Article

Error estimation and global fitting in transverse-relaxation dispersion experiments to determine chemical-exchange parameters

Rieko Ishima* & Dennis A. Torchia

Molecular Structural Biology Unit, National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, MD, 20892-4307, USA

Received 17 December 2004; Accepted 08 February 2005

Key words: conformational change, chemical exchange, CPMG, NMR, R2

Abstract

Off-resonance effects can introduce significant systematic errors in R_2 measurements in constant-time Carr-Purcell-Meiboom-Gill (CPMG) transverse relaxation dispersion experiments. For an off-resonance chemical shift of 500 Hz, ¹⁵N relaxation dispersion profiles obtained from experiment and computer simulation indicated a systematic error of ca. 3%. This error is three- to five-fold larger than the random error in R_2 caused by noise. Good estimates of total R_2 uncertainty are critical in order to obtain accurate estimates in optimized chemical exchange parameters and their uncertainties derived from χ^2 minimization of a target function. Here, we present a simple empirical approach that provides a good estimate of the total error (systematic + random) in ¹⁵N R_2 values measured for the HIV protease. The advantage of this empirical error estimate is that it is applicable even when some of the factors that contribute to the off-resonance error are not known. These errors are incorporated into a χ^2 minimization protocol, in which the Carver–Richards equation is used fit the observed R_2 dispersion profiles, that yields optimized chemical exchange parameters and their confidence limits. Optimized parameters are also derived, using the same protein sample and data-fitting protocol, from ¹H R_2 measurements in which systematic errors are negligible. Although ¹H and ¹⁵N relaxation profiles of individual residues were well fit, the optimized exchange parameters had large uncertainties (confidence limits). In contrast, when a single pair of exchange parameters (the exchange lifetime, τ_{ex} , and the fractional population, p_a), were constrained to globally fit all R_2 profiles for residues in the dimer interface of the protein, confidence limits were less than 8% for all optimized exchange parameters. In addition, F-tests showed that quality of the fits obtained using τ_{ex} , p_a as global parameters were not improved when these parameters were free to fit the R_2 profiles of individual residues. Finally, nearly the same optimized global τ_{ex} , p_a values were obtained, when the ¹H and ¹⁵N data sets for residues in the dimer interface, were fit independently; the difference in optimized global parameters, ca. 10%, was of marginal significance according to the F-test.

Introduction

The phenomenon of chemical exchange, manifested by an enhanced transverse nuclear spin relaxation rate, R_2 , often reflects a conformational equilibrium on the microsecond to millisecond time-scale. In the case of a protein, it is of interest to determine exchange parameters, such as the exchange lifetime, τ_{ex} , and populations of the exchanging species, in order to characterize the kinetics and thermodynamics of a conformational transition that may be relevant to function (Palmer et al., 2001; Korzhnev et al., 2004). Conformational exchange parameters are often derived from measurements of R_2 as a function of the effective field strength (v_{CP}) in a Carr-Purcell-

^{*}To whom correspondence should be addressed. E-mail: rishima@dir.nidcr.nih.gov

Meiboom-Gill (CPMG) experiment (Orekhov et al., 1994; Ishima et al., 1998; Loria et al., 1999; Mulder et al., 2001). In the case of R_2 experiments of uniformly labeled proteins, a 2D spectrum is typically recorded for each relaxation delay. However, the dispersion profile is most efficiently obtained by determining R_2 for each value of v_{CP} from only two spectra (a common reference spectrum acquired without a relaxation delay, and a spectrum acquired with a constant relaxation time delay, T_{CP} , for each value of v_{CP}). (Mulder et al., 2001; Skrynnikov et al., 2001; Tollinger et al., 2001; Ishima and Torchia, 2003). An added advantage of the constant time CPMG experiment is that the correct value of the relaxation rate due to chemical exchange, R_{es} , is obtained from the R_2 profile even when relaxation due to other mechanisms is multi-exponential (Mulder et al., 2002).

We previously reported relaxation profiles for amide ¹⁵N and ¹H spins of HIV-1 protease bound to an inhibitor, DMP323 (Figure 1) (Ishima and Torchia, 2003) using the constant time approach. The ¹⁵N data were obtained using a high sensitivity cryoprobeTM (Bruker Biospin), so that the measured R_2 values had small random errors (typically less than 1%). However, many dispersion profiles showed abrupt fluctuations in R_2 as function of v_{CP} that were inconsistent with theory and were well outside the small random errors. In Figure 1, large differences in R_2 values, at $v_{CP} = 300$ Hz as compared with 150 Hz, are seen in the dispersion profiles of residues 3 and 96. The magnitudes of these types of fluctuations in R_2 increased as the chemical shift of the ¹⁵N signal, relative to the RF carrier, increased. Although the off-resonance effects on R_2 , measured using the standard CPMG approach, have been thoroughly investigated (Czisch et al., 1997; Ross et al., 1997; Korzhnev et al., 2000), this has not yet been done in the case of the constant time experiment.

As has been shown (Ross et al., 1997) the intensity loss due to the off-resonance shift in the CPMG experiment is a function of angles Θ_{eff} and Φ_{eff} , which define the orientation of the



Figure 1. R_2 values plotted as a function of the effective field strength, v_{CP} , obtained for (a) amide ¹H and (b) amide ¹⁵N sites in residues in the dimer interfacial β -sheet (residue 2–4, 96–99), the adjacent loop (5–8), and tip of the flap regions (50, 51) of the HIV-1 protease D25N mutant bound to DMP323 (Ishima et al., 2004).

magnetization. In the conventional experiment using multiple delay times that increase with the number of CPMG pulses, n, the intensity oscillates as a function of $2n\Phi_{\rm eff}$. This causes offset dependent systematic fluctuations in the signal intensity and R_2 as a function of the decay time. Systematic error is revealed when the observed decay profile differs from the optimized exponential fit by amounts that are greater than can be accounted for by random error. In the two-point constant relaxation time experiment, systematic error in R_2 cannot be revealed by a two-point fit. However the off-resonance effect causes systematic fluctuations in the signal intensity and R_2 as a function of the number of CPMG pulses applied during the constant relaxation period. Therefore, in this case, the systematic error in R_2 is revealed by fluctuations in R_2 dispersion profiles, beyond those due to random noise, that increase with resonance off-set. Therefore, the source of error must be taken into account in order to extract reliable optimized exchange parameters from the data.

Here, we use a combination of computer simulation and an empirical experimental approach to investigate systematic error in R_2 dispersion data measured using the constant time CPMG approach. The computer simulations are used to calculate the effect on the intensity of transverse magnetization of imperfect CPMG refocusing pulses caused by (a) the off-resonance effect (b) pulse miscalibration and (c) RF inhomogeneity. The experimental approach used to estimate systematic error relies upon the observation that many sites in the protease do not undergo chemical exchange. The systematic error in R_2 for each of these sites is estimated by comparing the r.m.s. value of R_2 (calculated from the measured R_2 profile of the site) with the random error of R_2 due to noise. When this comparison is made as a function of the resonance offset, an empirical estimate of the total error in R_2 , due to random and off-resonance systematic error, is obtained.

Once estimates of total errors in the R_2 measurements are obtained, the Carver–Richards equation is used to fit the R_2 profiles. χ^2 -minimization is used to determine optimized values of the exchange parameters and their uncertainties. Data of residues in the dimer interface of the protease are fitted using individual (residue-specific) as well as global exchange parameters. The different fitting protocols are evaluated by calculating the *F* statistic, which validates the use of global exchange parameters to fit the data. This result is significant because global-fitting the R_2 profiles of residues in the dimer interface greatly reduces the uncertainties in the optimized exchange parameters compared with uncertainties obtained when the R_2 profile of each residue is independently fit.

Materials and methods

NMR samples and experiments

The NMR sample and experiments have been described in detail (Ishima and Torchia, 2003; Ishima et al., 2004) and are briefly discussed here. HIV-1 protease (MW ca. 22 kDa) used in this study has the active site mutation, D25N, which inactivates the protease but preserves its three dimensional structure, and five mutations that suppress autoproteolysis and aggregation (Louis et al., 2002; Katoh et al., 2003). In the presence of the inhibitor, DMP323, this construct forms a highly stable homodimer. The protein was uniformly labeled with ²H, ¹³C and ¹⁵N, and dissolved (ca. 0.25 mM concentration as a dimer) in 20 mM sodium phosphate buffer (pH 5.8, 95%H₂O/5%D₂O) in the presence of an excess of DMP323.

All NMR experiments were performed using Bruker DMX500 spectrometers, with the sample temperature maintained at 20 °C (Ishima and Torchia, 2003). The ¹⁵N relaxation dispersion experiments were recorded using a Bruker cryoprobe, while the amide ¹H relaxation dispersion experiments were recorded using an ambient temperature probe. A reference spectrum was acquired, without a CPMG period, together with either seven or eight spectra containing a constant CPMG period, T_{CP} in which the time between CPMG 180° pulses, $2\tau_{CP}$, was varied. The CPMG pulses produce an effective field, v_{CP} , that is defined by1/4 $\tau_{CP} = \gamma B_{eff}/(2\pi) = v_{CP}$ (Mulder et al., 2001; Skrynnikov et al., 2001; Tollinger et al., 2001; Ishima and Torchia, 2003). For each v_{CP} , the R_2 value was determined from the ratio of two signal intensities I_{CP} and I_0 , where I_0 is the reference measured without the CPMG period, and I_{CP} is measured at the end of the constant CPMG period.

All ¹H and ¹⁵N rectangular pulses were applied with RF fields of 25 kHz (90° pulse, 10 μ s) and 5 kHz (90° pulse, 50 μ s), respectively. Proton pulses were applied at the water resonance except during the proton CPMG period where the RF carrier position was switched to 8.5 ppm. ¹⁵N 90° pulses were applied at 117 ppm. The intensity distribution of the ¹⁵N B_1 field was nearly symmetric, with 2 kHz maximum value and a full width at half-maximum of ca. 120 Hz (Guenneugues et al., 1999). The signal intensity following a ¹⁵N 810° pulse was 81% of that following a 90° pulse, when the ¹⁵N magnetization was converted to proton magnetization and then detected. This method was used to estimate the RF spatial homogeneity because, in the ¹⁵N R_2 dispersion experiment, ¹⁵N magnetization was converted to proton magnetization and then detected.

Spectra were recorded with $T_{\rm CP} = 40$ ms and $v_{\rm CP}$ equal to 50, 100, 150, 200, 400, 800, 1000, 2000, 50 Hz and 50, 100, 150, 200, 250, 300, 500, 50 Hz for ¹H and ¹⁵N spins, respectively. Although theory shows that $T_{\rm CP}$ should be set equal to $1/R_2$ in order to minimize the error in R_2 determined by sampling an exponential decay at two points (Jones et al., 1996; Jones, 1997), $T_{\rm CP}$ was set to less than the average value of $1/R_2$ of amides in the sample, because spins that undergo chemical exchange have smaller than average $1/R_2$ values.

The ¹⁵N relaxation dispersion spectra were recorded with 100 and 1024 complex points in F_1 and F_2 dimensions, respectively, and with 16 scans per point. ¹⁵N experiments were performed three times using identical conditions: the second experiment was acquired immediately following the first, and the third was acquired 2 weeks later. Spectra were processed using NMRPipe and NMRdraw software (Garrett et al., 1991; Delaglio et al., 1995).

Evaluation of uncertainty of R_2 measurements

The uncertainty in R_2 was calculated in two ways. First the uncertainty, ΔR_2 , was calculated based on random noise of the spectra, using the following expression,

$$\Delta R_2 = \delta_{\rm e} / I_{\rm CP} T_{\rm CP} \tag{1}$$

where δ_e is the r.m.s. noise measured in the reference spectrum. The fractional error of R_2 is $\Delta R_2/R_2$. In this paper R_2 values in each dispersion profile were obtained from a two-point fit using a

single value of I^0 . For this reason the contribution stemming from the error in I^0 is not included in Equation 1, in contrast with the equation used previously (Ishima and Torchia, 2003). It should be noted that the error in I^0 does contribute a small systematic offset in R_2^0 , and this error is not included in uncertainty in the optimized R_2^0 obtained by χ^2 fitting each dispersion profile using the Carver–Richards equation, with $R_2^0 = R_{2a} = R_{2b}$ (Carver and Richards, 1972; Davis et al., 1994). Second, the uncertainty, R_2^{rmsd} , was calculated as the r.m.s.d. of the set of R_2 values measured in a relaxation dispersion profile. The fractional uncertainty is defined as $R_2^{\text{rmsd}}/\langle R_2 \rangle$, where $\langle R_2 \rangle$ is the average R_2 of the profile.

In the absence of chemical exchange and systematic error, η , defined as

$$\eta = (R_2^{\text{rmsd}} / \langle R_2 \rangle) / \langle \Delta R_2 / R_2 \rangle \tag{2}$$

for each residue, should have a value close to unity. Therefore, when η is significantly larger than unity for an individual residue, R_2 dispersion data contain systematic errors and/or exhibit chemical exchange.

In order to evaluate whether off-resonance effects introduced systematic error into the R_2 data, $R_2^{\text{rmsd}}/\langle R_2 \rangle$ was plotted against the chemical shift from RF carrier. R_2 data with large random errors $\langle \Delta R_2/R_2 \rangle > 0.8$ and those for which η is large ($\eta > 10$ for ¹⁵N and $\eta > 6$ for ¹H), indicating chemical exchange, were excluded from this plot.

R_2 dispersion curve fitting

Selected R_2 dispersion profiles were fit using χ^2 minimization by the Carver–Richards equation (Carver and Richards, 1972; Davis et al., 1994), which assumes two-site exchange. The equation contains four unknown physical parameters which are the free variables used to fit each R_2 dispersion curve: (1) the intrinsic transverse relaxation rate, R_2^0 , determined by modulation of dipolar interactions and chemical shift anisotropy (2) the difference in chemical shifts between the exchanging species, $\delta\omega$, (3) the population of the major exchange species, p_a , and (4) the correlation time for exchange, τ_{ex} . These unknown parameters were determined by minimizing χ^2 , where

$$\chi^2 = \sum_{i} \left(\frac{R_2^{i,exp} - R_2^{i,calc}}{\Delta R_2^{i,err}} \right)$$

Here, $R_2^{i,\text{exp}}$ and $R_2^{i,\text{cal}}$, are experimental and calculated R_2 values of *i*'th v_{CP} value, respectively, and $\Delta R_2^{i,\text{err}}$ is the estimated experimental uncertainty. For the ¹⁵N R_2 data, $\Delta R_2^{i,\text{err}}$ was estimated from the plot of $R_2^{\text{msd}}/\langle R_2 \rangle$ (which contains both systematic and average random error) against off-resonance chemical shift. For the ¹H R_2 data, $\Delta R_2^{i,\text{err}}$ was taken as ΔR_2 because the ¹H R_2^{rmsd} for residues not undergoing chemical exchange were in agreement with those obtained from random error. To ensure that the global χ^2 minimum was found, χ^2 minimization was repeated using a range of initial parameters, in both the slow and fast exchange regimes.

The unknown fitting parameters were optimized, for the same set of residues in a defined region of the protease, in three ways. First, the R_2 profile of each residue was individually fit and the optimized values of the four fitting parameters were obtained for each R_2 dispersion data set. Second, all of ${}^{15}N$ and ${}^{1}H R_2$ dispersion profiles were simultaneously fit using a single pair of optimized p_a and τ_{ex} values (i.e., p_a and τ_{ex} are global parameters) while the remaining parameters were free to vary from one residue to the next. Third, the ¹⁵N and the ¹H R_2 data sets were independently fit assuming that p_a and τ_{ex} are global parameters. Note that the second method uses the smallest number of free parameters to fit all of the data. The F-test was used to evaluate whether the use of additional parameters in the first and third methods was justified.

Uncertainties of the optimized parameters obtained using the first method (individual fit) were determined by repeating the optimization for 100 data sets simulated by the Monte Carlo method. Note that the best-fit parameters were used as initial parameters for every optimization of the simulated data. In the second and third methods, the optimized global parameters (p_a and τ_{ex} ,) were determined from a grid search in the following manner. At each grid point (defined by the pair of p_a , τ_{ex} values) χ^2 (as defined in equation 3) was minimized with respect to local parameters R_2^0 and $\delta\omega$ (independent variables for each site in the calculation). The uncertainties in p_a and τ_{ex} were obtained from the confidence ellipse region corresponding to values of χ^2 1.0

larger than the fitted minimum (Press et al., 1988). Although the R_2^0 and $\delta\omega$ were optimized together with the global parameters (p_a and τ_{ex}), uncertainties of $\delta\omega$ and R_2^0 were independently determined by the Monte-Carlo simulation, using the optimized global parameters as initial parameters.

Numerical simulations of density operator evolution during the constant time CPMG period

To estimate the magnitudes of systematic errors in the constant relaxation time experiments, the time evolution of the density operator, ρ , during the CPMG period was calculated using the Liouville-Von Neumann equation as described previously (Ishima et al., 2004). The time evolution of ρ was calculated by applying a sequence of unitary transformations to $\rho(0) = M_Y$, using time-independent Hamiltonians, consisting of terms for chemical shift precession and CPMG pulses, as propagators (Ernst et al., 1987), and neglecting relaxation. MATLAB software (Mathworks Inc, MA) was used for the calculation. The NMR signal was assumed to be 532 Hz off-resonance corresponding to the position of the amide ¹⁵N of residue 3, relative to the carrier frequency. The 180° pulse width was set to either 60 μ s or 100 μ s. In all calculations, T_{CP}, was set to 40 ms during which time from 2 to 160 CPMG pulses were applied, corresponding to v_{CP} values in the range of 25 Hz to 2 kHz. The results of the calculations were displayed as profiles in which the ratio $f_Y = \langle M_Y(T_{\rm CP}) \rangle / \langle M_Y(0) \rangle$, where $\langle Q \rangle$ designates expectation value of Q, was plotted as a function of either time or v_{CP} .

In calculations that took into account spatial B_1 inhomogeneity, we used a distribution consisting of five B_1 fields with relative strengths of 0.908, 0.954, 1.0, 1.1046, and 1.092, and weights of 0.0625, 0.25, 0.375, 0.25, and 0.0625, respectively (Geen and Freeman, 1991). The B_1 field strengths were chosen so that the intensity of transverse magnetization, M_Y , following an 810° pulse was 80% of that following a 90° pulse, i.e., $M_Y(810^\circ)/M_Y(90^\circ)=0.8$. An asymmetric B_1 profile was simulated using weights of 0.0825 0.25 0.375 0.25 0.0425 with B_1 field strengths of 0.85, 0.97, 1.0, 1.03, and 1.06, respectively.

Results and discussion

Uncertainty in R_2 dispersion experiments

 R_2 dispersion profiles were recorded for total of 87 residues. The fractional random error ($\Delta R_2/R_2$) of the ¹H data, as calculated from the signal-to-noise ratio, was typically in the range of 0.7–1.7% except for residues 4 and 6 ($\Delta R_2/R_2$, ca. 4–5%) whose dispersion amplitudes were exceptionally large (Figure 1a). The range of $\Delta R_2/R_2$ in the case of the ¹⁵N data was smaller, 0.4–1.3%, than that of the ¹H data. Although the ¹H data were accumulated with twice as many scans as the ¹⁵N data (Ishima et al., 2004), the signal-to-noise ratio of the ¹⁵N data measured was better because it was acquired using a Bruker cryoprobe.

One notable feature of the ¹⁵N R_2 profiles is that the ¹⁵N R_2 data points often fall outside of the range predicted by the random errors. For example, the ¹⁵N R_2 values of residue 3 measured for $\nu_{CP}\!=\!300$ Hz and of residue 96 for $\nu_{CP}\!=\!150$ Hz were approximately 4.5% and 3.5% smaller than the R_2 values measured at larger values of v_{CP} , 500 Hz and 200 Hz, respectively (Figure 2). These observations conflict with theory, which predicts that R_2 decreases as v_{CP} increases, except in the case of the slow exchange limit. The R_2 experiments were repeated, and the deviations of R_2 among the three experiments were 0.57% for residue 3 and 0.54% for residue 96 (Figure 2), in agreement with the uncertainty expected from the fractional random error, i.e., $\Delta R_2/R_2 \cong 0.5\%$. Because the observed fluctuations of R_2 values as a function of v_{CP} are well outside the range expected

from random error, they were tentatively ascribed to a systematic error.

In the experiments, the ¹⁵N chemical shifts of residues 3 and 96 were offset by 532 Hz and 426 Hz, respectively, from the RF carrier (117 ppm on the 500 MHz spectrometer), suggesting that off-resonance effect introduced systematic error. If this is the case, the same off-resonance effect should cause fluctuations in R_2 values that are larger than those expected from random noise in R_2 profiles for residues whose ¹⁵N signals are off-resonance but that do not exhibit chemical exchange. In Figure 3, R_2 dispersion profiles of five residues with chemical shifts within 200 Hz of the R.F. carrier (Figure 3a) are compared with profiles of the five residues whose signals are more than 500 Hz from the carrier (Figure 3b). It is clear from Figure 3 that residues with the larger offsets (Figure 3b) have a much greater variation in their R_2 values than the residues in Figure 3a. The $\Delta R_2/R_2$ values of all residues in Figure 3 are between 0.4% and 1%, irrespective of their chemical shift offsets from the carrier frequency. On the other hand, the fractional R_2^{rmsd} value in each profile was in the range of 0.8-1.4% for residues less than 200 Hz off-resonance (Figure 3a) but 2% to 5% for the residues more than 500 Hz off-resonance. These results leave no doubt that the ${}^{15}N$ R_2 data contain a systematic error stemming from an offresonance effect. In support of this conclusion, the systematic errors in the ¹⁵N dispersion profiles of residues 3 and 96 were suppressed when the data were recorded with the ¹⁵N carrier onresonance.



Figure 2. Three ¹⁵N R_2 relaxation profiles obtained for (a) residue 3 and (b) residue 96. The symbols, \bigcirc , \Box , \blacktriangle distinguish the three data sets. The profiles defined by the \bigcirc symbols are the same as those in Figure 1b for residues 3 and 96.



Figure 3. Comparison of ¹⁵N R_2 profiles of amides that do not undergo chemical exchange, whose signals are off-resonance from the RF carrier by (a) 100–200 Hz and (b) 500–600 Hz. The larger fluctuations of R_2 as a function of v_{CP} observed in (b) are due to the off-resonance effect.



Figure 4. Numerical simulations showing the fraction of transverse magnetization that remains after 40 ms of CPMG evolution, $\langle M'_y(T_{CP}) \rangle / \langle M'_y(0) \rangle$, plotted as a function of the effective field strength, v_{CP}. The simulations were performed assuming B_1 fields of 8.3 and 5 kHz corresponding to 180° pulse widths of (a) 60 μ s and (b) 100 μ s, respectively. The 100 μ s '180°' pulse rotated the magnetization by (c) 184° and (d) 189° due to assumed calibration errors of 2% and 5%, respectively. The simulation in (d) was repeated including B_1 spatial inhomogeneity (e) symmetric and (f) asymmetric, as defined in the Material and Methods.

Simulation of intensity fluctuations in the CT–CPMG experiment

To test quantitatively whether the magnitude of the systematic error in R_2 is explained by the off-resonance effect, we calculated the signal intensity, I_{CP} , remaining at the end of the CPMG constant time period using parameters similar to those in our ¹⁵N experiments (Figure 4). In agreement with previous results (Ross et al., 1997; Korzhnev et al., 2000), Figure 4 reveals that I_{CP} depends upon the 180° pulse width and v_{CP} , as well as the off-resonance chemical shift. The change in R_2 resulting from the change in I_{CP} is calculated according to Equation 1. For example, a 2% change of I_{CP} induces a 2.5–3.5% change in R_2 , and a 4% change of I_{CP} causes a 5.1–6.7% R_2 change. These error ranges are obtained assuming I_{CP}/I_0 is in the 0.47–0.56 range, which was observed for the ¹⁵N data of residue 3.

It is known that the fluctuations in I_{CP} due to the off-resonance effect increases as the 180° pulse lengthens (Ross et al., 1997; Korzhnev et al., 2000). Our simulations (Figures 4a and b) also demonstrate that when the 180° pulse width is increased, the abrupt changes of I_{CP} as a function of v_{CP} becomes more significant (Figures 4a and b). For example, in the case of a CPMG 180° pulse width of 100 μ s and an off-resonance frequency of 535 Hz (Figure 4b), a reduction of I_{CP} of ca. 2% occurs at $v_{CP}(1/(4\tau_{CP})) = 250$ Hz. However, these results are obtained assuming that the pulse calibration is perfect, and may therefore underestimate the reduction of I_{CP} . For example, if the R.F. pulse width is miscalibrated by 2%, corresponding to a CPMG pulse that rotates magnetization by 184° instead of 180°, the reduction in I_{CP} at v_{CP} near 800 Hz becomes significant (Figure 4c). Increasing the miscalibration to 5% results in larger fluctuations in the I_{CP} profile (Figure 4d). When B_1 inhomogeneity is incorporated into the calculation, the I_{CP} intensity profile changes and the amplitude of the fluctuations decreases (Figures 4e and 4f).

The simulations show that the off-resonance effect invariably reduces I_{CP} , which results in a systematic overestimate of the measured R_2 values. However, the reduction in I_{CP} varies abruptly as v_{CP} is varied, which causes the R_2 profile to exhibit sudden displacements (both positive and negative) from one v_{CP} value to next. We earlier pointed out that the ¹⁵N R_2 of residue 3 at v_{CP} =300 Hz and

 R_2 of residue 96 at $v_{CP} = 150$ Hz exhibited lower R_2 values than those at larger v_{CP} values. However, the data are more properly regarded as more-or-less positive displacements from their true values, i.e., their values in the absence of off-resonance effects.

It is difficult to calculate the precise error in R_2 due to the resonance offset because neither the pulse calibration error nor the shape of the B_1 inhomogeneity profile is accurately known. Nevertheless, the profiles in Figure 4 were calculated assuming conditions that are thought to approximate those of the ¹⁵N R_2 measurements and therefore provides a reasonable estimate of the systematic error in I_{CP} . From the I_{CP} profiles in Figure 4, the average fractional systematic error is estimated to be 1-2%. In the case of residue 3 whose ¹⁵N signal is located 535 Hz off resonance, a 2% error in I_{CP} corresponds to 2.9–3.9% R_2 error. Similarly, in the case of residue 96 whose ¹⁵N signal is 426 Hz off resonance, a 2% error in $I_{\rm CP}$ corresponds to 2.9–3.4% R_2 error. These errors caused by the off-resonance effect are nearly 10-fold larger than random errors ($\Delta R_2/R_2$, ca. 0.5% for both residues). On the other hand, these calculated off-resonance errors are close to those previously estimated from the R_2 dispersion curves in Figure 2 (4.5% and 3.5% for residue 3 at $v_{CP} = 300$ Hz, and residue 96 at $v_{CP} = 150$ Hz, respectively).

Observed R_2 profiles provide an estimate of systematic error

By acquiring spectra measured at two or more different carrier frequencies, it is possible to accumulate R_2 dispersion data that are nearly free from error due to the off-resonance effect. However, when we desire an initial estimate of conformational exchange parameters, R_2 measurements with errors of a few percent should suffice and it is not necessary to accumulate data sets acquired at different carrier frequencies. In this case, it is useful to estimate the average off-resonance systematic error in R_2 from the experimental R_2 profiles that do not exhibit significant dispersion. In this way one is able to check that the experimental errors derived using equation (1) from signal-to-noise are reasonable.

In the protocol applied to the current data, we calculated the quantity η for each dispersion



Figure 5. Histograms of η values for the (a) ¹H and (b) ¹⁵N R_2 dispersion data of the HIV-1 protease. Comparison of the histograms reveals a clear shift to larger values of η in the case of the ¹⁵N η histogram (b). The observed shift is caused by the presence of systematic error in ¹⁵N R_2 data. ¹⁵N profiles with $\eta > 10$ were found for residues 3, 4, 6, 40, 96, 97, 98, and 99 while ¹H profiles with $\eta > 6$ were found for residues 4, 6, 7, 95, and 97. These residues are in the N- and C-terminal regions of the protease except for residue 40, which is located in a flexible loop.



Figure 6. Values of $R_2^{\text{rmsd}}/\langle R_2 \rangle$ (*) and $\langle \Delta R_2 / R_2 \rangle$ (+), plotted against off-resonance chemical shift for (a) ¹H and for (b) ¹⁵N spins. The sites that undergo significant chemical exchange ($\eta > 10$ for ¹⁵N, and $\eta > 6$ for ¹H) or exhibit large random error ($\Delta R_2 / R_2 \gg 0.8\%$) were not included in the plot. In (b), values of $R_2^{\text{rmsd}}/\langle R_2 \rangle$ and $\langle \Delta R_2 / R_2 \rangle$ represented by (O) and (×), respectively, are sites that are in, or close to, the extended dimer interface (residues 2–11, 23–27, 49–53, 88–99).

profile. In the absence of chemical exchange and systematic error in R_2 , η should have an average value near unity. Histograms in Figure 5 show the correspondence between the number of data sets and the values of η in the case of ¹H (Figure 5a) and ¹⁵N (Figure 5b). The average value of η is 1.9 for the ¹H data with an asymmetric distribution in which 75% of η values are less than 2.0 (Figure 5a). When one excludes data with obvious R_2 dispersion, the average ¹H η becomes 1.4, indicating that ¹H data do not suffer from significant systematic errors. In contrast, the average η for the ¹⁵N data is 4.8, and more importantly, the average η is 3.7 even when datasets with obvious R_2 dispersion are excluded (Figure 5b).

In this calculation of the average η , a total 26 data sets, for sites that are located in or close to the dimer interface (residues 2–11, 23–27, 49–53, 88–99), were excluded. Among these sites,

 $\eta({}^{1}H) > 6$ was observed for residues 4, 6, 7, 95, and 97, which are in the N- and C-terminal strands of the protease. Except for residue 40, all eight sites that exhibit $\eta({}^{15}N) > 10$ are in the N- and C-terminal strands, and their R_2 profiles exhibit significant dispersion, Figure 1b.

The correlation between the systematic error in R_2 and resonance offset is more clearly seen in Figure 6, where $R_2^{\text{rmsd}}/10$ values are plotted against the resonance offset frequency. At ¹⁵N off-resonance frequencies greater than 200 Hz, ¹⁵N $R_2^{\text{rmsd}}/\langle R_2 \rangle$ values are larger than those of (Figure 6b), indicating the presence of systematic error in addition to random error. In contrast, the plot of ¹H $R_2^{\text{rmsd}}/\langle R_2 \rangle$ (Figure 6a) does not reveal a significant off-resonance dependence, presumably because the ¹H 180° pulses were 5 times shorter than those of ¹⁵N.

After deleting the data points represented by the open circles in Figure 6, which are the data points for the sites that are at or near the dimer interface, a third order polynomial function was used to fit the plot of $R_2^{\text{rmsd}}/\langle R_2 \rangle$ against off-resonance shift (Figure 6b). This simple fit is empirical and the theoretical dependence on the off-resonance error is more complex (Ross et al., 1997; Korzhnev et al., 2000). The third order polynomial function provides an estimate of the average total error (both systematic and random) in each R_2 dispersion data set caused by the off-resonance chemical shift. The advantage of this empirical error estimate is that it is applicable even when some of the factors that contribute to the off-resonance error are not known.

R_2 dispersion profiles selected for fitting

¹H R_2 dispersion data sets with $R_2^{\text{rmsd}} > 3\Delta R_2$ were selected for fitting by the Carver-Richards equation (Carver and Richards, 1972; Davis et al., 1994). ¹⁵N R_2 dispersion data sets with R_2^{rmsd} is three times larger than the values given by the polynomial function (which provides an estimate of the total error in R_2) in Figure 6 were selected for fitting. Of the data sets that satisfied these criterion, those of residues in the N- and C- terminal β -strands (residue numbers, 1-4 and 96-99) and the loop adjacent to the N-terminal strand (residue numbers, 5-8) were of interest because previous work had indicated that these residues executed a cooperative motion of functional significance (Ishima et al., 1999). Among these data sets, the ¹⁵N data of residues 3, 4, 6, 96, 97, 98, and 99, and ¹H data of residues 5, 6, 7, 97, and 98 were selected to optimize the exchange parameters by χ^2 -minimization. Another region that was of interest is the tip of the flaps that cover the active site. Residues 50 and 51 at the flap tip have been shown to undergo a significant conformational fluctuation related to a β I-turn/ β II-turn exchange (Nicholson et al., 1996; Ishima et al., 1999). The R_2^{rmsd} of the ¹H and ¹⁵N R_2 data for residues 50 and 51 marginally satisfied the above criteria, that is, R_2^{rmsd} was 2.4–3.2 times larger than ΔR_2 (for ¹H) or total (empirical) ΔR_2 (for ¹⁵N).

Data fitting using Carver–Richards equation

When each R_2 dispersion data set is individually fit with the Carver–Richards equation, four parameters, R_2^0 , $\delta\omega$, p_a , and τ_{ex} are optimized by χ^2 minimization as described in Materials and Methods. In fitting the ¹⁵N data, $\Delta R_2^{i,err}$ is the total error in R_2 , systematic plus random, obtained from the polynomial function, Figure 6. It is 1.2–6.1 times larger than the random error depending upon the resonance offset. This uniform scaling of the R_2 error in each dispersion data set does not shift the global minimum of the χ^2 function (Equation 3) of a residue specific fit. However, such an adjustment of $\Delta R_2^{i,err}$ values provides better estimate of uncertainties of the exchange parameters derived from Monte-Carlo or covariance matrix methods. In addition, the error scaling more fairly weighs the contribution of the individual R_2 data sets in the global fitting of a group of R_2 dispersion data sets.

Table 1 lists the values of the optimized parameters and the best-fit χ^2 values. Each dispersion profile consists of eight (for ¹⁵N) and nine (for ¹H) R_2 data points. Therefore, one expects χ^2 values of ca. 4 and 5 for the fits to the ¹⁵N and ¹H data sets, respectively, assuming a normal distribution of errors. As evident from Table 1, the χ^2 values are in approximate agreement with this expectation, except for the ¹⁵N R_2 data of residue 98. The former result indicates that our estimates of the total errors in the data are approximately correct. We think that the large χ^2 value of residue 98 in the ¹⁵N data is due to the fact that the signals of residues 98 and 22 overlap. It is noteworthy that, without scaling the total error for ${}^{15}N$ R_2 data sets, χ^2 values of some of the ¹⁵N sites exceed 20. Such large χ^2 values are inconsistent with the observation that random errors of ¹⁵N data are less than those of ${}^{1}H$ (Figure 6).

Examination of Table 1 shows that, except for R_2^0 , the optimized parameters have large uncertainties. For example, optimized pa values used to fit the Monte Carlo trial data set ranged from 0.5 to 0.99 for most residues, corresponding to variations in $p_a(1-p_a)$ ranging from 0.25 to 0.01. When exchange is near to, or in, the fast limit, R_{ex} , (the relaxation rate due to exchange) is proportional to the product $p_a(1 - p_a)\delta\omega^2$. Therefore in the fast exchange limit large correlated uncertainties in $p_a(1 - p_a)$ and $\delta \omega^2$, will be observed where as the uncertainty in their product, $p_a(1 - p_a)\delta\omega^2$ will much smaller (Table 2). The uncertainties of these parameters (Table 1) suggest that exchange is in the fast limit for most residues in the dimer interface of the protease. The optimized p_a and $\delta \omega$ values have small uncertainties only for the ¹⁵N

Residue	p _a		$\delta \omega$ (Hz)		$\tau_{\rm ex}~({\rm ms})$		R ₂ ⁰ (s)		χ^2
	Value	Error	Value	Error	Value	Error	Value	Error*	
¹⁵ N									
3	0.78		27		1.4		14		13.
	0.63	0.79-0.5	41	70–20	8.1	2.1 - 1.1	14	14.4–13.7	
4	0.98		98		3.3		15		1.8
	0.82	0.98-0.63	84	120-31	27.	76-1.5	15	15.2-14.9	
6	0.97		80		2.5		17		1.6
	0.83	0.97-0.64	64	97–27	7.7	4.2-1.6	17	17.2–16.9	
96	0.76		21		1.2		16		5.8
	0.65	0.79–0.5	38	48-17	1.1	2.6-0.79	16	16.4-15.6	
97	0.97		84		2.2		15		2.3
	0.88	0.97 - 0.74	79	110-35	8.2	4.8-1.5	16	15.7-15.2	
98	0.93		120		1.8		15		21.
	0.93	0.93-0.92	120	130-110	1.9	2.3-1.5	15	15.4-14.6	
99	0.98		110		3.6		18		2.5
	0.83	0.98-0.53	100	130-74	29	89-1.7	18	17.7–17.4	
^{1}H									
4	0.95		200		1.1		19		2.0
	0.93	0.96-0.93	310	330-280	1.8	1.8-0.93	19	19.8-18.7	
6	0.80		85		0.77		25		4.0
	0.76	0.94-0.5	110	140-68	1.0	1.1 - 0.74	25	25.9-24.8	
7	0.64		37		1.1		22		3.8
	0.56	0.69-0.5	37	38–36	1.1	1.2-1.0	22	21.8-21.5	
97	0.90		44		1.3		21		5.8
	0.69	0.96-0.5	64	120-26	1.1	5.5-1.1	21	21.2-20.8	
98	0.73		36		1.2		25		7.7
	0.56	0.63-0.5	71	130-30	14	42-0.98	25	25.6-24.7	

Table 1. Optimized values and uncertainties of exchange parameters obtained by independently fitting the R_2 data set of each residue.^{*}

^{*}For each data set, the optimized parameters are listed in the first row. The average values of these parameters and their uncertainties derived from 100 Monte-Carlo trials are listed in the second row. Each uncertainty is given as an error range, which indicates the 68% of the Monte-Carlo trials that yielded parameter values closest to the average value. The standard deviations of the parameter distributions derived from the Monte-Carlo calculations were not used to estimate the parameter uncertainties because the distributions were highly asymmetric.

data of residue 98 and the ¹H data of residue 4. In these two cases chemical exchange is not in the fast limit because their optimized $\delta \omega$ values are large, greater than 100 Hz. dispersion profiles of all the sites using a single pair of optimized p_a and τ_{ex} values.

Group fit of R_2 dispersion data

Table 1 also shows that optimized τ_{ex} values typically exhibit large uncertainties, some of which (the ¹⁵N data of residues 4 and 99, and the ¹H data of residue 98) span nearly two orders of magnitude. The uncertainty of τ_{ex} values is least in the case of R_2 profiles that have the largest dispersion amplitudes (e.g. residues 3 and 98 for ¹⁵N and residues 4, 6, and 7 for ¹H, see Figure 1). Despite their large uncertainties, all optimized τ_{ex} values lie in the range of 1–2 ms, suggesting that all the sites listed in Table 1 participate in the same motion. If this is the case, it should be possible to fit the

The R_2 profiles of all sites listed in Table 1 were fit with p_a and τ_{ex} as global parameters, i.e., their values were the same for all the sites (the second method). As before, $\delta\omega$ and R_2^0 were assumed to be local parameters whose values were free to vary from one site to the next. This fitting protocol yielded minimized total and normalized χ^2 values of $\chi^2 = 94.1$ and $\chi^2/N = 1.25$ (normalized with respect to the number of degrees of freedom) with $p_a = 0.94 \pm 0.003$ and $\tau_{ex} = 1.6 \pm 0.05$ ms (Figure 7c). The uncertainties of p_a and τ_{ex} ,

.6 ms. ^a										
Residue	$\delta\omega$ (Hz)		$\mathbf{R}_{2}^{0}\left(s ight)$	χ^2						
	Value	Error	Value	Error						
¹⁵ N										
3	46	2.3	14	0.23	12					
4	46	0.89	15	0.092	2.8					
6	56	1.1	17	0.13	3.6					
96	35	2.3	16	0.20	6.2					
97	56	0.96	15	0.11	2.9					
98	130	3.4	15	0.36	26.					
99	45	1.3	18	0.13	4.9					
^{1}H										
4	300	17	19	0.49	4.0					
6	170	9.1	25	0.51	6.9					
7	80	1.6	22	0.17	8.6					
97	55	1.5	21	0.16	7.7					

Table 2. Optimized values of $\delta \omega$ and R_2^0 , and their uncertainties

obtained by fitting the R_2 data with p_a and τ_{ex} fixed to 0.94 and

^aThe standard deviations of the parameter distributions derived from the Monte-Carlo calculations were used to estimate the parameter uncertainties.

25

0.45

7.9

4.3

optimized by fitting all R_2 profiles using p_a and τ_{ex} as global parameters, are much smaller than the uncertainties in these parameters individually optimized by fitting the R_2 profile of each residue. The latter protocol uses a larger number of parameters to fit the data than does the former; however, the *F*-test showed that an insignificant improvement in the fit (the probability that the improvement could have occurred by chance was 0.67) was gained using the larger number of fitting parameters. In other words, *F*-test justifies fitting the R_2 profiles using p_a and τ_{ex} as global parameters that characterize a cooperative motion of the dimer interface of the protease.

In order to test whether the ¹H and ¹⁵N data sets yield optimized values of p_a and τ_{ex} that are mutually consistent, the R_2 profiles of each type of spin were separately fit (the third method). Optimized values of p_a and τ_{ex} were obtained from fits of ¹H R_2 profiles of residues 4, 6, 7, 97 and 98 and of ¹⁵N R_2 profiles of residues of 3, 4, 6, 96, 97, 98, and 99. The minimum χ^2 values were 28.9 ($\chi^2/$ N = 0.89) and 53.7 ($\chi^2/N = 1.34$) for ¹H and ¹⁵N data sets, respectively, indicating that the fits of data sets of both types of spins are of high quality. The optimized values of p_a and τ_{ex} for ¹H data, 0.95 and 1.4 ms, differed slightly from those obtained



Figure 7. Contour plots of χ^2 as a function of p_a and τ_{ex} calculated by fitting (a) ¹H R_2 profiles of residues 4, 6, 7, 97, and 98, (b) ¹⁵N R_2 profiles of residues 3, 4, 6, 96, 97, 98, and 99, and (c) the combined ¹H and ¹⁵N data sets. The numbers on the plots are χ^2 values corresponding to the various contour levels, and the minimum χ^2 value is indicated by the + within the smallest contour. Figures (a) and (b) are calculated for fits obtained using the third method, and Figure (c) was calculated for fits obtained using the second method, as described in Materials and Methods.

for the¹⁵N data, 0.93 and 1.6 ms (Figures 7a and 7b). The *F*-test showed that the difference in the optimized parameters is at the margin of the statistical significance, with p = 0.01, indicating

98

68

the difference in the optimized parameters is barely greater than their uncertainties. Because these results indicate that the Carver–Richards equation provided good quality of fits to the experimental data, we did not attempt to fit the data using the exact, but complex, equation for two-site exchange (Tollinger et al., 2001) or a three-site exchange model (Trott and Palmer, 2004).

We also performed χ^2 minimization to determine optimized exchange parameters for the residues in the tips of the flaps of the protease (residues 50 and 51). Previous studies showed that these residues undergo slow dynamics (Ishima et al., 1999; Nicholson et al., 1996), which is also evident from the larger ¹⁵N R_2 values of these residues (Figure 1b) compared with the average R_2 values of residues seen in Figure 3 (which do not undergo exchange). The respective η values of residues 50 and 51 were 5.7 and 6.4 for ^{15}N and 3.1 and 2.7 for ¹H. The ¹H data of residue 51 nearly satisfied $R_2^{\text{rmsd}} > 3\Delta R_2$ condition that was the criterion used to select data sets for χ^2 minimization above. The χ^2 values of residue 50 were 2.7($^{15}N)$ and $2.8(^{1}\text{H})$ whereas those of residue 51 were 10 (^{15}N) and 8.9 (^{1}H) . Fitting all four data sets using the global protocol (method 2), resulted in an optimized p_a, with a large range of uncertainty, $0.55 \leq p_a \leq 0.98.$ This observation suggests that conformational exchange for these residues is in the fast limit.

As we have noted previously (Ishima and Torchia, 2003) one advantage of measuring both ¹H and ¹⁵N dispersion data sets is that one is able to measure R_2 ¹H dispersion for a particular amide when $\delta \omega_{\rm N}$ is small, and vice versa. As an example, the R_2 dispersion is more clearly evident in the ¹H data of the N-terminal residues, whereas the R_2 dispersion is more evident in the ¹⁵N data of the C-terminal residues. Taken together the data for the two spin types reveal that both C-terminal inner β -strands and N-terminal outer β -strands undergo a concerted conformational change. Another advantage of having two independent data sets is that the optimized parameters can be compared and validated. In the present case the optimized parameters determined by individual group fits of ¹⁵N and ¹H are consistent with each other, and are also consistent with the parameters determined by the global fit of both ¹⁵N and ¹H data sets.

Acknowledgements

This work was supported by the Intramural AIDS Targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health.

References

- Carver, J.P., and Richards, R.E. (1972) J. Magn. Reson., 6, 89–105.
- Czisch, M., King, G.C., and Ross, A. (1997) J. Magn. Reson., 126, 154–157.
- Davis, D.G., Perlman, M.E., and London, R.E. (1994) J. Magn. Reson. B, 104, 266–275.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biomol. NMR, 6, 277–293.
- Ernst, R.R., Bodenhausen, G., and Wokaun, A. (1987) Principles of Nuclear Magnetic Resonance in One and Two Dimensions, Clarendon Press, Oxford.
- Garrett, D.S., Powers, R., Gronenborn, A.M., and Clore, G.M. (1991) J. Magn. Reson., 95, 214–220.
- Geen, H., and Freeman, R. (1991) J. Mag. Reson., 93, 93-141.
- Guenneugues, M.P.B., and Desvaux, H. (1999) J. Magn. Reson., 136, 118–126.
- Ishima, R., Wingfield, P.T., Stahl, S.J., Kaufman, J.D., and Torchia, D.A. (1998) J. Am. Chem. Soc., 120, 10534–10542.
- Ishima, R., Freedberg, D.I., Wang, Y.X., Louis, J.M., and Torchia, D.A. (1999) *Structure*, 7, 1047–1055.
- Ishima, R., and Torchia, D.A. (2003) J. Biomol. NMR, 25, 243–348.
- Ishima, R., Baber, J., Louis, J.M., and Torchia, D.A. (2004) J. Biomol. NMR, 29, 187–198.
- Jones, J.A., Hodgkinson, P., Barker, A.L., and Hore, P.L. (1996) J. Magn. Reson. Ser. B., 113, 25–34.
- Jones, J.A. (1997) J. Magn. Reson, 126, 283-286.
- Katoh, E., Louis, J.M., Yamazaki, T., Gronenborn, A.M., Torchia, D.A., and Ishima, R. (2003) *Protein Sci.*, 12, 1376–1385.
- Korzhnev, D., Tischenko, E.V., and Arseniev, A.S. (2000) J. Biomol. NMR, 17, 231–237.
- Korzhnev, D.M., Salvatella, X., Vendruscolo, M., Di Nardo, A.A., Davidson, A.R., Dobson, C.M., and Kay, L.E. (2004) *Nature*, 430, 586–590.
- Loria, J.P., Rance, M., and Palmer, A.G. III (1999) J. Am. Chem. Soc., 121, 2331–2332.
- Louis, J.M., Ishima, R., Nesheiwat, I., Pannell, L.K., Lynch, S.M., Torchia, D.A. and Gronenborn, A.M. (2002) J. Biol. Chem.
- Mulder, F.A.A., Skrynnikov, N.R., Hon, B., Dahlquist, F.W., and Kay, L.E. (2001) J. Am. Chem. Soc., 123, 967–975.
- Mulder, F.A.A., Hon, B., Mittermaier, A., Dahlquist, F.W., and Kay, L.E. (2002) J. Am. Chem. Soc., 124, 1443–1451.
- Nicholson, L.K., Kay, L.E. and Torchia, D.A. (1996). In NMR Spectrocsopy and its Application to Biomedical Research. Sarkar, S.K. (Eds.), Elsevier, Amsterdam, pp. 241–279.
- Orekhov, V.Y., Pervushin, K.V., and Arseniev, A.S. (1994) Eur. J. Biochem., 219, 887–896.
- Palmer, A.G. III, Kroenke, C.D., and Loria, J.P. (2001) Methods in Enzymology, 339, 204–238.

- Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T. (1988) *Numerical Recipes in C*, Cambridge University Press, Cambridge U.K.
- Ross, A., Czisch, M., and King, G.C. (1997) J. Magn. Reson, 124, 355–365.
- Skrynnikov, N.R., Mulder, F.A.A., Hon, B., Dahlquist, F.W., and Kay, L.E. (2001) J. Am. Chem. Soc., 123, 4556–4566.
- Tollinger, M., Skrynnikov, N.R., Mulder, F.A.A., Forman–Kay, J.D., and Kay, L.E. (2001) *J. Am. Chem. Soc.*, **123**, 11341–11352.
- Trott, O., and Palmer, A.G. III (2004) J. Magn. Reson., 170, 104–112.