Available online at www.sciencedirect.com
ScienceDirect
Journal of Magnetic Resonance 186 (2007) 112–122
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Computer optimized spectral aliasing in the indirect dimension of ^{1}H–^{13}C heteronuclear 2D NMR experiments. A new algorithm and examples of applications to small molecules

Damien Jeannerat *

Department of Organic Chemistry, 30 Quai E. Ansermet, CH-1211 Geneva 4, Switzerland

Received 29 September 2006; revised 2 February 2007
Available online 4 February 2007

Abstract

A new algorithm for optimizing spectral width in the indirect dimension of heteronuclear 2D experiments is introduced. It takes a list of carbon chemical shifts as input and calculates the optimal spectral width and number of time increments to use in the carbon dimension of 2D experiments such as HSQC, HMBC, etc. When using optimized conditions where signals are better distributed along the carbon dimension, the number of time increments needed to resolve all of the signals is reduced by one to two orders of magnitude. This decreases the experimental time by the same factors and makes the acquisition of spectra such as HSQC–TOCSY, HSQC–NOESY, etc. more practical. The new algorithm allows users to limit the maximal t_1 evolution time when relaxation is a concern, and can take lists of carbons that do not need to be resolved. For any carbon, insights regarding the position of signals in the proton dimension increase the efficiency of the optimization by allowing the overlap of pairs of carbons with incompatible proton dispersions. The application of a second optimization using a fully-resolved spectrum as a source of proton dispersion for the carbons allows the number of time increments to be reduced further. Application to cyclosporine A shows that the time taken to acquire fully resolved HSQC spectra can be 126 times less than would be required in a full-width spectrum with the same resolution. The most interesting applications concern experiments where series of HSQC-based experiments have to be acquired, for example in relaxation time measurements. It is shown that the acquisition of quickly acquired series of selective-TOCSY–HSQC can facilitate assignment in carbohydrates. Computer-optimized spectral aliasing (COSA) generally requires no modification of the pulse sequence and can therefore be easily applied by non specialists.

Keywords: NMR fast methods; Computer-optimized spectral aliasing (COSA); Spectral folding; High resolution in indirect dimensions; Cyclosporine A; Melezitose

1. Introduction

In all but the simplest cases, heteronuclear 2D experiments are needed to determine the molecular structures of unknown compounds. Often, ^{1}H–^{13}C HSQC (or HMBC) and multiple-bounds HMBC experiments are sufficient, but when proton overlap becomes important, additional experiments such as the ADEQUATE or selective-TOCSY–HSQC, HSQC–TOCSY, and HSQC–NOESY, etc. can be very useful to eliminate ambiguities and facilitate assignment and structure determination. In all of these experiments, one needs to resolve the carbons in the indirect dimension in a minimal number of time increments. When a pair of carbons is only a few Hertz apart, the number of time increments needed to resolve them at high field may reach the tens of thousands when using standard methods. A powerful solution consists of using Hadamard techniques [1,2], but they involve the modification of pulse sequences, the generation of shaped pulses and present difficulties when close pairs of carbons have to be resolved. In this paper we will present a method that requires no changes in pulse sequences and that allows a reduction in the number of time increments by one to two

* Fax: +41 22 37 93215.
E-mail address: damien.jeannerat@chiorg.unige.ch

1090-7807/$ - see front matter
doi:10.1016/j.jmr.2007.02.003

orders of magnitude compared to classical acquisition when SNR is not limiting. The method can benefit from post-acquisition resolution enhancement such as linear prediction [3–5] and maximum entropy [6,7] if desired. The idea consists of taking a list of carbon chemical shifts stemming from a 1D carbon or DEPT spectrum and optimizing the acquisition parameters of the carbon dimension of 2D heteronuclear experiments. All one needs to do is to copy/paste the list of ^{13}C peaks into a web-interfaced optimization program and set the spectral width and number of time increments according to its results. A variant of this method takes into account the dispersion of signals in the proton dimension to further decrease the number of time increments.

1.1. Theory

For any list of carbon chemical shifts that are resolved in a 1D spectrum, one can calculate the number of time increments N_0 a full-width 2D spectrum would require to resolve all the signals. The frequency difference between the closest pair of carbons determines the maximal t_1 time evolution ($t_{1,\max}$) while the largest frequency difference (or spectral width) determines the dwell time. N_0 is simply the ratio of these two times. The Annex 1 provides the details of the calculations and discusses the criterion used to consider that signals are resolved. It also shows, if it is not obvious, that in order to resolve carbons that are a few Hz apart, thousands of time increments are normally necessary.

The application of linear prediction or any other processing technique allows one to increase the signal resolution and may diminish the number of time increments by a factor 2, 4, or more in very favorable cases, but the number of time increments required to resolve close pairs of carbons will still be in the thousands.

Reducing the spectral width below the value of the full spectrum increases the dwell time and therefore allows the number of time increments to be reduced. The resolution is maintained because $t_{1,\max}$ is unchanged, but this causes a violation of the Nyquist condition with the potential risk of introducing ambiguities in the frequency of the sampled oscillations. This is illustrated by Fig. 1 where the folding or aliasing of signals at the other end of the spectrum shows that the apparent frequency of aliased signals is different from the real one which is outside the spectrum.

The location of aliased signals is given by the modulo function

$$\nu_a = \mathrm{mod}(\nu_0 + \mathrm{SW_a}/2 - \mathrm{CF}, \mathrm{SW_a}) - \mathrm{SW_a}/2 + \mathrm{CF} \qquad (1)$$

where $\mathrm{SW_a}$ is the spectral window considered, ν_0 the "real" frequency and CF the carrier frequency in the F1 dimension. The less common case of folded signals obtained using the TPPI method will not be considered here. In order to avoid ambiguities during signal assignment, one should make sure that carbons do not overlap upon aliasing; i.e. avoid, for every combination of signals i and j, values for which $\nu_a^i = \nu_a^j$. By studying the overlap of carbon pairs upon aliasing we can therefore find the spectral width resolving all signals.

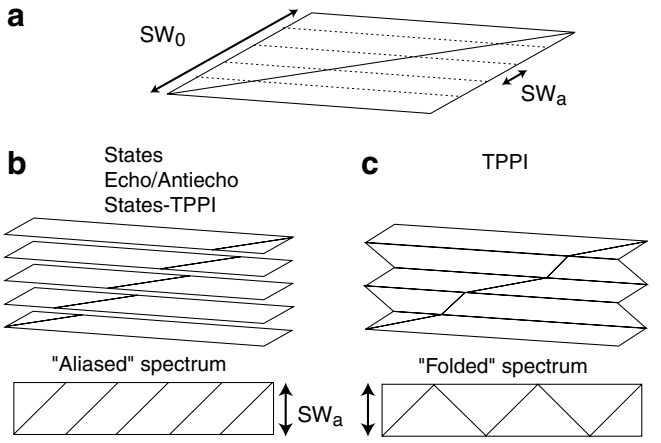


Fig. 1. (a) When the spectral width $\mathrm{SW_a}$ is set below the full spectral width $\mathrm{SW_0}$, signals are either aliased (b) or folded (c) depending on the quadrature detection used in the indirect dimension.

For any pair of signals, overlap occurs when the tentative spectral width

$$\mathrm{SW}^* = \Delta^{ij}/n \qquad (2)$$

where n is an integer and $\Delta^{ij} = (\nu_0^i - \nu_0^j)$. Fig. 2 illustrates the occurrences of overlap for different pairs of carbon of a four-spin system. The width of any up state of the overlap functions corresponds to the margin required to ensure separation. It takes into account the predicted width of the signal in the aliased spectrum.

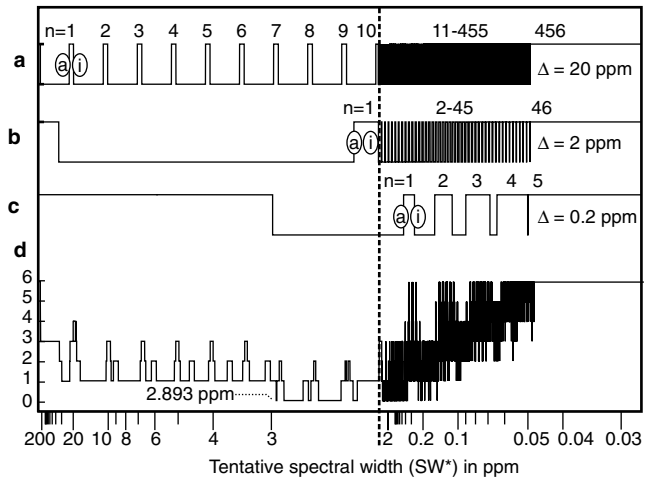


Fig. 2. Overlap as a function of tentative spectral widths SW^* for pairs of carbons distant by (a) 20, (b) 2 and (c) 0.2 ppm. (d) Total overlap of a four-spin system with 20, 2, 0.2, and 0 ppm chemical shifts. The dotted lines indicate a change of horizontal scale. The number of time increments was 32, the signal width 1.2 Hz and the ^{13}C Larmor frequency 125 MHz. Labels a and i indicate the boundaries of the first high state of the overlap functions $\mathrm{Max}_{ij}^{\mathrm{I}}$ and $\mathrm{Min}_{ij}^{\mathrm{I}}$.

Counting the number of overlap functions in the up state gives the total overlap shown in Fig. 2d. It reaches zero when conditions are found avoiding overlap upon folding and/or where there is an insufficient number of time increments for the resolution of all carbons. Computer-optimized spectral aliasing (COSA) consists of studying these functions in order to propose acquisition parameters that quickly obtain spectra with no carbon overlap. It determines the minimal number of time increments for which at least one spectral width fulfills the condition of no overlap. Note that when the predicted $t_{1,\max}$ exceeds a user-defined limit, the search continues in order to avoid experiments where relaxation would cause an excessively low signal-to-noise ratio (see Section 3.1 for more details). When more than one spectral width is satisfactory, the largest is proposed. All the smaller values are less interesting because they would correspond to acquisition conditions having $t_{1,\max}$ longer then needed and would therefore correspond to experiments with reduced signal-to-noise ratio.

The optimized conditions ensure an ideal spread of all the carbons along the F1 dimension and should be very satisfactory for molecules up to 20 carbons or so. With bigger molecules, the optimized number of time increments may be larger than the SNR imposes. One may therefore wish to exploit the fact that a quickly acquired low-resolution HSQC or HMBC spectrum can provide information to make the algorithm more effective by allowing the overlap, in the carbon dimension, of signals that are known to have different proton chemical shifts (see Fig. 3). The less ambiguous the proton dispersion list, the more effective the optimization since less constraints are imposed on the program. A low-resolution spectrum provides a relatively coarse list because, for any carbon, the proton chemical shift ranges may be quite broad or include signals from the nearby carbons. For example the second list in Fig. 3 shows three proton ranges because, at the chemical shift of C(13), three CH are overlapping in the carbon dimension. The finest dispersion list can be obtained by taking a computer-optimized aliased spectrum as a source of proton dispersion for the optimization, as illustrated in Fig. 3. This is very useful when one needs to acquire series of heteronuclear experiments, for example for relaxation measurements, 2D diffusion measurements [8–10], or in higher-dimensionality experiments.

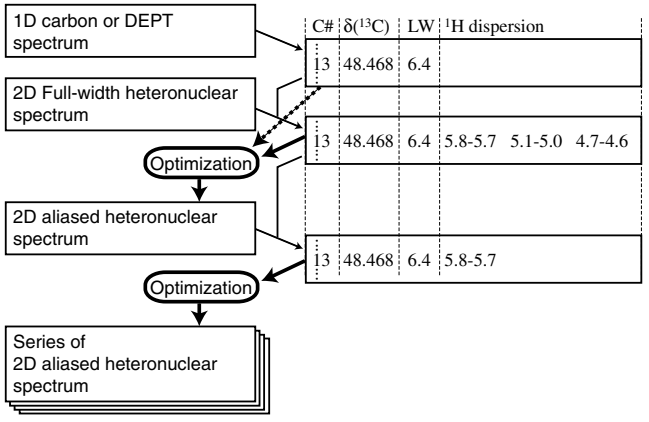


Fig. 3. Flow chart of two-step carbon dimension optimization of heteronuclear experiments. The goal of the first optimization is to obtain a fully resolved spectrum either using carbon data only (broad dotted arrow) or taking into account proton dispersion stemming from a full-width low-resolution spectrum (broad continuous arrow). The second optimization includes the data of the fully-resolved spectrum to further reduce the number of time increments used in the acquisition of series of spectra. The right-hand side of the figure shows the data submitted to the optimization program.

1.2. The optimization algorithm

In earlier developments, the approach for determining the optimal spectral width consisted of simulating aliased spectra [11] or setting spectral widths arbitrarily [12]. One inconvenience of simulations was that they could miss solutions when the calculated set of tentative spectral widths was not large enough. The second problem was that the simulations could last minutes or even hours. We have developed an algorithm that solves both problems. The logic is schematized in Fig. 4 and the manner in which it uses the overlap functions to quickly find conditions with no overlap is shown in Fig. 5.

The algorithm consists of an efficient study of the overlap functions shown in Figs. 2a–c with the goal of finding the largest possible spectral width for which all functions are in the low state. For any number of time increments (inner frame in Fig. 4), the search starts with a tentative spectral width value set to the full width, and the overlap tested for each pair of signals. When a pair of signals is found to overlap (for example s2 and s3 in Fig. 5A), SW^* is set to Min_{ij}^{n}, the position of the next low-state of the overlap function (Fig. 5B). This point corresponds to a spectral width for which the relevant carbons no longer overlap. The calculation takes into account the natural line width lw_n of the aliased signal (taken from the 1D carbon data or calculated on the basis of an estimate of the T_2 of the molecule) and the additional broadening due to apodization (see Annex 1). The procedure above is repeated for each pair of carbons until either a solution has been found, i.e. all the overlap functions are in low states (Fig. 5C), or until the resolution of all signals is determined to be impossible, i.e. when the overlap functions have no more low-state values (see right part of Fig. 2). In the latter case, a new search is started with a larger value of N^*.

The search for the optimal number of increments can be made even faster when realizing a property of the overlap functions. When no limits are set for $t_{1,\max}$, there is a threshold number of time increments N_t above which a spectral width separating all signals can be found for any number of time increments $N > N_t$. The existence of such a threshold allows us to quickly converge towards using the binary search illustrated in Fig. 6. In order to avoid conditions with excessive $t_{1,\max}$, our program uses the fast algorithm only to provide the starting point of the fine search for conditions that take into account relaxation as

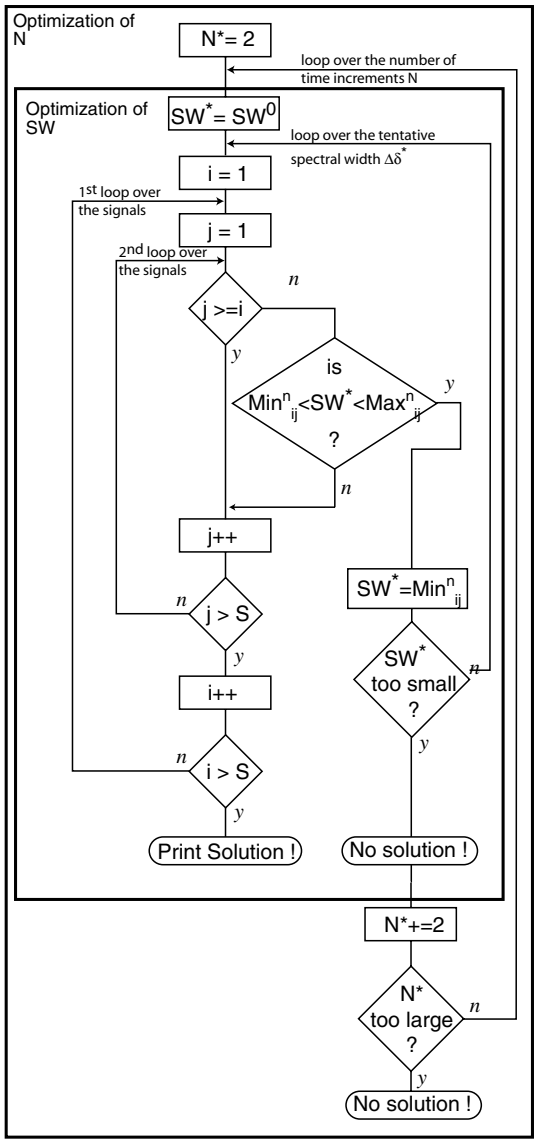


Fig. 4. Flowchart of the algorithm determining optimized spectral width (inner box) and number of time increments (outer box) avoiding signal overlap. When no number of time increments are specified, the optimization starts with a tentative value for $N^* = 2$ and will be incremented until a solution is found. For each pair of signals, overlap is tested positive when SW* is within the boundaries (Max^n_{ij} and Max^n_{ij}) of high states of the overlap functions (see also Fig. 2). When SW* is so small that farther reduction cannot reduce overlap, the number of time increments is incremented and a new search for an optimal spectral width started.

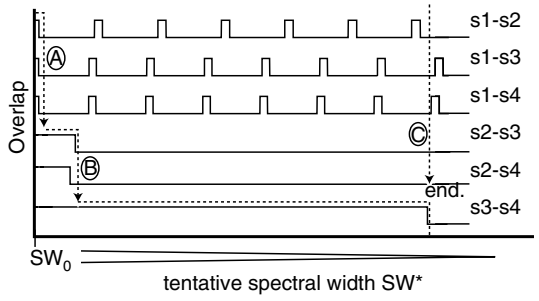


Fig. 5. Illustration of the algorithm used to determine optimal spectral width by successive reduction of tentative values of SW*.

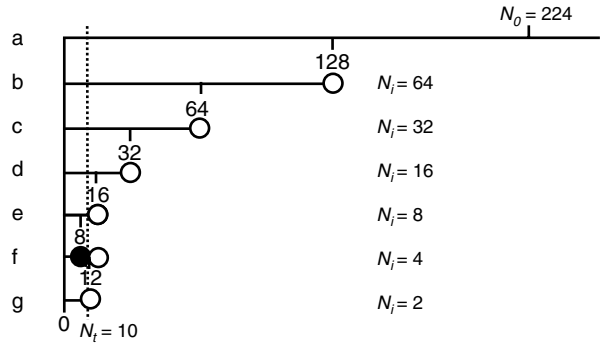


Fig. 6. Schematic representation of the binary search for the number of increments N_t (dotted line) above which resolution of all signals can be achieved using aliasing. (a) The first tentative number of time increments tested is $N^* = 2^n$ where n is the largest integer for which $2^n < N_0$, where N_0 is equal to the number of time increments a full spectrum would need to resolve all the signals. The increment for the convergence N_i is set to 2^{n-1} and divided by two at each calculation step until $N_i = 2$. (b–d) Open circles correspond to cases where conditions resolving all carbons exist meaning that $N^* > N_t$; N_i is therefore subtracted from N^* before the next cycle. Otherwise it is added to N^*.

illustrated in Fig. 4. Instead of using 2 as the starting value for the search (top of Fig. 4), N^* starts with N_t.

When a user-defined upper limit is set for $t_{1,max}$, the carbon pairs that are impossible to resolve are simply not tested and the optimization is run for all the other pairs. This simple modification ensures $t_{1,max}$ will not exceed the limit.

Another option that our program proposes is to permit the overlap of signals that do not need to be resolved. Overlap among all of these is allowed, but not with any others. This reduces the constraints for the optimization and usually results in a smaller number of time increments.

2. Results

2.1. Reduction in the number of time increments

When applied to an experimental ^{13}C database containing 484 DEPT spectra, the average reduction factor in the number of time increments is 83. The distribution of these factors as a function of the number of signals is given in Fig. 7. The large deviation from the average is explained by the fact that distances between the closest pairs of carbons in any dataset are broadly distributed. When this distance is small, signals are difficult to resolve and the optimization may be quite effective, resulting in large reduction factors.

As mentioned earlier, low-resolution heteronuclear experiments in which signal widths in the F1 dimension are too large for unambiguous assignment can serve as a source of information to improve the performance of COSA and reduce the experimental time for a second spectrum acquired using spectral aliasing ensuring resolution of all signals. When this data is included, the average reduction in the number of time increments increases from 83 to 274 (top of Fig. 7).

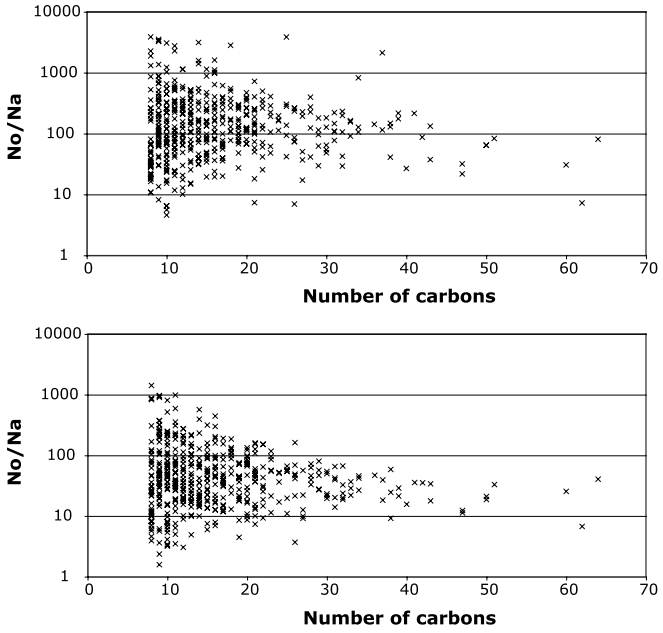


Fig. 7. Reduction factor in number of time increments when using spectral aliasing calculated for 484 experimental DEPT peak lists (bottom) and when considering the proton dispersion obtained from low-resolution HSQC spectra (top).

The procedure tending to achieve the maximum reduction in the number of time increments discussed earlier and schematized in Fig. 3, is illustrated by the HSQC spectra of cyclosporine A. Fig. 8 shows that carbons C(11), C(12), and C(13) could not be unambiguously assigned to the three signals designated by arrows in a low-resolution spectrum (see enlargements in Fig. 8). It is therefore expected that an optimized spectrum will place these three carbons at different positions in the carbon dimension. One can also note that these three carbons have a proton dispersion list which is unique, i.e. no other carbons have signals at these proton chemical shifts. This means that an overlap of any of these three carbons with any other would cause no assignment problems.

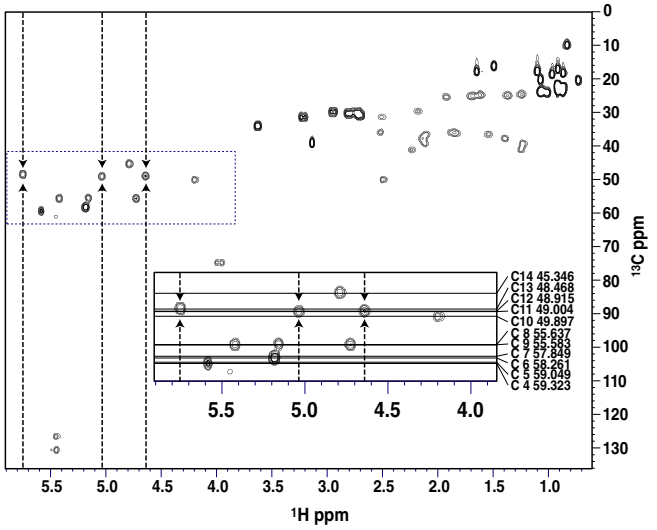


Fig. 8. Low-resolution full-width HSQC spectrum of cyclosporine A acquired using 256 time increments. The signal widths in the carbon dimension are about 2.7 ppm broad. Only a few signals can be unambiguously assigned to carbons. For example, those of C(11), C(12), and C(13) are too broad and can be attributed to any of the three signals indicated by arrows. They have the same proton dispersion lists (5.8–5.7, 5.1–5.0, and 4.7–4.6 ^{1}H ppm) but these proton regions are occupied by no other carbons as shown by the absence of signals below the dotted arrows. The algorithm will therefore have to resolve these three carbons independently from the others.

When submitting carbon data along with the "crude" proton dispersion lists to the optimization program, the acquisition parameters result in the spectrum shown in Fig. 9. One can appreciate that each signal can be unambiguously assigned to individual carbons, as the automatically generated labels show.

In order to provide "fine" proton dispersion lists to the optimization program, one must first analyze the spectrum of Fig. 9. The second optimization is then applied, and the spectrum obtained represents another 3-fold reduction in experimental time (Fig. 10). The time taken for the acquisition of the 120 time increments resulting in the fully resolved spectrum is 126 times less than that required by a full-width spectrum to reach the same resolution. This number of time increments and spectral widths are therefore optimal for the acquisition of series of HSQC-based experiments such as relaxation rate measurements, etc.

2.2. Examples of applications

Facilitated carbon assignment is the most simple application of COSA. It can be applied to "one-bound experiments" like the HSQC experiment. It allows one to obtain, in 2D spectra, the ideal resolution usually only enjoys in 1D carbon spectra. The optimization of one-bound experiments should use peak lists stemming from DEPT spectra, so that quaternary carbons are not taken into account.

A more interesting application is molecular structure determination using long-range experiments like HMBC. Provided a small modification of the sequence is made to avoid gradient coding during half of the t_1 delay [13], HMBC signals can reach 1 Hz/pt resolution in the carbon dimension (see also Section 3.3). Note that one can usually observe resolved coupling structures due to homo- and heteronuclear couplings which may be structurally informative using the Karplus rule for $^{3}J_{HH}$ and $^{3}J_{HC}$ [14].

Aliased HMBC spectra turned out to be very useful in confirming the structure of nearly symmetrical molecules with quasi degenerated carbon resonances [13,15] and in the study of complex natural products [16]. COSA has also been applied to the ADEQUATE experiment [17] used to confirm the carbon connectivity of a newly discovered natural product [18]. Another application was found during the study of a mixture of three coexisting forms of a synthesis intermediate of erythronolide A. The classical 2D experiments (DQF–COSY, HSQC and HMBC) turned out to be insufficient to permit signal assignment and structure deter-

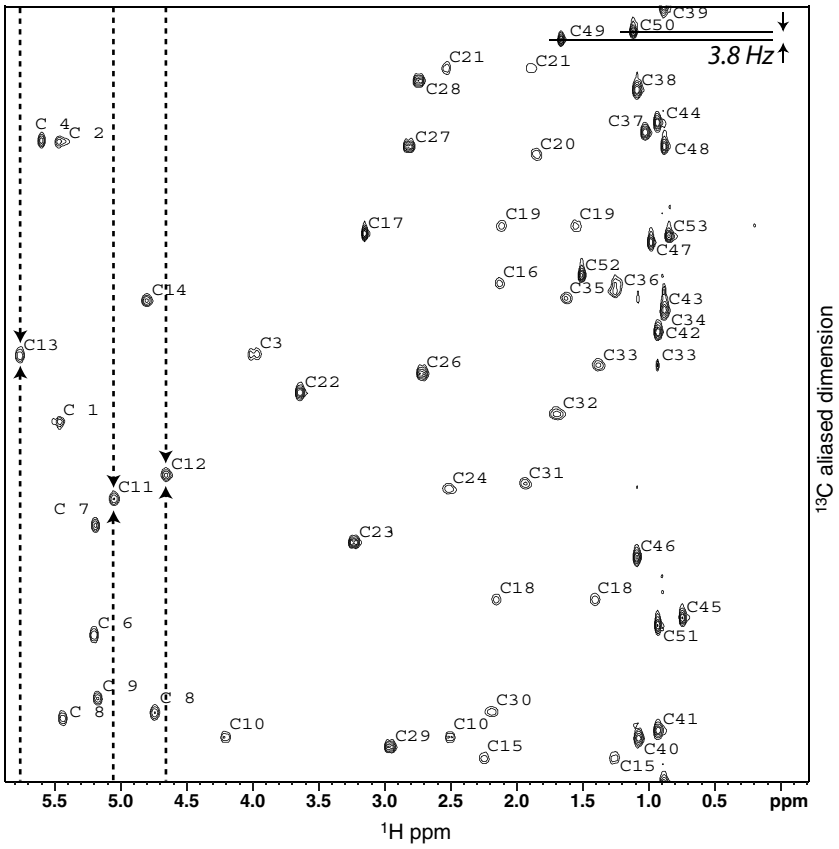


Fig. 9. Aliased HSQC spectrum of cyclosporine A obtained using COSA based on the list of signals of a DEPT spectrum and data from the full-width spectrum of Fig. 8. The highly resolved 2.92 ppm spectrum is obtained with only 336 time increments. In contrast to the full spectrum, all the signals can be unambiguously assigned. The arrows highlight the resolution of carbons C(11), C(12), and C(13). Even C(49) and C(50), distant by a mere 3.8 Hz in the carbon dimension are well resolved in F1. The maximal t_1 of the experiment was 456.9 ms.

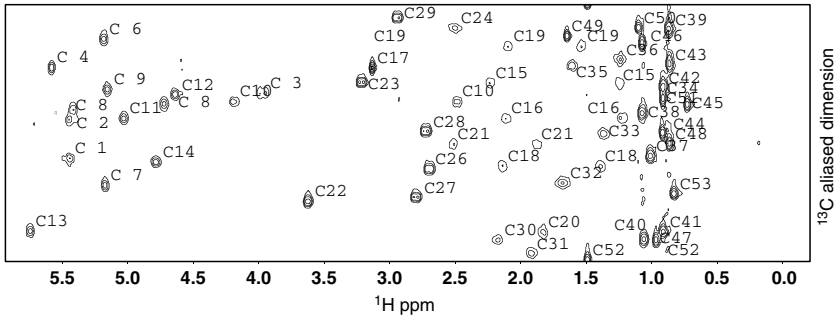


Fig. 10. Aliased HSQC spectrum of cyclosporine A obtained using COSA based on a list of signals of a DEPT spectrum and data from the full-width (Fig. 8) and aliased spectrum (Fig. 9). The 0.96 ppm spectrum is obtained with only 120 time increments. The maximal t_1 was 499.5 ms.

mination. HSQC–TOCSY and HSQC–NOESY experiments acquired using spectral aliasing were of great help in eliminating ambiguities due to overlap in the COSY and NOESY spectra [19]. The full potential of these experiments could only be exploited as a result of the ability of COSA to permit the acquisition of maximal carbon resolution spectrum within reasonable acquisition times.

The acquisition of a small series of selective-TOCSY–HSQC [20] becomes quite practical when a highly resolved HSQC map can be quickly acquired. This allows one to identify different coupling networks in the different sugar-ring units of carbohydrates, or residues in polypeptides when an entry point for magnetization is available. Fig. 11 shows how easy the assignment of the three-ring units of melezitose (see Scheme 1) [21] can be when using COSA applied to selective-TOCSY–HSQC experiments.

The potentially most interesting application of COSA is the acquisition of an extended series of 2D experiments. Relaxation measurements based on HSQC edition which is very common in biomolecular NMR can be applied much more easily to small molecules using COSA. The same is true for diffusion measurements in mixtures when

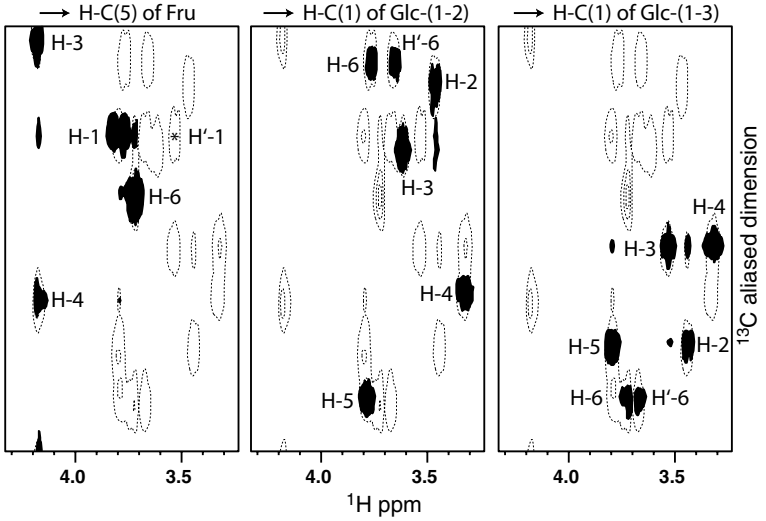
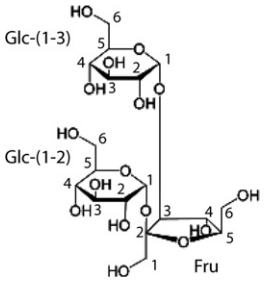


Fig. 11. Illustration of the signal assignment of the different sugar rings of melezitose using series of selective-TOCSY–HSQC spectra. The superposition of selective-TOCSY–HSQC spectra (filled contours) with selective excitation of isolated protons over the aliased HSQC spectrum (dotted contours) makes it obvious which sugar unit each signal belongs to. The spectral width in the carbon dimension was 1.195 ppm (150.36 Hz at 125 ^{13}C Larmor frequency) and all signals were resolved with only 26 time increments. The star indicates the only proton which did not receive magnetization after the 200 ms TOCSY mixing because of unfavorable coupling constants with other protons and remoteness from the source of magnetization.



Scheme 1.

1D spectra show too much overlap. We have developed a CT–HSQC–IDOSY sequence where the diffusion and t_1-evolution times are combined [10]. For these applications, acquisition times are reduced from many days, a duration which is unacceptable in most NMR laboratories, to a few hours.

3. Discussion

3.1. The compromise between the number of time increments and maximal t_1

It was mentioned earlier that the spectral width SW_a associated with the smallest value of $N_a = N_t$ may correspond to longer than needed maximal t_1. Fig. 12 illustrates, in the case of the two anomers of glucose, a trend which is general. The values of SW_a, determined for different numbers of time increments $N_a > N_t$ increases discretely. Each value of SW_a corresponds to a different aliasing pattern that is often the same for more than one value of N_a. Fig. 12b shows that the maximal t_1 for a given set of acquisition conditions $t_{1,max} = N_a/(2SW_a)$ increases linearly with

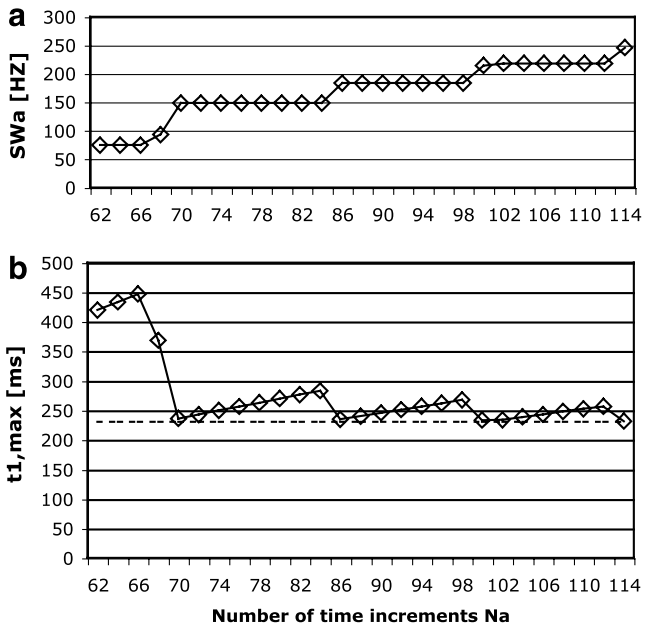


Fig. 12. (a) Comparison of the optimal spectral width SW_a calculated for a number of time increments $N_a > N_t$. (a) Each diamond indicates a set of acquisition conditions (SW_a; N_a) separating all signals. (b) Maximal t_1 of each set of acquisition parameters shown in (a).

N_a for a given overlap pattern, but decreases abruptly when the overlap patterns change. For larger values of N_a, the changes are less and less important and $t_{1,max}$ converges towards the lower boundary (dotted line in Fig. 12b) corresponding to $t_{1,max}$ for the full spectrum resolving all signals.

When relaxation is a concern, the most interesting acquisition conditions therefore correspond to those with sufficiently small $t_{1,max}$. The choice among them will depend on the T_2 relaxation rate acting during t_1. Relaxa-

tion can be neglected until $t_{1,\max} > -\ln(\ 80\%\ )\ T_2 = 0.22\ T_2$ because this corresponds to a signal reduction of less than 20% in the last FID compared to the first. The SNR will be severely diminished when $t_{1,\max} > \ -\ln(10\%)\ T_2 = 2.3\ T_2$, that is when only 10% of magnetization is left in the last FID. A rule of thumb is therefore to keep $t_{1,\max}$ below T_2. When no limit is specified in the optimization program, the upper limit for $t_{1,\max}$ is 10% above the $t_{1,\max}$ of the full spectrum resolving all signals, but this value can be changed. Note that when specifying an upper limit to $t_{1,\max}$, close pairs of signals may not be resolvable. Given the criterion for overlap stated in Annex 1, pairs with less than $1/t_{1,\max}$ Hz free space between their half-height cannot be resolved. In that case, the optimization program gives, as a warning message, a list of the pairs of signals for which resolution will not be attempted during optimization.

A brief comment is merited on situations where the number of time increments proposed by the optimization program is insufficient for the experiment to provide sufficiently high signal-to-noise ratio (SNR). The obvious solution would be to increase the number of times each increment is acquired. An alternative consists of acquiring the spectra with $N = nN_a$ and $SW = nSW_a$, where n is an integer increasing the number of acquired FID sufficiently for the SNR to suffice. Note than when the value of n is larger then the reduction factor of the number of time increments proposed by the optimization, aliasing is simply useless.

3.2. How to handle aliased spectra?

The manipulation of aliased spectra may seem more difficult than the case where spectra have the F1 axis corresponding to true ^{13}C chemical shifts. Unfolding aliased spectra should solve the problem; but besides the fact that it results in very large data matrices, the unfolding process needs, for correct signal positioning, criteria that risk artifacts or weak signals being misidentified and placed in the wrong position. In fact, aliased spectra turn out to be quite easy to handle and in some respects are even preferable to unfolded ones. First, a better spread of signals along the indirect dimension requires much fewer zooms into the signal-rich regions than normal spectra usually require. Moreover, the replacement of chemical shift with direct labeling of the carbon in the F1 dimension is quite simple when input of the optimization is accessible to the NMR display software. From the deceptive signal position v_a the correct chemical shift is $v_0 = v_a \pm nSW_a$. The value of n can be unambiguously determined since the optimization ensures that *the* correct value of n is the one for which $v - v_0^i$ is minimal, with i ranging through all the carbons. This means that synchronized displacement of a mouse-driven pointer in an aliased spectrum would make jumps in the 1D spectrum as the value of n changes. The simplest solution would be for the NMR software simply to replace the chemical shift scale in the F1 dimension, with ticks indicating the position of each carbon when spectra acquired using spectral aliasing are detected. All that the display software needs is the list of carbon chemical shifts used for the optimization.

In the case where proton chemical-shift information has been used for the optimization, i.e. when some positions in the carbon dimension may have more than one assignment candidate, the NMR software should indicate *into* the spectrum the carbon each signal corresponds to. To do this, it should be given the data used for optimization including the proton dispersion list, in order to determine, for each carbon, the proton chemical shift ranges they may cover. The program could then simply draw a line running along the relevant position in the carbon dimension, covering the range of chemical shifts in the proton dimension and having a label to indicate its carbon number. When the NMR software is not implemented for aliased spectra, one can overlap a mask giving the signal position of each carbon in the aliased spectrum on top of the resulting spectrum. Such a mask can be generated by saving the main window of the optimization program in an image file and superimposing it over the aliased spectrum using a drawing program.

3.3. Generating proton dispersion lists

In order to take advantage of proton dispersion to increase the efficiency of COSA, our program must obtain information about 2D spectra. This should be done by a routine in the NMR display software. It should read the 2D spectrum data matrix and, for each carbon found in the 1D spectrum, list the parts of the cross-section along the proton dimension that are above the lowest contour level. Note that if the spectrum has a low resolution in F1, carbons with small differences in chemical shifts may have equal or quite similar proton dispersion. This is not a problem; it only indicates that these carbons will need to be resolved in the aliased spectrum. Such a list could also be made using the combination of a list of carbon chemical shifts stemming from a 1D carbon spectrum and a two-dimensional peak-picking list providing signal position and widths in both dimensions of the low-resolution spectrum. But if the 2D peak list does not describe the full content of the spectrum, any missed signals may result in incomplete proton dispersion lists and assignment problems when working on the aliased spectrum.

3.4. Pulse sequence compatibility with spectral aliasing

Spectral aliasing can be applied with the vast majority of NMR sequences. Two potential problems should be mentioned. First, sequences with double-quantum evolution during t_1 should be avoided when possible. For example, HMQC and HMBC spectra acquired at high resolution reveal a coupling structure in the carbon dimension. This may not be a problem as long as the coupling patterns do not overlap with other signals, but as a general rule, one should choose sequences with single-quantum evolution

during t_1 (HSQC, etc.) so that carbons appear as singlets with the advantage of higher amplitude. Note that we did not encountered any problems due to extended coupling structures of HMBC spectra because their center of symmetry are easily found and are tilted in a manner that avoids overlap.

The second problem concerns sequences in which a significant part of the t_1 evolution time is gradient coded. Molecular diffusion during the relatively long t_1 may significantly reduce signal amplitude and therefore reduce both signal-to-noise ratio and resolution. This is especially problematic with small molecules in low-viscosity solvents like $CDCl_3$. Modifications of the pulse programs to avoid this problem should pose no difficulty. Modification of the HMBC has been given elsewhere [13].

3.5. Compatibility with other processing techniques

Aliased spectra can, in principle, be processed using any special technique. Whether it is used to further improve resolution, increase signal-to-noise ratio, eliminate solvent, etc. the fact that spectral widths are reduced should cause no difficulties. One only has to bear in mind that experiments acquired using spectral aliasing have longer $t_{1,\max}$ than full-width spectra, making the signal amplitude more likely to show a clear reduction of amplitude along the t_1 dimension. This may reduce the performance of the method but will make it ineffective only if it cannot accommodate damped signals along t_1. The method including a list of carbon chemical shifts as "prior knowledge" should take values calculated using Eq. (1) in the calculation and the original values when displaying the results.

The case of forward linear prediction may be discussed as an example. One must first estimate the expected resolution improvement factor p by taking into account to the signal-to-noise ratio, relaxation properties and the $t_{1,\max}$ of the optimized conditions. If one uses linear prediction to decrease the experimental time, one simply acquires a number of time increments equal to the number proposed by the optimization program divided by p. Potentially more interesting would be to further increase the resolution of the aliased spectrum. In such a case, the optimized conditions can be used directly keeping in mind the above-mentioned problem of signal damping in t_1. A resolution enhancement factor of two has been obtained in aliased CT–HSQC–IDOSY experiments[10]. The third possibility consists of changing the signal-width scaling factor of the optimization program and using $k = 0.5/p$ [see Annex 1]. This allows more signals to be resolved with even less time increments, but it comes with the inconvenience that the optimization program does not attempt to resolve better signals that are more easily resolvable than pairs of close signals. This means that the aliased spectrum may in fact require linear prediction to resolve signals that are far apart in a full-width spectrum; a situation probably to be avoided. The second possibility should therefore be preferred to the third because the additional inconvenience does not compensate for the reduction in acquisition time.

Two-dimensional spectra with resolutions higher than 1D carbon spectra could be obtained using the following procedure. COSA applied to a peak list including broad signals hiding pairs of unresolved signals would provide a spectral width resulting in an ideal spread along F1. The number of time increments should then be set to the maximum that relaxation allows. This value is usually larger than TD_a. Complete sampling of the t_1 dimension should provide ideal conditions for special processing techniques to resolve overlapping signals.

4. Experimental

The optimization program consists in a java application accessible from http://rmn.unige.ch/simplealias. The ^{13}C chemical shifts should be given to three significant digits in order to provide sufficient precision; two digits, corresponding to a precision close to ± 0.5 Hz, biases the optimization and may cause difficulties in identifying carbons when the resolution is maximal. The signal widths entered in the optimization program can either be given as a single average value or in the form of a list for each carbon. In the latter case, one should make sure that the spectra contain a sufficiently large number of points for the signal widths to be reasonably well defined, and subtract the line-broadening parameter used for the processing. The optimization program assumed cosine squared processing and used the additivity of line widths (see Eq. (A.2)). The spectra used as the source of carbon chemical shifts and the optimized two-dimensional spectra should be acquired at the same temperature in order to avoid chemical shift changes due to temperature differences. With respect to the problem of sample heating, low-power decoupling. [22,23] is usually not necessary, but recommended when working with samples with temperature-sensitive chemical shifts like carbohydrates. No particular attention was paid to the position of the carrier frequency in the carbon dimension. The absence of artifacts means that even signals located at the edges of the spectra can be quantified after window scrolling with NMRpipe for example [24]. Should artifacts become a problem, one could easily adjust the carrier position to avoid having signals in regions with artifacts.

The ^{13}C database was generated using spectra measured on a Bruker 400 MHz ^1H Larmor frequency spectrometer equipped with a direct-detection probe. The lists of carbons stem from DEPT-135 spectra. The low-resolution HSQC spectra were acquired using 256 time increments and 166 ppm spectral widths. The experiments on the 50 mg cyclosporine in 700 μl C_6D_6 were recorded on a Bruker 500 MHz ^1H Larmor frequency spectrometer equipped with an inverse-detection probe. The full-width spectrum was acquired using 256 time increments, 240 ppm carbon spectral width and two scans per increment. The selective HSQC–TOCSY were acquired on an 80 mg melezitose in 700 μl D_2O using 200 ms TOCSY mixing time.

Acknowledgments

The author wishes to thank André Pinto and Jean-Paul Saulnier for spectrometer management and the *Department de l'Instruction Publique* of the State of Geneva for funding and Straco Translation team, especially Mr.Gordon Thomson for editing assistance.

Annex 1

The line widths at half-height lw of any pair of signals i and j determine the minimal difference in resonance frequency

$$\Delta\nu^{ij}_{\min} = k\ (w_i + w_j) \tag{A.1}$$

that the signals should have in order to be resolved. The factor k determines the severity of the criterion for "being resolved". When $k = 0.5$, the value used in this paper and default value of the optimization program, the pair of signals are fairly distinguishable. Assuming a pair of normalized lorentzian shape of equal line widths, the height between the peaks is 80% of the maximal height. Note that other line shapes or mixed line shapes caused by the processing function would display a similar resolution. Depending on the application, one may wish to choose other values for k. When signals have to be quantified, a value of $k = 4$ corresponds to a signal amplitude between the signals of 0.03 and ensures a negligible contribution by the integral of one signal to the integral of the second when regions are cut at the mid point. One can also use values of k smaller than 0.5 if linear prediction or other processing methods are used to increase resolution.

The line widths can be approximated to be the sum of two contributions:

$$\mathrm{lw} \cong \mathrm{lw}^n + \mathrm{lw}^p \tag{A.2}$$

where lw^n is the natural line width which is determined by relaxation properties and field inhomogeneity and lw^p is the additional broadening due to processing in the indirect dimension. Assuming squared cosine processing and no attempts are made to increase resolution using linear prediction,

$$\mathrm{lw}^p = 1.000/t_{1,\max} \tag{A.3}$$

where $\mathrm{t}_{1,\max}$ is the longest t_1 evolution time of the experiment. For the cosine function, the numerator is 0.820 and with the sinc function (corresponding to no processing) it is 0.603. The additivity of line widths is only strictly correct for Lorentzian shapes; when using the above mentioned window functions, $\mathrm{lw}^n + \mathrm{lw}^p$ overestimates lw if $\mathrm{lw}^n \cong \mathrm{lw}^p$.

Combining Eqs. (A.1)–(A.3) allows one to determine the evolution time

$$t_{1,\max} = 1.0/(\Delta\nu^{ij} - \mathrm{lw}^n_i/2 - \mathrm{lw}^n_j/2) \tag{A.4}$$

needed to resolve signals i and j. Note that the content within parenthesis corresponds to the free space between the two peaks. In order to resolve signals with a 2 Hz gap between them, $t_{1,\max}$ has to be 500 ms.

For a full carbon spectrum, say 200 ppm at 125 MHz ^{13}C Larmor frequency, the spectral width SW corresponds to 25 kHz. The Nyquist condition imposes a dwell time

$$\Delta_t = 1/(2\mathrm{SW}) \tag{A.5}$$

i.e. $\Delta_t = 0.02$ ms in the case considered. The number of time increments is given by the ratio:

$$N = t_{1,\max}/\Delta_t. \tag{A.6}$$

One therefore needs 25,000 time increments to resolve carbons with a 2 Hz space between them in a 200 ppm spectrum acquired on a 500 MHz ^{1}H Larmor frequency spectrometer.

Other considerations about the interdependence between spectral width, number of time increments and maximal t_1 evolution time can be found elsewhere [11].

References

[1] E. Kupče, T. Nishida, R. Freeman, Hadamard NMR spectroscopy, Prog. Nucl. Magn. Reson. Spectrosc. 42 (2003) 95–122.
[2] R. Freeman, E. Kupče, New methods for fast multidimensional NMR, J. Biomol. NMR 27 (2003) 101–114.
[3] P. Koehl, Linear prediction spectral analysis of NMR data, Prog. Nucl. Magn. Reson. Spectrosc. 34 (1999) 257–299.
[4] W.F. Reynolds, M. Yu, R.G. Enriquez, I. Leon, Investigation of the advantages and limitations of forward linear prediction for processing 2D data sets, Magn. Reson. Chem. 35 (1997) 505–519.
[5] W.F. Reynolds, S. McLean, L.-L. Tay, M. Yu, R.G. Enriquez, D.M. Estwick, K.O. Pascoe, Comparison of ^{13}C resolution and sensitivity of HSQC and HMQC sequences and application of HSQC-based sequences to the total ^{1}H and ^{13}C spectral assignment of clionasterol, Magn. Reson. Chem. 35 (1997) 455–462.
[6] M. Groth, J. Malicka, C. Czaplewski, S. Oldziej, L. Lankiewicz, W. Wiczk, A. Liwo, Maximum entropy approach to the determination of solution conformation of flexible polypeptides by global conformational analysis and NMR spectroscopy—application to DNS1-c-[D-A2bu2, Trp4,Leu5]-enkephalin and DNS1-c-[D-A2bu2, Trp4, D-Leu5]enkephalin, J. Biomol. NMR 15 (1999) 315–330.
[7] K.M. Wright, Maximum entropy methods in NMR data processing, in: D.N. Rutledge (Ed.), Signal Treatment and Signal Analysis in NMR, Elsevier, Amsterdam, 1996, pp. 25–43.
[8] H. Barjat, G.A. Morris, A.G. Swanson, A three-dimensional DOSY–HMQC experiment for the high-resolution analysis of complex mixtures, J. Magn. Reson. 131 (1998) 131–138.
[9] A.V. Buevich, J. Baum, Residue-specific real-time NMR diffusion experiments define the association states of proteins during folding, J. Am. Chem. Soc. 124 (2002) 7156–7162.
[10] B. Vitorge, D. Jeannerat, NMR diffusion measurements in complex mixtures using constant-time HSQC–IDOSY and computer-optimized spectral aliasing for high resolution in the carbon dimension, Anal. Chem. 78 (2006) 5601–5606.
[11] D. Jeannerat, High resolution in heteronuclear ^{1}H–^{13}C NMR experiments by optimizing spectral aliasing with one-dimensional carbon data, Magn. Reson. Chem. 41 (2002) 3–17.
[12] A.J. Dunn, P.J. Sidebottom, Fast ^{1}H–^{13}C correlation data for use in automatic structure confirmation of small organic compounds, Magn. Reson. Chem. 43 (2005) 124–131.
[13] D. Jeannerat, D. Ronan, Y. Baudry, A. Pinto, J.-P. Saulnier, S. Matile, NMR characterization of complex *p*-oligophenyl Scaffolds by means of aliasing techniques to obtain resolution-enhanced two-dimensional spectra, Helv. Chim. Acta 87 (2004) 2190–2207.

[14] M. Karplus, Vicinal proton coupling in nuclear magnetic resonance, J. Am. Chem. Soc. 85 (1963) 2870–2871.
[15] D. Ronan, D. Jeannerat, A. Pinto, N. Sakai, S. Matile, New staves for old barrels: regioisomeric (12,22,33,42,53,62,73,82)-*p*-octiphenyl rods with an NMR tag, New J. Chem. 30 (2006) 168–176.
[16] A. Zumbuehl, D. Jeannerat, S.E. Martin, M. Sohrmann, P. Stano, T. Vigassy, D.D. Clark, S.L. Hussey, M. Peter, B.R. Peterson, E. Pretsch, P. Walde, E.M. Carreira, An amphotericin B-fluorescein conjugate as a powerful probe for biochemical studies of the membrane, Angew. Chem. Int. Ed. 43 (2004) 5181.
[17] B. Reif, M. Kock, R. Kerssebaum, H. Kang, W. Fenical, C. Griesinger, ADEQUATE, a new set of experiments to determine the constitution of small molecules at natural abundance, J. Magn. Reson. Series A 118 (1996) 282–285.
[18] H.L. Abdeljebbara, M. Humam, D. Jeannerat, B. Vitorge, S. Amzazi, P. Christen, A. Benjouad, K. Hostettmann, K. Bekkoucheb, Withanolides from Withania adpressa, Helv. Chim. Acta 90 (2007) 346–352.
[19] D. Jeannerat, C. Corminboeuf, D. Muri, E.M. Carreira, Structure determination of slowly exchanging conformers in solution using high-resolution NMR, computational modeling and GIAO chemical shieldings. Application to a synthesis intermediate of erythronolide, (in preparation).
[20] H. Sato, Y. Kajihara, An unambiguous assignment method by 2D selective-TOCSY–HSQC and selective-TOCSY–DQFCOSY and structural analysis by selective-TOCSY–NOESY experiments of biantennary undecasaccharide, Carb. Res. 340 (2005) 469–479.
[21] L. Mäler, J. Lang, G. Widmalm, J. Kowalewski, Multiple-field carbon-13 NMR relaxation investigation on melezitose, Magn. Reson. Chem. 33 (1995) 541–548.
[22] E. Kupče, G. Wagner, Wideband homonuclear decoupling in protein spectra, J. Magn. Reson. Ser. B 109 (1995) 329–333.
[23] R. Fu, G. Bodenhausen, Broadband decoupling in NMR with frequency-modulated "chirp" pulses, Chem. Phy. Lett. 245 (1995) 415–420.
[24] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (1995) 277–293.