A New Linear Prediction Model Method for the Determination of Slow Amide Proton Exchange Rates from a Series of One-Dimensional ^{1}H NMR Spectra

René Moss, Henrik Gesmar, and Jens. J. Led*

Contribution from the Department of Chemistry, University of Copenhagen, The H. C. Ørsted Institute, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark

Received June 2, 1993®

Abstract: A NMR method for quantitative determination of the exchange rates of slowly exchanging amide protons in proteins is presented, with the des-[Phe(B25)] mutant of human insulin as an example. The exchange rates are obtained by fitting a model that includes exponential decay to the parameters produced by a linear prediction analysis of a series of one-dimensional free induction decays recorded during the exchange of the amide protons with solvent deuterium. Because the linear prediction model method is able to combine information from independent linear prediction calculations and to exploit differences in chemical shift and exchange rate simultaneously, it allows exchange rates of amide protons with closely spaced signals to be determined. The method covers exchange rates from 0.1 h^{-1} to well above 10 h^{-1}.

Introduction

The exchange of amide protons in biomacromolecules is a potential source of information about the structure and dynamics of these molecules in solution and, thereby, about their function and activity. NMR spectroscopy, in various forms, is the leading technique for studying these exchanges[1-3] both because of its high resolution and its inherent precision and because of the span of exchange rates that it covers ranging from reciprocal milliseconds to reciprocal months. Here we present a new and simple, quantitative approach for the determination of the exchange rates of slowly exchanging amide protons that covers a range of rates (0.1–10 h^{-1}) not easily accessible by the classical NMR methods.

The exchange of relatively slowly exchanging amide protons with solvent deuterium can be followed in series of 1D or 2D NMR spectra recorded at different times during the exchange process, and the rates of exchange can be determined from the time dependence of the intensities of the corresponding amide proton signals. Recently[4] it was demonstrated that amide proton exchange rates can be determined quantitatively from the exchange-induced F_1 line broadening of the crosspeaks in a *single* 2D NMR spectrum recorded during the exchange process, and it was shown that this approach allows the determination of exchange rates that are normally too fast to be measured by the classical 2D method.[3]

The approach presented here uses a series of 1D NMR spectra as in the classical 1D method. However, unlike this method the data analysis used here relies on (1) a combination of independent linear prediction (LP) analyses[5,6] of the individual FID's in the series of spectra and (2) a new method that allows a subsequent simultaneous analysis of all the resulting spectral parameters under the constraints of an appropriate model, *in casu* exponentially decaying intensities, constant frequency and signal width for the individual resonances, and phases that depend linearly on the frequency. Although the resolution of 1D methods is, in general, inherently lower than the resolution of 2D methods, the combined LP-model approach proposed here has a significantly higher resolution than the classical 1D approach and provides more accurate exchange rates. This is because it includes not only the estimated intensities as the classical approach but also the frequencies, signal widths, and phase parameters. Thus, it is able to exploit differencies in frequency and exchange rate simultaneously. In addition it allows exchange rates to be determined that are even faster than those accessible by the 2D F_1 line broadening method,[4] *i.e* exchange rates that broaden the crosspeaks in the 2D spectra beyond recognition. A more general description of the LP-model method will be published elsewhere. Only those features of the LP-model method that are pertinent to the determination of amide proton exchange rates presented here will be outlined below.

Method

A series of 1D NMR FID's obtained at different times, τ, during the exchange of the amide hydrogen with solvent deuterium is acquired. The subsequent data analysis by the LP model method is performed in two steps:

Step one: The first step of the LP model approach consists of independent LP analyses of the individual FID's, recorded at different times during the exchange process. This analysis produces an estimate of the intensity, the frequency, the signal width, and the phase for each signal found in each FID. The number of LP coefficients must be large enough to assure that there are no residual signals in the difference between an experimental FID and the FID recalculated from the estimated spectral parameters. The recently developed fast LP method (FLP),[7,8] based on the Toeplitz algorithm,[9] renders it feasible to apply a sufficiently large number of LP coefficients in the analysis. Thus, for signals that are well-resolved by the FLP procedure, one can determine the exchange rates, k_j, by fitting exponential decays to the corresponding intensities. However, the number of signals found by the LP calculation in each individual FID is not exactly the same. Closely spaced signals may not be resolved in each FID, and even when they are resolved, their corresponding spectral parameters are heavily correlated. This has the effect

* Author to whom correspondence should be addressed.

® Abstract published in *Advance ACS Abstracts*, December 15, 1993.

(1) Glickson, J. D.; Phillips, W. D.; Rupley, J. A. *J. Am. Chem. Soc.* **1971**, *93*, 4031–4038.
(2) Richarz, R.; Sehr, P.; Wagner, G.; Wüthrich, K. *J. Mol. Biol.* **1979**, *130*, 19–30.
(3) Wagner, G.; Wüthrich, K. *J. Mol. Biol.* **1982**, *160*, 343–361.
(4) Olsen, H. B.; Gesmar, H.; Led, J. J. *J. Am. Chem. Soc.* **1993**, *115*, 1456–1460.
(5) Gesmar, H.; Led, J. J. *J. Magn. Reson.* **1988**, *76*, 183–192.
(6) Gesmar, H.; Led, J. J.; Abildgaard, F. *Prog. NMR Spectrosc.* **1990**, *22*, 255–288.
(7) Hansen, P. C.; Gesmar, H. *Numerical Algorithms* **1993**, *4*, 151–166.
(8) Gesmar, H.; Hansen, P. C. *J. Magn. Reson.* In press.
(9) Cybenko, G. *SIAM J. Sci. Stat. Comput.* **1987**, *8*, 734–740.

0002-7863/94/1516-0747$04.50/0

that the estimated intensities of closely spaced signals are often distributed in an unpredictable manner among the involved signals. Furthermore, a variable number of noise-related signals are found in all the spectra as described previously.[5,6] Consequently, a direct fit of the intensities with an exponential decay that describes the exchange would be very difficult and probably lead to erroneous results for closely spaced signals.

Step two: In order to overcome the above-mentioned problems of closely spaced signals and to improve the general quality of the exchange rate estimations, the second step of the LP model approach is carried out. An *a priori* model is set up for the expected number of resonances and a set of model parameters that describe the evolution of these resonances during the exchange process is constructed. This confines the number of model parameters to one frequency (ν_j), one line width ($\Delta\nu_{1/2j}$), one initial intensity (I_j^0), and one exchange rate (k_j) for each resonance and two general phase parameters (ϕ_0 and ϕ_1) that describe the linear dependence of the phases upon the frequency. These are the parameters that describe the physics of the system. Thus, the entire set of FID's as a function of the sampling time (t) and the time (τ) that has passed since the exchange was initiated is given by

$$\mathrm{FID}\ (\tau,t) = \sum_{j=1}^{p} I_j^0 \exp[-k_j\tau] \exp[(2\pi\mathcal{J}\nu_j - \pi\Delta\nu_{1/2j})t + \mathcal{J}(\phi_0 + \phi_1\nu_j)] \quad (1)$$

Here ϕ_0 and ϕ_1 are the general zero- and first-order phase parameters, respectively, and p is the number of resonances that are included in the model calculation. How the small set of model parameters is determined from the large set of spectral parameters is briefly described in the following.

From the large number of spectral parameters determined in the first step a number of "auxiliary" FID's are constructed—one for each experimental FID. The number of points of these auxiliary FID's, *i.e.* the number of different values of the parameter t, is equal to the number of spectral parameters that describe the spectral region to be analyzed. Therefore, the content of information in these FID's is exactly the same as in the corresponding spectral parameters. Thus, the model parameters can be determined from the spectral parameters by a nonlinear LSQ fit of eq 1 to the set of auxiliary FID's. In this manner, all the information of the large number of spectral parameters is concentrated in the relatively small set of model parameters. The auxiliary FID's are noiseless, but based on the Cramér Rao lower bounds as determined by the FLP procedure,[10] standard deviations and correlations are calculated for the points of the auxiliary FID's. Because these standard deviations and correlations are included in the nonlinear LSQ fit of eq 1, reliable LSQ estimates of not only the model parameters but also their standard deviations are assured. However, these estimates are based on the assumption of uncorrelated Gaussian noise, and systematic errors in the experiment are not accounted for. Therefore, because of magnetic field inhomogeneity, instrumental instabilities, and variation of the experimental conditions (*e.g* the temperature) the deviation of the result from the "true" values may be somewhat larger than what would be expected from the estimated standard deviations alone.

When the above procedure is applied, all the information of the independent LP calculations is combined, in spite of the fact that it may not be directly comparable, because of the variation in number of signals found in the individual FID's, as discussed above. This results in an improved determination of the physical relevant parameters, as compared to what could be obtained by direct fitting of exponential decays to the intensities. As a consequence the procedure is able to benefit from differences in exchange rate and chemical shift simultaneously, which results in increased resolution and, thereby, a determination of more exchange rates.

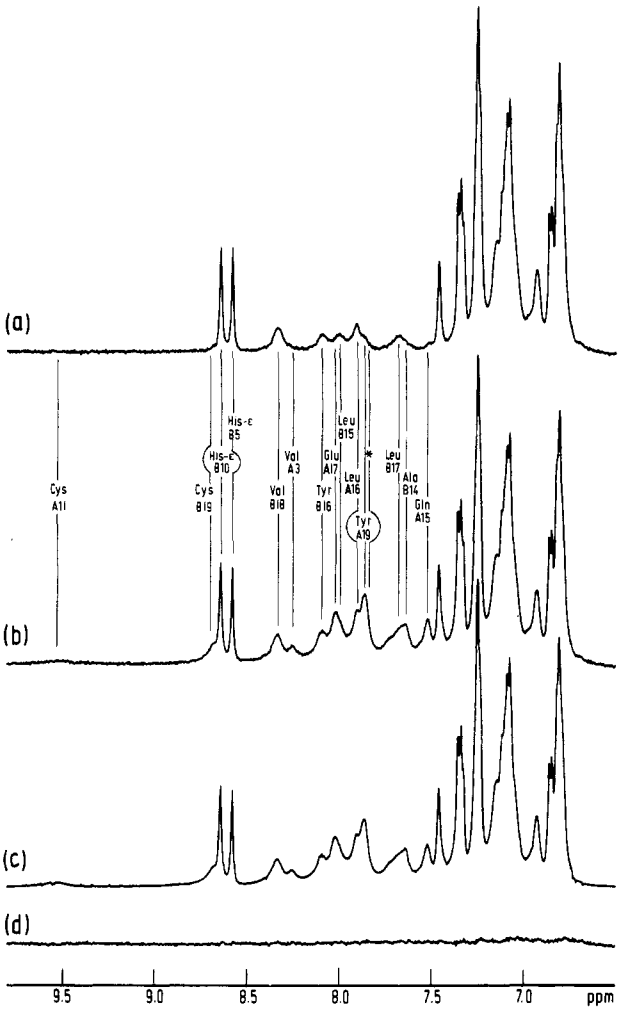

Figure 1. The amide region of the ^{1}H NMR spectrum of the des-[Phe(B25)] mutant of human insulin in D_2O. Spectrum b corresponds to the 1st and spectrum a to the 80th FID of a series of 240 FIDs recorded during the exchange of the amide hydrogen with solvent deuterium; the FID's of b and a were recorded 5 and 65 min, respectively, after dissolution. Spectrum c is the LP spectrum obtained by Fourier transformation of the FID calculated from the result of a 3000-order FLP calculation on the first FID. This recalculated FID should not be confused with the auxiliary FID's (see text) which have considerably fewer points and different spectral widths. Spectrum d is the difference between b and c. The asterisk indicates one or more slowly exchanging amide signals that could not be identified due to severe overlap. Possible candidates are Ser(A9), Ile(A10), and Thr(B30) according to an assignment (to be published) based on 2D spectra in H_2O.

Experimental and Computational Procedures

The LP-model approach was applied to NMR data from the 50-residue des-[Phe(B25)] mutant of human insulin (Figure 1), previously studied by the 2D F_1 line broadening method.[4] The ^{1}H FIDs were sampled on a Bruker AM500 NMR spectrometer at 500 MHz and 310 K. The lyophilized protein was dissolved (5 mM) in D_2O at pD 3.49 and the data acquisition was started as soon as possible (≈5 min after dissolution). The residual water signal was suppressed by a DANTE pulse.[11] The FID's corresponding to the first 120 spectra were recorded with a time interval of 45 s between consecutive FID's while the last 120 spectra were recorded with a corresponding time interval of 12 min. All experimental FIDs consisted of 8192 data points and the applied sweep width was 10 000 Hz. No zero filling or window function was applied before the data analysis.

(10) Barkhuijsen, H.; de Beer, R.; van Ormondt, D. *J. Magn. Reson.* **1986**, *67*, 371–375.

(11) Bodenhausen, G.; Freeman, R.; Morris, G. A. *J. Magn. Reson.* **1976**, *23*, 171–175.

As the first step of the data analysis independent LP calculations were performed on the 240 FID's. For each FID 3000 LP coefficients were used. As illustrated in Figure 1 for the first FID, this number was sufficiently large to assure that no signal residuals were found in the difference between the experimental and the recalculated FID's. Because of noise and closely spaced signals, *e.g.* B14 and B17 (cf. Figure 1), the number of signals found in each FID varied as discussed above, ranging from 10 to 30 for the part of the amide region studied here (7.46–9.53 ppm). Thus a total of approximately 16 000 spectral parameters were needed to describe this region in all 240 FID's.

A complete assignment of the NMR spectrum of the des-[Phe(B25)] mutant will be published elsewhere. However, the resonances from 13 slowly exchanging amide protons together with two non-exchanging ($k_j = 0$) histidine resonances were identified in the region between 7.46 and 9.53 ppm ($p = 15$). Thus, a total of only 60 parameters was required by the model expressed in eq 1. As the second step of the procedure, these 60 model parameters were determined from the ≈16 000 spectral parameters by the procedure described in the Method section. Initial values for the nonlinear LSQ analysis were chosen by inspection of the spectral parameters as determined by the LP calculations of step one. For the region 7.46–9.53 ppm the corresponding subset of model parameters was determined by a nonlinear LSQ fit of eq 1 to 240 auxiliary FID's constructed from the spectral parameters with frequencies in the proper interval. The resulting exchange rates are given in Table 1 together with those obtained previously[4] by the 2D F_1 line broadening method.

The results obtained by the LP model method and the F_1 line broadening method[4] are in good agreement when the difference between the pH values applied in the two experiments (3.0 in the 2D experiment vs 3.49 in this work) is taken into account. In both cases the values are close to the pH value that corresponds to the minimum of the exchange rate of the peptide group hydrogen.[12] Therefore, the pH dependence of the exchange rates is also minimum. Moreover, in the cases where the difference goes beyond the uncertainty of the exchange rates, the deviations agree qualitatively with Molday et al.[12] according to whom the minimum of the exchange rate of a given amide proton, and thus the exchange rate itself, depends slightly on the primary structure around the exchanging amide group.

Conclusion

As it appears from Table 1, the method presented here allows the determination of exchange rates that are considerably faster than those obtainable from the F_1 line broadenings in 2D spectra.[4] Because the LP model method is able to combine information from independent LP calculations, it provides reliable estimates of the amide hydrogen exchange rates in proteins. It is also able to benefit from differences in exchange rates of the NH protons, thereby allowing a distinction between signals as closely spaced as those from Ala(B14) and Leu(B17) or from Leu(B15) and Glu(A17) (cf. Figure 1) despite the lower resolution in 1D spectra in general.

Table 1. Exchange Rate Parameters for Exchanging Amide Protons[a] in des-[Phe(B25)] Human Insulin

δ, ppm	k_{exch}, h^{-1}	k_{exch},[b] h^{-1}	NH
7.4969 ± 0.0001	2.37 ± 0.02	1.032 ± 0.056	Gln(A15)
7.6158 ± 0.0003	2.34 ± 0.04		Ala(B14)
7.6554 ± 0.0003	0.558 ± 0.006	0.310 ± 0.083	Leu(B17)
7.8355 ± 0.0004	7.5 ± 0.3		*c*
7.8424 ± 0.0005	1.43 ± 0.03	0.508 ± 0.039[d]	Tyr(A19)
7.8847 ± 0.0002	0.520 ± 0.005	0.373 ± 0.050	Leu(A16)
7.9752 ± 0.0003	0.77 ± 0.01	0.564 ± 0.039	Leu(B15)
8.0023 ± 0.0003	2.97 ± 0.08		Glu(A17)
8.0698 ± 0.0001	0.559 ± 0.005	0.399 ± 0.015	Tyr(B16)
8.229 ± 0.0007	14 ± 0.9		Val(A3)
8.3103 ± 0.00006	0.251 ± 0.001	0.218 ± 0.019	Val(B18)
8.6490 ± 0.0004	2.19 ± 0.03		Cys(B19)
9.503 ± 0.005	15 ± 2		Cys(A11)

[a] Including 1 σ standard deviations. [b] From line width in 2D ^{1}H spectra in ref 4. [c] The amide proton signals from Ser(A9), Ile(A10), and Thr(B30) are indistinguishable in the 1D spectra. The calculated number is the rate constant of one of these amide protons or a combination of two or all of them. [d] Assigned incorrectly to Phe(B24) in ref 4.

Acknowledgment. This work was supported by the Danish Technical Research Council, J. No. 16-4679.H, the Danish Natural Science Research Council, J. No.'s 11-8977-1 and 16-5027, the Ministry of Industry, J. No. 85886, Direktør Ib Henriksens Fond, and Julie Damms Studiefond. We thank Mrs. Helle B. Olsen and Mrs. Else Philipp for recording the spectra and Dr. Per Balschmidt Novo Nordisk A/S for providing the des-[Phe(B25)] insulin. Helpful discussions with Dr. Georg Ole Sørensen are greatfully acknowledged.

(12) Molday, R. S.; Engelander, S. W.; Kallen, R. G. *Biochemistry* **1972**, *11*, 150–158.