Backbone Dynamics for the Wild Type and a Double H52R/T56W Mutant of the vnd/NK-2 Homeodomain from *Drosophila melanogaster*

Simone Fausti,[‡,§] Solly Weiler,[||,⊥] Carla Cuniberti,[§] Kae-Jung Hwang,[@] Kyoung Tai No,[@] James M. Gruschus,[||] Angelo Perico,[‡] Marshall Nirenberg,[#] and James A. Ferretti*[,||]

Istituto di Studi Chimico-Fisici di Macromolecole Sintetiche e Naturali (IMAG), National Research Council, Via De Marini, 6, 16149 Genova, Italy, Dipartimento di Chimica e Chimica Industriale, Universita' di Genova, Via Dodecaneso, 31, 16146 Genova, Italy, Computer Aided Molecular Design Research Center, Soong Sil University, 1-1 Sangdo-5-Dong, Dongjak-Ku, Seoul 156-743, Republic of Korea, Laboratory of Biochemical Genetics, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892-4036, and Laboratory of Biophysical Chemistry, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892-0380

Received February 26, 2001; Revised Manuscript Received June 12, 2001

ABSTRACT: The ^{15}N relaxation behavior and heteronuclear Overhauser effect data for the wild type and an H52R/T56W double mutant protein that encompasses the vnd/NK-2 homeodomain from *Drosophila melanogaster* were used to characterize and describe the protein backbone dynamics. This investigation, which includes a description of a model structure for the H52R/T56W double mutant vnd/NK-2 homeodomain, was carried out for the two proteins in both the free and DNA-bound states. The double residue replacement at positions 52 and 56 within the DNA recognition helix of vnd/NK-2 has been shown to lead to a significant secondary structural modification resulting in an increase in the length of the recognition helix for the unbound protein. These structural changes are accompanied by corresponding changes in the T_1 and $T_{1\rho}$ relaxation times as well as in the heteronuclear Overhauser effect (XNOE) values that show that the structural stability of the protein is enhanced by the two residue replacements. The values of the rotational anisotropy, $D_{||}/D_{\perp}$, derived from analysis of the ^{15}N T_1 and $T_{1\rho}$ relaxation values are small (1.189 for the unbound homeodomain and 1.110 for the bound homeodomain; both analyzed as prolate ellipsoids of revolution). A comparison of the T_2 values of the wild type and double mutant homeodomain reveals the presence of a low-frequency exchange contribution for the wild type analogue. These relaxation studies show that the motional behavior of the protein primarily reflects the tertiary structure and stability of the homeodomain backbone as well as the respective changes induced upon site-directed residue replacement or DNA binding.

The homeodomain is the highly conserved DNA-binding domain of a class of proteins that regulate transcription in the commitment of embryonic cells to specific developmental pathways (*1*−*3*). Homeobox-containing genes that encode the homeodomain are found in all eukaryotic species from plants and yeast through humans. The homeobox gene often is expressed at very early stages in embryonic development and is responsible for the regulation of other downstream target genes. The ability of the homeodomain to function properly is dependent upon its capacity to adopt the correct three-dimensional structure and to properly bind a specific sequence or set of sequences of DNA (*4*, *5*). Members of the homeodomain family show a high degree of homology in their tertiary structures, possessing three helical segments, including a helix−turn−helix DNA-binding motif (*6*−*15*). The N-terminal arm and helix III contact the DNA and make specific interactions with the bases and phosphate backbone.

Conservation of residues in a given homeodomain appears to be critical for proper function. Single amino acid residue replacements, i.e., mutations, in the homeodomain result in or cause altered embryonic development or genetic disease (*16*, *17*). Such single residue replacements can result in structural and dynamic changes in the homeodomain, a modification of the physical properties such as the "melting" temperature, or an alteration in the affinity and target specificity for the DNA (*4*, *18*, *19*). For example, in the case of the homeodomain eve, a single R52H mutation has been shown to result in temperature-dependent developmental abnormalities (*20*). To help in understanding functional studies on transcriptional control in development, it is useful to characterize the structural and dynamic behavior of both the conserved and variable amino acid residues in the homeodomain in both the free and DNA-bound states. The role of the dynamics of the homeodomain in sequence specific DNA binding, while not well-understood, may be critical, since in numerous instances the functional homeodomain may either be unfolded or show poorly defined

* To whom all communication should be directed. Fax: (301) 402-3405. E-mail: jafer@helix.nih.gov.

[‡] National Research Council.

[§] Universita' di Genova.

[||] Laboratory of Biophysical Chemistry, National Heart, Lung and Blood Institute, National Institutes of Health.

[⊥] Present Address: Howard Hughes Medical Institute, Dana-Farber Cancer Institute, 1 Jimmy Fund Way, SM 758, Boston, MA 01772.

[@] Soong Sil University.

[#] Laboratory of Biochemical Genetics, National Heart, Lung and Blood Institute, National Institutes of Health.

10.1021/bi010398r CCC: $20.00
Published on Web 09/12/2001

secondary structure in the absence of DNA.

The homeodomain of particular interest in this study is encoded by the *vnd/NK-2* gene from *Drosophila melanogaster*. The *vnd/NK-2* gene is responsible for development of part of the central nervous system of *Drosophila* (*21*). The three-dimensional structures in the free and DNA-bound states have been reported previously (*15*, *22*). An interesting feature of the vnd/NK-2 homeodomain is that it has a short recognition helix, helix III, which elongates upon binding to DNA (*23*). In addition, the secondary structures, DNA binding properties, and thermal denaturation behavior of several site-directed mutants of the homeodomain encoded by the *vnd/NK-2* gene have been reported (*18*). Mutations of variable residues in two positions in the vnd/NK-2 homeodomain, namely, H52R and T56W, result in a significant increase in the length of helix III in the absence of DNA. Also, circular dichroism studies suggest that these residues make an important contribution to the structural stability and thus perhaps also to the dynamics of the three-helix bundle that characterizes the homeodomain (*18*).

To better understand the dynamic behavior of the homeodomain and to gain insight into the role of variable residues in the C-terminal region of the DNA recognition helix, we performed ^{15}N relaxation studies on the wild type and H52R/T56W mutant vnd/NK-2 homeodomain, in both the free and DNA-bound state. Although these mutations in positions 52 and 56 do not alter DNA binding affinity, as shown both by gel assays (*18*) and by ITC (*24*), arginine in position 52 is close enough to form a salt bridge with glutamate in position 17 and tryptophan in position 56 is spatially close to the hydrophobic core and to residues 19−25 in helix I and the loop region (*6*, *14*). Thus, it is of interest to examine any modifications in the motional behavior introduced by the two mutations. In addition, the relaxation behavior of the double mutant is studied in the absence and presence of sodium phosphate, the presence of which has a significant effect on the melting temperature and the related thermodynamic parameters (*24*).

MATERIALS AND METHODS

Protein Preparation. The uniformly ^{15}N-labeled wild type vnd/NK-2 homeodomain used in these studies was prepared twice using two slightly different protocols. The first protocol used a pET11d vector and produced a 77-amino acid protein that encompasses the homeodomain. This procedure was described in an earlier publication (*23*). The second protocol used a pET15b vector and involves the same protocol used to express the H52R/T56W double mutant 80-amino acid homeodomain. The procedure used to introduce the mutations, and the expression and purification steps, were described in a recent publication (*18*).

For the free proteins (i.e., no DNA), lyophilized samples were dissolved in a 90% H_2O/10% ^2H_2O mixture and the pH was adjusted to 4.5 using dilute NaOH. For the buffered sample, the protein was dissolved directly into a 50 mM phosphate solution (again a 90% H_2O/10% ^2H_2O mixture) at pH 4.7 with the pH subsequently adjusted to 4.5 with a NaOH solution. In all cases, the final protein concentrations were ~1 mM. The preparation of the wild type vnd/NK-2 homeodomain−DNA complex was described previously (*15*). For the preparation of the H52R/T56W double mutant complex with the 16-base pair double-stranded DNA, separate dilute solutions of the mutant homeodomain (0.2 mM) and the DNA (0.2 mM) were prepared in a 90% H_2O/10% ^2H_2O mixture. Initial pH values were 5.5 for the protein and 6.4 for the DNA. The protein solution was then added dropwise to that containing the DNA to a final protein:DNA stoichiometry of approximately 0.95:1.0. Finally, the resulting solution containing the homeodomain−DNA complex was concentrated by centrifugation in a Centricon 3 (Amicon) filter to approximately 1 mM.

NMR Spectroscopy. Spectra were recorded on Bruker Instruments Inc. DMX360, DRX600, and DRX800 NMR spectrometers using triple-resonance TXI probes operating at the appropriate frequencies. The pulse sequence programs for the two-dimensional (2D) ^{15}N HSQC experiments used to measure longitudinal (T_1), rotating frame ($T_{1\rho}$), and transverse (T_2) relaxation times have been described previously (*25*, *26*). Spectra for both the wild type and H52R/T56W double mutant in the absence of DNA were recorded at 282 K and pH 4.5, where the homeodomain is known to be in the folded conformation. The temperature of the sample in the probe was calibrated using the chemical shift of the water resonance relative to that for TSP. Spectra for the wild type and double mutant homeodomains bound to DNA were recorded at 309 K and pH 6.0. Resonance assignments for the free and DNA-bound wild type ^{15}N-labeled vnd/NK-2 homeodomain in the absence of salt were redone using the procedure described in an earlier study (*15*). Also, the assignments of the unbound H52R/T56W double mutant are given in a previous study (*18*). By contrast, the resonance assignments for the double mutant bound to DNA were made in the current study using primarily a three-dimensional (3D) ^{15}N NOESY-HSQC spectrum (*27*). The spectral parameters for this experiment at 60.8 MHz with flipback and WATERGATE were as follows: 56 (t_1) × 48 (t_2) × 768 (t_3) data points and spectra widths (acquisition times) of 6001.0 Hz for F_1 (9.3 ms), 1733.0 Hz for F_2 (27.7 ms), and 8333.0 Hz for F_3 (92.2 ms) with TPPI and States−TPPI quadrature detection in the t_1 and t_2 dimensions, respectively. The spectral parameters (acquisition times) for the T_1 and $T_{1\rho}$ relaxation measurements for the backbone ^{15}N made from the 2D HSQC data in all cases were as follows: 128 (t_1) × 768 (t_2) data points and spectral widths (acquisition times) of 1136.0 Hz for F_1 (112.7 ms) and 5040.0 Hz for F_2 (152.4 ms) at 36.5 MHz (360 MHz ^{1}H), 1852.0 Hz for F_1 (69.1 ms) and 8993.0 Hz for F_2 (85.4 ms) at 60.8 MHz (600 MHz ^{1}H), and 2110.0 Hz for F_1 (60.7 ms) and 10 417.0 Hz for F_2 (73.7 ms) at 81.1 MHz (800 MHz ^{1}H). The $T_{1\rho}$ measurements were carried out at 60.8 MHz with a spin-lock field of 1800 Hz so that $T_{1\rho} = T_2$ for molecular motions at frequencies higher than 1800 Hz. A modified Carr−Purcell pulse sequence was used for the T_2 measurements, which were carried out on the Bruker DRX600 spectrometer. The spectral parameters were 128 (t_1) × 512 (t_2) data points with the same spectral widths and acquisition times as described above and States−TPPI quadrature detection in t_1. Here a 1 ms delay was inserted between the six ^{15}N refocusing π pulses.

For all T_1 experiments, the shortest delay interval τ that was used was 8 ms. For the unbound homeodomains, the longest τ interval was 1200, 2000, or 2600 ms at 36.5, 60.8, or 81.1 MHz, respectively. For the homeodomain−DNA systems, the longest interval was 2160, 2600, or 3000 ms at

the respective field strength. For the $T_{1\rho}$ measurements on the unbound homeodomains, values of the τ intervals were 4.01, 12.60, 22.39, 39.82, 70.78, 125.88, and 223.90 ms. For the DNA-bound homeodomains, the τ interval values were 3.83, 10.072, 16.79, 32.15, 61.91, 101.75, 158.39, and 242.39 ms. For the Carr−Purcell T_2 measurements, the τ delays were 8, 16, 24, 32, 40, 48, 56, 80, and 160 ms. For each relaxation experiment, spectra were collected in randomized order to minimize the effects of systematic errors.

H52R/T56W Double Mutant Structure Modeling. Two different initial sets of coordinates were used to generate a model structure for the H52R/T56W double mutant vnd/NK-2 homeodomain using molecular dynamics (MD) with the cvff force field. The atomic coordinates for the wild type homeodomain (PDB entry 1vnd; structure 1) were used as the starting point for the double mutant initial structures. This approach is justified since the 2D homonuclear NOESY spectra for the wild type and H52R/T56W double mutant homeodomains are almost superimposable, except for the resonances corresponding to the residues after position 50 in the C-terminal region of the protein (*18*). For the structure of the double mutant, the two side chains were mutated, and for the first initial structure, the coordinates were used with no further modification. The second initial structure also included the extension of helix III, introduced manually by setting the backbone torsional angles of the C-terminal end of helix III (residues 53−60) at typical α-helical values ($\phi = -57°$, $\psi = -47°$, and $\omega = -180°$). All initial structures were solvated with a 7 Å layer of approximately 1000 TIP3P water molecules. Simulations were carried out after introducing either 24 or 30 helical αN(i, $i + 3$) and αN(i, $i + 4$) proton distance restraints based on observed (24 well-defined or 30 well-defined plus ambiguous) NOE cross-peaks (*18*). No restraints other than for the extension of helix III, residues 53−60, were employed for these simulations. Upper bounds were set at 4.2 and 5.0 Å for the proton−proton αN(i, $i + 3$) and αN(i, $i + 4$) distances characterized by well-defined cross-peaks, respectively, and at 5.5 and 6.0 Å for the corresponding distances represented by ambiguous cross-peaks. The MD simulations consisted of heating from 0 to 298 K during the first 400 ps and then performing dynamics for a period of 1 ns before reminimization to allow the protein structure to relax to a low-energy conformation without undue distortion. Structure calculations and modeling were performed with the INSIGHT_II program running on a Silicon Graphics Origin 2400 workstation.

Relaxation Parameter Analysis. The NMR data were processed and analyzed using nmrPipe and nmrDraw (*28*). Heteronuclear dimensions were extended by linear prediction and zero-filled prior to Fourier transformation. The program pipp (*29*) was used to extract the resonance intensities from the spectra. Relaxation rates were determined by nonlinear fits of the time dependence of the peak heights using a two-parameter exponential fit to the data. Errors in the relaxation parameters were estimated from the standard deviations of the same parameters using fittings with half-sized subsets from the full set of τ values. These subsets were obtained by pairwise sampling of the respective complete sets. Errors in the estimate of the heteronuclear NOE values (XNOE) were determined from the rms values of the noise from the measurement and reference spectra.

From the experimental measurements, the transverse relaxation components of $T_{1\rho}$ (i.e., T_2^{calc}) were calculated from the equations (*30*)

$$\omega_{eff} = 2\Pi(B_1^2 + \Delta^2)^{1/2}$$

$$(T_{1\rho})^{-1} = (T_1)^{-1}\cos^2\theta + (T_2^{calc})^{-1}\sin^2\theta \qquad (1)$$

where θ is the tilt angle between the effective spin-lock field axis and that of the main B_1 magnetic field and Δ is the offset for the spin-lock field (both in hertz) and was calculated from the chemical shifts relative to the ^{15}N carrier frequency. These transverse relaxation components of $T_{1\rho}$ were then compared with the experimentally measured T_2 values, T_2^{exp}.

Diffusion Tensor Analysis. The rotational diffusion tensors for the wild type protein in both the free and DNA-bound states (complete three-dimensional structures are available only for the wild type homeodomain) were obtained from the T_1/T_2^{calc} ratios. T_2^{calc} values were preferred to the T_2^{exp} ones, since they are less affected by chemical exchange. Moreover, only amide N−H groups showing no significant internal motion, as judged from the values of the individual heteronuclear NOE values (XNOE), were used for the calculations. The diffusion tensor analysis was carried out with the program Quadric Diffusion (*31*). A local correlation time (τ_{loc}) for each residue is evaluated from the corresponding T_1/T_2^{calc} ratio using the model-free expression for the isotropic spectral density function. From this correlation time, the local diffusion coefficient for each residue can easily be calculated [$D_{loc} = (6\tau_{loc})^{-1}$]. Diffusion tensor parameters are then fit to the experimental data using the relevant experimental three-dimensional structure [30 structures are available for the unbound wild type vnd/NK-2 homeodomain (PDB entry 1VND) and 20 structures are available for the homeodomain−DNA complex (PDB entry 1NK2)]. Various motional models were considered: an isotropic model with one parameter (D_{iso}), an axially symmetric model with four parameters ($D_{||}$, $D_{\perp}$, ϕ, and θ), and a fully anisotropic model with six parameters (D_{zz}, D_{yy}, D_{xx}, ϕ, θ, and φ). To establish the appropriateness of each of the models, a figure of merit, F, is implemented in Quadric Diffusion. An improved fit to the highly parametrized model (going from the isotropic model to either an axially symmetric or a fully asymmetric model) is considered to be statistically significant at a confidence level of >99% if the value of F is greater than 4.5.

RESULTS AND DISCUSSION

NMR Chemical Shifts. Two-dimensional NOESY-HSQC spectra were analyzed for the wild type and H52R/T56W double mutant vnd/NK-2 homeodomains in both the free (282 K) and DNA-bound states (309 K). The experiments on the homeodomain in the absence of DNA were carried out at the low temperatures to minimize any effects due to the vicinity of the thermal folding-unfolding transition (298 K for the wild type and 306 K for the double mutant). The homeodomain−DNA complex is known to be stable up to at least 330 K. A weighted average of the ^1H and ^{15}N chemical shift differences between the backbone amide nitrogen atoms of both the wild type vnd/NK-2 and the H52R/T56W double mutant homeodomains bound to DNA

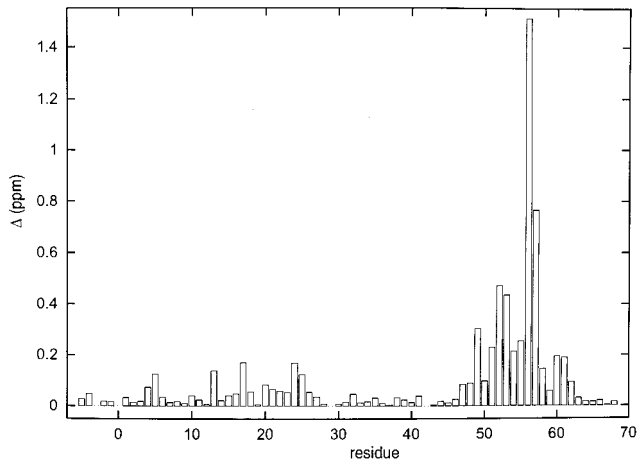

FIGURE 1: Chemical shift differences between the wild type and H52R/T56W double mutant vnd/NK-2 homeodomains bound to DNA for backbone amide resonances. The displayed differences are averaged over both ^{15}N and ^{1}H chemical shifts and are given in ^{1}H parts per million by $\Delta = [(\delta^{1}\mathrm{H})^2 + (\delta^{15}\mathrm{N}/5)^2]^{1/2}$.

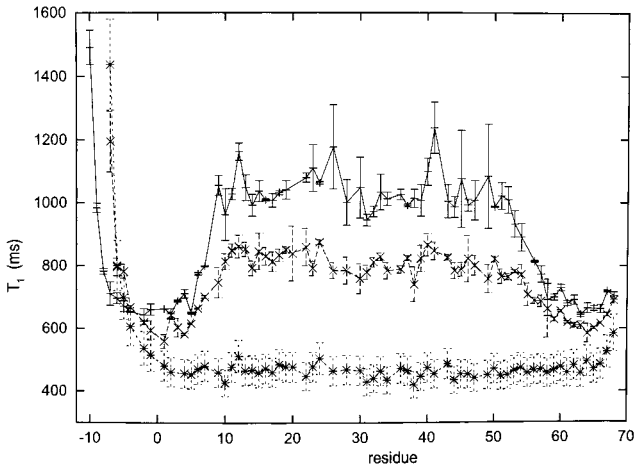

FIGURE 2: ^{15}N T_1 relaxation times (milliseconds) for the unbound wild type vnd/NK-2 homeodomain at 81.8 (solid line with +), 60.8 (dashed line with ×), and 36.5 MHz (dotted line with ∗).

is summarized graphically in Figure 1. Significant chemical shift differences are seen from residue 50 through residue 62 (*18*) that are likely due to a global stabilization of helix III. In addition to these expected differences in helix III, especially in the area close to residues 52 and 56, small but observable chemical shift changes are seen in the region between residues 15 and 26. These latter changes may occur as a result of some stabilization of helix I and the loop region due to the possible formation of a salt bridge between residues 17 and 52 or because of the absence of the histidine side chain in position 52 and the presence of the tryptophan side chain in position 56, since this region of helix III is spatially close to residues ∼19−25. The pK_a of the glutamate is likely to be ∼4.3, or lower if the salt bridge is actually formed (the sample pH was measured to be 4.5), thus indicating that the interaction of the side chains of residues 17 and 52 would be primarily electrostatic.

NMR Relaxation Data. Most of the relaxation data were obtained on an 80-residue recombinant protein that encompasses the homeodomain. A small fraction of the data for the wild type analogue in the free state at 60.8 and 36.5 MHz was obtained on an older 77-amino acid protein. In practice, differences in the relaxation parameters between the two wild type proteins are inconsequential. For the wild type homeodomain, the relaxation data were obtained for 69 of the 75 possible resonances (67 of 72 in the case of the older sample) in the HSQC spectrum expected for the recombinant protein that encompasses the homeodomain. For the H52R/T56W double mutant, relaxation data for 70 of 75 possible residues were obtained. These data consist of the longitudinal (T_1), rotating frame ($T_{1\rho}$), and transverse (T_2^{exp}) relaxation times as well as the heteronuclear ^{15}N−^{1}H NOE (XNOE) values. The ^{15}N T_1 values at the three field strengths (36.5, 60.8, and 81.1 MHz) are shown in the absence of DNA for the wild type homeodomain in Figure 2 (analogous data not shown for the double mutant). The wild type analogue was shown to adopt the three-helix fold characteristic of the homeodomain from residue 10 through 52, whereas the double mutant with an elongated DNA recognition helix is structured from residue 10 through 60. The sensitivity of the NMR relaxation rates to the dynamics of the structured region of the protein as well as to the internal dynamics of the N-terminal region before residue 10 and the C-terminal region of the protein past residue 60 is emphasized by the field dependence of the T_1 values. This motion-dependent manifestation is best seen from the measurements at 81.1 MHz (Figure 2), whereas from the T_1 measurements at 36.5 MHz, the discrete dependence on local internal dynamics and on overall rotational reorientation is not readily apparent. The interpretation here is that the internal and overall rotational reorientation correlation times fall on opposite sides of but nearly equally distant from the T_1 minimum for measurements at 36.5 MHz. The weak dependence of the T_1 values on the Larmor frequency for residues in the unstructured N- and C-terminal regions of the protein shows that, except for one or two residues at the very beginning and very end of the protein, the fluctuations of the backbone N−H vectors for the N- and C-terminal regions occur on a time scale on the order of 10^{-9} s (*32*). Although the average values of T_1 for residues in helix I differ slightly from those in helix II and helix III, residues in the loop region between helix I and helix II (i.e., residues 23−27) show no evidence for internal motion on the picosecond to nanosecond time scale. A comparison of the T_1 and XNOE values for the wild type and double mutant homeodomains (81.1 MHz for the T_1 values and 60.8 MHz for the XNOE values) in the free or unbound state is shown in panels a and b of Figure 3. The primary differences in the T_1 and XNOE values between the wild type and double mutant homeodomains is that for the double mutant, values of these relaxation parameters that are characteristic of the structured part of the protein are observed out to residue 57 or 58 rather than only to residue 52. A model structure of the H52R/T56W double mutant vnd/NK-2 homeodomain based on the experimental structure of the wild type analogue (i.e., the first initial structure without manual preelongation of helix III) together with the additional helical restraints from the extension of helix III (*18*) is presented in Figure 4. The coordinates for the model structure are available by contacting K.-J. Hwang at kjh@camd.ssu.ac.kr. The results of the simulations also suggest that helix III in the double mutant is fully maintained only to approximately residue 57 or 58, even though cross-peaks associated with a helical structure are seen out to residue 60 (*18*).

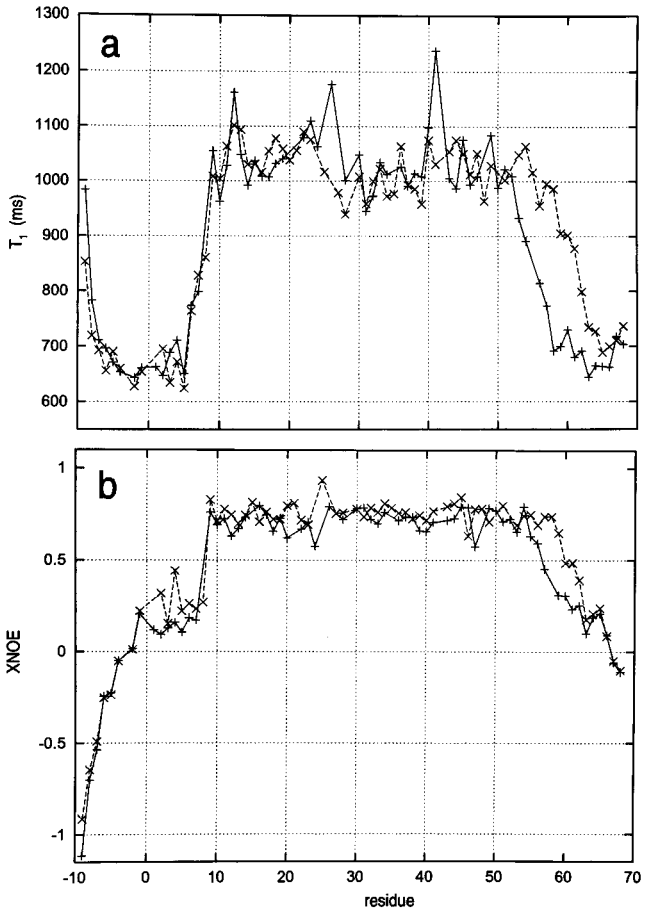

FIGURE 3: Comparison of (a) ^{15}N T_1 relaxation times (milliseconds) at 81.1 MHz and (b) the ^{15}N−^{1}H XNOE values at 60.8 MHz for the wild type (solid line with +) and H52R/T56W double mutant (dashed line with ×) vnd/NK-2 homeodomains in the unbound state.

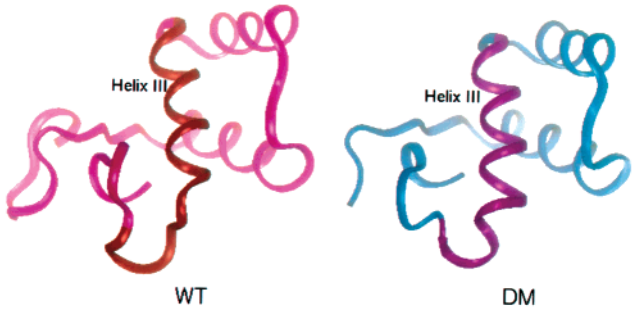

FIGURE 4: Representation of the model structure of the H52R/T56W double mutant vnd/NK-2 homeodomain (DM) compared with the wild type analogue (WT). The structures are oriented to emphasize the elongation of helix III for the double mutant. Helix III is shown in dark red for WT and dark magenta for DM.

The vnd/NK-2 homeodomain binds to an unusual sequence of DNA (*15*, *23*) that contains CAAGTG as its core. The binding affinities of both the wild type and double mutant analogue for the DNA are on the order of 10^{-10} M. A comparison of the T_1 values at 800 MHz and the XNOE values at 600 MHz between the wild type homeodomains free and bound to DNA is presented in panels a and b of Figure 5, respectively. For the wild type homeodomain bound to DNA, the relaxation parameters at 60.8 and 81.1 MHz for four residues and at 36.5 MHz for two additional residues were eliminated as a result of assignment ambiguities. A comparison of the Lipari−Szabo order parameter, S^2, for the

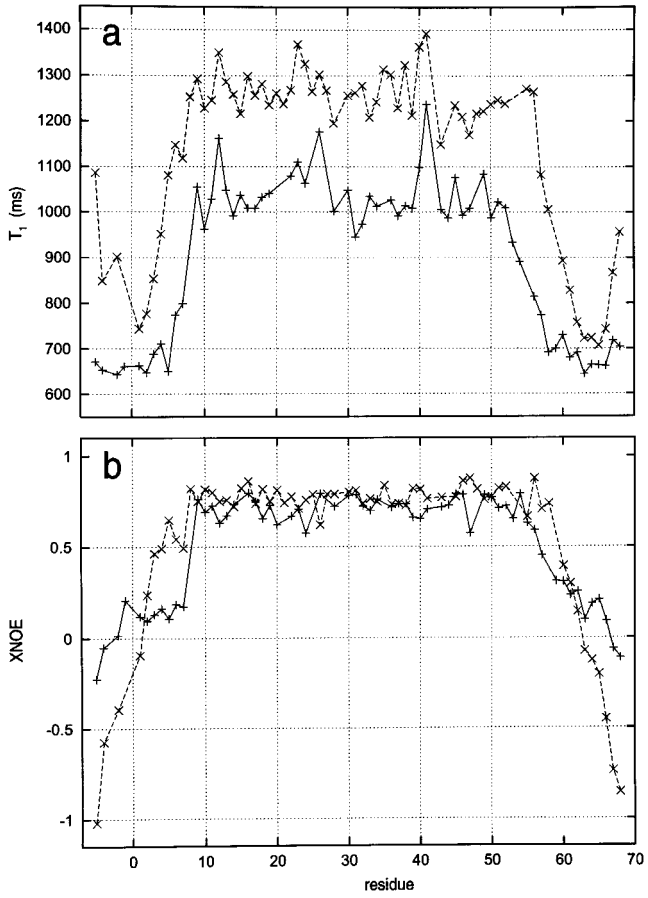

FIGURE 5: Comparison of (a) ^{15}N T_1 relaxation times (milliseconds) at 81.1 MHz and (b) the ^{15}N−^{1}H XNOE values at 60.8 MHz for the wild type vnd/NK-2 homeodomain in the free state (solid line with +) and bound to DNA (dashed line with ×).

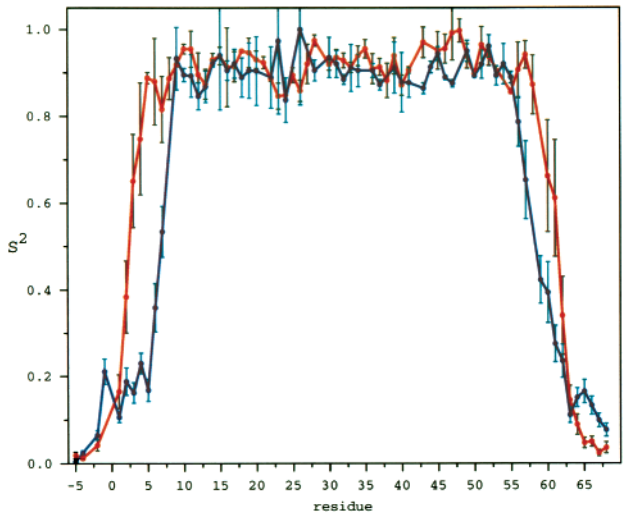

FIGURE 6: Comparison of the order parameter, S^2, for the wild type vnd/NK-2 homeodomain in the free (blue) and DNA-bound states (red).

wild type homeodomain in the free and DNA-bound states is depicted in Figure 6. For the H52R/T56W double mutant bound to DNA at 60.8 and 81.1 MHz, three residues were eliminated because of low intensity or resonance overlap, and one additional residue was eliminated from the data at 36.5 MHz. Careful examination of the T_1 values for both the wild type and double mutant homeodomains bound to DNA shows, in contrast to the behavior of the unbound

homeodomains, a clear distinction between the structured and unstructured parts of the proteins at all field strengths. Such behavior is expected for the homeodomain−DNA complex, where the molecular masses are ∼20 kDa, compared to ∼9.5 kDa for the free homeodomain, thus leading to longer correlation times for both internal and overall motion in the complex: ∼10.5 ns at 285 K for the free homeodomain and ∼11.5 ns at 308 K for the complex. The T_1 and XNOE values for the homeodomain bound to DNA evidence structure in the N-terminal arm before helix I from approximately residue 3 and in the C-terminal arm to approximately residue 58. It has been shown that the N-terminal arm of the homeodomain is measurably more structured due to direct interactions with the DNA. Furthermore, the interaction of the homeodomain with the DNA results in an elongation of the DNA recognition helix, helix III, that is maintained at least to residue 60 analogous to that observed for the H52R/T56W double mutant homeodomain in both the free and DNA-bound state.

To examine lower-frequency motions, the chemical exchange contribution to T_2^{exp}, for both the wild type and double mutant homeodomains, was evaluated from the difference between the experimentally measured T_2 and the T_2^{calc} values (transverse relaxation component of $T_{1\rho}$). It is important to emphasize that the relaxation data on both proteins were collected at 282 K, where at pH 4.5 and without added salt the thermal denaturation occurs at ∼298 K for the wild type homeodomain and ∼306 K for the double mutant. The strength of the spin-lock field for all experiments was ∼1800 Hz. The comparisons in Figure 7 show that, for the wild type homeodomain, there is a contribution to T_2 for motions below the spin-lock frequency for many of the residues, whereas for the double mutant homeodomain, $T_{1\rho} = T_2$ for all residues. Since it has been shown that the thermal unfolding transition is broader and occurs at a lower temperature for the wild type analogue (*18*), a plausible origin of these low-frequency contributions to the experimental T_2 values could be motions associated with the initiation of unfolding. An interesting consequence of this hypothesis is that, al least for this homeodomain, the unfolding phenomenon begins as a low-frequency exchange process.

The availability of the structural (*22*) and dynamic data together with the thermodynamic characterization (*24*) of both the wild type and the double mutant vnd/NK-2 homeodomain provides some unique insights into the folding and DNA binding properties of the protein. Since extension of the third helix of the wild type analogue upon binding must cost free energy, one would expect the double mutant to bind DNA with higher affinity, although experimentally the wild type and double mutant bind with similar affinities (*18*). Since the T_2 results indicate the presence of low-frequency flexibility for the wild type homeodomain, one should consider the possibility that the wild type might be better able to accommodate DNA binding (i.e., less internal energy or lower ΔH) than the more rigid double mutant, thus potentially counterbalancing at least some of the free energy cost to extend helix III.

Analysis of Rotational Diffusion. The ^{15}N relaxation behavior of small proteins is determined by both the overall rotational reorientation rate and local internal fluctuations of each N−H vector and its orientation with respect to the axes of the diffusion tensor. Calculation of the rotational

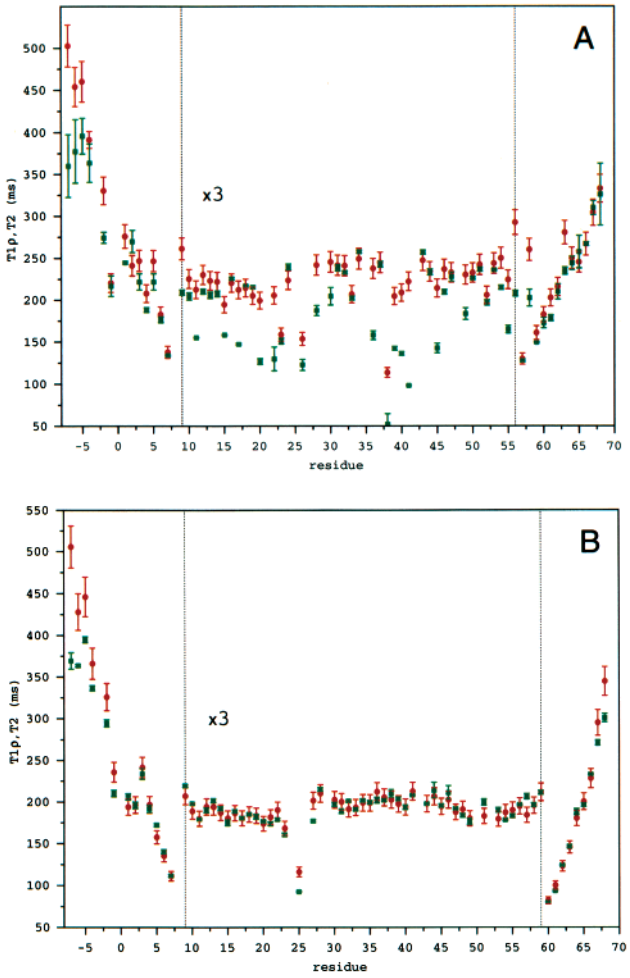

FIGURE 7: Comparisons of the experimentally measured T_2 (green) and transverse relaxation component of $T_{1\rho}$ (red) for the unbound (A) wild type and (B) the H52R/T56W vnd/NK-2 homeodomains, at 60.8 MHz.

diffusion coefficients, D_{xx}, D_{yy}, and D_{zz}, requires the structural coordinates of the protein in addition to the values of T_1/T_2^{calc}. As an initial step in the calculation of the rotational diffusion tensors, residues for which internal motion contributes significantly to the observed T_1/T_2^{calc} ratios were removed from the calculations. Residues whose amide XNOE values were <0.65 were eliminated from the analysis. Consequently, the analysis was carried out using 27 data points for the wild type free protein: residues 10−22, 28−38, and 42−53, where E15, R21, P29, H33, A35, I38, P42, I47, and W48 also were excluded from the calculations. For the homeodomain−DNA complex, 34 data points were available: residues 10−22, 28−38, and 42−57, where L16, P29, P42, Q44, R53, and Y54 were not used. Each of the conformers shows significant variation in the orientation of its inertial tensor with respect to the molecular frame such that the components of the principal axes show relatively large standard deviations, namely, 1.00:0.87(0.06):0.48(0.13) for the free protein and 1.00:0.94(0.03):0.73(0.03), where the standard deviations are given in parentheses.

Rotational diffusion tensors were analyzed using isotropic, axially symmetric, and fully anisotropic diffusion models with the program Quadric Diffusion (*31*) for both the unbound (27 amide N−H vectors) or DNA-bound (34 amide N−H vectors) homeodomains. The results, which are based

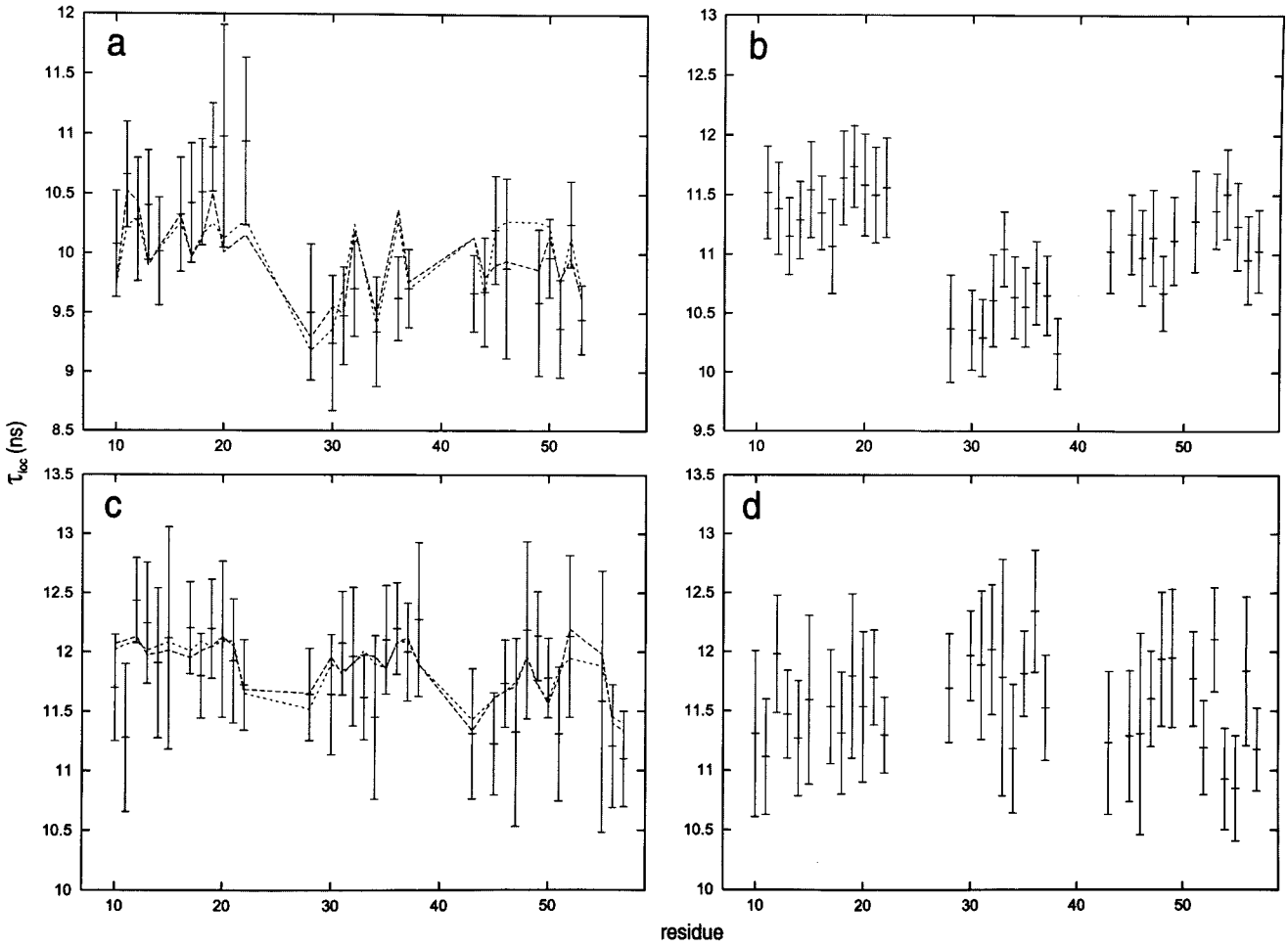

FIGURE 8: Local correlation times, τ_{loc} (nanoseconds), with error bars evaluated from T_1/T_2^{calc} ratios at 60.8 MHz for the unbound wild type vnd/NK-2 homeodomain (a), the unbound H52R/T56W double mutant vnd/NK-2 (b), the wild type vnd/NK-2 homeodomain−DNA complex (c), and the DNA-bound double mutant homeodomain (d). In panels a and c, the lines represent the expected τ_{loc} values calculated from the fully anisotropic model (dashed lines) and from the unconstrained axially symmetric model for rotational diffusion (dotted lines).

Table 1: Diffusion Tensor Parameters[a]

	D_{iso} ($\times 10^{-7}$ s^{-1})	$2D_{zz}/(D_{xx}+D_{yy})$ or $D_{\|}/D_{\perp}$	D_{xx}/D_{yy}	F
		Unbound Protein		
isotropic	1.666 (0.008) [−]	−	−	−
axial prolate	1.682 (0.016) [0.013]	1.189 (0.062) [0.024]	−	3.05 [0.69]
axial oblate	1.705 (0.017) [0.012]	0.755 (0.120) [0.037]	−	3.08 [0.83]
axial	1.692 (0.017) [0.022]	0.962 (0.096) [0.240]	−	3.50 [0.61]
anisotropic	1.695 (0.021) [0.016]	1.212 (0.076) [0.022]	1.196 (0.125) [0.064]	1.07 [0.51]
		DNA Complex		
isotropic	1.405 (0.007) [−]	−	−	−
axial prolate	1.427 (0.010) [0.001]	1.110 (0.052) [0.005]	−	4.13 [0.14]
axial oblate	1.413 (0.010) [0.001]	0.864 (0.070) [0.005]	−	5.80 [0.38]
axial	1.413 (0.010) [0.001]	0.864 (0.070) [0.005]	−	5.80 [0.38]
anisotropic	1.395 (0.033) [0.008]	1.173 (0.174) [0.045]	1.152 (0.088) [0.008]	0.26 [0.19]

[a] Five different motional models were analyzed. The axially symmetric case was determined as an unconstrained fit (axial) as well as under the conditions $D_{\|}/D_{\perp} > 1$ (axial prolate) and $D_{\|}/D_{\perp} < 1$ (axial oblate). Each parameter is the average over the values obtained from the individual structures, where the standard deviation is given in square brackets and the averaged value of the errors resulting from the analysis for each individual structure is given in parentheses. For the unbound protein, one of the 30 conformers was neglected because it resulted in a singularity.

on the average of the 30 structures for the homeodomain in the absence of DNA and on the average of the 20 structures for the homeodomain−DNA complex, are summarized in Table 1. For the free homeodomain, the large standard deviation is due primarily to the flexibility associated with the unstructured ends of the protein.

For the DNA complex, the calculations of the figure of merit, *F*, indicate that the data fit slightly better to an axially symmetric model rather than to an isotropic or fully anisotropic model. Since the level of asymmetry is low, the data can be fit to both a prolate and an oblate ellipsoidal tensor. Analysis of the orientations of the diffusion axes (hereafter called *x*, *y*, and *z*) shows that the principal axis for the axially symmetric prolate ellipsoidal tensor (the *z*-axis about which rotational diffusion is fastest) is nearly parallel to the long axis of the double-stranded DNA, whereas that

for the oblate ellipsoid is nearly perpendicular to the long axis. In addition, in this latter case both helix I and helix II lie almost in the *xy* plane. On the other hand, for the fully anisotropic model, these helices are out of the plane nearly aligned to the *z*-axis with the double-stranded DNA lying again close to the *xy* plane, as in the case of the oblate ellipsoid. For a slightly asymmetrical molecule, it is not possible to reliably distinguish between the two possible solutions to the diffusion equation for axially symmetrical rotors. The critical point here is that the results are consistent with the oblate as well as with the prolate ellipsoid for the entire set of 20 conformers of the complex both for the diffusion and for the inertial tensors. Also, the standard deviations relative to those found for the free protein are lower since the flexible ends of the protein represent a smaller relative contribution to the diffusion tensor.

In attempting a comparison of the orientation of the principal axes of the symmetry properties of the inertial tensor with respect to the diffusion tensor, we should note several important features. First, the inertial tensor is calculated from the distribution of the masses in a fixed structure, whereas the diffusion tensor is related to the friction distribution and includes hydrodynamic interactions. Moreover, the diffusion tensor is derived from a fit to the experimental relaxation times that involve rotational diffusion of a fluctuating structure (*32*) as well as an experimentally derived NMR structure with its associated errors.

From the data shown in panels a and b of Figure 8, we observe that the τ_{loc} values (derived from the T_1/T_2^{calc} ratios; cf. Material and Methods) for all the residues between positions 10 and 53 in the case of the wild type analogue in the free state and between positions 10 and 60 for the H52R/T56W double mutant homeodomain show limited dispersion. The values for τ_{loc} in helix I are, on average, slightly longer than those observed in helix II and helix III. The implication here is that helix I is slightly less flexible. This lack of variation in τ_{loc} occurs since all amide N−H bond vectors within a given helix that have similar orientations are nearly parallel with respect to the axis of the helix in a molecule and, in a molecule with a low level of asymmetry (Table 1), have no preferential rotational diffusion axis. Consequently, all amide N−H bond vectors in the structured (i.e., helical) regions of the homeodomain sample similar orientations with equal probabilities. Analogous behavior is observed for both the wild type and double mutant homeodomains in the DNA-bound state (Figure 8c,d). Here also, the lack of asymmetry provides a plausible explanation for the limited dispersion for τ_{loc} values for the homeodomain in the DNA-bound state.

Conclusions. The relaxation parameter measurements and the ensuing analysis describe the motional behavior of the wild type and the H52R/T56W double mutant vnd/NK-2 homeodomain in both the free and DNA-bound state. This study demonstrates that the N- and C-terminal regions of the wild type and mutant homeodomains are flexible in the absence of DNA. However, the H52R/T56W double mutant shows an increased level of order in the C-terminal portion of helix III in the absence of DNA over that found for the wild type analogue, in accord with structural evidence. In the DNA-bound state, both the wild type and double mutant show an increased level of order in both the N-terminal and C-terminal portions of the homeodomain. The N-terminal arm binds in the minor groove of the DNA, and the C-terminal end of helix III extends to residue 62, outside the classical homeodomain. An important effect of the conservative H52R mutation in position 52 is stabilization of the tertiary structure of the homeodomain, at least in part, by the formation of a salt bridge with E17. The T56W mutation is critical to the elongation and stabilization of helix III, possibly by virtue of the hydrophobic effect, i.e., by minimizing solvent exposure. The functional effect of these substitutions remains unknown, and thus, the corresponding transgenic experiments become very important.

ACKNOWLEDGMENT

We thank Dr. Nico Tjandra for advice and help with the experimental procedures and for many stimulating and useful discussions on the motional properties of proteins.

REFERENCES

1. Gehring, W. J., Affolter, M., and Burglin, T. (1994) Homeodomain proteins, *Annu. Rev. Biochem. 63*, 487−526.
2. McGinnis, W., and Krumlauf, R. (1992) Homeobox genes and axial patterning, *Cell 68*, 283−302.
3. Scott, M. P., Tamkun, J. W., and Hartzell, G. W., III (1989) The structure and function of the homeodomain, *Biochim. Biophys. Acta 989*, 25−48.
4. Xiang, B., Weiler, S., Nirenberg, M., and Ferretti, J. A. (1998) Structural basis of an embryonically lethal single Ala−Thr mutation in the vnd/NK-2 homeodomain, *Proc. Natl. Acad. Sci. U.S.A. 95*, 7412−7416.
5. Schier, A. F., and Gehring, W. (1992) Direct homeodomain-DNA interaction in the autoregulation of the *fushi tarazu* gene, *Nature 356*, 804−807.
6. Kissinger, C. R., Liu, B. S., Martinblanco, E., Kornberg, T. B., and Pabo, C. O. (1990) Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions, *Cell 63*, 579−590.
7. Wolberger, C., Vershon, A. K., Liu, B. S., Johnson, A. D., and Pabo, C. O. (1991) Crystal structure of MAT α-2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions, *Cell 67*, 517−528.
8. Klemm, J. D., Rould, M. A., Aurora, R., Herr, W., and Pabo, C. O. (1994) Crystal structure of the OCT-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding molecules, *Cell 77*, 21−32.
9. Wilson, D. S., Gunther, B., Desplan, C., and Kuriyan, J. (1995) High-resolution crystal structure of a paired (PAX) class cooperative homeodomain dimer on DNA, *Cell 82*, 709−719.
10. Li, T., Stark, M. R., Johnson, A. D., and Wolberger, C. (1995) Crystal structure of the MAT-A1-MAT-α-2 homeodomain heterodimer bound to DNA, *Science 270*, 262−290.
11. Hirsch, J. A., and Aggarwal, A. K. (1995) Structure of the even-skipped homeodomain complexed to AT-rich DNA: new perspectives on homeodomain specificity, *EMBO J. 14*, 6280−6291.
12. Jacobson, E. M., Li, P., Leon del Rio, A., Rosenfeld, M. G., and Aggarwal, A. K. (1997) Structure of the Pit-1 POU domain bound to DNA as a dimer: unexpected arrangement and flexibility, *Genes Dev. 11*, 198−212.
13. Tan, S., and Richmond, T. J. (1998) Crystal structure of the yeast MAT α-2/MCM1/DNA ternary complex, *Nature 391*, 660−666.
14. Qian, Y. Q., Otting, G., Billeter, M., Muller, M., Gehring, W., and Wuthrich, K. (1993) Nuclear-magnetic-resonance spectroscopy of a DNA complex with the uniformly C-13 labeled *Antennapedia* homeodomain and structure determination of the DNA-bound homeodomain, *J. Mol. Biol. 234*, 127−138.
15. Gruschus, J. M., Tsao, D. H. H., Wang, L.-H., Nirenberg, M., and Ferretti, J. A. (1999) The three-dimensional structure of

the vnd/NK-2 homeodomain-DNA complex by NMR spectroscopy, *J. Mol. Biol. 289*, 529−545.
16. Boncinelli, E. (1997) Homeobox genes and disease, *Curr. Opin. Genet. Dev. 7*, 331−337.
17. Ferretti, J. A. (2001) The Homeobox and Genetic Disease: Structure and Dynamics of Wild Type and Mutant Homeodomain Proteins, *J. Biochem. Mol. Biol. 34*, 1−7.
18. Weiler, S., Gruschus, J. M., Tsao, D. H. H., Yu, L., Wang, L.-H., Nirenberg, M., and Ferretti, J. A. (1998) Site-directed mutations in the vnd/NK-2 homeodomain: basis of variations in structure and sequence-specific DNA binding, *J. Biol. Chem. 273*, 10994−11000.
19. TuckerKellogg, L., Rould, M. A., Chambers, K. A., Ades, S. E., Sauer, R. T., and Pabo, C. O. (1997) Engrailed (Gln50−Lys) homeodomain-DNA complex at 1.9 Å resolution: structural basis for enhanced affinity and altered specificity, *Structure 5*, 1047−1057.
20. Frasch, M., Warrior, R., Tugwood, J., and Levine, M. (1988) Molecular analysis of *even-skipped* mutants in *Drosophila* development, *Genes Dev. 2*, 1824−1838.
21. Kim, Y., and Nirenberg, M. (1989) *Drosophila* NK-homeobox genes, *Proc. Natl. Acad. Sci. U.S.A. 86*, 7716−7720.
22. Tsao, D. H. H., Gruschus, J. M., Wang, L.-H., Nirenberg, M., and Ferretti, J. A. (1995) The three-dimensional solution structure of the NK-2 homeodomain from *Drosophila*, *J. Mol. Biol. 251*, 297−307.
23. Tsao, D. H. H., Gruschus, J. M., Wang, L.-H., Nirenberg, M., and Ferretti, J. A. (1994) Elongation of helix III of the NK-2 homeodomain upon binding to DNA: A secondary structure study by NMR, *Biochemistry 33*, 15053−15060.
24. Gonzalez, M., Weiler, S., Ferretti, J. A., and Ginsburg, A. (2001) *Biochemistry 40*, 4923−4931.
25. Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W., and Bax, A. (1992) Backbone dynamics of calmodulin studied by N-15 relaxation using inverse detected 2-dimensional NMR spectroscopy: The central helix is flexible, *Biochemistry 31*, 5269−5278.
26. Tjandra, N., Kuboniwa, H., Ren, H., and Bax, A. (1995) Rotational dynamics of calcium free calmodulin studied by N-15 NMR relaxation measurements, *Eur. J. Biochem. 230*, 1014−1024.
27. Gruschus, J. M., and Ferretti, J. A. (1999) Signal enhancement using 45° water flipback for 3D ^{15}N-edited ROESY and NOESY HMQC and HSQC, *J. Magn. Reson. 140*, 451−459.
28. Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectra processing system based on UNIX pipes, *J. Biomol. NMR 6*, 279−293.
29. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) A common sense approach to peak picking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams, *J. Magn. Reson. 95*, 214−220.
30. Meekhof, A. E., and Freund, S. M. V. (1999) Separating the contributions to ^{15}N transverse relaxation in a fibronectin type III domain, *J. Biomol. NMR 14*, 13−22.
31. Lee, L. K., Rance, M., Chazin, W. J., and Palmer, A. G. I. (1997) Rotational diffusion anisotropy of proteins from simultaneous analysis of ^{15}N and ^{13}C nuclear spin relaxation, *J. Biomol. NMR 9*, 287−298.
32. La Penna, G., Fausti, S., Perico, A., and Ferretti, J. A. (2000) Smoluchowski dynamics of the vnd/NK-2 homeodomain from *Drosophila melanogaster*: Second-order maximum correlation approximation, *Biopolymers 54*, 89−103.

BI010398R