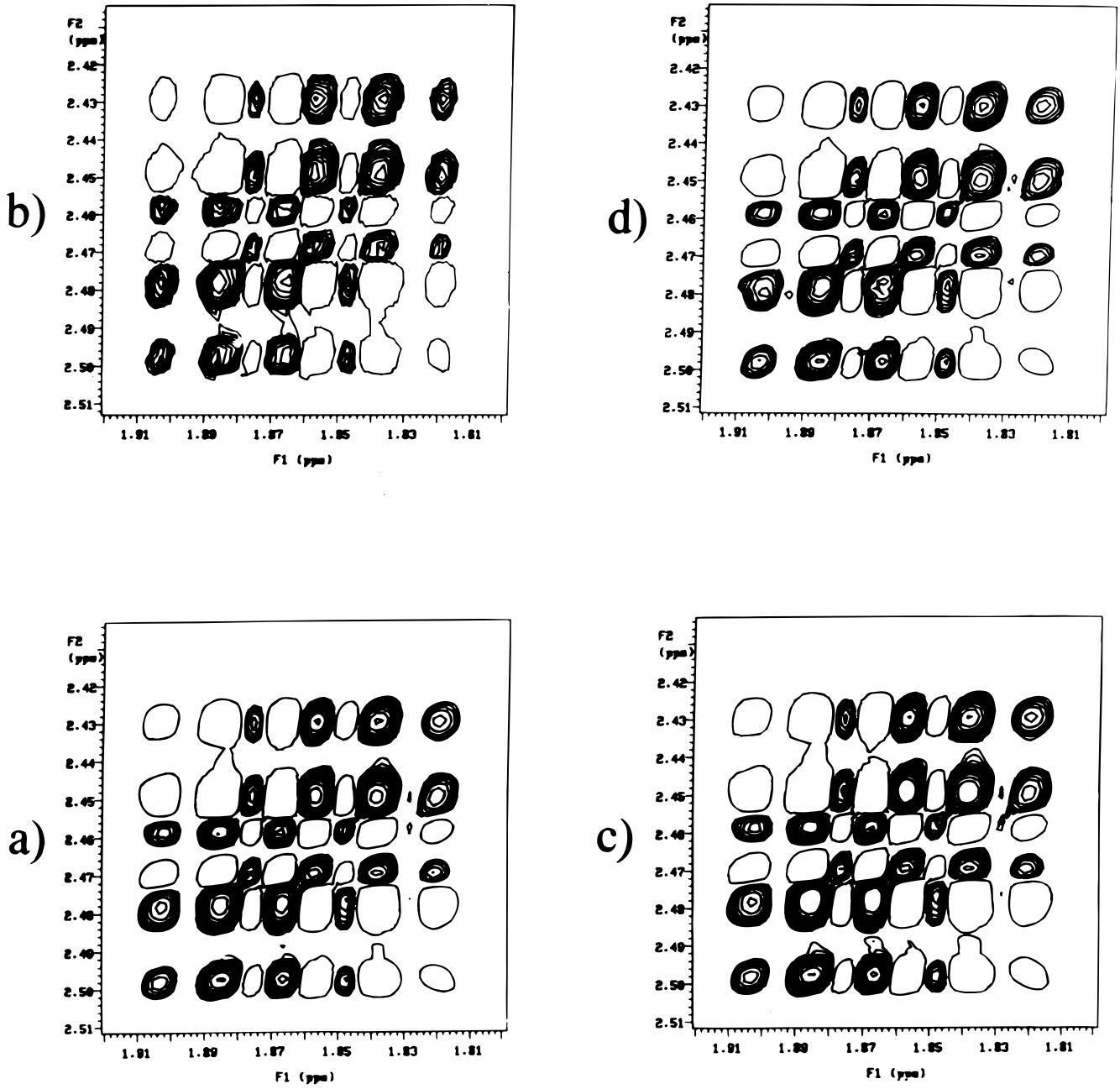


Figure 7. Typical DQ COSY cross peak (corresponding to H-6_α/H-6_β) for **1** with (a) experiment 12, (b) experiment 13, (c) experiment 12a and (d) experiment 12b. Note the close similarity of the spectra in (a) and (d), corresponding to linear prediction to 2048 and 512, respectively, with zero filling to 4096 in each case.

still be estimated for crowded spectral regions with the aid of linear prediction.

Consideration of other 2D experiments

Unlike HSQC (and hybrid sequences such as HSQC TOCSY), all of the other commonly used 2D experiments for organic structure elucidation show ^{1}H multiplet structure along f_1. Consequently, as we have pointed out for HMQC,[12] these sequences generally do not show increased sensitivity with increasing linear prediction beyond data point resolution comparable to multiplet width. Nevertheless, our experience has been that, for a given desired data point resolution, it is still better to use linear prediction in order to save time or improve sensitivity. To illustrate this, we obtained two HMQC spectra for **1**, using the full spectral windows and respectively using 16 transients per increment with 256 increments predicted to 1024 and using four transients per increment with 1024 increments and no linear prediction (experiments 10 and 11). All other parameters were identical with those used for experiments 4–9. The average CH_2 cross-section signal-to-noise ratio was 160:1 for experiment 10 and 77:1 for experiment 11, consistent with the expected twofold difference in sensitivity. These values are *ca.* 70–80% of those obtained for the corresponding HSQC spectra on the same sample (see above), reflecting the detrimental impact of the increased HMQC f_1 multiplet width on signal-to-

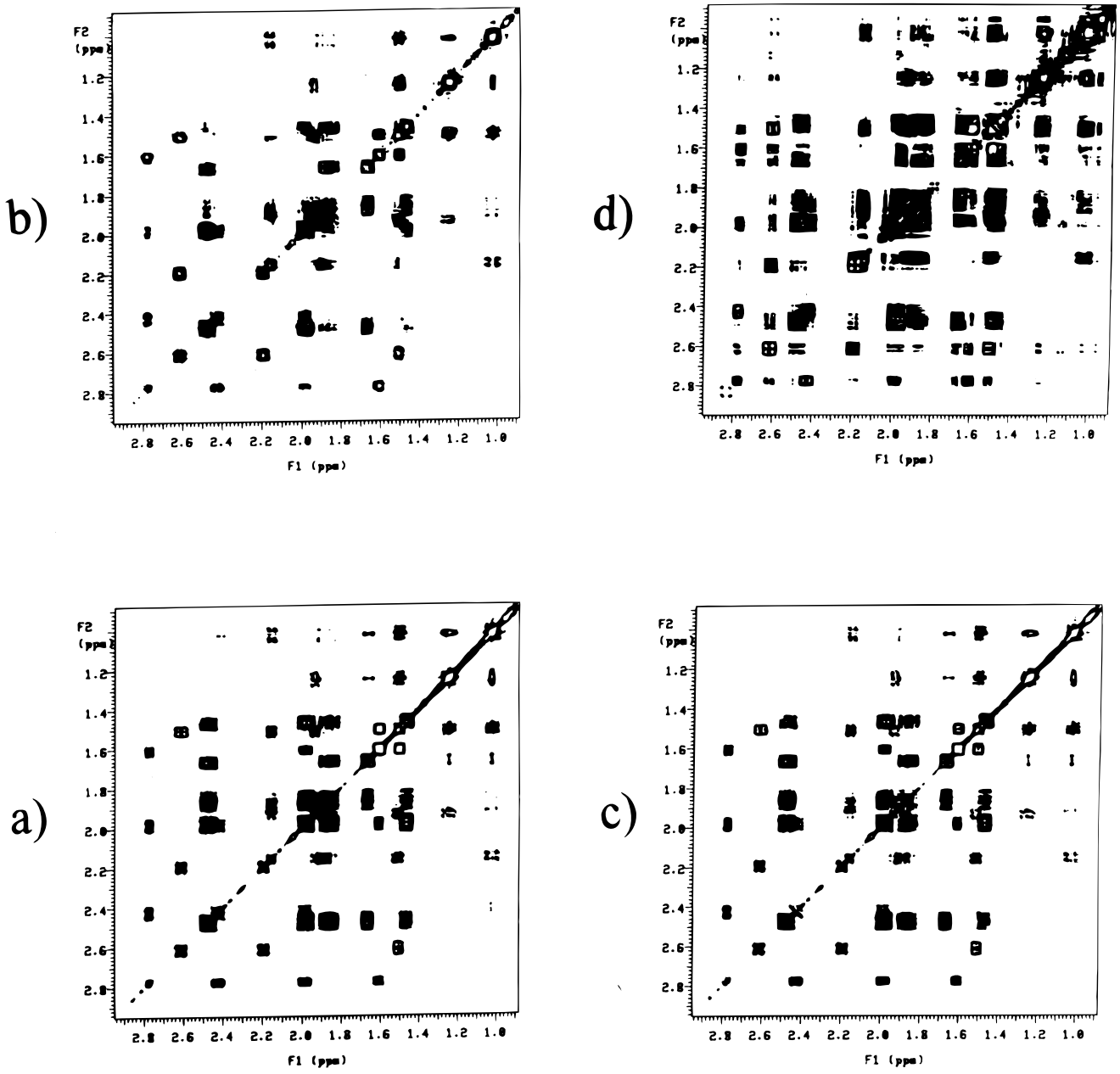


Figure 8. Absolute value COSY spectra for **1** obtained with (a) experiment 14, (b) experiment 14a, (c) experiment 14b and (d) experiment 14c. The spectrum in (b) (256 increments predicted to 1024) is clearly inferior to that in (a) (1024 increments), while that in (c) (512 increments predicted to 1024) is much closer to that in (a). Spectrum (d) illustrates the artifacts generated by extensive f_1 zero-filling (256 increments with zero filling to 2048).

noise ratio. This difference becomes more pronounced with increased linear prediction since the signal-to-noise ratio continues to increase for the HSQC spectrum (see Fig. 3) but actually decreases slightly for the HMQC spectrum.

The HMBC experiment[21] with its very wide ^{13}C spectral window, along with ^{1}H multiplet structure along f_1, represents a special case. Unlike other cross peaks, cross peaks involving methyl singlets have no f_1 multiplet structure and consequently increase in resolution and sensitivity with increasing linear prediction up to at least 16-fold prediction. These cross peaks are often important for structural and spectral assignments of natural products, particularly in the case of triterpenes.[22–24] Consequently, it may be useful to process HMBC spectra twice, once with fourfold linear prediction for most CH and CH_2 cross peaks but with more extensive prediction for CH_3 groups and other singlets.

Many of the 2D homonuclear correlation experiments also benefit from linear prediction, e.g. NOESY.[4] One sequence of this type which provides a particularly rigorous test of linear prediction is the phase-sensitive DQ COSY sequence.[25] Cross peaks obtained with this sequence consist of complex two-dimensional arrays of closely spaced peaks, half positive in phase and half negative. In order to test the ability of linear prediction to reproduce these patterns, two double quantum COSY spectra were obtained for the aliphatic region of **1**, using the dilute (1.5 mg ml^{-1}) sample. The spectral

widths were identical along both axes (1770 Hz) and 2048 points were collected and zero-filled to 4096. Experiment 12 used eight transients per time increment with 256 increments and linear prediction to 2048 (total acquisition time 2 h 7 min). Experiment 13 also had eight transients per time increment but 2048 increments and no linear prediction (total acquisition time 21 h 35 min). f_1 zero filling to 4096 was used in each case (i.e. better than 1 Hz data point resolution along each axis). The two spectra were similar in quality, even though the linearly predicted spectra took well under one eighth of the time of the experiment without linear prediction. Typical examples of complex cross peaks obtained for the two spectra are illustrated in Fig. 7(a) and (b). It is clear that eightfold linear prediction is capable of reproducing even complex, tightly spaced two-dimensional multiplet patterns.

By contrast, linear prediction does not work as well for absolute COSY spectra. This is illustrated in Fig. 8. Experiment 14 [Fig. 8(a)] used f_1 and f_2 spectral windows of 3500 Hz, 1024 data points and 1024 time increments (both zero filled to 2048). The spectrum was processed using the recommended[26] pseudo-echo weighting followed by triangular folding. Experiment 14a [Fig. 8(b)] was identical except that only the first 256 time increments were used and then linear predicted to 1024. The spectrum in Fig. 8(b) is clearly inferior to that in Fig. 8(a). Pseudo-echo processing produces a weighted interferogram which has zero intensity at the beginning and end of the evolution time with a maximum near the middle. Thus, when only the first quarter of the time increments are experimental, the predicted time increments contribute far more to the interferogram that the experimental points, in contrast to earlier spectra. This problem could be overcome by using the first 512 time increments, with linear prediction to 1024. This spectrum [Fig. 8(c)] is almost identical with that in Fig. 8(a). Some cross peaks are slightly weaker than in Fig. 8(a) but can be clearly observed by a small increase in the vertical scale. Fourfold linear prediction does work better if one uses a less extreme weighting function such as Gaussian multiplication, but at the expense of loss of resolution. Hence we suggest that no more than twofold linear prediction should be used for absolute value COSY spectra. This still represents a useful time saving.

Comparison of f_1 forward linear prediction with zero filling

In all of the spectra to date, linear prediction has been accompanied with zero filling to twice the number of points, to allow use of all linear predicted points as real points. Many reported studies used only f_1 zero filling without any linear prediction. To compare with the results presented above, a number of the spectra were reprocessed using no linear prediction, but with zero filling to yield the same total number of points as used previously. To begin, we reprocessed experiment 6 with zero filling in place of linear prediction. The initial processing (experiment 6e) used all 256 increments with zero filling to 2048. Typical ^{13}C cross-sections through a specific ^{1}H signal (H-13) for experiment 6 with linear prediction and for experiment 6e with only zero filling are compared in Fig. 9(a) and (b). The spectrum with only zero filling showed broader and less intense peaks than that with linear prediction and individual f_1 cross-sections through specific f_2 frequencies of experiment 6e showed significant truncation wiggles [e.g. see Fig. 9(b)]. The average signal-to-noise ratio for $^{13}CH_2$ cross-sections was 79:1 with zero filling compared with 122:1 with linear prediction. Hence linear prediction is clearly superior in this case. In assessing this observation it should be noted that we used the same f_1 Gaussian weighting for both spectra (see Experimental). Reprocessing experiment 6e with more severe Gaussian weighting (0.009 s time constant in place of 0.036 s) eliminated the wiggles but gave significantly broader and less intense lines [see Fig. 9(c)]. Next, the same data set was reprocessed using only the first 128 or the first 64 time increments with zero filling to 2048 (experiments 6f and 6g). The spectral deterioration was even more noticeable in these cases and very pronounced truncation wiggles were observed. Similar observations were made on reprocessing the spectrum from experiment 9, using 256, 128 and 64 time increments with zero filling to 2048 (experiments 9a–c). Finally, the reprocessing of a number of stored HMBC data sets which used 256 increments in conjunction with 25 000–28 000 Hz f_1 spectral windows consistently showed significant improvement in resolution and sensitivity for fourfold linear prediction over zero filling.

By contrast, the difference between linear prediction and zero filling is significantly less for more extensive linear prediction. For example, when experiment 6 was linearly predicted to 1024 with zero filling to 8192 (experiment 6h), the sensitivity loss was less than 20% (signal-to-noise ratio 202:1 compared with 239:1) relative to linear prediction to 4096 followed by zero filling to 8192 (experiment 6a). Typical ^{13}C cross-sections for experiments 6a and 6h are illustrated in Fig. 9(d) and (e), respectively. Furthermore, the two spectra obtained by the reprocessing of experiment 9 with linear prediction to 4096 and zero filling to 8192 (experiment 9d) and with linear prediction to 1024 with zero filling to 8192 (experiment 9e) had nearly identical signal-to-noise ratios. The latter two spectra also showed only a one third increases in sensitivity relative to experiment 9, much less than the doubling of sensitivity for spectrum 6a compared with spectrum 6. Thus, for spectra with marginal signal-to-noise ratios, it appears that little is gained by linear prediction beyond a factor of four.

The most spectacular failure of zero filling was observed for the absolute value COSY spectrum. Reprocessing of experiment 14, using the first 256 time increments with zero filling to 2048 gave the spectrum illustrated in Fig. 8(d). It contains many spurious peaks. The nature of the weighting function used (see above) meant that the weighted interferogram was severely truncated. This led to intense truncation wiggles which appeared as t_1 ridges in the absolute value display. In turn, these lead to the spurious cross peaks after triangular folding. Hence zero filling beyond a factor of two is not recommended for absolute value COSY spectra. By contrast, reprocessing of the DQ COSY spectrum from experiment 12 using zero filling from 256 to 4096 with no linear prediction (experiment 12a) gave

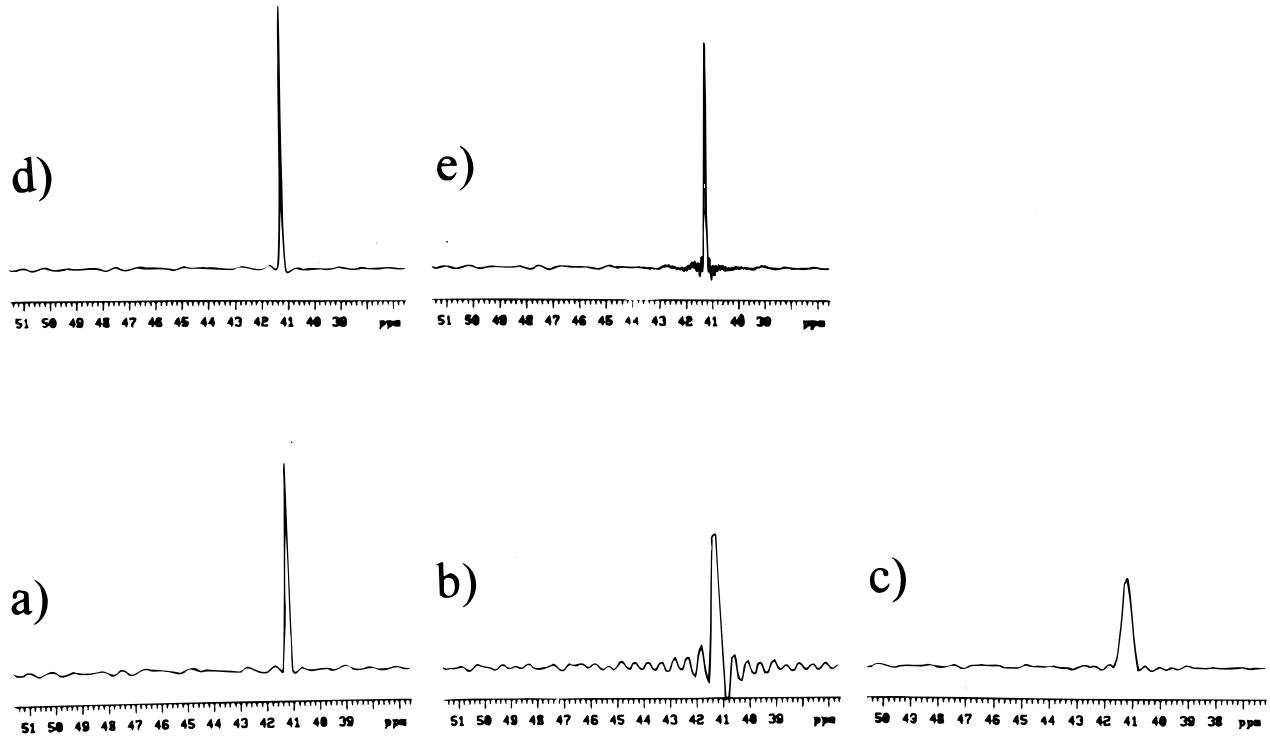


Figure 9. ^{13}C cross-sections through the chemical shifts of H-13 of **1** for spectra with varying degrees of linear prediction, zero filling and Gaussian multiplication: (a) 256 increments with linear prediction to 1024 and zero filling to 2048, Gaussian time constant = 0.036 s (experiment 6); (b) 256 increments with zero filling to 2048, Gaussian time constant = 0.036 s (experiment 6e); (c) repeat of experiment 6e with Gaussian time constant = 0.009 s; (d) 256 increments with linear prediction to 4096 and zero filling to 8192, Gaussian time constant = 0.144 s (experiment 6b); (e) 256 increments with linear prediction to 1024 and zero filling to 8192, Gaussian time constant = 0.144 s (experiment 6h).

a spectrum [Fig. 7(c)] which was only marginally inferior to that from experiment 12 with linear prediction to 2048 [Fig. 7(a)]. When experiment 12 was reprocessed, using linear prediction to only 512 points, followed by zero filling to 4096 (experiment 12b), the resultant spectrum [Fig. 7(d)] was essentially indistinguishable from that in Fig. 7(a). Inspection of the raw interferograms showed that most of the signal intensity had decayed away by 256 increments, reflecting the narrow spectral window (1770 Hz) and consequent long evolution time (0.15 s). Thus, with 2048 real points along each axis, the data point resolution is significantly less than the linewidths.

Based on these observations, one can conclude that the advantages of linear prediction over zero filling are greatest when the interferograms are significantly truncated but that these advantages become smaller as the signal level at the end of the evolution time (with or without linear prediction) approaches the noise level. This limit, which depends on the maximum f_1 evolution time, the T_2 relaxation time and the inherent signal-to-noise ratio of the experiment, is usually reached when the data point resolution approaches the natural f_1 linewidths, but sooner for weak spectra such as in experiment 9. Hence it is best to use linear prediction up to this limit with zero filling used beyond this point if better data point resolution is required.

f_2 linear prediction and linear prediction along both axes

There are obvious advantages to f_1 linear prediction since doubling the number of acquired increments automatically doubles the total acquisition time. However, this is not generally true for f_2. The actual acquisition time is usually only a small fraction of the total experiment time and thus the number of f_2 data points can be increased significantly with minimal impact on the total experiment time. The exception to this is provided by HSQC and HMQC spectra, which require ^{13}C decoupling during acquisition. In these cases, the necessity for keeping the decoupler duty cycle low to minimize thermal gradients in the sample means that an increased acquisition time must usually be accompanied by an increased preacquisition delay and consequent significant increase in the total experiment time. To test whether there would be advantages in f_2 linear prediction under these circumstances, the data set from experiment 5 (four transients per increment and 1024 time increments) was reprocessed using f_2 linear prediction to 1024 points from 512 (experiment 5a). The CH_2 cross-sections for the latter data set showed increased f_2 resolution but with a small loss in overall sensitivity. However, f_2 linear prediction without simultaneous f_1 linear prediction sacrifices the advantages of the latter. Since the spectrometer software allows linear prediction along both axes, this was tested for the data set from experiment 6 using linear prediction to 1024 along both axes (experiment 6i). A contour plot showed noticeable losses of sensitivity, resolution and precision of peak position compared with a spectrum with only f_1 linear prediction (experiment 6). There are two main problems with f_2 linear prediction. First, the narrow f_2 spectral window results in a sufficiently long acquisition time that most of the signal has decayed away before the

start of linear prediction. The second and more serious problem is that each FID contains cosine and sine components corresponding to all f_2 frequencies. By contrast, the f_1 interferograms have far fewer frequencies components, i.e. only those associated with one specific f_2 frequency. Hence f_2 forward linear prediction is inherently more difficult than f_1 linear prediction. This suggested that additional zero filling might be a viable alternative in this case. To test this, experiment 6 was reprocessed with the 512 data points zero filled to 2048 and with 256 time increments linear predicted to 1024 and zero filled to 2048 (experiment 6j). A contour plot for a expanded region of the spectrum is illustrated in Fig. 10 with a cross-sectional plot for C-1 above the contour plot. This shows somewhat improved resolution compared with a similar plot from experiment 6, owing to the improved definition of multiplet structure with additional points combined with the use of a less severe weighting function. The average signal-to-noise ratio also improved from 122:1 to 152:1. Thus, in contrast to f_1 processing, modest zero filling seem to be better than linear prediction for f_2 processing. However, if excellent f_2 resolution is required for an HMQC or HSQC spectrum, it is probably better

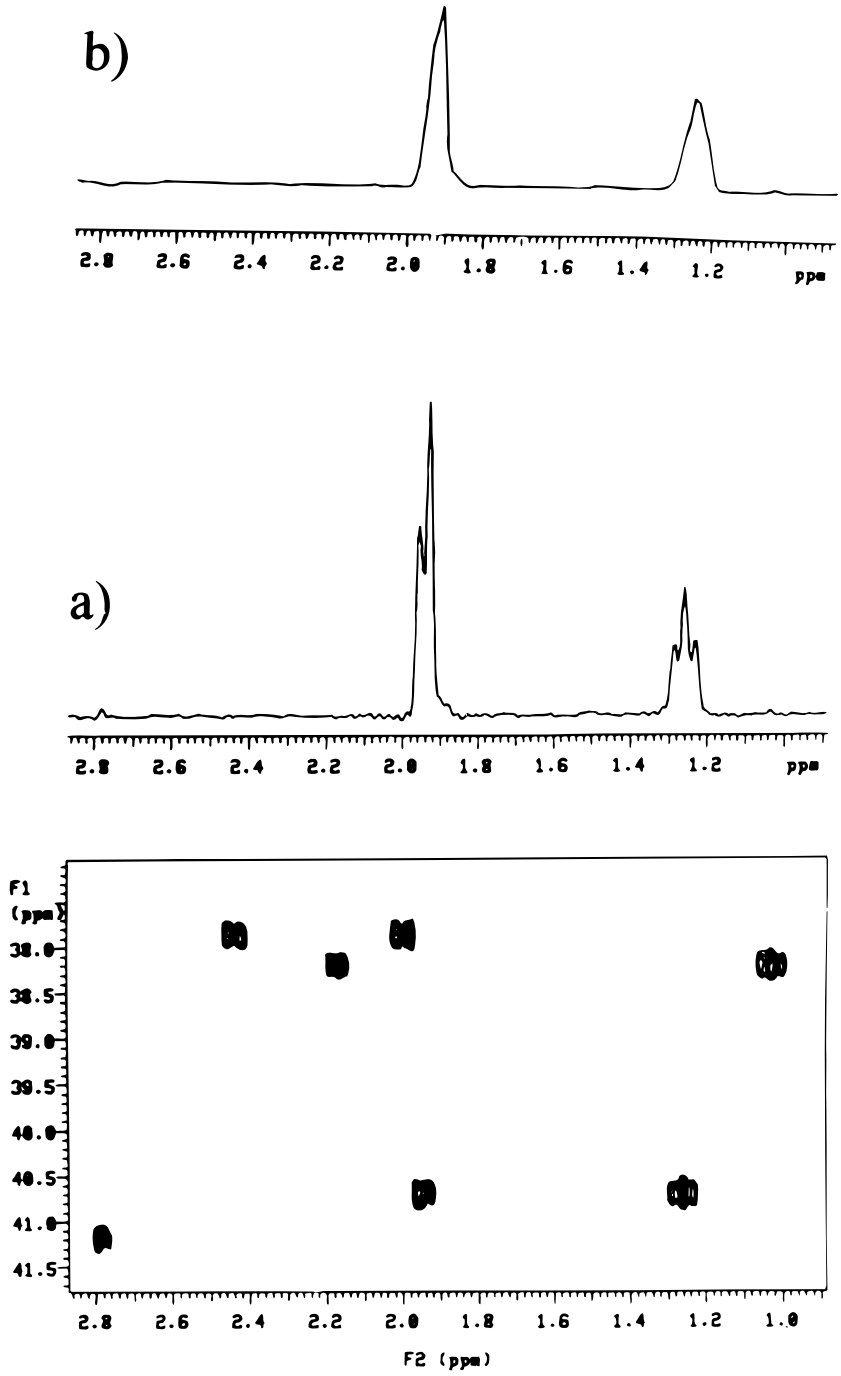


Figure 10. Contour plot of a region of the HSQC spectrum of **1** obtained with experiment 6j. The cross-sectional plots above the contour plot are for C-1, showing the improved f_2 resolution with zero filling to 2048 [cross-section (a), experiment 6j] *vs.* zero filling to only 1024 [cross-section (b), experiment 6].

to run a ^{13}C coupled spectrum[12,27] since more data points can be acquired with a relatively short preacquisition delay.

CONCLUSION

It has been confirmed that f_1 forward linear prediction is a powerful technique for improving resolution and/or sensitivity of routine 2D spectra, and also as for minimizing the time needed to acquire spectra of the desired quality. For ^{1}H-detected ^{13}C–^{1}H shift correlation sequences, the optimum number of increments appears to be 256 in most cases. With this number of increments, f_1 linear prediction seems surprisingly robust, with HSQC spectra showing improved resolution and sensitivity, and also accurate prediction of ^{13}C chemical shifts, up to 16-fold linear prediction. With the exception of absolute value COSY spectra, other 2D spectra show improvements out to at least fourfold linear prediction. Fourfold forward f_1 linear prediction gives better results than the corresponding amount of zero filling in almost all cases. In view of the obvious advantages of f_1 forward linear prediction for 2D NMR, we strongly encourage its regular use. By contrast, f_2 forward linear prediction is not advantageous but modest f_2 zero filling may yield improved results.

EXPERIMENTAL

All spectra were obtained on a Varian UNITY 500 spectrometer equipped with a 5 mm inverse detection probe (^{1}H 90° pulse width = 9.5 μs, ^{13}C decoupler pulse width = 9.4 μs) and operating at 25 °C. All spectra were acquired without sample spinning, using the hypercomplex method[28] to acquire phase-sensitive spectra. HSQC and HMQC spectra were acquired using a preacquisition delay of 1.6 s and a BIRD nulling delay of 0.4 s. GARP ^{13}C decoupling[29] was used during acquisition. The coupled HSQC spectrum for the saponin was acquired with a BIRD nulling delay of 0.25 s and a preacquisition delay of 0.6 s. The DQ COSY spectra was acquired using a preacquisition delay of 1.6 s with a homospoil pulse applied during the delay to destroy residual magnitization.

Gaussian weighting functions were used for both axes of the phase-sensitive 2D spectra, with the functions matched to the acquisition time or the evolution time (with or without linear prediction and/or zero filling). This was done using the interactive software which displays the FID or interferogram with the shape of the weighting function superimposed. The latter was adjusted to approach zero at the end of the acquisition or evolution time. This typically involved using a Gaussian time constant which was 0.6 times the acquisition or evolution time. Linear prediction was carried out using standard Varian software which is based on the algorithm of Barkhuisjen *et al.*[1] The latter is a linear least-squares procedure based on singular value decomposition. Linear prediction for HSQC and HMQC spectra assumed up to eight peaks per f_2 frequency, while the linear predicted DQ COSY and absolute value COSY spectra allowed for up to 32 peaks per f_2 frequency.

Acknowledgements

Financial support from NSERCC (Canada) and DGAPA (UNAM) (Mexico) is acknowledged by W.F.R. and R.G.E., respectively. Helpful discussions with Professor Lewis Kay (Toronto), Professor A. D. Bain (McMaster) and Dr V. V. Krishnamurthy (Varian Associates) are acknowledged.

REFERENCES

1. H. Barkhuisjen, R. deBeer, W. M. M. J. Bovee and D. van Ormondt, *J. Magn. Reson.* **61**, 465 (1985).
2. C. F. Tirendi and J. F. Martin, *J. Magn. Reson.* **81**, 577 (1989).
3. G. Zhu and A. Bax, *J. Magn. Reson.* **90**, 405 (1990); **98**, 192 (1992).
4. D. M. Babcock, P. V. Saharabachbre and W. H. Gneimer, *Magn. Reson. Chem.* **34**, 851 (1996).
5. W. F. Reynolds, M. Yu, B. Ortiz, A. Rodriguez, F. Yuste, F. Walls, R. G. Enriquez and D. Gnecco, *Magn. Reson. Chem.* **33**, 3 (1995).
6. W. F. Reynolds, A. Maxwell, B. Telang, K. Bedaisie and G. Ramcharan, *Magn. Reson. Chem.* **33**, 412 (1995).
7. S. M. Kristensen and J. J. Led, *Magn. Reson. chem.* **33**, 461 (1995).
8. W. F. Reynolds, J.-P. Yang and R. G. Enriquez, *Magn. Reson. Chem.* **33**, 705 (1995).
9. M. H. M. Sharaf, P. L. Schiff, A. N. Tackie, C. H. Phoebe, L. Howard, C. Meyers, C. E., Hadden, S. K. Wrenn, A. O. Davis, C. W. Andrews, D. Minick, R. L. Johnson, J. P. Schockcor, R. C. Crouch and G. E. Martin, *Magn. Reson. Chem.* **33**, 761 (1995).
10. H. Hani, A. Elgamal, H. S. M. Soliman, G. Toth, J. Halasz and H. Duddeck, *Magn. Reson. Chem.* **34**, 697 (1996).
11. G. Toth, J. Halasz, S. Boros, A. Levai, C. Nemes and T. Patonay, *Magn. Reson. Chem.* **34**, 932 (1996).
12. W. F. Reynolds, S. McLean, L.-L. Tay, M. Yu, R. G. Enriquez, D. M. Estwick and K. O. Pascoe, *Magn. Reson. Chem.* **35**, 455 (1997).
13. G. Bodenhausen and D. J. Ruben, *Chem. Phys. Lett.* **69**, 185 (1980).
14. W. F. Reynolds, R. G. Enriquez, L. I. Escobar and X. Lozoya, *Can. J. Chem.* **62**, 2421 (1984).
15. A. Bax and S. Subramanian, *J. Magn. Reson.* **67**, 565 (1986).
16. A. Bax, M. Ikura, L. E. Kay, D. A. Torchia and R. Tschudin, *J. Magn. Reson.* **68**, 304 (1990).
17. A. F. Mehlkopf, D. Korbee, T. A. Tiggelman and R. Freeman, *J. Magn. Reson.* **58**, 315 (1984).
18. G. A. Morris, *J. Magn. Reson.* **100**, 316 (1992).
19. W. F. Reynolds, M. Yu, R. G. Enriquez, H. Gonzalez, I. Leon, G. Magos and M. L. Villareal, *J. Nat. Prod.* **58**, 1730 (1995).
20. W. F. Reynolds, M. Yu, R. G. Enriquez and I. Leon, to be published.
21. A. Bax and M. F. Summers, *J. Am. Chem. Soc.* **108**, 2093 (1986).
22. W. F. Reynolds, S. McLean, J. Poplowski, R. G. Enriquez, L. I. Escobar and I. Leon, *Tetrahedron* **42** 3419 (1986).
23. S. McLean, M. Perpick-Dumont, W. F. Reynolds, H. Jacobs and S. S. Lachmansing, *Can. J. Chem.* **65**, 2519 (1987).
24. S. McLean, W. F. Reynolds, J.-P. Yang, H. Jacobs and L.-L. Jean-Pierre, *Magn. Reson. Chem.* **32**, 422 (1994).
25. U. Piantini, O. W. Sorensen and R. R. Ernst, *J. Am. Chem.*

Soc. **104**, 6800 (1982).
26. A. Bax and R. Freeman, *J. Magn. Reson*. **44**, 542 (1981).
27. D. Yang, X. Xu and C. Ye, *Magn. Reson. Chem*. **30**, 711 (1992).
28. D. J. States, R. A. Haberkorn and D. J. Ruben, *J. Magn. Reson*. **48**, 286 (1982).
29. A. J. Shaka, P. B. Barber and R. Freeman, *J. Magn. Reson*. **64**, 547 (1985).