

## Coupled Decomposition of Four-Dimensional NOESY Spectra

Sebastian Hiller,<sup>†</sup> Ilghis Ibraghimov,<sup>‡</sup> Gerhard Wagner,<sup>†</sup> and  
Vladislav Y. Orekhov<sup>\*,†,§</sup>

*Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School,  
240 Longwood Avenue, Boston Massachusetts 02115, Elegant Mathematics Ltd., Hanauer Mühle 2,  
66564 Ottweiler, Germany, and Swedish NMR Centre, University of Gothenburg, Box 465,  
Gothenburg, SE 405 30, Sweden*

Received March 15, 2009; E-mail: orov@nmr.gu.se

**Abstract:** Four-dimensional (4D) NOESY spectra provide unambiguous distance information at a resolution that cannot be achieved in fewer dimensions and thus increase the quality of biomolecular structure determination substantially. Since the degree of chemical shift degeneracy increases with protein size, the use of 4D NOESY spectra is particularly important for large proteins. The potential high resolution in 4D spectra cannot be achieved in a reasonable time with conventional acquisition routines that sample the Nyquist grid uniformly. It can, however, be obtained with nonuniform sampling of the data grid, but optimal processing of such data has not yet been established. Here we describe a processing method for a pair of sparsely sampled 4D NOESY spectra, a methyl–methyl and an amide–methyl NOESY, recorded on a perdeuterated protein with protonated isoleucine, leucine, and valine methyl groups. The coupled multidimensional decomposition (Co-MDD) of these two spectra together with a 2D template spectrum results in a substantial increase in sensitivity, evidenced by 50–100% additional cross peaks, when compared to alternative processing schemes. At the same time, Co-MDD allows the use of low sparse levels of 10–15% of the full data grid for NOESY spectra. For the 283-residue integral human membrane protein VDAC-1, which has a rotational correlation time of about 70 ns in detergent micelles, the two 4D Co-MDD NOESYs yielded a total of 366 NOEs, resulting in 139 unambiguous upper limit distance constraints for the structure calculation.

### Introduction

The nuclear Overhauser effect provides the core data for the structure determination of large biomolecules.<sup>1</sup> Cross peaks in NOESY spectra reveal short distances in the spatial structure of the molecule and can thus be used to derive constraints for structure calculations.<sup>2</sup> It is crucial for these calculations that cross peaks are well-resolved for unambiguously identifying the connected protons. Three-dimensional NOESY spectra are widely used, but for large systems it is advantageous to resolve the NOEs in four dimensions, since this drastically reduces the ambiguity of the constraints, leading to better structures.<sup>3–6</sup>

With conventional acquisition and processing schemes, the achievable resolution in 4D NOESYs is low, due to the long acquisition times of 50–100 days that would be needed to sample the time domain at the high resolution achievable with modern high-field spectrometers. This obstacle can be circum-

vented by sampling data only at a fraction of the grid points and using processing techniques other than the standard fast Fourier transformation.<sup>7–18</sup> Multidimensional decomposition (MDD) has been used to process a single 4D methyl–methyl NOESY of a protein in D<sub>2</sub>O solution, sampled nonuniformly at a sparse level of 30%.<sup>19</sup>

<sup>†</sup> Harvard Medical School.

<sup>‡</sup> Elegant Mathematics Ltd.

<sup>§</sup> University of Gothenburg.

- (1) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
- (2) Güntert, P.; Mumenthaler, C.; Wüthrich, K. *J. Mol. Biol.* **1997**, *273*, 283–298.
- (3) Kay, L. E.; Clore, G. M.; Bax, A.; Gronenborn, A. M. *Science* **1990**, *249*, 411–414.
- (4) Clore, G. M.; Gronenborn, A. M. *Science* **1991**, *252*, 1390–1399.
- (5) Güntert, P. *Q. Rev. Biophys.* **1998**, *31*, 145–237.
- (6) Morshauer, R. C.; Zuiderweg, E. R. *J. Magn. Reson.* **1999**, *139*, 232–239.

- (7) Hoch, J. C.; Stern, A. S.; Donoho, D. L.; Johnstone, I. M. *J. Magn. Reson.* **1990**, *86*, 236–246.
- (8) Zhu, G. A.; Bax, A. *J. Magn. Reson.* **1992**, *98*, 192–199.
- (9) Kupce, E.; Freeman, R. *J. Am. Chem. Soc.* **2003**, *125*, 13958–13959.
- (10) Orekhov, V. Y.; Ibraghimov, I.; Billeter, M. *J. Biomol. NMR* **2003**, *27*, 165–173.
- (11) Rovnyak, D.; Hoch, J. C.; Stern, A. S.; Wagner, G. *J. Biomol. NMR* **2004**, *30*, 1–10.
- (12) Armstrong, G. S.; Mandelshtam, V. A.; Shaka, A. J.; Bendiak, B. *J. Magn. Reson.* **2005**, *173*, 160–168.
- (13) Hiller, S.; Fiorito, F.; Wüthrich, K.; Wider, G. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10876–10881.
- (14) Malmodin, D.; Billeter, M. *J. Am. Chem. Soc.* **2005**, *127*, 13486–13487.
- (15) Kazimierczuk, K.; Kozminski, W.; Zhukov, I. *J. Magn. Reson.* **2006**, *179*, 323–328.
- (16) Snyder, D. A.; Xu, Y.; Yang, D.; Brüschweiler, R. *J. Am. Chem. Soc.* **2007**, *129*, 14126–14127.
- (17) Hyberts, S. G.; Heffron, G. J.; Tarragona, N. G.; Solanky, K.; Edmonds, K. A.; Luithardt, H.; Fejzo, J.; Chorev, M.; Aktas, H.; Colson, K.; Falchuk, K. H.; Halperin, J. A.; Wagner, G. *J. Am. Chem. Soc.* **2007**, *129*, 5108–5116.
- (18) Coggins, B. E.; Zhou, P. *J. Biomol. NMR* **2008**, *42*, 225–239.
- (19) Tugarinov, V.; Kay, L. E.; Ibraghimov, I.; Orekhov, V. Y. *J. Am. Chem. Soc.* **2005**, *127*, 2767–2775.

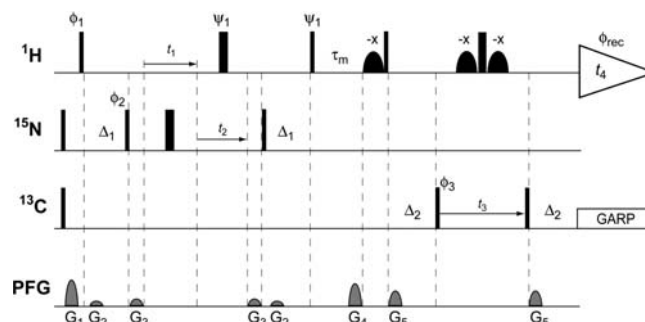
In a previous work, we have shown that coupled (Co)-MDD processing of several spectra can be used to build a model of a hyperdimensional (HD) spectrum.<sup>20</sup> An HD spectrum has a dimensionality higher than that of any of the input experiments. Each additional spectrum in the input HD set contributes at least one new spectral dimension. Thus, in addition to the direct experimentally established correlations, HD spectra contain indirect spin correlations.

In the present work, we explore the ability to use Co-MDD for improving the quality of individual input spectra. By exploiting prior knowledge about shared dimensions, the coupled processing allows high sensitivity in one spectrum to enhance the sensitivity in other spectra. We describe the use of Co-MDD for processing a nonuniformly sampled 4D methyl–methyl NOESY and a nonuniformly sampled 4D amide–methyl NOESY, together with a regularly sampled 2D [<sup>13</sup>C,<sup>1</sup>H]-HMQC. The Co-MDD scheme also enables the use of very low sparse levels of 10–15% for the individual experiments. The experiments are recorded in H<sub>2</sub>O solution, since the exchangeable amide signals are needed, and consequently, proper handling of the water magnetization is important in the pulse sequence. We describe the technical details of acquisition and processing with the new method on the background of the previously published de novo structure determination of the integral human membrane protein VDAC-1 in detergent micelles, during which these experiments were used (PDB code 2K4T, BMRB code 16381).<sup>21,22</sup> In LDAO micelles in aqueous solution, VDAC-1 has a rotational correlation time of 70 ns, as determined by the TRACT method.<sup>23</sup> In the meantime, two crystal structures of VDAC-1 have become available, providing an unbiased reference point for analyzing the 4D NOESY experiments.<sup>24,25</sup> We compare the Co-MDD NOESYs with spectra resulting from the same raw data using alternative processing schemes and analyze the completeness and correctness of the observed NOE cross peaks with the three-dimensional protein structure.

## Methods

**Sample Preparation.** Human VDAC-1(1–283)-Leu-Glu-His<sub>6</sub> was recombinantly expressed in BL21(DE3) cells on [U-99%-<sup>2</sup>H,<sup>15</sup>N]-M9 medium containing 50 mg/L [4-<sup>13</sup>C;3,3-<sup>2</sup>H<sub>2</sub>]-α-ketobutyrate and 100 mg/L [4,4'-<sup>13</sup>C<sub>2</sub>;3-<sup>2</sup>H]-α-ketoisovalerate.<sup>26</sup> The protein was isolated, purified, and refolded into perdeuterated lauryldimethylamine oxide (LDAO) micelles.<sup>21</sup> The final sample conditions were 25 mM Na·PO<sub>4</sub>, 5 mM DTT, pH 6.8, 310 mM [U-99%-<sup>2</sup>H]-LDAO, 0.5 mM [U-99%-<sup>2</sup>H,<sup>15</sup>N; 99%-<sup>1</sup>H<sup>δ</sup>,<sup>13</sup>C<sup>δ</sup>-IL; 99%-<sup>1</sup>H<sup>γ</sup>,<sup>13</sup>C<sup>γ</sup>-V]-VDAC-1. Perdeuterated LDAO was purchased from FB Reagents, Cambridge, MA (www.fbreagents.com), all other isotopes from Cambridge Isotopes, Andover, MA.

**Acquisition.** All experiments were recorded on a Bruker Avance 900 spectrometer equipped with a cryogenic probe. A



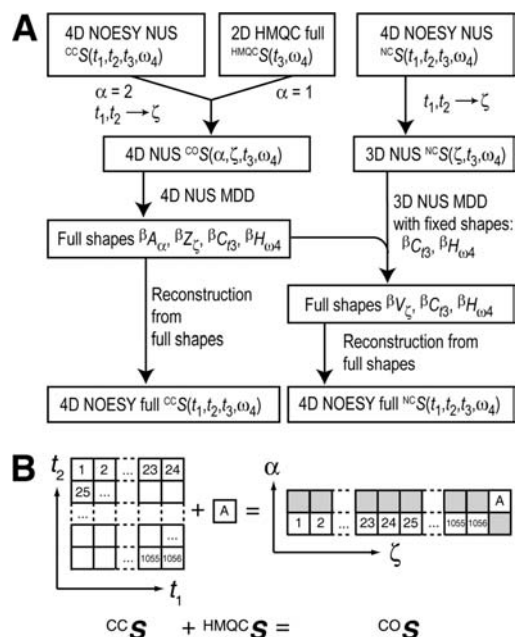
**Figure 1.** Pulse sequence for the 4D NUS-Co-MDD-<sup>15</sup>N-HMQC-[<sup>1</sup>H,<sup>1</sup>H]-NOESY-<sup>13</sup>C-HMQC experiment. Radio-frequency pulses were applied at 4.7 ppm for <sup>1</sup>H, 118.0 ppm for <sup>15</sup>N, and 17.3 ppm for <sup>13</sup>C. Narrow and wide bars represent 90° and 180° pulses with rectangular shape and high power at maximal power, respectively. Pulses with a bell shape on <sup>1</sup>H are water flip-back pulses, applied as Gaussian shapes, with 1 ms duration. The last three <sup>1</sup>H pulses, a WATERGATE element,<sup>47</sup> are centered with respect to the sequence Δ<sub>2</sub>–90°–t<sub>3</sub>–90°–Δ<sub>2</sub> on the <sup>13</sup>C channel, so that its water flip pulses overlap with the delays Δ<sub>2</sub> for short t<sub>3</sub> times. Decoupling using GARP<sup>48</sup> on <sup>13</sup>C is indicated by a rectangle. t<sub>4</sub> represents the acquisition period. τ<sub>m</sub> is the NOESY mixing time. On the line marked PFG, curved shapes indicate sine bell-shaped pulsed magnetic field gradients applied along the z-axis with the following durations and strengths: G<sub>1</sub>, 1000 μs, 35 G/cm; G<sub>2</sub>, 200 μs, 6 G/cm; G<sub>3</sub>, 300 μs, 9 G/cm; G<sub>4</sub>, 800 μs, 30 G/cm; G<sub>5</sub>, 600 μs, 20 G/cm. Phase cycling: φ<sub>1</sub> = x, φ<sub>2</sub> = {x, -x}, φ<sub>3</sub> = {x, x, -x, -x}, φ<sub>rec</sub> = {-x, x, x, -x}, Ψ<sub>1</sub> = x + 45°, all other pulses = x. Fixed delays were Δ<sub>1</sub> = 3.8 ms and Δ<sub>2</sub> = 3.5 ms. Quadrature detection with States-TPPI<sup>49</sup> for the indirect dimensions was achieved with the angles φ<sub>1</sub>, φ<sub>2</sub>, and φ<sub>3</sub> for t<sub>1</sub>, t<sub>2</sub>, and t<sub>3</sub>, respectively.

3D [<sup>1</sup>H,<sup>1</sup>H]-NOESY-<sup>13</sup>C-HMQC and a 2D [<sup>13</sup>C,<sup>1</sup>H]-HMQC were recorded with conventional sampling and processing. The 4D <sup>13</sup>C-HMQC-[<sup>1</sup>H,<sup>1</sup>H]-NOESY-<sup>13</sup>C-HMQC<sup>27,19</sup> was collected following a random sampling schedule corresponding to 11.5% of the full time domain grid defined by (24, 44, 96, 1024) complex points (11620 of 101376 time domain points) for (ω<sub>1</sub>(<sup>1</sup>H), ω<sub>2</sub>(<sup>13</sup>C), ω<sub>3</sub>(<sup>13</sup>C), ω<sub>4</sub>(<sup>1</sup>H)) with sweep widths of (1200, 3400, 3400, 12626 Hz). The 4D <sup>15</sup>N-HMQC-[<sup>1</sup>H,<sup>1</sup>H]-NOESY-<sup>13</sup>C-HMQC (Figure 1) was collected following a random sampling schedule corresponding to 15% of the full data grid defined by (28, 28, 96, 1024) complex points (11200 of 75264 time domain points) for (ω<sub>1</sub>(<sup>1</sup>H), ω<sub>2</sub>(<sup>15</sup>N), ω<sub>3</sub>(<sup>13</sup>C), ω<sub>4</sub>(<sup>1</sup>H)) with spectral widths of (3600, 2900, 3400, 12626 Hz). The sample schedules were created using the *nussampler* routine from the MDDNMR software package with nonbiased probability density.

Since the two spectra are recorded here for a large protein in H<sub>2</sub>O solution, water flip-back pulses were used.<sup>28</sup> The preferred way for water handling during the long NOE mixing time of 200 ms in both NOESY experiments is to utilize radiation damping. This is achieved by phase shifting two <sup>1</sup>H-pulses by 45° relative to the first one (Figure 1), ensuring that the water magnetization is never in the -z state at the onset of the NOESY mixing time.<sup>29</sup> For both pulse sequences, the intrinsic water flip back rates on a 900 MHz spectrometer with a cryogenic probe were found to be 80–90%, resulting in a steady state water flip back of 60–70%.<sup>30</sup> Throughout the experiments, the <sup>1</sup>H carrier frequency remained on the water resonance and the indirect proton dimension ω<sub>1</sub> with the chemical shift ranges of 0.0–1.3 ppm and 6.7–10.7 ppm were aliased three and one time around the water resonance for the <sup>13</sup>C-HMQC-[<sup>1</sup>H,<sup>1</sup>H]-NOESY-<sup>13</sup>C-HMQC and the <sup>15</sup>N-HMQC-[<sup>1</sup>H,<sup>1</sup>H]-NOESY-<sup>13</sup>C-HMQC, respectively. Further, we found that it was not beneficial to use the available constant time periods on <sup>1</sup>H, since

- (20) Jaravine, V. A.; Zhuravleva, A. V.; Permi, P.; Ibraghimov, I.; Orekhov, V. Y. *J. Am. Chem. Soc.* **2008**, *130*, 3927–3936.  
 (21) Hiller, S.; Garces, R. G.; Malia, T. J.; Orekhov, V. Y.; Colombini, M.; Wagner, G. *Science* **2008**, *321*, 1206–1210.  
 (22) Hiller, S.; Malia, T. J.; Garces, R. G.; Orekhov, V. Y.; Wagner, G. Manuscript in preparation.  
 (23) Lee, D.; Hilty, C.; Wider, G.; Wüthrich, K. *J. Magn. Reson.* **2006**, *178*, 72–76.  
 (24) Bayrhuber, M.; Meins, T.; Habeck, M.; Becker, S.; Giller, K.; Villinger, S.; Vonrhein, C.; Griesinger, C.; Zweckstetter, M.; Zeth, K. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 15370–15375.  
 (25) Ujwal, R.; Cascio, D.; Colletier, J. P.; Faham, S.; Zhang, J.; Toro, L.; Ping, P.; Abramson, J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17742–17747.  
 (26) Goto, N. K.; Gardner, K. H.; Mueller, G. A.; Willis, R. C.; Kay, L. E. *J. Biomol. NMR* **1999**, *13*, 369–374.

- (27) Tugarinov, V.; Hwang, P. M.; Ollerenshaw, J. E.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 10420–10428.  
 (28) Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 12593–12594.  
 (29) Talluri, S.; Wagner, G. *J. Magn. Reson. B* **1996**, *112*, 200–205.  
 (30) Hiller, S.; Wider, G.; Etezady-Esfarjani, T.; Horst, R.; Wüthrich, K. *J. Biomol. NMR* **2005**, *32*, 61–70.



**Figure 2.** Data flow and data generation in the Co-MDD processing scheme. (A) Data flowchart. The conventional Fourier transform steps are not shown. See the Methods section for details. (B) Generation of the dimensions  $\alpha$  and  $\zeta$  in the 4D data set  $^{CC}S$  from the dimensions  $t_1$  and  $t_2$  in the 4D data set  $^{CC}S$  and the 2D data set  $^{HM}QS$ . Squares correspond to complex points in the data grids. Corresponding points are marked by index numbers or the letter “A”. Gray points have zero intensity. Dimensions  $t_3$  and  $\omega_4$  have been omitted in this drawing.

placing a pair of  $180^\circ$  pulses on  $^1H$  and  $^{15}N$  in the  $[^{15}N, ^1H]$ -HMQC element converts CRINEPT to INEPT type transfer, resulting in a signal loss of 20–30%.<sup>31,32</sup>

**Coupled Processing of 2D and 4D Spectra.** Coupled MDD processing was performed using the MDDNMR software package,<sup>33</sup> which is available on request from V.Y.O. The data flowchart for the coupled processing of two 4D NOESY and one 2D HMQC spectra is shown in Figure 2. First, the acquisition dimension is processed using the nmrPipe software<sup>34</sup> for all three spectra: The time domain signal is multiplied by a shifted squared sine-bell weighting function, the number of points is doubled by zero-filling, Fourier transform is performed and the region of interest is extracted (for VDAC-1 the range 1.3–0.2 ppm). The resulting data sets are  $^{CC}S(t_1, t_2, t_3, \omega_4)$  for the 4D  $^{13}C$ -HMQC- $[^1H, ^1H]$ -NOESY- $^{13}C$ -HMQC,  $^{NC}S(t_1, t_2, t_3, \omega_4)$  for the 4D  $^{15}N$ -HMQC- $[^1H, ^1H]$ -NOESY- $^{13}C$ -HMQC, and  $^{HM}QS(t_3, \omega_4)$  for the 2D  $[^{13}C, ^1H]$ -HMQC. The two dimensions of the 2D  $[^{13}C, ^1H]$ -HMQC spectrum are termed  $t_3$  and  $\omega_4$  to indicate that the line shapes observed in these two dimensions correspond to dimensions  $t_3$  and  $\omega_4$  of the two 4D NOESY spectra. The spectra span four evolution times  $t_1, t_2, t_3$ , and one frequency domain  $\omega_4$ .

For coprocessing of the  $^{HM}QS(t_3, \omega_4)$  and  $^{CC}S(t_1, t_2, t_3, \omega_4)$  spectra, these experimental data sets are approximated by an MDD model:

$$\begin{aligned} ^{HM}QS(t_3, \omega_4) &\approx \sum_{\beta} \theta_{\beta} L_{C\beta}(t_3) L_{H\beta}(\omega_4) \\ ^{CC}S(t_1, t_2, t_3, \omega_4) &\approx \sum_{\beta, \gamma} \lambda_{\gamma, \beta} L_{H\gamma}(t_1) L_{C\gamma}(t_2) L_{C\beta}(t_3) L_{H\beta}(\omega_4) \end{aligned} \quad (1)$$

The summations run over all diagonal and cross-peaks, with the indices  $\gamma$  and  $\beta$  distinguishing methyl groups of origination and destination of magnetization, respectively. The normalized one-dimensional functions  $L$  describe the time or frequency domain signals of the methyl nuclei  $^1H$  or  $^{13}C$  and thus encode their peak positions, phases, and line-profiles. The peak intensities are subsumed in the coefficients  $\theta_{\beta}$  and  $\lambda_{\gamma, \beta}$ .

Substituting,  $Z_{\beta}^{1,1,2} = \sum_{\gamma} \lambda_{\gamma, \beta} L_{H\gamma}(t_1) L_{C\gamma}(t_2)$ ,  $C_{\beta}^{1,3} = L_{C\beta}(t_3)$ , and  $H_{\beta}^{1,4} = L_{H\beta}(\omega_4)$ , eq 1 takes the form

$$^{HM}QS(t_3, \omega_4) \approx \sum_{\beta} \theta_{\beta} C_{\beta}^{1,3} H_{\beta}^{1,4} \quad (2a)$$

$$^{CC}S(t_1, t_2, t_3, \omega_4) \approx \sum_{\beta} Z_{\beta}^{1,1,2} C_{\beta}^{1,3} H_{\beta}^{1,4} \quad (2b)$$

In eqs 2a the factors  $C_{\beta}^{1,3}$  and  $H_{\beta}^{1,4}$  correspond to  $^{13}C$  and  $^1H$  profiles, or “shapes”, associated with methyl group  $\beta$ . A shape  $Z_{\beta}^{1,1,2}$  can be thought of as a 2D  $[^1H, ^{13}C]$  time domain correlation spectrum. A 2D Fourier transform of such a shape would contain a diagonal peak and cross-peaks from all methyl protons  $\gamma$  that are proximal to methyl  $\beta$ .

The evolution times and spectral frequencies are defined on a grid with regular intervals  $\Delta t_1, \Delta t_2, \Delta t_3$ , and  $\Delta \omega_4$

$$t_1 = (p-1)\Delta t_1, t_2 = (q-1)\Delta t_2, t_3 = (m-1)\Delta t_3, \omega_4 = (n-1)\Delta \omega_4 \quad (3)$$

where indexes  $p, q, m, n$  run from 1 to the maximal values  $P, Q, M, N$  defined by the data sizes in dimensions 1–4, respectively. Thus, the shapes  $C_{\beta}^{1,3}, H_{\beta}^{1,4}$  are vectors and  $Z_{\beta}^{1,1,2}$  are matrices with elements  $C_{\beta}^m, H_{\beta}^n$ , and  $Z_{\beta}^{p,q}$ , respectively. In practice, the NMR time domain signal is comprised of both cosine- and sine-modulated components, facilitating quadrature detection. Thus, sizes of the real, time domain shapes  $P, Q$ , and  $M$  have twice as many points as the number of time intervals in corresponding spectral dimensions. MDD operates only with one-dimensional shapes and, as we described it in our previous publications,<sup>19</sup> it is convenient to think about the matrix  $Z$  as a vector with elements  $Z_k$ . The index  $k$  spans all combinations of indexes of  $p$  and  $q$  and runs from 1 to  $K = PQ$ . Similar to  $Z_k$ , we can write  $Z_{\beta}^{1,1,2}$  as a function of one argument  $\zeta$ , which spans the combination of evolution times  $t_1, t_2$  on a defined grid (Figure 2B). After normalizing vectors  $Z_{\beta}^{\zeta}$  and putting the normalization factor to element  $\alpha = 2$  of new vector  $A^{\beta}$ , eq 2b converts to

$$^{CC}S(\zeta, t_3, \omega_4) \approx \sum_{\beta} A_{\alpha=2}^{\beta} Z_{\zeta}^{\beta} C_{\beta}^{1,3} H_{\beta}^{1,4} \quad (4a)$$

The new index  $\alpha = 1$  and  $\alpha = 2$  distinguishes normalization factors for the 2D and 4D experiments, respectively (Figure 2B). Equation 2a is multiplied by an arbitrary element of the shape  $Z_{\beta}^{\zeta}$ , using a new normalization  $A_{\alpha=1}^{\beta} = \theta_{\beta} / Z_{\zeta K1}^{\beta}$  and can thus be rewritten as

$$^{HM}QS(t_3, \omega_4) \approx \sum_{\beta} A_{\alpha=1}^{\beta} Z_{\zeta K1}^{\beta} C_{\beta}^{1,3} H_{\beta}^{1,4} \quad (4b)$$

where in the  $Z$  dimension, all data points from the  $^{HM}QS(t_3, \omega_4)$  spectrum are assigned to point  $\zeta_{K1}$ . In theory, any point from the  $Z_{\beta}^{\zeta}$  shapes can be used. However, for stability of the MDD calculations, we take the point with the index  $K + 1$ , which is not present in the  $^{CC}S(\zeta, t_3, \omega_4)$  data set. Later on for the reconstruction of the complete 4D spectrum, this point in the  $Z_{\beta}^{\zeta}$  shapes is discarded. The two experiments on the left sides of eqs 4a and corresponding models on the right sides can be combined so that the resulting four dimensional data set  $^{CC}S$  is described by a common model:

(31) Wimperis, S.; Bodenhausen, G. *Chem. Phys. Lett.* **1987**, *140*, 41–45.

(32) Riek, R.; Wider, G.; Pervushin, K.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 4918–4923.

(33) Jaravine, V. A.; Orekhov, V. Y. *J. Am. Chem. Soc.* **2006**, *128*, 13421–13426.

(34) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277–293.

$${}^{\text{CO}}S(\alpha, \zeta, t_3, \omega_4) \approx \sum_{\beta} A_{\alpha}^{\beta} Z_{\zeta}^{\beta} C_{t_3}^{\beta} H_{\omega_4}^{\beta} \quad (5)$$

The experimental data  ${}^{\text{CO}}S(\alpha, \zeta, t_3, \omega_4)$  has the size of  $2 \times (K + 1) \times M \times N$  and is sparse because of the nonuniform sampling of the 4D spectrum and because index  $\alpha$  separates the two experiments. It serves as input for MDD calculations, which are performed independently for overlapping subregions along  $\omega_4$ , resulting in the calculation of the complete vectors  $A$ ,  $Z$ ,  $C$ , and  $H$ .<sup>35</sup> For VDAC-1, nine subregions of about 0.11 ppm each, cover the range 1.3–0.2 ppm. For each subregion, MDD is performed with the number of components  $N_c = 1.5N_p + 4$ , where  $N_p$  is the number of peaks automatically picked in this subregion in the 2D [<sup>13</sup>C,<sup>1</sup>H]-HMQC spectrum using the nmrPipe peak picking routine. The number of component for the individual regions varied from 5 to 70. All MDD calculations are performed with the default  $\lambda$  value of 0.005 and with 200 iterations. For each subregion the four-dimensional decomposition results in a set of components enumerated by the index  $\beta$ . Each component consists of four complete real one-dimensional shapes:  $A_{\alpha}^{\beta}$ , shapes contain scaling factors for the peaks in the 4D and 2D spectra;  $Z_{\zeta}^{\beta}$ ,  $C_{t_3}^{\beta}$  time domain shapes correspond to the dimensions  $t_{1,2}$ (<sup>1</sup>H, <sup>13</sup>C) and  $t_3$ (<sup>13</sup>C) of the 4D [<sup>13</sup>C,<sup>1</sup>H]-HMQC-[<sup>1</sup>H,<sup>1</sup>H]-NOESY-<sup>13</sup>C-HMQC spectrum; and  $H_{\omega_4}^{\beta}$ , frequency domain shapes for the dimension  $\omega_4$ (<sup>1</sup>H). Using the procedure described above, the 4D <sup>15</sup>N-HMQC-[<sup>1</sup>H,<sup>1</sup>H]-NOESY-<sup>13</sup>C-HMQC spectrum can be arranged to data set  ${}^{\text{NC}}S(\zeta, t_3, \omega_4)$  and described by the MDD model as

$${}^{\text{NC}}S(\zeta, t_3, \omega_4) = \sum_{\beta} V_{\zeta}^{\beta} C_{t_3}^{\beta} H_{\omega_4}^{\beta} \quad (6)$$

where shapes  $V_{\zeta}^{\beta}$  correspond to the dimensions  $t_{1,2}$ (<sup>1</sup>H, <sup>15</sup>N); i.e., they are analogous to the shapes  $Z_{\zeta}^{\beta}$  in eq 4a. In the decomposition of eq 6, we used the shapes  $C_{t_3}^{\beta}$  and  $H_{\omega_4}^{\beta}$  obtained in the co-MDD decompositions of the data set in eq 5. These shapes were fixed and only shapes  $V_{\zeta}^{\beta}$ , were allowed to be adjusted. With all shapes defined from the MDD calculations, the complete 4D NOESY data sets are reconstructed using eqs 4a and 6 for the 4D [<sup>13</sup>C-HMQC-[<sup>1</sup>H,<sup>1</sup>H]-NOESY-<sup>13</sup>C-HMQC and the 4D [<sup>15</sup>N-HMQC-[<sup>1</sup>H,<sup>1</sup>H]-NOESY-<sup>13</sup>C-HMQC, respectively. Reconstructions for the complete region of interest are produced by combining reconstructions for all subregions. The remaining time domain dimensions of the reconstructed 4D spectra are processed conventionally using Fourier transformation in dimensions 1–3 and linear prediction of 20 points in dimensions 2 and 3, resulting in sizes of  $64 \times 128 \times 256 \times 162$  points.

**Alternative Processing.** Two other processing schemes, discrete Fourier transformation (NUS-DFT)<sup>15</sup> and the original MDD,<sup>35,19</sup> were applied to the same nonuniformly sampled raw data. For the processing with DFT, zero values were assigned to the missing data points. The MDD calculations were performed using the same subregion selection, number of components and  $\lambda$  as for the Co-MDD scheme. For both NUS-DFT and MDD, weighting functions, zero-filling, linear prediction and Fourier transformation parameters for processing of the resulting data set were the same as for the Co-MDD scheme.

## Results and Discussion

**Selection of Magnetization Pathways.** A first step for the recording of 4D NOESYs in a structure determination is the identification of suitable magnetization transfer pathways. For large proteins, the selective labeling of methyl groups in Ile, Leu, and Val residues is well established,<sup>36,26</sup> and together with the amide moieties this labeling scheme enables four possible

magnetization transfer pathways for 4D NOESY spectra, the amide–amide, methyl–methyl, methyl–amide, and amide–methyl. For large systems, such as VDAC-1, the methyl–methyl pathway is the most sensitive experiment among these. The amide–amide pathway is the least sensitive, and it did yield hardly any detectable cross peaks in 4D NOESYs in the case of VDAC-1. Note that in 3D NOESYs, the amide–amide NOEs were above the sensitivity limit and were crucial to establish the backbone topology of the channel (see below). From the 4D amide–methyl and 4D methyl–amide experiment, only one is required, since these spectra are transposed relative to each other. Processing with Co-MDD requires that the experiments share the dimensions  $\omega_3$  and  $\omega_4$ , corresponding to the second heteronuclear element, and thus, the 4D NOESYs methyl–methyl and amide–methyl, together with the 2D [<sup>13</sup>C,<sup>1</sup>H]-HMQC, form a suitable set of experiments for Co-MDD.

The implementation of a desired 4D NOESY pathway into a pulse sequence requires the selection of the optimal 2D heteronuclear editing elements before and after the NOESY mixing time. For methyl groups in large molecular systems, the [<sup>13</sup>C,<sup>1</sup>H]-HMQC experiment has been identified as the most sensitive experiment.<sup>27,19</sup> For the amide groups, we compared the [<sup>15</sup>N,<sup>1</sup>H]-HMQC with the [<sup>15</sup>N,<sup>1</sup>H]-TROSY experiment.<sup>37,32</sup> With the digital resolutions used in the present work, the sensitivity of the [<sup>15</sup>N,<sup>1</sup>H]-HMQC experiment was higher than the [<sup>15</sup>N,<sup>1</sup>H]-TROSY by factors in the range of 2.0–3.0 for different residues of VDAC. We thus found the <sup>15</sup>N-HMQC-[<sup>1</sup>H,<sup>1</sup>H]-NOESY-<sup>13</sup>C-HMQC experiment to be the most favorable implementation for the amide–methyl pathway for VDAC-1 (Figure 1).

**Coprocessing of 2D HMQC and 4D NOESYs.** In the following section, we describe why the presence of the strong diagonal signals in the NOESY spectrum helps to detect small cross peaks in the same spectrum and why a similar effect can be achieved by coprocessing with a reference spectrum or by using precalculated reference shapes. We restrict the present discussion to rationalizing the resulting MDD spectrum. A description of the mathematical MDD algorithm can be found elsewhere.<sup>38</sup> The MDD spectrum is an approximation of a three- or higher dimensional spectrum by sum of a given number of components. Each component is the outer product of three one-dimensional shapes. No assumptions are made about functional forms of the shapes. In other words, each element of a shape is considered unknown by the MDD algorithm. The algorithm ascribes components for describing the most intense features in the data. In addition to signals, experimental data may contain strong artifacts. Moreover, even genuine signals may slightly deviate from the model assumption of the outer shapes product. In general, the MDD algorithm is thus likely to allocate available components to describe the strongest artifacts and the noise intensities instead of finding the weak signals. The weak peaks may be recovered, however, when several signals share line-shapes along all but one spectral dimension. All these peaks are described by a single component. The algorithm has to allocate a component for describing at least the strongest cross-peaks in the group. The strong peaks also define the shapes, notably including the shared ones. The remaining shape, which corresponds to the dimension, where peaks in the group are distinct, contains the strong peaks and all intensities matching

(35) Luan, T.; Jaravine, V.; Yee, A.; Arrowsmith, C. H.; Orekhov, V. Y. *J. Biomol. NMR* **2005**, *33*, 1–14.

(36) Goto, N. K.; Kay, L. E. *Curr. Opin. Struct. Biol.* **2000**, *10*, 585–592.

(37) Pervushin, K. V.; Wider, G.; Wüthrich, K. *J. Biomol. NMR* **1998**, *12*, 345–348.

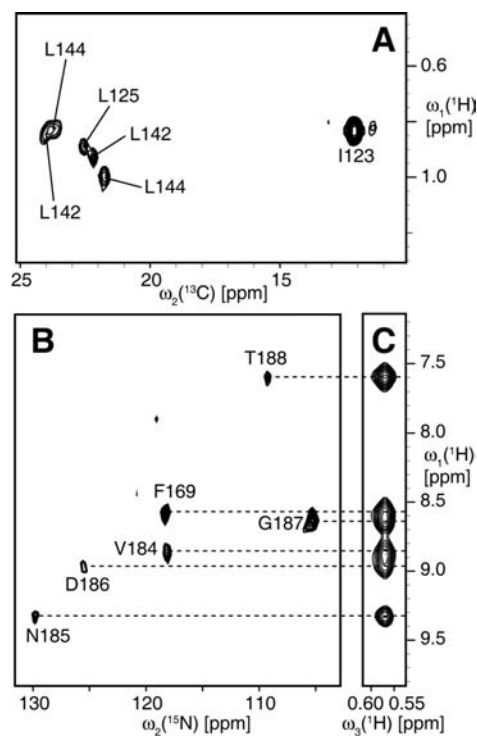
(38) Ibragimov, I. *Numer. Linear Algebra Appl.* **2002**, *9*, 551–565.

the shared shapes, including the weak peaks and noise baseline. A good example for this feature is the 3D  $^{15}\text{N}$ -resolved  $[\text{H}, \text{H}]$ -NOESY-HSQC spectrum,<sup>10</sup> where the diagonal signal and all the cross-peaks to the corresponding amide proton are described by one component. Thus, the key to detect a weak peak is to organize the MDD calculations so that weak cross peak signals are grouped with strong (e.g., diagonal) signals. If the peaks are distinct in more than one MDD dimension, they cannot be grouped in to one component and the weak peaks are lost. Such a case occurs if the 4D NOESY spectrum would be solved directly using four-dimensional MDD. In this work, similar to our previous publications,<sup>19,35</sup> the above formulated “all but one dimension” rule for grouping the peaks is implemented by combining the first two dimensions from the 4D NOESY spectra into one, that is  $Z_\zeta$  and  $V_\zeta$  shapes for the carbon and nitrogen resolved spectra, respectively.

The strong peaks help to detect the weak ones by allocating a component to the right place in the frequency spectrum and defining all but one line shape for these peaks. There are cases, however, when peak groups do not contain sufficiently strong signals. Examples are peak constellations related to the weakest diagonals in the 4D  $^{13}\text{C}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC spectrum or the signals in the 4D  $^{15}\text{N}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC, where diagonal signals are not available at all. Data in these spectra alone are then not sufficient to correctly determine the shared shapes. Previously, we showed that weak peaks can be recovered using predefined shapes obtained in a reference experiment<sup>20</sup> and we use this approach in the present work for processing the 4D  $^{15}\text{N}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC. In order to obtain good results with the fixed shapes technique, the spectra need to be fully compatible with each other, i.e., grids in time and frequency domains, spectral widths, carrier frequencies, signal phases, etc. must be the same in both spectra for the shared dimensions. Moreover, the reference set of shapes should be able to describe all strong features in the processed spectrum, including the artifacts (if any). In this work, the 4D  $^{15}\text{N}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC spectrum was successfully solved using  $C_{13}$  and  $H_{\omega 4}$  shapes from 4D  $^{13}\text{C}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC spectrum. The two 4D spectra matched very well, because their pulse sequences had equivalent  $[\text{H}, \text{H}]$ -HMQC blocks for the signal detection.

For processing of 4D  $^{13}\text{C}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC, the more sensitive 2D  $[\text{H}, \text{H}]$ -HMQC spectrum served as a reference. The MDD algorithm is only applicable to data sets of three and more dimensions and we thus coprocessed the 2D and 4D spectra by combining the two experiments into one four-dimensional data set as described in the Methods section. In this way, methyls with weak diagonal peaks in the 4D experiment are grouped with their stronger signals in the 2D. The 4D  $^{15}\text{N}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC in turn was processed with the fixed shapes from the above processing. In principle, it would be possible to coprocess the 4D  $^{15}\text{N}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC together with the 2D  $[\text{H}, \text{H}]$ -HMQC, or together with the 2D  $[\text{H}, \text{H}]$ -HMQC and the 4D  $^{13}\text{C}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC. However, due to the relatively low sensitivity of the 4D  $^{15}\text{N}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC, we found the fixed shape approach to yield the best result among these possibilities.

**Use of the 4D NOESYs in the VDAC Structure Determination.** The 4D NOESY spectra described here were used during the de novo structure determination of the protein VDAC-1 and, thus, at a time when the correct 3D structure of the molecule was yet unknown. Prior to the analysis of the NOESY



**Figure 3.** 3D and 4D NOESY spectra of the integral membrane protein VDAC-1 in LDAO micelles. (A) Plane from the NUS-Co-MDD- $^{13}\text{C}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC at the position of I123  $\text{C}^9\text{H}^0$ , (B) Plane from the NUS-Co-MDD- $^{15}\text{N}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC at the position of V184  $\text{H}^{21}/\text{C}^{21}$ . Resonance assignments are indicated. (C) Strip from the 3D- $^{13}\text{C}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY at the position of V184  $\text{H}^{21}/\text{C}^{21}$ . Corresponding peaks in (B) and (C) are connected by dashed lines.

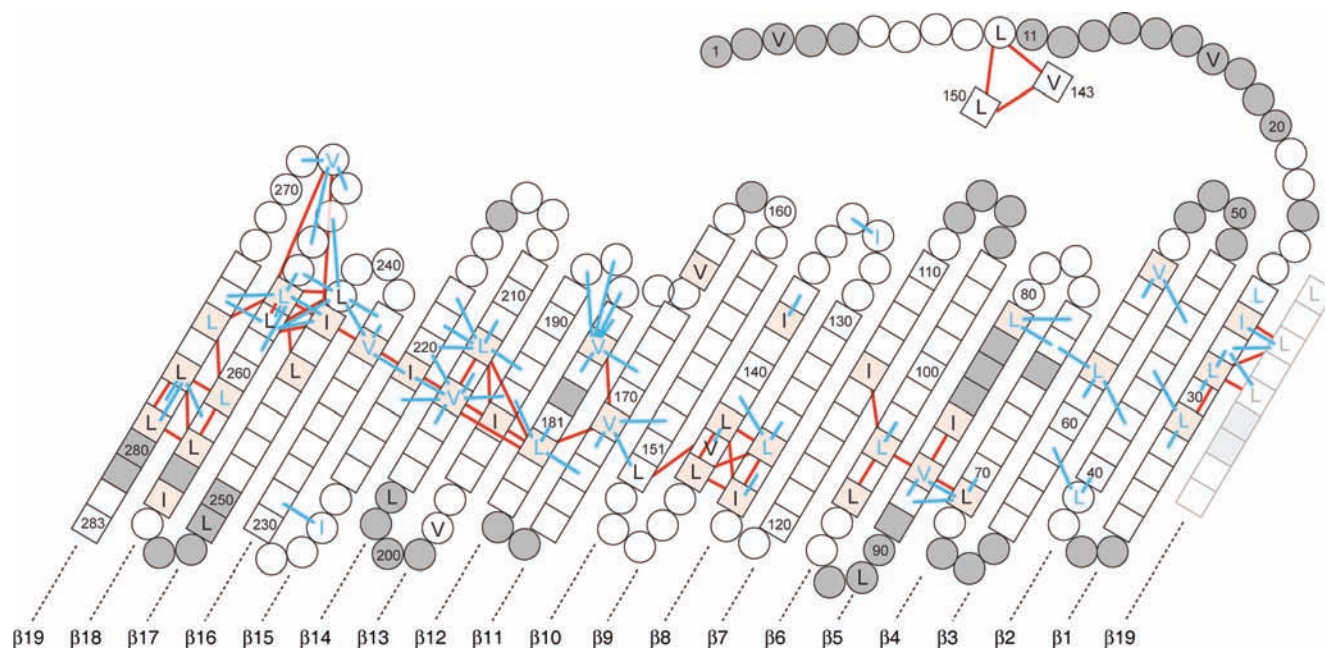
spectra, the sequence-specific assignment of the backbone had been established using HNCA and HNCACB experiments, supported by amino acid specific labeling.<sup>39,21</sup> For the assignment of the ILV methyls, a HMCM[CG]CBCA-COSY experiment was used,<sup>40</sup> but due to the large number of ILV residues and the resulting chemical shift degeneracy, only about 20% of the methyl groups could be assigned using this covalent correlations, while the other methyl assignments remained ambiguous.

In this situation, the two high-resolution 4D NOESY spectra, together with two conventional 3D NOESY spectra, one resolved for the methyls and one for the amides, were used. Strong interstrand  $^1\text{H}^{\text{N}}-^1\text{H}^{\text{N}}$  NOEs were observed in the 3D  $^{15}\text{N}$ -resolved- $[\text{H}, \text{H}]$ -NOESY, establishing a preliminary  $\beta$ -sheet topology, but many NOE cross peak assignments remained ambiguous from the 3D NOESYs alone due to frequent peak overlap (Figure 3). On the basis of this initial structural model, the ambiguous side chain assignments could be resolved, mainly from the 4D NOESY spectra, leading to 96% complete sequence-specific resonance assignment of the methyl groups and subsequently to a large number of methyl–methyl and methyl–amide NOE contacts (Figure 4). In particular, most methyl–methyl NOEs were resolved only in the 4D  $^{13}\text{C}$ -HMQC- $[\text{H}, \text{H}]$ -NOESY- $^{13}\text{C}$ -HMQC, due to the small chemical shift dispersion of 0.7 ppm for the methyl  $^1\text{H}$  resonances, which for most cross peaks led to signal overlap with the strong diagonal peaks in the 3D  $^{13}\text{C}$ -resolved- $[\text{H}, \text{H}]$ -NOESY.

For VDAC-1, the 28 leucine, 12 valine, and 11 isoleucine residues span an NOE network around the  $\beta$ -barrel (Figure 4).

(39) Malia, T. J.; Wagner, G. *Biochemistry* **2007**, *46*, 514–525.

(40) Tugarinov, V.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 13868–13878.



**Figure 4.** Display of all methyl–methyl and methyl–amide NOEs observed in the two 4D Co-MDD-NOESYs for the integral membrane protein VDAC-1 in LDAO micelles. Boxes are amino acid residues in  $\beta$ -sheet secondary structure, and circles denote residues in loops or helical secondary structure. Selected residue numbers, the strand numbers  $\beta_1$ – $\beta_{19}$ , and all residues of type I, L, and V are indicated. Gray: residues with an unassigned amide moiety. White and yellow: residues with an assigned amide moiety. The side chains of yellow and white I, L, or V residues point to the outside and inside of the barrel, respectively. Blue lines are NOEs observed in the 4D  $^{15}\text{N}$ -HMQC- $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC. A blue letter shows an NOE in this experiment from an I, L, or V methyl group to its own amide moiety. Red lines denote NOE contacts observed in the 4D  $^{13}\text{C}$ -HMQC- $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC. Note that a red line can represent up to eight actually observed NOE cross peaks. Residues L150 and V143, which have NOE contacts to residue L10, are plotted twice. Strand  $\beta_{19}$ , which closed the  $\beta$ -barrel by parallel pairing to  $\beta_1$ , is plotted twice.

For all except one  $\beta$  strand pairing ( $\beta_7$ – $\beta_8$ ), the relative orientation of the strands was supported by at least one unambiguous NOE from the 4D NOESY spectra. Of particular structural interest in this protein is the parallel pairing of strands  $\beta_1$ – $\beta_{19}$ , a previously unknown feature of integral  $\beta$ -barrel membrane proteins, and the contact between the N-terminal tail and the inner barrel wall at residues 10, 143, and 150. The parallel pairing of  $\beta_1$ – $\beta_{19}$  was supported by six different 4D NOE contacts and the contact between residues 10, 143, and 150 by eight 4D NOEs (Figure 4; each line in this Figure may represent more than one NOE cross peak). The availability of unambiguous high-resolution 4D NOEs thus had a high impact on confirming the topology and establishing the three-dimensional structure of VDAC-1.

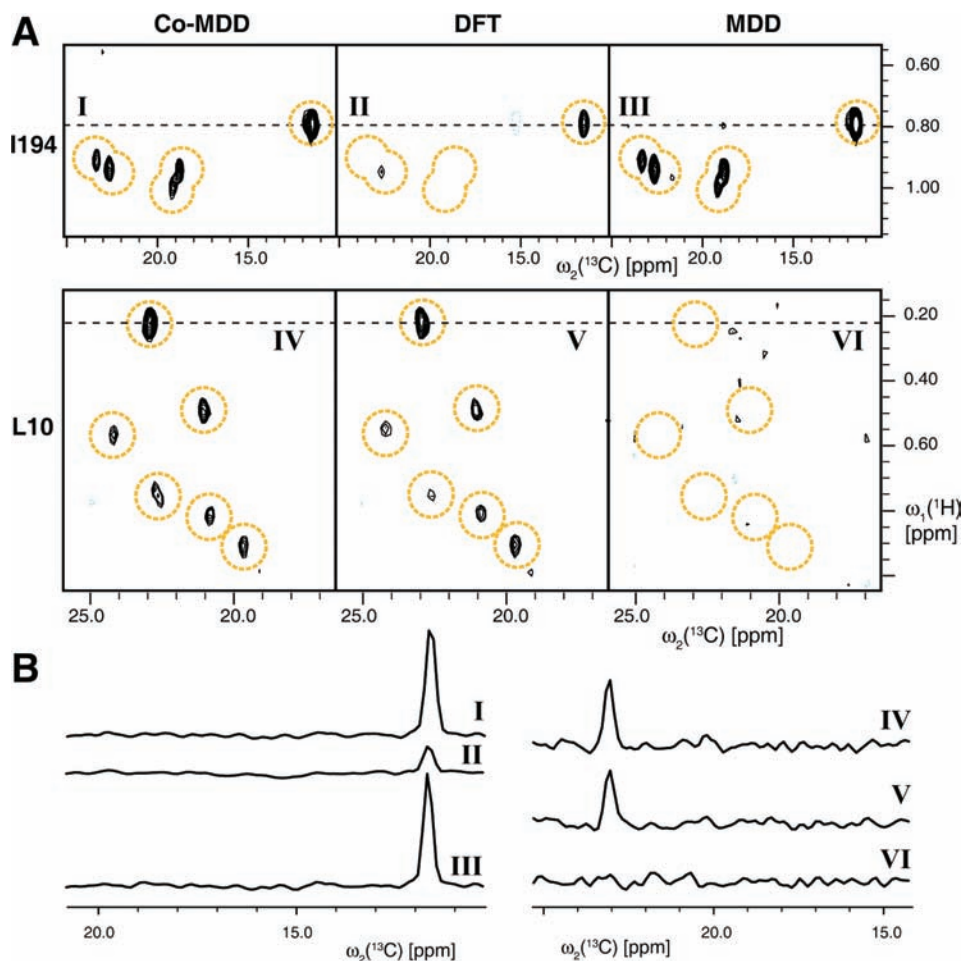
**Assessment of Spectral Quality (1): Sensitivity.** For an evaluation of the sensitivity of the 4D NOESYs, we compare the Co-MDD spectra with spectra obtained using two alternative processing methods of the same nonuniformly sampled raw data: (i) a Fourier transformation of the sparse data, with all missing points set to zero (NUS-DFT), and (ii) processing using MDD for each of the two NOESY data sets individually, i.e., without coupling of the decompositions. It should be noted that methods are currently being developed that might improve the performance of the NUS-DFT method by nonlinear artifact cleaning.<sup>41,18</sup> The current comparison, however, uses the performance of the linear NUS-DFT method as a benchmark. It would be desirable also to compare the performance of the co-MDD analysis with recently introduced method of forward maximum entropy (FM) reconstruction.<sup>17</sup> The comparison, however, is currently not

feasible due to long computational times of the FM reconstruction for the presented data set.

The methyl groups of VDAC-1 can be distinguished in two groups, M1 and M2, according to the result from the three processing methods in the 4D  $^{13}\text{C}$ -HMQC- $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC spectrum (Figure 5). Group M1 was found to contain about 85% of the methyls and for these the 2D planes in the MDD and the Co-MDD spectra are of similar spectral quality, with a slight sensitivity advantage of the Co-MDD data. In the NUS-DFT spectrum however, the planes of these residues have a substantially lower sensitivity, so that many low intensity cross peaks are not distinguishable from the noise. For group M2, which comprises the remaining 15% of the methyl groups, the relative outcome of the three processing schemes is different than for group M1. The 2D planes from group M2 residues did not show peaks in the MDD spectrum, but similar sensitivity in the planes of the Co-MDD and the NUS-DFT spectra (Figure 5). Overall, only 53% of the cross peaks that are observed by Co-MDD are resolved by NUS-DFT in the 4D methyl–methyl spectrum.

The occurrence of the two groups can readily be rationalized with the additional observation that the residues from group M2 feature weaker diagonal peak intensities than those of group M1. The MDD method alone relies on strong diagonal peaks and thus cannot detect the group M2 peaks. By coupling to the high-sensitivity 2D  $^{13}\text{C}$ , $^1\text{H}$ -HMQC spectrum in the Co-MDD processing, these planes are recovered and were in each case at least as sensitive as the NUS-DFT. In contrast, the weak performance of the NUS-DFT method for group M1 peaks with strong diagonal signals can be understood since the Fourier transformation of sparsely sampled data set corresponds to a Fourier transform of the hypothetically fully sampled data grid

(41) Kazimierczuk, K.; Zawadzka, A.; Kozminski, W.; Zhukov, I. *J. Magn. Reson.* **2007**, *188*, 344–356.



**Figure 5.** Comparison of different processing schemes for the 4D NUS  $^{13}\text{C}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC data. (A) 2D planes obtained by the three methods Co-MDD, DFT and MDD for I194  $\text{H}^{\text{O}}\text{C}^{\text{O}}$ , a representative from group M1, which contains 85% of all methyls (see text), and for L10  $\text{H}^{\text{O}}\text{C}^{\text{O}}$ , a representative from group M2. Potential peak positions are indicated by orange circles. All spectral planes are plotted with a base level of 3.0 times their spectral noise. (B) 1D cross sections along the dashed lines in A, as indicated by corresponding roman numbers. I–III and IV–VI were scaled according to the noise level of the 2D planes.

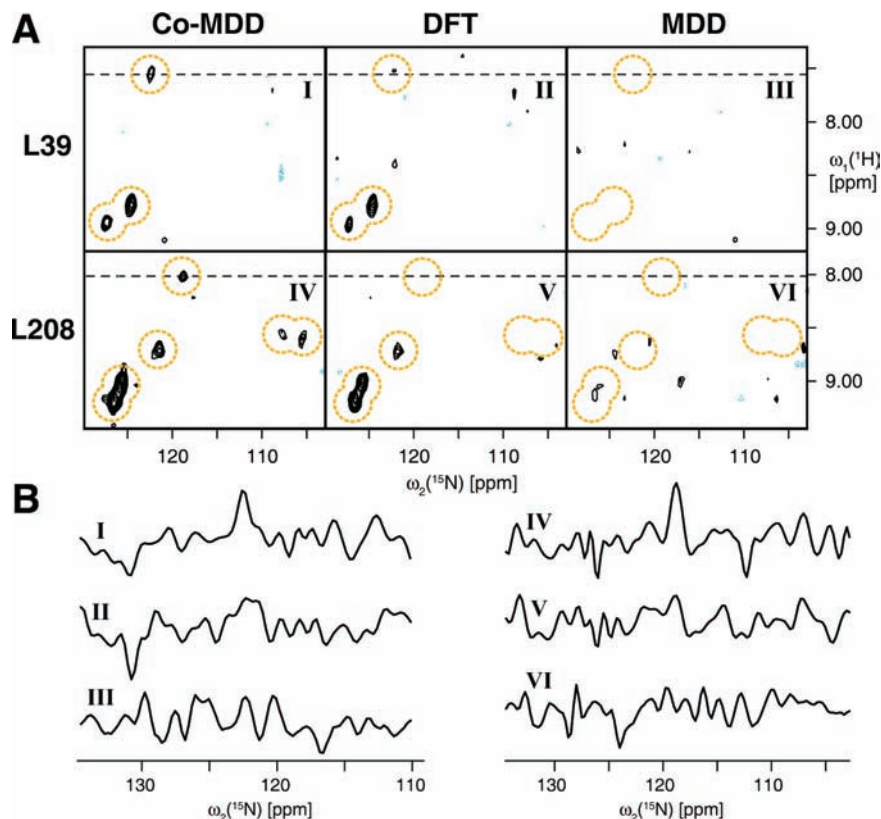
convoluted with a Fourier transformation of the sampling schedule. This convolution creates substantial noise from the strong diagonal peaks and other intense spectral features, such as residual water or detergent signals.

In the 4D  $^{15}\text{N}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC spectrum, which does not have any diagonal peaks, the result for all methyls resembles the outcome of the group M2 in the 4D  $^{13}\text{C}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC spectrum (Figure 6). When processed with the MDD method alone, the  $\omega_1/\omega_2$ -planes from all methyl resonances are essentially void of cross peaks. The NUS-DFT spectrum shows a number of cross peaks for many residues (Figure 6). If the Co-MDD method is used, which fixes the  $\omega_3(^{13}\text{C})$  and  $\omega_4(^1\text{H})$  shapes from the 4D  $^{13}\text{C}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC, about 20% additional cross peaks compared to the NUS-DFT are recovered from the noise (Figure 6). For none of the residues was a decrease in sensitivity in the Co-MDD observed when compared to the NUS-DFT.

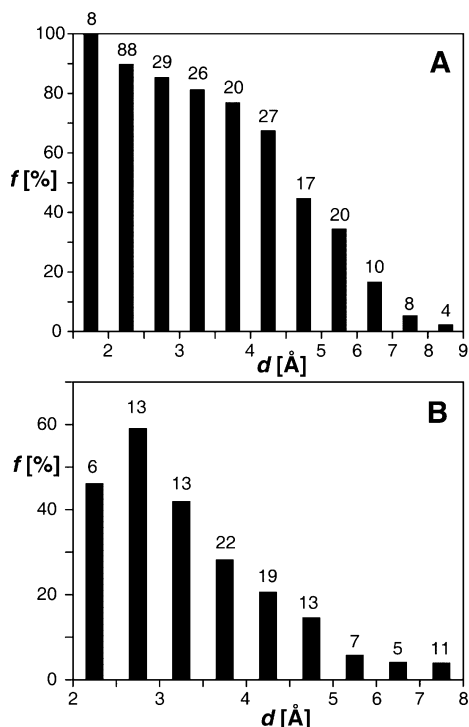
Thus, for both the 4D  $^{13}\text{C}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC and the 4D  $^{15}\text{N}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC spectra, processing with the Co-MDD method yielded for each methyl the best achievable results.

**Assessment of Spectral Quality (2): Completeness and Correctness.** Other criteria for the spectral quality beside the sensitivity are the completeness and the correctness of the observed peaks. For an evaluation of the completeness of the NOEs, we use the 2.3 Å crystal structure of mouse VDAC-1 as a reference,<sup>25</sup> which has been determined after the NMR structure of human VDAC-1. The NMR structure has been calculated using the 4D NOEs and is thus not a neutral reference point for these NOEs. Even though the crystal structure of mouse VDAC-1 was determined in bicelles and features three single amino acid mutations as compared to the NMR structure, a 3D comparison of the structures shows that the 19-stranded barrel, the relative orientation of the strands, and the contact of the N-terminal tail to the barrel wall inside at residues 10, 143, and 150 are identical in both structures.

For an assessment of the completeness we consider how well the observed NOEs compare to the NOEs expected from the 3D structure (Figure 7). For the 4D  $^{13}\text{C}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC, the completeness is between 80% and 100% for short contacts and decreases for longer distances. A drop in percentage can be seen at 4.5 Å. This distribution is comparable with previous MDD applications and with



**Figure 6.** Comparison of different processing schemes for the 4D NUS  $^{15}\text{N}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC. (A) 2D planes obtained by the three methods Co-MDD, DFT, and MDD for L39  $\text{H}^{91}\text{C}^{91}$  and L208  $\text{H}^{91}\text{C}^{91}$ . Potential peak positions are indicated by orange circles. All spectral planes are plotted with a base level of 3.0 times their spectral noise. (B) 1D cross sections along the dashed lines in A, as indicated by corresponding roman numbers. I–III and IV–VI were scaled according to the noise level of the 2D planes.



**Figure 7.** Peak statistics of the two Co-MDD 4D NOESYs. The fraction  $f$  is the number of observed peaks divided by the number of expected peaks from the 2.3 Å crystal structure,<sup>25</sup> plotted as a histogram of the  $^1\text{H}$ – $^1\text{H}$  distance  $d$ . The absolute number of observed peaks in each distance category is indicated above each bar: (A) 4D  $^{13}\text{C}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC; (B) 4D  $^{15}\text{N}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC.

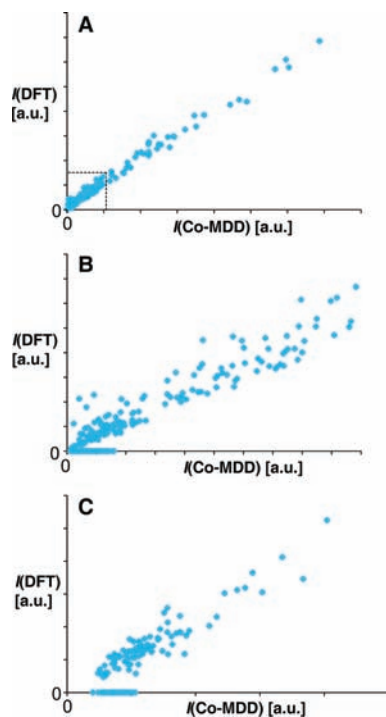
those observed in conventional Fourier transform 3D NOESYs.<sup>42</sup> For the 4D  $^{15}\text{N}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC, the average coverage of expected NOES in the range 2.0–4.5 Å is 31% (Figure 7). In this spectrum, the observation of cross peaks correlated with the sensitivity of these peaks in the 3D spectra and the distribution of the distance dependence of the completeness indicates that the absence of peaks is indeed sensitivity dependent. The large effective molecular weight of VDAC-1 in detergent micelles thus approximately represents the size limit of this experiment, and for systems with a smaller rotational correlation time an even higher coverage of peaks in the 4D  $^{15}\text{N}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC can be expected.

To assess the linearity of the peak intensities, we compare the intensities of the spectra with those in the DFT spectra mentioned above. Starting from the same nonlinearly sampled data set, the data processing with DFT is a linear function. We observe that for both 4D spectra, a linear correlation between the two processing methods is obtained and the remaining differences are within the spectral noise (Figure 8). It can thus be concluded that Co-MDD does not strongly disturb peak intensities, which is in line with previous similar investigations.<sup>35</sup>

In terms of correctness we note that all observed NOE peaks were picked before assignment and structure were known. They fit the structure (Figure 4) and are thus confirmed as real peaks and not artifacts induced by the MDD method. Overall, we

(42) Tugarinov, V.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 5701–5706.





**Figure 8.** Correlations of peak intensities  $I$  in nonuniformly sampled 4D NOESY spectra processed with either discrete Fourier transformation (DFT) and Co-MDD (see text). Unobserved peaks (intensities below noise level) were assigned a value of 0. (A) Cross and diagonal peaks in the 4D  $^{13}\text{C}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC; the region in the dashed box is shown enlarged in (B). (C) Cross peaks in the 4D  $^{15}\text{N}$ -HMQC- $^1\text{H}$ , $^1\text{H}$ -NOESY- $^{13}\text{C}$ -HMQC.

found that the Co-MDD processing did not create artifacts that could be mistaken for real peaks.

### Conclusion

Here, we have described a new method for the coupled processing of high-resolution 4D NOESY experiments, have

analyzed its performance, and discussed its use in a structure determination. Co-MDD is more sensitive than DFT NMR and MDD alone and recovers a substantial fraction of low-intensity peaks that are not detected by these methods. The experiments can be recorded on protein systems up to a rotational correlation time of at least 70 ns. Following the same lines of experiment design, the application of the Co-MDD method is readily applicable to other 4D NOESY data sets and to related pulse sequences.<sup>43–46</sup> For example, on systems tumbling faster than VDAC-1, the intensity of amide–amide and methyl–amide NOEs will be above the noise limit, and these spectra can then also be processed with the Co-MDD method. Due to these features and the general strength of the 4D NOESY approach, we anticipate that the method presented here will be widely used in structural studies of biomacromolecules and their complexes.

**Acknowledgment.** This work was supported by grants from the Swedish Research Council (2005-2951, 2008-4299), the NIH (GM075879, GM066360, GM47467, EB002026), and the Wenner-Gren Foundation, Stockholm. S.H. was supported in part by the Swiss National Science Foundation.

JA902012X

- (43) Farmer, B. T.; Mueller, L. *J. Biomol. NMR* **1994**, *4*, 673–687.
- (44) Frueh, D. P.; Vosburg, D. A.; Walsh, C. T.; Wagner, G. *J. Biomol. NMR* **2006**, *34*, 31–40.
- (45) Xu, Y.; Long, D.; Yang, D. *J. Am. Chem. Soc.* **2007**, *129*, 7722–7723.
- (46) Ying, J.; Chill, J. H.; Louis, J. M.; Bax, A. *J. Biomol. NMR* **2007**, *37*, 195–204.
- (47) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661–665.
- (48) Shaka, A. J.; Barker, P. B.; Freeman, R. *J. Magn. Reson.* **1985**, *64*, 547–552.
- (49) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393–399.