Journal of Biomolecular NMR, **15:** 271–288, 1999.
KLUWER/ESCOM

Refined solution structure and backbone dynamics of ^{15}N-labeled C12A-p8^{MTCP1} studied by NMR relaxation

Philippe Barthe[a], Laurent Chiche[a], Nathalie Declerck[a], Marc-André Delsuc[a], Jean-François Lefèvre[b,*], Thérèse Malliavin[a], Joel Mispelter[c], Marc-Henri Stern[d], Jean-Marc Lhoste[a] & Christian Roumestand[a,**]

[a]*Centre de Biochimie Structurale, CNRS-UMR 9955, INSERM-U414, Université de Montpellier I, Faculté de Pharmacie, 15 Avenue Charles Flahault, F-34060 Montpellier Cedex, France*
[b]*ESBS, Université Louis Pasteur, CNRS UPR-9003, Blvd S. Brant, F-67400 Illkirch-Graffenstaden, France*
[c]*INSERM-U350, Institut Curie, Biologie, Centre Universitaire Bât 112, F-91405 Orsay Cedex, France*
[d]*Unité INSERM-U462, Hôpital Saint-Louis, F-75475 Paris, France*

Received 2 June 1999; Accepted 1 October 1999

Key words: leukemia, NMR structure, protein dynamics, translocations

Abstract

MTCP1 (for Mature-T-Cell Proliferation) was the first gene unequivocally identified in the group of uncommon leukemias with a mature phenotype. The three-dimensional solution structure of the human p8^{MTCP1} protein encoded by the *MTCP1* oncogene has been previously determined by homonuclear proton two-dimensional NMR methods at 600 MHz: it consists of an original scaffold comprising three α-helices, associated with a new cysteine motif. Two of the helices are covalently paired by two disulfide bridges, forming an α-hairpin which resembles an antiparallel coiled-coil. The third helix is orientated roughly parallel to the plane defined by the α-antiparallel motif and appears less well defined. In order to gain more insight into the details of this new scaffold, we uniformly labeled with nitrogen-15 a mutant of this protein (C12A-p8^{MTCP1}) in which the unbound cysteine at position 12 has been replaced by an alanine residue, thus allowing reproducibly high yields of recombinant protein. The refined structure benefits from 211 additional NOEs, extracted from ^{15}N-edited 3D experiments, and from a nearly complete set of ϕ angular restraints allowing the estimation of the helical content of the structured part of the protein. Moreover, measurements of ^{15}N spin relaxation times and heteronuclear ^{15}N{^{1}H}NOEs provided additional insights into the dynamics of the protein backbone. The analysis of the linear correlation between J(0) and J(ω) was used to interpret relaxation parameters. It appears that the apparent relative disorder seen in helix III is not simply due to a lack of experimental constraints, but associated with substantial contributions of sub-nanosecond motions in this segment.

Abbreviations: 2D (3D), two-dimensional (three-dimensional); NOE, nuclear Overhauser enhancement; ^{15}N{^{1}H}, heteronuclear ^{15}N nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; DQF-COSY, double-quantum-filtered scalar-correlated spectroscopy; *z*-TOCSY, *z*-filtered total correlation spectroscopy; HSQC, heteronuclear single quantum correlation; HMQC, heteronuclear multiple quantum correlation; $R_N(N_z)$ (R_1), heteronuclear ^{15}N longitudinal relaxation rate; $R_N(N_{xy})$ (R_2), heteronuclear ^{15}N transverse relaxation rate; $R_N(H_z \rightarrow N_z)$, cross-relaxation rate between ^{15}N and its attached amide proton; rmsd, root mean square deviation; PCR, polymerase chain reaction; DTT, 1,4-dithiothreitol.

Introduction

T-cell lymphoproliferative diseases are often associated with recurrent chromosomal translocations involving T-cell receptor genes (TCR) and genes that are thought to play a role in the pathogenesis of these diseases. These chromosomal translocations have been proposed to occur during recombination events of the TCR genes, leading to genetic aberrations which deregulate proto-oncogenes that become juxtaposed

*Deceased March 1999.
**To whom correspondence should be addressed.

with the TCR genes. The *MTCP1* gene (also called *c6.1B*) (Fisch et al., 1993; Stern et al., 1993), located in the Xq28 chromosomal region, was the first gene to be unequivocally identified in the heterogeneous group of uncommon leukemias presenting a mature phenotype. It is involved in the translocation t(X;14)(q28;q11) recurrently associated with a rare subset of mature T-cell proliferations (Dallapiccola et al., 1984; Witzig et al., 1986; Goyns et al., 1993; Stern et al., 1993; Madani et al., 1996).

The two *MTCP1* major splicing forms, A and B, encode two entirely different proteins. The longer transcript, B1, codes for a 107-residue (13 kDa) protein known as p13^{MTCP1} (Madani et al., 1996). This protein shows high sequence similarity (40% identity) with p14^{TCL1} (Fu et al., 1994), the product of the 14q32.1 oncogene *TCL1*. Type A transcripts code for a small, 8 kDa (68 residues) cysteine-rich protein (seven cysteines out of 68 amino acids), p8^{MTCP1} (Soulier et al., 1994) localized in mitochondria (Madani et al., 1995). Whereas the expression of p13^{MTCP1} is restricted to mature T-cell proliferation with t(X,14) translocations, p8^{MTCP1} is expressed at low levels in most human tissues and is over-expressed in the proliferating T-cells (Soulier et al., 1994; Madani et al., 1996). Thus, even though p8^{MTCP1} may participate in oncogenesis, it is more likely to be associated with a function common to many cell types.

A previous report on the structure determination of human p8^{MTCP1} by two-dimensional (2D) homonuclear NMR reveals an original scaffold consisting of three α-helices, associated with a new cysteine motif (Barthe et al., 1997). In the present paper, we present the refined structure of the ^{15}N-labeled mutant protein C12A-p8^{MTCP1}, using double-resonance heteronuclear NMR. In this mutant, the non-paired cysteine at position 12 has been replaced by an alanine residue: this has been shown to ensure reproducibly high yields of protein, whereas the structure remains unchanged. To further investigate the original scaffold of p8^{MTCP1}, we complemented the structural study by an analysis of the backbone dynamics, using experimental ^{15}N relaxation parameters measured on the NH vectors. Since the spectral density functions J(0), J(ω_N) and $<$J(ω_H)$>$ can be straightforwardly calculated from only T_1, T_2, and ^{15}N{^{1}H}NOE relaxation parameters using the so-called reduced spectral density matrix relation (Lefèvre et al., 1996; Markus et al., 1996), we choose to use this approach to analyse our relaxation data. Quantitative results on the motions involved in the NH vector dynamics were obtained from the linear correlation experimentally observed between J(0) and J(ω_N) (Lefèvre et al., 1996).

Materials and methods

Sample preparation

The gene encoding p8^{MTCP1} had originally been cloned into pGEX-2T (Pharmacia Biotech.), allowing high expression of the protein fused to gluthation-S-transferase (GST). However, after thrombin cleavage and further purification, the protein was found to rapidly undergo oligomerisation and aggregation, probably because of a free cysteinyl side-chain at position 12 as observed in p8^{MTCP1} native structure (Barthe et al., 1997). Cys12 was thus mutated to alanine by means of PCR-based site-directed mutagenesis. Moreover, a more precise cloning of the p8^{MTCP1} gene was required in order to eliminate 4 and 5 extra residues introduced at the N- and C-termini, respectively, of the protein expressed in the original pGEX-p8^{MTCP1} construct. The pGEX-C12A-p8^{MTCP1} construct was then introduced into *E. coli* strain BL21(DE3). The mutant protein was purified as previously described for the native protein obtained by expression of the original pGEX-p8^{MTCP1} construct (Barthe et al., 1997), except that DTT was omitted during purification steps. Uniform ^{15}N labelling was achieved by growing the cells in minimal medium with $^{15}NH_4Cl$ as sole nitrogen source.

NMR measurements

All NMR experiments were carried out at 14.1 Tesla, on a Bruker AMX600 spectrometer equipped with a 5 mm *z*-gradient ^{1}H-^{13}C-^{15}N triple resonance probe and with the conventional Bruker gradient linear amplifier (maximum output power: 50 G/cm). Protein samples were dissolved in a 25 mM phosphate buffer with 50 mM NaCl, with 10% ^2H_2O for the lock and pH adjusted to 6.5. The temperature was carefully adjusted using a calibration sample (80% glycol in d_6-DMSO) and set to 20 °C, except where otherwise noted. Note that this temperature corresponds to the 25 °C reading used for the native protein p8^{MTCP1} in our previously published results (Barthe et al., 1997). In all experiments, the ^{1}H carrier was centered on the water resonance, and the WATERGATE sequence (Piotto et al., 1992; Sklenar, 1995) was used to suppress the solvent resonance. The spectral widths in the ^{1}H and ^{15}N dimensions were 6900 Hz and 1200 Hz,

respectively. Quadrature detection in the indirect dimensions was achieved using States-TPPI (Marion et al., 1989b), and the initial evolution period in all experiments was delayed by exactly half the chosen dwell in order to remove baseline distortion and optimize aliasing characteristics in the processed spectra of the [^{1}H,^{15}N] double resonance experiments (Bax et al., 1991). All NMR spectra were processed and analysed with Gifa (version 4.22) (Pons et al., 1996).

2D homonuclear spectroscopy. Homonuclear 2D experiments (NOESY (Jeener et al., 1979; Kumar et al., 1980), *z*-TOCSY (Braunschweiler and Ernst, 1983; Davis and Bax, 1985; Rance, 1987), and DQF-COSY (Rance et al., 1983)) were carried out using similar experimental conditions and parameters as previously described for p8^{MTCP1} (Barthe et al., 1997), except that there was no need to add (^{2}H)DTT to the sample to prevent protein dimerization. Experiments were recorded on two samples with different protein concentration (4 mM and 0.4 mM).

Diffusion measurements. Diffusion experiments were recorded using a modified LED sequence (Gibbs and Johnson, 1991) using bipolar gradients (Wider et al., 1994; Wu et al., 1995) to prevent any parisitic effect arising from eddy currents. Special care was taken for solvent signal suppression: the WATERGATE sequence was used in association with low power irradiation of the water signal during both the relaxation delay (1 s) and the constant diffusion delay (150 ms). In order to account for any variation of the sample viscosity upon dilution, Tris (4 mM) was added to the 4 mM sample of C12A-p8^{MTCP1}, as an internal standard. The self-diffusion constants were then measured on this sample, as well as on several dilutions obtained from it. To determine the self-diffusion constants, 36 1D experiments with an identical diffusion delay were performed with pulse field gradients of 1.8 ms duration and variable gradient strength ranging from 1 G/cm to 47.5 G/cm (2% to 95% of the maximum output power of the amplifier). Each individual experiment has a time domain of 2K complex points, and was acquired with a number of transients ranging from 32 (minimum phase cycling) to 1080, depending on the sample protein concentration. The diffusion coefficients were obtained from the intensity decays of selected non-exchangeable peaks when increasing the gradient intensity by Maximum Entropy processing (Delsuc and Malliavin, 1998) with an inverse Laplace transform computed on 128 points. The corresponding molecular weights were deduced, without taking into account any shape factor, from a standard curve previously established under the same experimental conditions from a series of proteins of known molecular weights ranging from 4 kDa to 50 kDa (to be published elsewhere).

$^3J_{NH-H\alpha}$ measurement. To complete our previous set of $^3J_{NH-H\alpha}$ coupling constants (Barthe et al., 1997), a series of 12 J-modulated [^{1}H,^{15}N] COSY spectra (Neri et al., 1990) was performed on a 0.4 mM sample (20 °C, pH 6.5), using different evolution delays in the range 10–100 ms for homonuclear J-modulation and a time domain data size of 64 t_1 × 2K t_2 complex points and 32 transients per complex t_1 increment. The conventional experiment was modified in order to introduce the WATERGATE sequence in the last reverse INEPT step for water suppression purposes (Santoro et al., 1996). The peak heights were then fitted accordingly (Billeter et al., 1992) to extract the individual coupling constant values.

Amide proton exchange experiments. A sample of 400 µl of a fully protonated 1 mM ^{15}N protein (pH 6.5) was lyophilized. The protein was then redissolved in the same amount of ^2H_2O and a series of 24 [^{1}H,^{15}N] 2D HSQC (Bodenhausen and Ruben, 1980; Bax et al., 1990) was recorded using a time domain data size of 64 t_1 × 512 t_2 complex points and 4 transients per complex t_1 increment. The total measuring time for individual experiments was 15 min. The first experiment was started 5 min after preparation of the ^2H_2O solution, the following ones added up to a total recording time of about 2 days. Two series of experiments were recorded at 5 °C and 20 °C, respectively. Protection factors for each amide proton were calculated according to Bai et al. (1993).

[^{1}H,^{15}N] 3D experiments. The following double resonance [^{1}H,^{15}N] 3D spectra were acquired with a uniformly ^{15}N-labeled 4 mM sample of protein (10% ^{2}H2O, 20 °C, pH = 6.5): HMQC-NOESY-HMQC (Frenkiel et al., 1990; Ikura et al., 1990), NOESY-HSQC and TOCSY-HSQC (Marion et al., 1989a; Bax et al., 1990). The NOESY experiments were recorded with a mixing time of 150 ms and the TOCSY experiment with an isotropic mixing period (TOWNY (Kadkhodaei et al., 1993)) of 50 ms. WALTZ-16 modulation (Shaka et al., 1983) was used to decouple ^{15}N during acquisition. For all the 3D experiments, 4 tran-

sients (16 for HMQC-NOESY-HMQC) were collected for each hypercomplex t_1,t_2 pair, with 512 points in the t_1 (^{1}H) dimension (80 points for HMQC-NOESY-HMQC), 80 points in the t_2 (^{15}N) dimension, and 1K complex points in the acquisition (t_3) dimension. The measuring time for one 3D experiment was about 3 days.

Relaxation rate constant and $^{15}N\{^1H\}$NOE measurements. Two series of relaxation rate constant measurements were performed, either on a 4 mM or a 0.4 mM protein sample. The pulse sequences used to determine ^{15}N $R_N(N_z)$ (R_1), $R_N(N_{xy})$ (R_2), and ^{15}N{^{1}H}NOE values were similar to those described (Peng and Wagner, 1992a,b; Kay et al., 1992). To minimize artifacts, pulse field gradients were inserted during the intervals when the spin system is in the H_zN_z state (Bax and Pochapsky, 1992), for which H_z and N_z denote *z* components of the $^1H^N$ and ^{15}N magnetization, respectively. R_1 data sets were recorded in such a way that the signal intensity decayed to zero as a function of the relaxation delay, allowing a two-parameter exponential fit. A recycle delay of 4 s was employed and ^{15}N decoupling during acquisition was performed using a WALTZ sequence. R_1 experiments were performed with 16 relaxation delays T (18, 54, 102, 150, 210, 258, 306, 402, 498, 606, 810, 1002, 1506, 2010, 2502 and 3006 ms). The delay between the 180° ^{1}H pulses used to suppress the DD-CSA cross-relaxation was 3 ms. To ensure that water magnetization is minimally perturbed by the application of ^{1}H pulses during the T_1 delays, on-resonance 3-9-19 pulse trains of 180° global flip angle were used with the excitation minimum positioned at the carrier frequency. R_2 experiments were recorded employing a Carr–Purcell–Meiboom–Gill (CPMG) pulse train (Carr and Purcell, 1954; Meiboom and Gill, 1958) consisting of four 180° ^{15}N pulses and a centred ^{1}H pulse, each cycle with 4 ms duration and the spin-echo period being approximately 1 ms. The ^{15}N pulse duration was 36 μs. Thirteen experiments were acquired with relaxation delays T of 16, 32, 48, 80, 112, 128, 160, 192, 256, 320, 384, 512 and 768 ms. In these two sets of experiments, the points corresponding to different relaxation delays were acquired in an interleaved manner to avoid any bias that could arise from possible shim degradation. To permit the estimation of noise levels, duplicate spectra were recorded for T = 18 ms and 150 ms (R_1 spectra) and T = 16 and 112 ms (R_2 spectra). All experiments were recorded with a time domain data size of 96 $t_1 \times$ 2K t_2 complex points and 4 or 8 transients per complex t_1 increment for the 4 mM and the 0.4 mM protein sample, respectively.

For heteronuclear ^{15}N{^{1}H}NOEs, special care was taken to avoid large errors that can occur when dealing with protons in fast exchange with the solvent. Accordingly, a carefully optimized water flip-back pulse (Grzesiek and Bax, 1993) was added before the last proton 90° pulse in the experiment without saturation. Proton saturation was achieved by application of high-power 120° pulses spaced at 20 ms intervals for 3 s prior to the first pulse on ^{15}N (Kay et al., 1989). A relaxation delay equal to 6 s between each scan was used in order to obtain a complete relaxation of water magnetization and to reduce effects arising from amide proton exchange. Moreover, the two experiments with and without proton saturation were acquired in an interleaved manner, FID by FID. A relaxation delay of 30 s was used before the FIDs of the experiment without saturation. Experiments were recorded with the same time domain data size as for the R_1 and the R_2 experiments, and 32 or 96 transients per complex t_1 increment for the 4 mM and the 0.4 mM protein sample, respectively. The NOE enhancements were obtained as the ratio of the peak heights in the spectra recorded with and without saturation of protons during the relaxation delay. Uncertainties for these NOEs were the sum of the uncertainties in the peak intensities of each experiment.

Molecular modeling calculations

All calculations were carried out following the standard protocol previously described for p8^{MTCP1} (Barthe et al., 1997). Briefly, 1142 distance restraints (212 intra-residue, 340 sequential, 404 medium-range ($i - j < 5$), and 186 long-range upper bound restraints, 882 lower bound restraints) were obtained from the volume of cross peaks measured on 2D NOESY and [^{1}H-^{15}N] 3D NOESY-HSQC spectra and 35 angular restraints on the φ dihedral angles were obtained from the coupling constants $^3J_{NH-H\alpha}$ measured on the J modulated [^{1}H-^{15}N] COSY, using the empirically calibrated Karplus relation (Karplus, 1963; Pardi et al., 1984). In addition, 21 χ_1 angles were obtained from the analysis of the $^3J_{H\alpha-H\beta}$ coupling constants and intra-residue NOEs (Hyberts et al., 1987), and the usual distance constraints were added to enforce the three disulfide bridges previously determined (Barthe et al., 1997). No additional restraints were used for hydrogen bonds. From these restraints, a set of 50 structures was generated with the variable target function program DYANA (Güntert and Wüthrich, 1991; ver-

sion 1.3, Güntert et al., 1997). Each run started from 999 randomized conformers. When no stereospecific assignment was possible, pseudoatoms were defined and corrections added as described by Wüthrich et al. (1983). Among the 50 preliminary structures generated by DYANA, 30 presented a value of the target function smaller than 4 Å^2 and no distance violation larger than 0.7 Å. The ϕ and χ_1 angles showed no violation greater than 7°.

The 30 best structures (based on the final target penalty function values) were further submitted to molecular mechanics energy refinement with the SANDER module of AMBER 4.1 (Pearlman et al., 1995) using the 1994 force field (Cornell et al., 1995) as previously described for p8^{MTCP1}. Briefly, 5000 cycles of restrained energy minimisation were first carried out followed by a 50 ps long simulated annealing procedure in which the temperature was raised to 900 K for 30 ps, then gradually lowered to 300 K. During this stage, the force constants for the NMR distance constraints and for the angular constraints were gradually increased from 3.2 to 32 kcal mol^{-1} Å^{-2} and from 0.5 to 50 kcal mol^{-1} rad^{-2}, respectively. A final restrained minimisation led to the refined structures discussed below. The structures were displayed and analyzed using the INSIGHT program (version 97.0, Biosym Technologies, San Diego, CA).

Relaxation data analysis

When the relaxation of the ^{15}N nucleus is predominantly caused by the dipolar interaction with its attached amide proton and by the anisotropy of its chemical shift, the relaxation data can be interpreted in terms of the motion of the ^{15}N-^{1}H vector. Given that the three experimentally determined parameters, $R_N(N_z)$, $R_N(N_{xy})$ and NOE, depend on the spectral density function at five different frequencies (Abragam, 1961), it is not possible to calculate the spectral density values at these frequencies without an assumption about the form of the spectral density function. This problem can be approached by the application of the so-called reduced spectral density mapping, in which the relaxation rates are directly translated into spectral density at three different frequencies. Reduced spectral density mapping, introduced by Farrow et al. (1995), Ishima and Nagayama (1995a,b), and Lefèvre et al. (1996), makes use of the finding that at high frequencies the spectral density function is quite flat, i.e., in the region of 400–800 MHz the J(ω) function varies very little. Under these conditions, the three spectral densities $J(\omega_H+\omega_N)$, $J(\omega_H)$, and $J(\omega_H-\omega_N)$ may be averaged to one average spectral density $<J(\omega_H)>$. Then, the three relaxation rates $R_N(N_z)$, $R_N(N_{xy})$ and $R_N(H_z \rightarrow N_z)$ are sufficient to determine the spectral density values at J(0), $J(\omega_N)$ and $<J(\omega_H)>$ according to:

$$\begin{bmatrix} J(0) \\ J(\omega_N) \\ <J(\omega_H)> \end{bmatrix} = \begin{bmatrix} \frac{-3}{4(3d^2+c^2)} & \frac{3}{2(3d^2+c^2)} & \frac{-9}{10(3d^2+c^2)} \\ \frac{1}{(3d^2+c^2)} & 0 & \frac{-7}{5(3d^2+c^2)} \\ 0 & 0 & \frac{1}{5d^2} \end{bmatrix} \times \begin{bmatrix} R_N(N_z) \\ R_N(N_{x,y}) \\ R_N(H_z \rightarrow N_z) \end{bmatrix} \quad (1)$$

in which $d^2 = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{h^2\gamma_N^2\gamma_H^2}{(4\pi)^2 r_{NH}^6}$

and $c^2 = \frac{1}{3}(\gamma_N B_0)^2(\Delta\sigma)^2$

where μ_0 is the permeability of vacuum, *h* is Planck's constant, γ_H ($2.6752 \cdot 10^8$ rad s^{-1} T^{-1}) and γ_N ($-2.711 \cdot 10^7$ rad s^{-1} T^{-1}) are the gyromagnetic ratios of the ^{1}H and ^{15}N nuclei, respectively, and ω_H and ω_N are the ^{1}H and ^{15}N Larmor frequency, respectively; r_{NH} is the internuclear ^{15}N-^{1}H distance (1.02 Å), B_0 is the magnetic field strength, and $\Delta\sigma$ is the difference between the parallel and perpendicular components of the axially symmetric ^{15}N chemical shift tensor, estimated to be -160 ppm (Hiyama et al., 1988). For a magnetic field of 14.1 Tesla, $d = 1.2986 \cdot 10^9$ rad^2 s^{-2} and $c = 1.2452 \cdot 10^9$ rad^2 s^{-2}.

The cross-relaxation rate $R_N(H_z \rightarrow N_z)$ between ^{15}N and its attached amide proton is correlated with NOE and is calculated using $NOE = 1 + (\gamma_H/\gamma_N) \cdot R_N(H_z \rightarrow N_z)/R_N(N_z)$. The frequency in the average spectral density, $<J(\omega_H)>$, may be taken equal to $0.87\omega_H$ (Farrow et al., 1995). The combination of the chosen rates has the advantage that it is independent of proton relaxation due to dipolar interaction with other protons. The three spectral densities J(0), $J(\omega_N)$ and $<J(\omega_H)>$ represent the proportion of the total energy used for motion at each corresponding frequency. A comparative analysis of the spectral densities along the sequence gives a straightforward idea of the different distributions of the frequencies of NH bond motions along the backbone. Lefèvre et al. (1996) analyzed several relaxation data obtained on various proteins and often found the existence of a linear correlation between J(0) and $J(\omega_N)$:

$$J(\omega_N) = \alpha_N J(0) + \beta_N \quad (2)$$

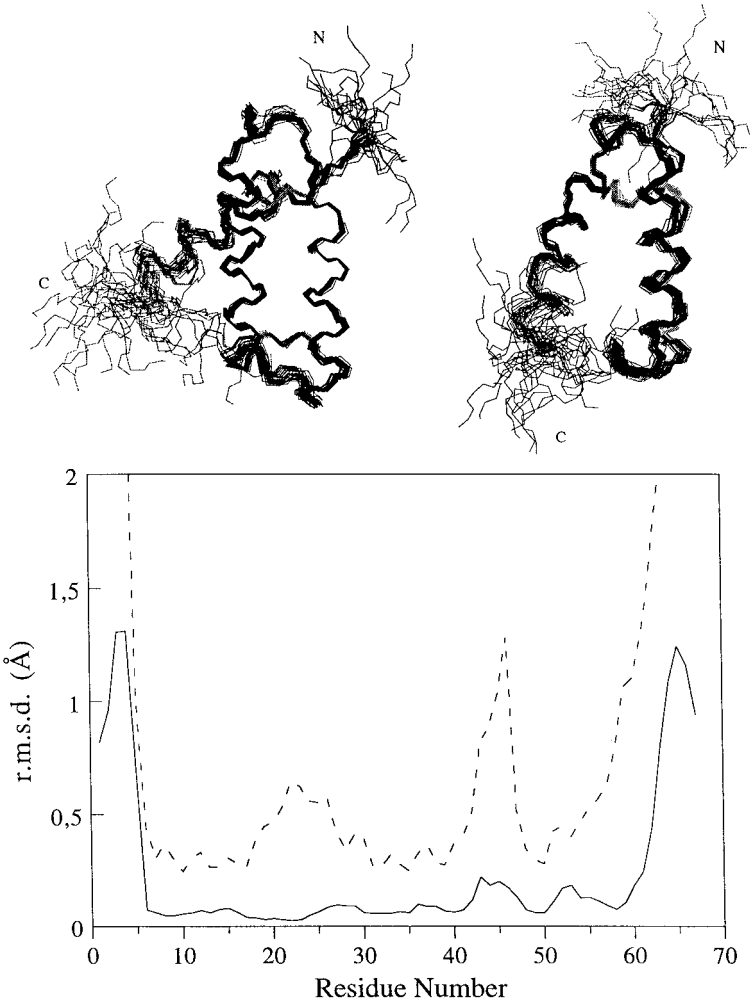

Figure 1. The refined structure of C12A-p8^{MTCP1}. (Top) Two views of the 30 best structures superimposed over the backbone heavy atoms N, Cα and C for residues 5–63. Only backbone atoms are shown, except for the three disulfide bridges (7–38, 17–28, and 39–50) (shaded grey). The two views are related by a ∼90° rotation about the vertical axis. The energy-minimized average structure calculated from the above family of 30 refined structures was used to check the quality of the structure in a Ramachandran plot through the program PROCHECK (Laskowski et al., 1993). Among 61 'meaningful' residues (i.e., non-glycine, non-proline, and non-terminal residues), 50 residues (i.e., 82%) fall in the most favored regions of the Ramachandran plot; 10 residues (i.e. 16.4%) are in the additional allowed regions; only 1 residue (1.6%), Ala 66, located in the disordered C-terminal end of the protein, falls in the generously allowed region and no residue falls in the disallowed region. (Bottom) Plots versus the amino acid sequence of the mean of the global rms differences calculated for the backbone atoms (see text) superimposed over the peptidic segment 5–63 (broken line), and the mean local rms deviation for the backbone superposition of all tripeptide segments along the sequence relative to the mean NMR structure (solid line). The rmsd values for the tripeptide segments are plotted at the position of the central residue.

A similar correlation was also observed between J(0) and J(ω_H). For a more profound interpretation of this observed correlation, a minimal number of assumptions need to be made about the shape of the spectral density function, and thus about the different motions of the NH bonds in proteins. The first usual assumption is that the global reorientation of the whole protein in the nanosecond range is not correlated with the internal motions. These motions can themselves consist of several uncorrelated motions (Brooks et al., 1987). The spectral density function is then decomposed as a weighted sum of independent contributions $J_k(\omega)$, in which the first component ($k = 0$) corresponds to the overall motion:

$$J(\omega) = a_0 J_0(\omega) + \sum\nolimits_k a_k J_k(\omega) \qquad (3)$$

a_k is the scaling factor of the spectral density function $J_k(\omega)$ characterizing each independent motion and $\Sigma_k a_k = 1$.

Results

NMR structural analysis and structure calculation

In a previous paper, almost all the assignments of proton resonances were reported (Barthe et al., 1997) for the wild-type protein p8^{MTCP1}. The relative positions of the NH/Hα cross peaks in the fingerprint region of C12A-p8^{MTCP1} remain roughly similar to those of the unmodified protein: the only significant differences in chemical shifts in the backbone are located either around the mutated position or at the N- and C-terminus, due to the deletion of the supernumerary N- and C-terminal residues present in our previous construction of p8^{MTCP1}. Thus, in the 3D [^{1}H-^{15}N] double-resonance experiments recorded on the ^{15}N-labeled C12A-p8^{MTCP1}, all the sequential assignments made previously on the wild-type construct were confirmed easily and extended in the ^{15}N planes of the 3D spectrum. The complete list of ^{1}H and ^{15}N chemical shifts is available as Supplementary material. In some cases, when considerable amide resonance overlap occurs, d_{NN} connectivities have been observed in the [^{1}H-^{15}N] HMQC-NOESY-HMQC spectrum. Careful inspection of the 3D [^{1}H-^{15}N] NOESY-HSQC spectrum yields 112 additional medium- and 28 long-range effects which were included in the structure calculations. In addition, most of the $^3J_{NH-H\alpha}$ coupling constants were measured on J modulated [^{1}H-^{15}N] COSY spectra (Supplementary material), and converted as angular restraints for the 3D structure calculation.

Using this improved restraint set, DYANA and AMBER calculations (see Materials and Methods) led to a virtually identical 3D structure as previously found for the wild-type protein p8^{MTCP1} using homonuclear NMR only (Figure 1). The main structural motif of C12A-p8^{MTCP1} consists of two antiparallel amphipathic helices spanning residues 8 to 20 (helix I) and 29 to 40 (helix II), strapped in an α-hairpin motif by the two left-handed disulfides 7–38 and 17–28. The angle between the axes of helices I and II in the α-hairpin is about 5°. The two antiparallel helices are linked by two interlocking turns, which adopt the same conformation as in the wild-type protein. The third helix, spanning residues 48 to 63, is connected to the double-helix motif by a relatively well defined loop (from residue Gln41 to Arg46). The third right-handed disulfide bridge 39–50 links the top of helix III to the tip of helix II. Helix III is contained in a plane approximately parallel to the plane defined by the antiparallel helical motif. Its axis forms an angle of 57° ± 10° with respect to the main axis of the antiparallel helical motif. As found for the wild type protein, the relative orientation of helix III with respect to the double-helix motif remains ill-defined, due to the lack of long-range NOEs observed between helix III and the α-hairpin. The N- (residues 1–4) and C-termini (residues 64–68) appear essentially disordered. The survey of the structural statistics and of the residual violations of experimental constraints for 30 conformers is shown in Table 1. Discarding residues 1–4 and 64–68, for which no long-range NOEs were available, the low average pairwise rmsd values calculated on

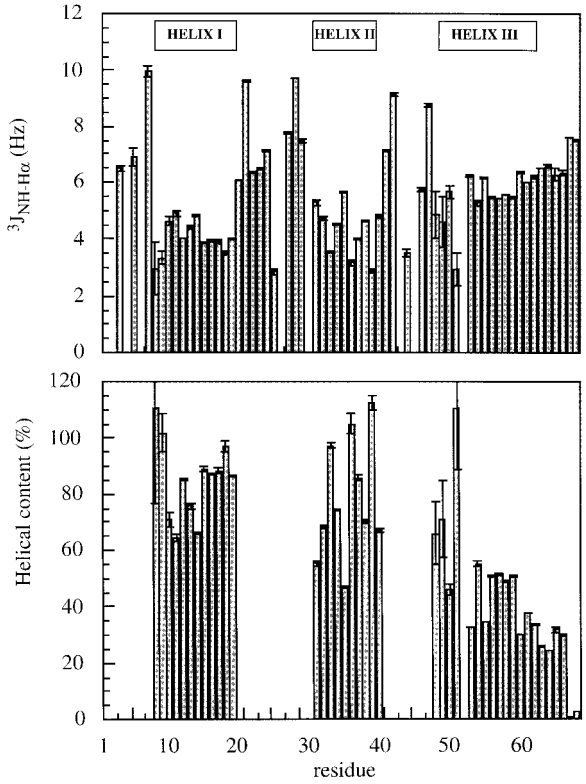

Figure 2. (Top) Bar graph of the $^3J_{NH-H\alpha}$ coupling constant values versus protein sequence. These values have been measured from J modulated [^{1}H-^{15}N] COSY spectra on a 0.4 mM sample of protein at 20 °C. Blank slots indicate proline residues (Pro2, Pro6 and Pro43) and residues for which $^3J_{NH-H\alpha}$ could not be accurately measured due to either overlap (Glu56/Leu 61 and Lys55/Ser67) or very short T_2. The locations of the three α-helices are indicated in the white boxes. (Bottom) Bar graph of the time averaged α-helical content (f_{hlx}) calculated from $^3J_{NH-H\alpha}$ for residues belonging to the three helices of C12A-p8^{MTCP1}, using Equation 4.

Table 1. Experimental constraints and refinement statistics of the 30 conformers representing the solution structure of C12A-p8^{MTCP1} before and after molecular mechanics energy refinement.

Distance constraints			
Intraresidue		212	
Sequential		340	
Medium-range		404	
Long-range		186	
Disulphide bonds		3	
Constraints per residue		30	
Dihedral constraints			
Phi (Φ)		35	
Chi1 (χ_1)		21	
Parameter		DYANA	DYANA + AMBER
Target function (Å^2)		3.54 ± 0.58	
Upper limit violations			
Number >0.2 Å		7.50 ± 3.17	4.80 ± 1.52
Sum of violations (Å)		11.68 ± 1.14	10.45 ± 0.68
Maximum violation (Å)		0.47 ± 0.13	0.27 ± 0.03
Dihedral angle violations			
Number >5°		0	1.63 ± 1.00
Sum of violations (°)		3.57 ± 2.53	16.64 ± 5.50
Maximum violation (°)		2.36 ± 1.71	6.24 ± 1.00
van der Waals violations			
Number >0.2 Å		3.77 ± 1.98	
Sum of violations (Å)		7.09 ± 1.06	
Maximum violation (Å)		0.33 ± 0.07	
AMBER energies (kcal.mol^{-1})			
Bond energy			23.9 ± 1.1
Valence angle energy			129.8 ± 5.8
van der Waals energy			-406.0 ± 15.0
Electrostatic energy			-1942.2 ± 26.6
Constraint energy			60.0 ± 3.9
Total non-bonding energy			-1489.2 ± 28.1
Total energy			-977.9 ± 25.2
rmsd values (Å)			
Residues			
1–68 (all)	BA[a]/HA[b]	3.38 ± 0.84 / 3.89 ± 0.74	3.22 ± 0.83 / 3.62 ± 0.74
5–63	BA[a]/HA[b]	1.03 ± 0.25 / 1.79 ± 0.20	0.82 ± 0.28 / 1.43 ± 0.24
8–40	BA[a]/HA[b]	0.54 ± 0.15 / 1.40 ± 0.18	0.26 ± 0.12 / 0.88 ± 0.19
48–63	BA[a]/HA[b]	1.01 ± 0.31 / 1.96 ± 0.31	0.73 ± 0.30 / 1.60 ± 0.32

[a] Backbone atoms.
[b] All heavy atoms.

the backbone atoms (0.82 ± 0.28 Å) indicate that all conformers converge to the same overall fold.

As indicated by a slightly higher local rmsd (Wagner et al., 1987) (Figure 1), helix III appears less well defined than helix I and II in the α-hairpin. This suggests that, in addition to the poor definition of its orientation with respect to the α-hairpin, helix III has more internal flexibility. This is supported by the values of the $^3J_{NH-H\alpha}$ coupling constants measured for each amide proton (Figure 2a): average values of 4.02 $\pm$ 0.2 Hz, 4.3 $\pm$ 0.3 Hz, and 5.5 $\pm$ 0.2 Hz are found in helix I, II and III, respectively. It is known that helical ϕ angles result in coupling constants of 3 to 5 Hz, while the random-coil conformation has coupling constants in the range of 7 to 8 Hz (Pardi et al., 1984; Wüthrich, 1986). The apparent coupling constants

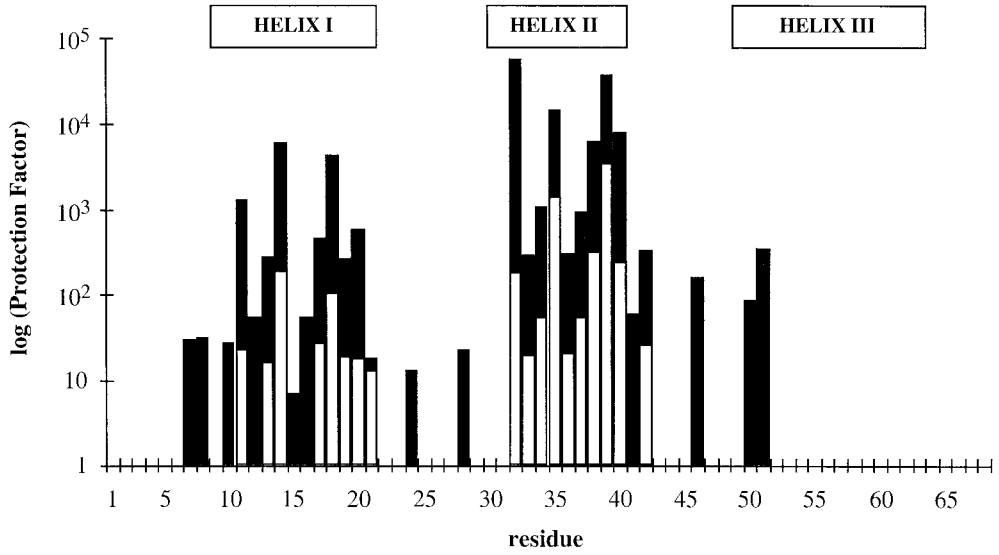

Figure 3. Bar graph of the protection factor estimated from the measured rates of backbone amide proton exchange with the solvent at pH 6.5 and at 20 °C (open bars) or 5 °C (filled bars). The locations of the three α-helices are indicated at the top.

measured on the J-modulated [^{1}H-^{15}N] COSY spectra are weighted population averages and depend on the distribution of angles over the populations (Kessler et al., 1988). It is thus possible to estimate the population of molecules containing helical ϕ angles from the $^3J_{NH-H\alpha}$ coupling constants, assuming a two-state model where a given residue is at equilibrium between an helical state and a random-coil state. Thus, one can determine a time averaged α-helical content (f_{hlx}) from the measured value of $^3J_{NH-H\alpha}$ using Equation 4 below (Bradley et al., 1990):

$$f_{hlx} = \frac{J_{meas} - J_{rc}}{J_{hlx} - J_{rc}} \quad (4)$$

where J_{meas} is the measured coupling constant, and J_{hlx} and J_{rc} are the ideal coupling constant for α-helix and random-coil, respectively. It is generally assumed that the ideal α-helix and random-coil have coupling constants of $J_{hlx} = 3.9$ Hz and $J_{rc} = 7.4$ Hz, respectively (Wüthrich, 1986; Bradley et al., 1990). Nevertheless, some of the coupling constants measured in the helices of C12A-p8^{MTCP1} have values lower than the ideal values previously established on model peptides (Wüthrich et al., 1986), leading to aberrant values for helical populations ($\gg$ 100%). We have tried to estimate a more accurate coupling constant for the protein, calculated as the mean value of the lowest coupling constant found in the three helices of C12A-p8^{MTCP1}, all the values retained for the calculation being $\leq$ 3.9 Hz (the 'theoretical' value of coupling constant for α-helix). These low coupling constants are measured for residues which always fall in the α-helical region of the Ramachandran plot (Ramachandran et al., 1971) and display a regular H-bond pattern in all the computed structures. This resulted in an average value of 3.4 Hz which was used to estimate the helical population. In Figure 2b, the helical population estimates are shown for each residue in the three helices of C12A-p8^{MTCP1}. From coupling constant data, helices I, II and III have time-averaged helical populations at 20 °C of 85%, 78% and 47%, respectively. Inspection of the individual helical population estimate for each residue in helix III shows a progressive decrease from the top (residue Val48, 65%) to the tip (residue Arg63, 22%) of the helix, the C-terminal residues being essentially in a random-coil conformation.

Increased flexibility in helix III is further supported upon analysing the protection factors calculated from the amide proton exchange rates. It is generally accepted that hydrogen exchange in proteins can be analysed assuming that the slowly exchanging amide proton exists at equilibrium between two states through unfolding and folding reactions: the native state, where the H-bond is formed, and an open state, where the H-bond is broken. In this two-step model, slowly exchanging hydrogens exchange only in the open state, after becoming exposed to the solvent by global or local fluctuations around the average native structure (Hvidt and Nielsen, 1966; Englander and Kallenbach, 1983). Thus, this parameter is sensi-

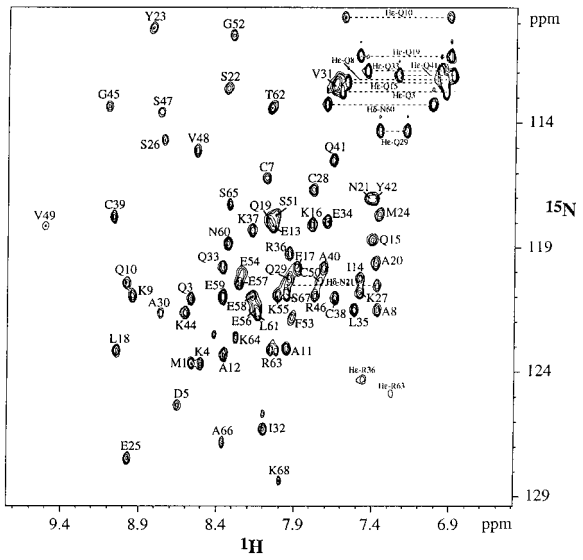

Figure 4. [^{1}H-^{15}N] heteronuclear single quantum coherence spectrum of the ^{15}N-labeled C12A-p8^{MTCP1}. Distinct ^{1}H-^{15}N correlation signals were observed for all backbone N-H pairs except for residues Glu56 and Leu61. Horizontal dashed lines connect the side-chain NH_2 frequencies of asparagine and glutamine residues.

tive to both the structure and the dynamics of proteins (Pedersen et al., 1991, 1993). At 20 °C, measurable exchange rates are primarily found for the hydrophobic residues deeply buried between the amphipathic helices I and II of C12A-p8^{MTCP1} (Figure 3). Upon lowering the temperature to 5 °C, in order to obtain information on exchange rates for more rapidly exchanging backbone protons, amide protons from the solvent-exposed surface of the α-hairpin also appear significantly protected, whereas no exchange rate can be measured for amide protons in helix III, except in the vicinity of the Cys39-Cys50 disulfide bond where a low value of $^{3}J_{NH-H\alpha}$ has been measured. This suggests a shifting of the folded–unfolded equilibrium toward the unfolded state in helix III. To characterize more thoroughly the increased flexibility of helix III, we further measured ^{15}N T_1, T_2 and ^{15}N{^{1}H}NOEs.

Determination of the experimental conditions for the relaxation study

Figure 4 shows the [^{1}H-^{15}N] HSQC spectrum recorded at 600 MHz on a 4 mM sample of protein at 20 °C. Nearly all the possible correlations involving backbone amide protons can be identified, with only one severe overlap for cross peaks corresponding to Glu56 and Leu61. Thus, a nearly complete relaxation data set has been obtained for C12A-p8^{MTCP1}, using

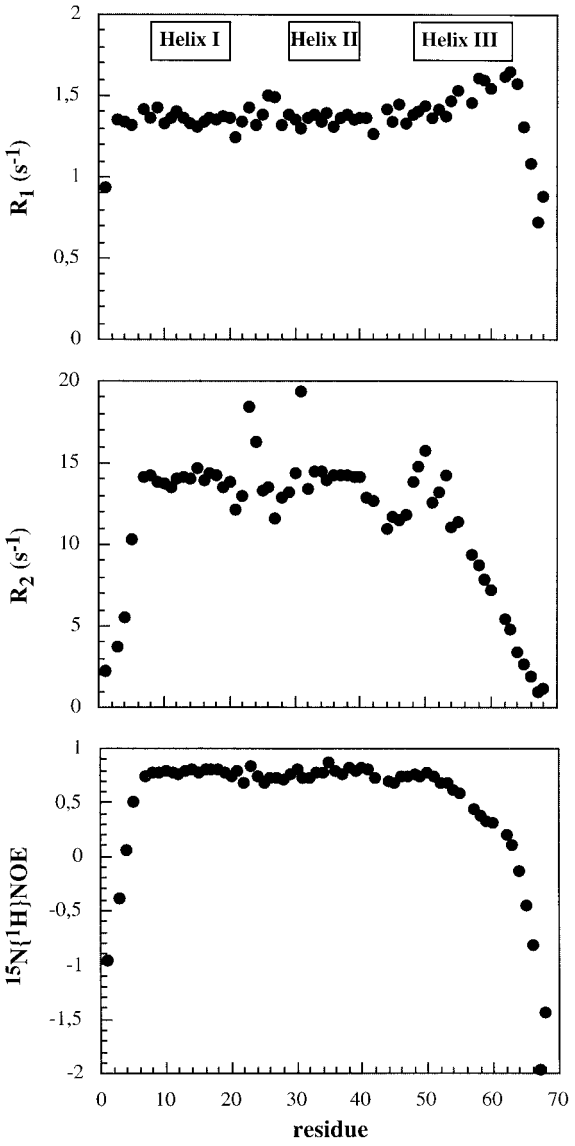

Figure 5. Relaxation rates and ^{15}N{^{1}H}NOEs as a function of the sequence obtained for C12A-p8^{MTCP1} (4 mM, pH 6.5, 20 °C). Relaxation rates were measured for all residues except for the three prolines (Pro2, Pro6 and Pro43) and for the two residues Glu56 and Leu61, due to superposed correlation peaks in the HSQC spectra. The locations of the three α-helices are indicated in white boxes at the top of the figure. The relaxation rate constants R_1 and R_2 were obtained from non linear fits of peak heights (Skelton et al., 1993) to monoexponential functions (Press et al., 1986). The uncertainties were determined from 500 Monte Carlo simulations: error bars, if shown, would be smaller than the size of the characters used to indicate the data points. ^{15}N{^{1}H}NOEs are given by ratios of peak heights in the experiments with and without proton saturation.

the regular heteronuclear experiments for R_1, R_2, and $^{15}N\{^1H\}$NOEs (see Materials and methods). The values of the relaxation parameters R_1, R_2, and NOE are shown in Figure 5. The high NOE values in helix I and II already suggest only a weak contribution of rapid internal flexibility to the dynamics of the α-hairpin, whereas, in helix III, the regular decrease in NOE values from the top to the tip of the helix suggests the existence of a 'gradient of flexibility' in this element of structure. A clear correlation exists between the transverse relaxation rate and the NOE data. However, this correlation fails with the longitudinal relaxation rate data in the third helix. Except in the first turns, presumably due to additional exchange contributions, the R_2 and NOE values decrease monotonically as a function of the residue number, whereas the R_1 values increase.

Using the assumption that the protein reorients isotropically in solution, estimates of the apparent overall tumbling time τ_c can be determined from the R_2/R_1 ratio for each residue when certain precautions are taken (Kay et al., 1989). In a first analysis, we used only R_1 and R_2 of residues belonging to helices I and II. As suggested by NOEs greater than 0.6, these residues belong to well-defined secondary structures and their relaxation is expected to be driven predominantly by the overall tumbling motion, while exchange processes should be negligible. This approach leads to a suprisingly high value for the correlation time of 9.76 ± 0.23 ns. Since the overall correlation time of a molecule should be approximately proportional to its molecular weight, a correlation time of about 6.5 ns is expected for C12A-p8^{MTCP1} at 20 °C. The ratio of the principal components of the average inertia tensor for the backbone atom on the average structure of C12A-p8^{MTCP1} was determined to be (1.9: 1.3: 1), suggesting that the overall rotation of C12A-p8^{MTCP1} is expected to have a small degree of anisotropy that might be detected by NMR relaxation measurements (Tjandra et al., 1995). But alone, the small effects due to small variations among the rates of reorientation around different molecular axes are unlikely to account for the higher than expected correlation time. The most likely explanation is that the sample contains an equilibrium of monomeric and oligomeric forms of the protein, and therefore, the correlation time of 9.76 ns is that of a population-weighted average.

Hydrodynamic measurements were thus performed in order to verify this assumption. Among other techniques, NMR self-diffusion coefficient (D_s) measurements report on the size of the molecule under conditions identical to those used for structure or dynamics determination. A self-diffusion coefficient of 92.1 ± 1.2 μm^2 s^{-1} was measured for C12A-p8^{MTCP1} at 4 mM, which appears more compatible with a dimeric state of the protein when considering a standard curve established under similar experimental conditions (see Materials and methods). Since no intermolecular NOEs were found, a fast exchange probably exists between monomeric and dimeric (or oligomeric) states. Although rapid partial monomer–dimer exchange can be taken into account in the analysis of relaxation rates (Fushman et al., 1997), we chose to shift this equilibrium toward the monomeric form, by varying the experimental conditions (pH, salt and protein concentration). Conditions were found upon dilution of the sample at a concentration below 500 μM in which the equilibrium is almost fully shifted toward the monomeric form ($D_s = 107.5 \pm 1.5$ μm^2 s^{-1}, Figure 6).

Comparison of the [^{1}H-^{15}N] HSQC spectrum recorded for the 4 mM and for a 400 μM sample (data not shown) indicates that the two spectra are very similar, most of the cross peaks exhibit a chemical shift perturbation upon dilution weaker than 0.05 ppm and 0.2 ppm – never exceeding 0.1 and 0.5 ppm – in the ^{1}H or ^{15}N dimensions, respectively. Likewise, the Hα resonances measured on 2D TOCSY spectra recorded on the two samples do not exhibit perturbation greater than 0.05 ppm. The locations of these perturbed residues in the protein structure are essentially random, suggesting a non-specific character for the protein association. This is further reinforced by the concentration dependence of the equilibrium.

Dynamics study of C12A-p8^{MTCP1}

Based on the results obtained from the diffusion measurements presented above, we decided to perform the dynamics study on a 400 μM sample of C12A-p8^{MTCP1}, at pH 6.5 and 20 °C. At this concentration, the protein is almost completely in the monomeric state, and the relaxation rates can be obtained within a reasonable measuring time. Figure 7 shows the two relaxation rate constants and the heteronuclear NOEs at stationary state that were measured on the diluted sample of C12A-p8^{MTCP1}, as well as the three corresponding reduced spectral densities J(0), J(ω_N) and $<J(\omega_H)>$ obtained from the reduced relaxation matrix. The relaxation parameters measured in the 400 μM sample exhibit a roughly similar dependence on the residue position in the backbone as the 4 mM data, but, due to a difference in the overall correla-

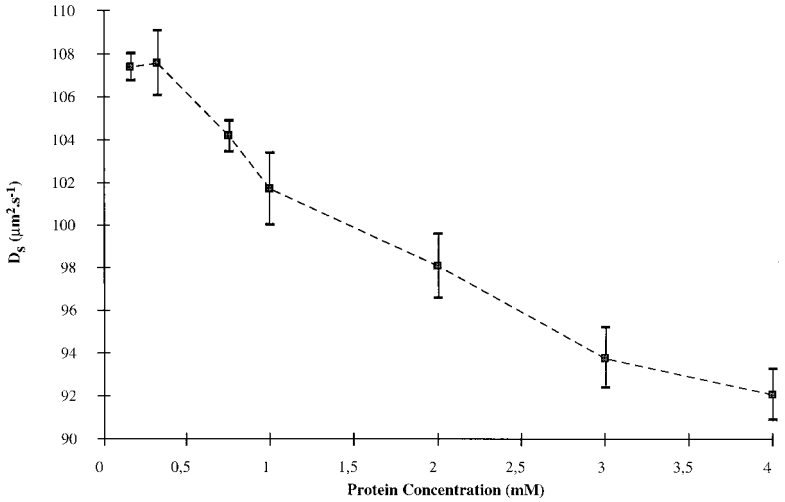

Figure 6. Evolution of the self-diffusion coefficient (D_S) value upon dilution of a sample of C12A-p8^{MTCP1} (pH 6.5, 20 °C).

tion time, an overall shift in the magnitude of R_1, R_2 and NOEs is observed, especially in helices I and II. Indeed, the overall tumbling time estimated from the trimmed mean R_2/R_1 ratio was 6.58 ± 0.13 ns for the diluted 400 μM protein sample, compared with 9.76 ± 0.23 ns for the concentrated 4 mM sample. Interestingly, the 'unusual' regular increase of R_1 as a function of the residue number in helix III is virtually absent in the diluted sample, all R_1 rates remaining around a plateau value nearly identical to that of helices I and II.

The spectral densities for the NH bonds clearly show a different behavior in the α-hairpin than in helix III. In helices I and II, mean values are 2.41 ns rad^{-1}, 0.327 ns rad^{-1}, and 5.4 ps rad^{-1} with standard deviations equal to 0.07 ns rad^{-1}, 0.009 ns rad^{-1} and 1.2 ps rad^{-1}, respectively, for J(0), J(ω_N) and <J(ω_H)>. Most of the residues in the two interlocking turns (residues 21–28) connecting helix I to helix II show lower J(0) values than in the rest of the α-hairpin, with a concomitant increase of the J(ω_N) and <J(ω_H)> values. This reflects the higher flexibility of these turns as compared to the helices in the α-hairpin. Tyr23 and, to a lesser extent, Met24 show significantly higher values of J(0), while J(ω_N) and <J(ω_H)> for these residues are not smaller than the mean values. This supports the hypothesis that slow movements in the micro- to millisecond range exist in this loop. In helix III, mean values for J(0), J(ω_N) and <J(ω_H)> are 1.94 ns rad^{-1}, 0.31 ns rad^{-1} and 12.2 ps rad^{-1}, with standard deviations equal to 0.66 ns rad^{-1}, 0.014 ns rad^{-1} and 5.8 ps rad^{-1}, respectively. The increase in the standard deviation is due simply to the fact that spectral density values on the third helix do not remain on a plateau but display a monotonic evolution from the N- to the C-terminal end of the helix. This indicates an increasing contribution of high-frequency motions. As suggested by lower values for J(0), the loop connecting the α-hairpin to helix III is expected to be more flexible than helices I and II, but appears more rigid than the C-terminal turns of helix III.

The search for correlation between J(0) and J(ω_N) allowed us to analyze the spectral density functions in terms of motion of the N-H vectors. Figure 8 shows experimental values of J(ω_N) plotted against the corresponding values of J(0) for each available NH bond. We observe that these values are linearly correlated. Note that the N-terminal residue (Met1) and C-terminal residues (Ala66 and Lys68), which exhibit strong negative $^{15}N\{^1H\}$NOE values, have been discarded for the fit since they would have biased the result. For reasons explained below, the residue Tyr23 has also been excluded for the calculation of the parameters. The combination of Equations 2 and 3 (see Materials and methods) leads to (Lefèvre et al., 1996):

$$J_k(\omega_N) = \alpha_N J_k(0) + \beta_N \quad (5)$$

The existence of a linear correlation indicates that the components $J_k(\omega)$ are the same for all residues, each component being related to specific motions that can be defined by a time scale. The dynamics of each NH bond is differentiated through the different weights α_N of each motion. These differences are brought to light through the position of the experimen-

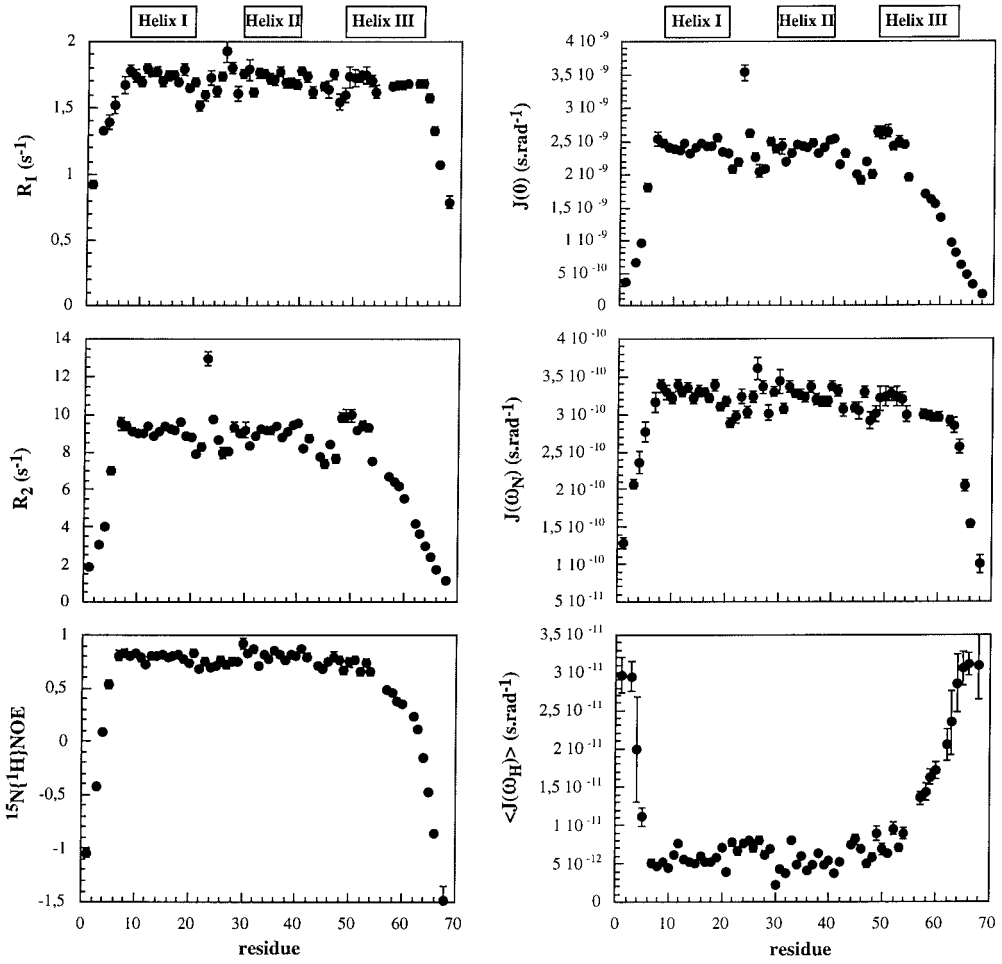

Figure 7. Relaxation rates and $^{15}N\{^1H\}$NOEs (left) as a function of the sequence obtained for C12A-p8^{MTCP1} (400 μM, pH 6.5, 20 °C), and corresponding values of the three spectral densities (right) calculated from Equation 1. Relaxation rates were measured for all the same residues as for the 4 mM protein sample, except for residues Lys55 and Ser67, due to superposed correlation peaks in the HSQC spectra of the diluted spectra. The locations of the three α-helices are indicated in white boxes at the top of the figure. The error bars on relaxation rates represent standard deviations in the estimated parameters obtained from 500 Monte Carlo simulations. Owing to the linearity of Equation 1, these uncertainties have been propagated to the spectral density values (Peng and Wagner, 1992).

tal points J(0),J(ω_N) along the correlation line. The points corresponding to N- and C-terminal residues, which were excluded from the fit, do not exhibit such correlation: this indicates that they undergo additional motions compared to the rest of the protein.

To get a better idea about the number of contributions that should be taken into account, and their corresponding time scale, a Lorentzian function can be tentatively used to describe each component of the spectral density in Equation 5, assuming that they are limited by an isotropic Brownian motion (Lefèvre et al., 1996; van Heijenoort et al., 1998). The corresponding theoretical curve corresponding to this model is shown in Figure 8. This leads to a third-degree equation in τ_c:

$$2\alpha_N\omega_N^2\tau_c^3 + 5\beta_N\omega_N^2\tau_c^2 + 2(\alpha_N - 1)\tau_c + 5\beta_N = 0 \quad (6)$$

Note that the NH bond of residue Tyr23, which was discarded from the fit, lies off the theoretical curve, indicating the limitation of this model for describing the motions for this residue. Tyr23 displays anomalously high J(0) values, indicating an additional contribution from an exchange mechanism in the micro- to millisecond range. As we discuss further, this leads to the contribution of low-frequency motions that are nonetheless of higher frequency than that of the overall tumbling ($a_0 > 1$ (vide infra)), which has no

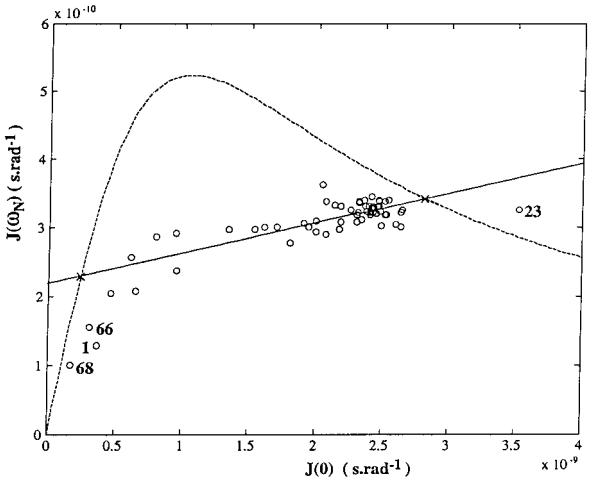

Figure 8. Plot of $J(\omega_N)$ as a function of J(0). NH bonds are represented as open circles. The fit (dark line) was obtained by linear regression. The residues that have been excluded for the fit are labelled (see text). The dashed curve was calculated for J(0) and $J(\omega_N)$ as a function of τ_c, using a single Lorentzian.

physical meaning within the context of the theoretical model. The roots of Equation 6 are correlation times connected to the various motions contributing to the low-frequency part of spectral density functions. The coefficients α_N and β_N being determined from the linear fit, this equation can be solved straightforwardly. One of the roots is negative and will not be considered since it cannot have any physical meaning. The two other positive roots give the values of the overall tumbling time (τ_c = 7.05 ± 0.04 ns) and of a global internal correlation time called τ_i (Dayie et al., 1996; Lefèvre et al., 1996) (τ_i = 0.610 ± 0.014 ns). Note that the calculated internal correlation time should not be considered as a real physical parameter since the overall spectral density function for the internal motion may itself consist of several contributions. This value should only be considered as indicative of the presence of complex internal movements at an intermediate time scale around τ_i.

From Equation 3, where the contributions of the overall tumbling and internal motions are explicitly separated, a system of three equations is obtained for the three spectral densities that allows the evaluation of the contribution of global and internal motions:

$$\begin{aligned} J(0) &= a_0 J_0(0) + \sum\nolimits_k a_k J_k(0) \\ J(\omega_N) &= a_0 J_0(\omega_N) + \sum\nolimits_k a_k J_k(\omega_N) \\ < J(\omega_H) > &= a_0 < J_0(\omega_H) > + \sum\nolimits_k a_k < J_k(\omega_H) > \end{aligned} \quad (7)$$

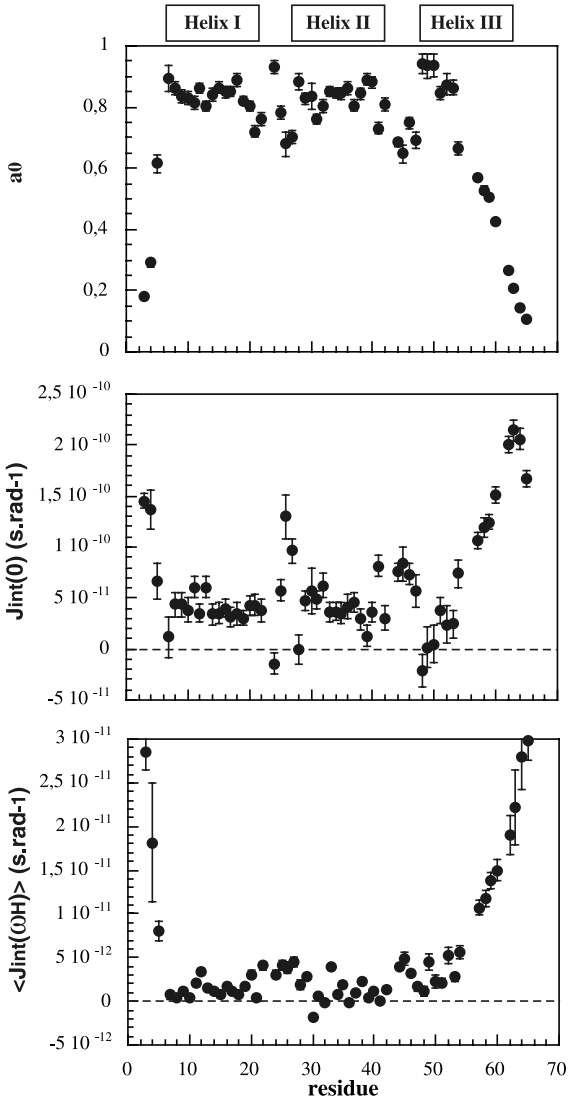

Figure 9. Plots of the motion parameters obtained from Equation 7 versus the protein sequence. From top to bottom: a_0, $J_{int}(0)$ in s/rad, and $<J_{int}(\omega_H)>$ in s/rad. An overall tumbling correlation time of 7.05 ns was used for the calculation.

As the second term of the spectral density function is supposed to concern internal fast motions, it can be considered to be very flat from zero to ω_N. It is thus reasonable to equalize $\Sigma_k\ a_k J_k(0)$ and $\Sigma_k\ a_k J_k(\omega_N)$. The three unknown values a_0 – corresponding to the weight of the overall tumbling spectral density function – $\Sigma_k\ a_k J_k(0)$ and $\Sigma_k\ a_k\ <J_k(\omega_H)>$ – corresponding to the contributions of the spectral densities of internal motions at the two related frequencies – can then be evaluated for each NH vector of the backbone from the three experimental densities. In

the following, $\Sigma_k\ a_k J_k(0)$ and $\Sigma_k\ a_k\ <J_k(\omega_H)>$, both related to internal motions, will be denoted as $J_{int}(0)$ and $<J_{int}(\omega_H)>$, for the sake of clarity. A Lorentzian form was used for $J_0(0)$ with τ_c equal to 7.05 ns. The results are shown in Figure 9, excepting residues which were previously discarded from the linear fit. In helices I and II, mean values are 0.84 $\pm$ 0.03, 40 $\pm$ 12 ps rad^{-1}, and 1.3 $\pm$ 1 ps rad^{-1}, respectively, for a_0, $J_{int}(0)$ and $<J_{int}(\omega_H)>$. In helix III, mean values given in the same order are 0.66 $\pm$ 0.26, 81 $\pm$ 21 ps rad^{-1} and 8.9 $\pm$ 1.9 ps rad^{-1}. Upon examining more carefully the α-hairpin, it appears that residues located in the hydrophobic core – buried between the two amphipathic helices – display higher a_0 values than the mean, whereas the $<J_{int}(\omega_H)>$ and, to a lesser extent, $J_{int}(0)$ values are lower. This indicates that the motion of these hydrophobic residues is dominated by the overall tumbling of the molecule, with reduced participation of internal motions. In the turns joining the two helices, a concomitant a_0 decrease with a $J_{int}(0)$ and $<J_{int}(\omega_H)>$ increase is generally observed, indicating increasing contributions of internal motions in this relatively flexible segment. A similar behavior is observed in the loop joining helix II to helix III. Nevertheless, residues Met24, Cys28, and, to a lesser extent, Ala30, have an opposite behavior: high values of a_0 give rise to slightly negative unphysical values of $J_{int}(0)$ or of $<J_{int}(\omega_H)>$. Note that Tyr23 was discarded for the calculations (vide supra), otherwise it would have an a_0 value greater than 1. A similar phenomenon is observed in the N-terminal turn of helix III for residues Val48,Val49, and Cys50. As stated before, this supports the hypothesis that some slow processes in the microsecond to millisecond range exist in these peptidic segments. As expected from the experimental spectral densities, the decrease of the values of a_0 in helix III concurs with an increase of the values of $J_{int}(0)$ and $<J_{int}(\omega_H)>$, consistent with an increasing contribution of high-frequency motions from residue Glu54 to residue Arg63. Note that this anticorrelation is not observed for the $J_{int}(0)$ values of residues Lys64 and Ser65. This is a strong indication that these residues (as well as residues Met1, Ala66 and Lys68 which were excluded from the calculations (vide supra)), which exhibit negative $^{15}N\{^1H\}$NOE values, have a different dynamic behavior than the rest of the protein, with probable contributions of additional motions. On the other hand, at the N-terminal end of helix III, residues Ser51, Gly52, and Phe53 have both low- and high-frequency contribution values similar to those observed in helices I and II.

Discussion

If the replacement of Cys12 by an alanine residue allows us to obtain reproducibly high yields of ^{15}N-labeled protein (> 20 mg/l), it does not bring any relevant change in the 3D structure of p8^{MTCP1}: the differences observed between the two proteins are of the same order as those observed between two conformations in a given population calculated for one protein. Despite a substantial increase of the restraint data set, the relative position of helix III with respect to the α-hairpin motif remains less well defined. In addition, higher local rmsd values suggest that this third helix is less well defined than helices I and II. Indeed, the $^3J_{NH-H\alpha}$ values, as well as the results of amide proton exchange rate measurements, indicate an equilibrium between folded and unfolded populations in helix III, the ratio of the unfolded population increasing gradually from the N- to the C-terminal end of the helix. In addition, owing to ^{15}N relaxation measurements, sub-nanosecond motions have been observed for the NH vectors for residues in helix III.

As a first result, the relaxation study indicates the presence of an equilibrium between monomeric and oligomeric states at high concentration (4 mM) of C12A-p8^{MTCP1}. Diffusion experiments were used in order to define experimental conditions where the equilibrium is shifted to the monomeric state. Best results were obtained on the dilute sample: the trimmed R_2/R_1 ratio yields a value of 6.58 $\pm$ 0.13 ns (9.76 $\pm$ 0.23 ns at 4 mM) for the overall tumbling time, compatible with the monomer size, at a protein concentration of 400 μM. At this step, the existence of a significant contribution to the motion at an intermediate frequency in helix III is already suggested by the comparison of the relaxation parameters obtained from the concentrated 4 mM and the diluted 400 μM protein sample. In the two samples, a clear correlation exists between the transverse relaxation rate and the NOE data. This correlation still exists with the R_1 data, except in the third helix: R_2 and NOE values decrease monotonically as a function of the residue number, whereas the R_1 values increase. This increase in R_1 values is strong in the concentrated sample, slight to virtually absent in the dilute sample. This is a strong indication of the existence of slow local motions in this helix. Indeed, Fushman et al. (1997) have demonstrated that such local motions lead necessarily to a decrease in both the transverse relaxation rate and the NOE compared to the case when they are absent, whereas they affect the rate of longitudinal relaxation

in a more complex way. Under certain conditions, an increase in R_1 is expected. This 'unusual' behavior of R_1 depends on the overall tumbling time, and is usually not observed for small proteins having τ_c on the order of a few nanoseconds. Thus, in the concentrated sample of C12A-p8^{MTCP1} ($\tau_c \approx 10$ ns), an increase in the R_1 values is seen in the third helix, whereas in the diluted sample ($\tau_c \approx 7$ ns) this 'unusual' behavior is considerably attenuated, since the effect of the local motions on R_1 becomes almost negligible for this value of τ_c.

The relaxation parameters R_1, R_2 and $^{15}N\{^1H\}$-NOE were then analyzed using the linear correlation observed between spectral densities (i.e. J(0) and J(ω)) obtained from the reduced matrix approach. The residues excluded from the linear correlation analysis are either residues exhibiting a very large contribution from exchange processes (Tyr23), leading to anomalously high values of J(0), or residues exhibiting significant negative values of NOEs (Met1, Ala66, and Lys68). These latter residues are located in the highly flexible N- and C-terminal ends of the protein, and the negative values found for NOEs suggest that they do not participate in the protein fold (Fushman et al., 1997). The global correlation time extracted from the linear correlation between J(0) and J(ω) was 7.05 $\pm$ 0.04 ns, and the existence of this linear correlation implies that all residues are subject to internal motions having significant low-frequency contributions at an intermediate time scale around τ_i = 0.610 $\pm$ 0.014 ns. Significant contributions of these slow internal motions are found essentially in helix III.

Motions on the microsecond to millisecond time scale are detected in loops joining the different helices where anomalously high values of J(0) (essentially centred on residue Tyr23 and, to a lesser extent, Phe53) are indicative of exchange processes. For the present relaxation data set obtained at one magnetic field strength, only qualitative information can be derived from this phenomenon. These fluctuations in the slow frequency regime are known to occur essentially either around aromatic or cysteine residues: small displacements of the aromatic rings or slow isomerization of the cysteine side chains engaged in disulfide bonds (Szyperski et al., 1993) could give rise to large chemical shift variations, leading to an efficient adiabatic relaxation pathway. This relaxation pathway contributes to transverse relaxation (the $R2_{ex}$ term) and leads to an increase in the spectral density J(0). In this case, the assumption that $\Sigma_k\ a_k J_k(0)$ and $\Sigma_k\ a_k J_k(\omega_N)$ are equal, in the linear correlation approach, is not warranted and yields erroneous values for $J_{int}(0)$. Such a situation is observed around residues Cys28 and Cys50, where exchange processes are indicated by negative values of $J_{int}(0)$ or $J_{int}(\omega_H)$.

Concluding remarks

In conclusion, it appears that the dynamic parameters obtained for C12A-p8^{MTCP1} are in good agreement with the protein fold. The NH vectors located in the well-structured α-hairpin exhibit restraint motions dominated by the overall tumbling of the protein. Significant exchange processes take place in the two interlocking turns joining helix I to helix II, as well as in the loop joining helix II to helix III, which thus appear more flexible. The dynamics of the NH vectors in helix III is more complex, with increasing additional contribution of sub-nanosecond motions from the N- to the C-terminal end of the helix. This helix shows a gradient in both structural definition and dynamic characteristics. Nevertheless, the relationship between the folded–unfolded equilibrium experienced by the NH vectors in this helix, as assessed by the analysis of the $^3J_{NH-H\alpha}$ coupling constant values or the exchange measurement, and the existence of an additional slow internal motion on the sub-nanosecond time-scale is unwarranted. Indeed, slower motions are expected from such equilibria (> 1 ms). Long molecular dynamics simulations may be used to gain a better atomic description of the movements involved in the dynamics of this third helix.

Acknowledgements

This work was supported by grants from the Ligue Nationale contre le Cancer and from the Association pour la Recherche sur le Cancer. The authors are indebted to Carine van Heijenoort for her help in implementing the relaxation microprograms on our spectrometer. The authors thank Marie-Paule Strub for her help in protein expression, and Cathy Royer for critical reading of the manuscript.

References

Abragam, A. (1961) *Principles of Nuclear Magnetism*, Oxford Science Publication, Clarendon Press, Oxford.

Bai, Y., Milne, J.S., Mayne, L. and Englander, S.W. (1993) *Proteins Struct. Funct. Genet.*, **17**, 75–86.

Barthe, P., Yang, Y.-S., Chiche, L., Hoh, F., Strub, M.P., Guignard, L., Soulier, J., Stern, M.H., van Tilbeurgh, H., Lhoste, J.M. and Roumestand, C. (1997) *J. Mol. Biol.*, **274**, 801–815.

Bax, A., Ikura, M., Kay, L.E., Torchia, D.A. and Tschudin, R. (1990) *J. Magn. Reson.*, **86**, 304–318.

Bax, A., Ikura, M., Kay, L.E. and Zhu, G. (1991) *J. Magn. Reson.*, **91**, 174–178.

Bax, A. and Pochapsky, S.S. (1992) *J. Magn. Reson.*, **99**, 638–643.

Billeter, M., Neri, D., Otting, G., Qian, Y.-Q. and Wüthrich, K. (1992) *J. Biomol. NMR*, **2**, 257–274.

Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–189.

Bradley, E.K., Thomason, J.F., Cohen, F.E., Kosen, P.A. and Kuntz, I.D. (1990) *J. Mol. Biol.*, **215**, 607–622.

Braunschweiler, L. and Ernst, R.R. (1983) *J. Magn. Reson.*, **53**, 521–528.

Brooks III, C.L., Karplus, M. and Pettitt, B.M. (1987) *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*, John Wiley and Sons, New York, NY, pp. 1–259.

Carr, H.Y. and Purcell, E.M. (1954) *Phys. Rev.*, **94**, 630–632.

Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M. Jr, Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. and Kollman, P.A. (1995) *J. Am. Chem. Soc.*, **117**, 5179–5197.

Dallapiccola, B., Alimena, G., Chessa, L., Gastaldi, R., De Rossi, G., Semenzato, G., Quinti, I. and Pandolfi, F. (1984) *Int. J. Cancer*, **34**, 171–176.

Davis, D.G. and Bax, A. (1985) *J. Am. Chem. Soc.*, **107**, 2820–2821.

Dayie, K.T., Wagner, G. and Lefèvre, J.F. (1996) *Annu. Rev. Phys. Chem.*, **47**, 243–282.

Delsuc, M.A. and Malliavin, T. (1998) *Anal. Chem.*, **70**, 2146–2148.

Englander, S.W. and Kallenbach, N.R. (1983) *Quart. Rev. Biophys.*, **16**, 521–655.

Farrow, N.A., Zhang, O., Szabo, A., Torchia, D.A. and Kay, L.E. (1995) *J. Biomol. NMR*, **6**, 153–162.

Fisch, P., Forster, A., Sherrington, P.D., Dyer, M.J. and Rabbitts, T.H. (1993) *Oncogene*, **8**, 3271–3276.

Frenkiel, T., Bauer, C., Carr, M.D., Birdsall, B. and Feeney, J. (1990) *J. Magn. Reson.*, **90**, 420–425.

Fu, T.-B., Virgilio, L., Narducci, M.G., Facchiano, A., Russo, G. and Croce C.M. (1994) *Cancer Res.*, **54**, 6297–6301.

Fushman, D., Cahill, S. and Cowburn, D. (1997) *J. Mol. Biol.*, **266**, 173–194.

Gibbs, S.J. and Johnson, C.S. (1991) *J. Magn. Reson.*, **93**, 395–402.

Goyns, M.H., Hammond, D.W., Harrison, C.J., Menasce, L.P., Ross, F.M. and Hancock, B.W. (1993) *Leukemia*, **7**, 848–852.

Grzesiek, S. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 12593–12594.

Güntert, P. and Wüthrich, K. (1991) *J. Biomol. NMR*, **1**, 447–456.

Güntert, P., Mumenthaler, C. and Wüthrich, K. (1997) *J. Mol. Biol.*, **273**, 283–298.

Hiyama, Y., Niu, C.H., Silverton, J.V., Bavoso, A. and Torchia, D.A. (1988) *J. Am. Chem. Soc.*, **110**, 2378–2383.

Hvidt, A. and Nielsen, S.O. (1966) *Adv. Protein Chem.*, **21**, 287–386.

Hyberts, S.G., Märki, W. and Wagner, G. (1987) *Eur. J. Biochem.*, **164**, 625–635.

Ikura, M., Bax, A., Clore, M. and Gronenborn, A. (1990) *J. Am. Chem. Soc.*, **112**, 9020–9022.

Ishima, R. and Nagayama, K. (1995a) *Biochemistry*, **34**, 3162–3171.

Ishima, R. and Nagayama, K. (1995b) *J. Magn. Reson.*, **B108**, 73–76.

Jeener, J., Meier, B.H., Bachman, P. and Ernst, R.R. (1979) *J. Chem. Phys.*, **71**, 4546–4553.

Kadkhodaei, M., Hwang, T.-L., Tang, J. and Shaka, A.J. (1993) *J. Magn. Reson.*, **A105**, 104–107.

Karplus, M. (1963) *J. Am. Chem. Soc.*, **85**, 2870–2871.

Kay, L.E., Torchia, D.A. and Bax, A. (1989) *Biochemistry*, **28**, 8972–8979.

Kay, L.E., Nicholson, L.K., Delaglio, F., Bax, A. and Torchia, D.A. (1992) *J. Magn. Reson.*, **97**, 359–375.

Kessler, H., Griesinger, C., Lautz, J., Muller, A., van Gunsteren, W.F. and Berendsen, H.J.C. (1988) *J. Am. Chem. Soc.*, **110**, 3393–3396.

Kumar, A., Ernst, R.R. and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun.*, **95**, 1–6.

Lefèvre, J.-F., Dayie, K.T., Peng, J.W. and Wagner, G. (1996) *Biochemistry*, **35**, 2674–2686.

Madani, A., Soulier, J., Schmid, M., Plichtova, R., Lermé, F., Gateau-Roesch, O., Garnier, J.P., Pla, M., Sigaux, F. and Stern, M.-H. (1995) *Oncogene*, **10**, 2259–2262.

Madani, A., Choukroun, V., Soulier, J., Cacheux, V., Claisse, J.F., Valensi, F., Daliphard, S., Cazin, B., Levy, V., Leblond, V., Daniel, M.T., Sigaux, F. and Stern, M.H. (1996) *Blood*, **87**, 1923–1927.

Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989a) *Biochemistry*, **28**, 6150–6156.

Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989b) *J. Magn. Reson.*, **85**, 393–399.

Markus, M.A., Dayie, K.T., Matsudaira, P. and Wagner, G. (1996) *Biochemistry*, **35**, 1722–1732.

Meiboom, S. and Gill, D. (1958) *Rev. Sci. Instrum.*, **29**, 688–691.

Neri, D., Otting, G. and Wüthrich, K. (1990) *J. Am. Chem. Soc.*, **112**, 3663–3665.

Pardi, A., Billeter, M. and Wüthrich, K. (1984) *J. Mol. Biol.*, **180**, 741–751.

Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham III, T.E., Fergusson, D.M., Seibel, G.L., Chandra Singh, U., Weiner, P.K. and Kollman, P.A. (1995) AMBER 4.1, University of California, San Francisco, CA.

Pedersen, T.G., Sigurskjold, B.W., Andersen, K.V., Kjær, M., Poulsen, F.M., Dobson, C.M. and Redfield, C. (1991) *J. Mol. Biol.*, **218**, 413–426.

Pedersen, T.G., Thomsen, N.K., Andersen, K.V., Madsen, J.C. and Poulsen, F.M. (1993) *J. Mol. Biol.*, **230**, 651–660.

Peng, J.W. and Wagner, G. (1992a) *Biochemistry*, **31**, 8571–8586.

Peng, J.W. and Wagner, G. (1992b) *J. Magn. Reson.*, **98**, 308–332.

Piotto, M., Saudek, V. and Sklenar, V. (1992) *J. Biomol. NMR*, **2**, 661–665.

Pons, J.L., Malliavin, T.E. and Delsuc, M.A. (1996) *J. Biomol. NMR*, **8**, 445–452.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1986) *Numerical Recipes*, Cambridge University Press, Cambridge.

Ramachandran, G.N., Chandrasekharan, R. and Kopple, K.D. (1971) *Biopolymers*, **10**, 2113–2131.

Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, **117**, 479–495.

Rance, M. (1987) *J. Magn. Reson.*, **74**, 557–564.

Santoro, J., Bruix, M. and Rico, M., Thirteenth European Experimental NMR Conference, May 19–24, 1996, Paris.

Shaka, A.J., Keeler, J. and Freeman, R. (1983) *J. Magn. Reson.*, **53**, 313–340.

Skelton, N.J., Palmer, A.G., III, Akke, M., Kördel, J., Rance, M. and Chazin, W.J. (1993) *J. Magn. Reson.*, **B102**, 253–264.

Sklenar, V. (1995) *J. Magn. Reson.*, **A114**, 132–135.

Soulier, J., Madani, A., Cacheux, V., Rosenzwajg, M., Sigaux, F. and Stern, M.-H. (1994) *Oncogene*, **9**, 3565–3570.

Stern, M.H., Soulier, J., Rosenzwajg, M., Nakahara, K., Canki-Klain, N., Aurias, A., Sigaux, F. and Kirsch, I.R. (1993) *Oncogene*, **8**, 2475–2483.

Szyperski, T., Luginbühl, P., Otting, G., Güntert, P. and Wüthrich, K. (1993) *J. Biomol. NMR*, **3**, 151–164.

Tjandra, N., Feller, S.E., Pastor, R.W. and Bax, A. (1995) *J. Am. Chem. Soc.*, **117**, 12562–12566.

van Heijenoort, C., Penin, F. and Guittet, E. (1998) *Biochemistry*, **37**, 5060–5073.

Wagner, G., Braun, W., Havel, T.F., Schaumann, T., Go, N. and Wüthrich, K. (1987) *J. Mol. Biol.*, **196**, 611–639.

Wider, G., Dötsch, V. and Wüthrich, K. (1994) *J. Magn. Reson.*, **A108**, 255–258.

Witzig, T.E., Phyliky, R.L., Li, C.-Y., Homburger, H.A., Dewald, G.W. and Handwerger, B.S. (1986) *Am. J. Hematol.*, **21**, 139–155.

Wu, D., Chen, A. and Johnson Jr., C.S. (1995) *J. Magn. Reson.*, **A115**, 260–264.

Wüthrich, K., Billeter, M. and Braun, W. (1983) *J. Mol. Biol.*, **169**, 949–961.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York, NY.