# Local Folding and Misfolding in the PBX Homeodomain from a Three-State Analysis of CPMG Relaxation Dispersion NMR Data

Patrick J. Farber,[†] Jelle Slager,[‡] and Anthony K. Mittermaier*,[†]

[†]Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec, Canada H3A 0B8
[‡]Department of Biophysical Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands

S *Supporting Information*

ABSTRACT: NMR Carr−Purcell−Meiboom−Gill (CPMG) relaxation dispersion experiments represent a powerful approach for characterizing protein internal motions and for gaining insight into fundamental biological processes such as protein folding, catalysis, and allostery. In most cases, CPMG data are analyzed assuming that the protein exchanges between two different conformational states. Systems exchanging among more than two states are far more challenging to characterize by CPMG NMR. For example, in the case of three-state exchange in the fast time scale regime, it is difficult to uniquely connect the parameters extracted from CPMG analyses with the physical parameters of most interest, intercoversion rates, populations, and chemical shift differences for exchanging states. We have developed a grid search selection procedure that allows these physical parameters to be uniquely determined from CPMG data, based on additional information, which in this study comprises ligand-induced chemical shift perturbations. We applied this approach to the PBX homeodomain (PBX-HD), a three-helix protein with a C-terminal extension that folds into a fourth helix upon binding to DNA. We recently showed that the C-terminal extension transiently folds, even in the absence DNA, in a process that is likely tied to the cooperative binding of PBX-HD to DNA and other homeodomains. Using the grid search selection procedure, we found that PBX-HD undergoes exchange between three different conformational states, a major form in which the C-terminal extension is unfolded, the previously identified state in which the C-terminal extension forms a fourth helix, and an additional state in which the C-terminal extension is misfolded.
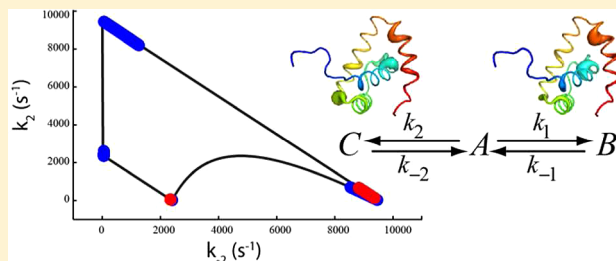
■ INTRODUCTION

Many biological functions of proteins involve conformational fluctuations that occur on the millisecond to microsecond time scales. Carr−Purcell−Meiboom−Gill (CPMG) relaxation dispersion NMR spectroscopy is a powerful method for quantifying these types of motions at atomic resolution.[1,2] The approach is based on measuring the contributions to transverse spin relaxation rates (R_2) that are produced by fluctuations in chemical shift associated with exchange among different conformational states. CPMG experiments quench exchange contributions to R_2 by applying trains of radio frequency pulses. An analysis of the dependence of R_2 on the pulse repetition rate provides quantitative thermodynamic (population) and kinetic (rate) information regarding exchanging states and yields structural information in the form of NMR chemical shift differences between the states. The first step in analyzing CPMG data is the selection of a suitable model for the exchange process. In many cases, a kinetic model involving two exchanging states (A ↔ B) is sufficient. CPMG studies of protein folding,[3] molecular recognition,[4] allostery,[5] and catalysis[6] have successfully employed two-state kinetic models involving exchange between folded/unfolded, free/bound, or active/inactive states, for example.

Nevertheless, more complex motions involving millisecond−microsecond time scale exchange among three or more conformational states are often observed. For example, several mutants of the Fyn SH3 domain have been shown to undergo exchange between a folded state, an unfolded state, and an on-pathway folding intermediate.[7] Multiple disulfide bonds in the trypsin inhibitor BPTI undergo simultaneous isomerization dynamics.[8] Several protein complexes, including those formed by the CREB pKID domain with a KIX domain target,[9] the zymogen prothrombin with an antithrombic peptide,[10] and ubiquitin with an SH3 domain,[11] experience exchange between multiple binding modes.

Although CPMG methods can give valuable insight into multistate exchange, the analysis is significantly more difficult than that for two-state exchanging systems. Among the challenges presented by such systems is ascertaining that exchange involves more than two states in the first place; individual CPMG profiles are usually well-fit by two-state exchange equations even when exchange involves more than two states,[12] except under certain favorable conditions.[8] In addition, the number of kinetic and thermodynamic parameters describing the system grows rapidly with the number of exchanging states. It is increasingly difficult to extract a unique

set of kinetic and thermodynamic parameters from CPMG data for systems of increasing complexity. For example, in the case of linear three-state exchange (C ↔ A ↔ B) that is fast on the NMR chemical shift time scale, CPMG analyses yield two global kinetic exchange parameters, while the system is formally described by four global rate constants.[8]

Here, we present a two-step approach for characterizing fast time scale three-state protein conformational exchange. In the first step, we employ a suite of recently developed ^{1}H/^{15}N NMR spin relaxation experiments that yield exchange-free transverse relaxation rates[13] together with ^{1}H-decoupled ^{15}N CPMG[14] and traditional ^{15}N R_1 and {^{1}H}−^{15}N ssNOE measurements to determine robust multisite exchange parameters. As an added benefit, these experiments provide information on nanosecond to picosecond time scale protein motions. In the second step, we use a novel grid search selection procedure to obtain full sets of multisite exchange parameters by combining the results of step 1 with additional chemical shift information.

We applied this approach to characterize millisecond to microsecond dynamics in the PBX homeodomain (PBX-HD), an 81 residue protein consisting of 3 α helices that adopt a consensus homeodomain structure[15] and an unstructured 18-residue C-terminal extension.[16] PBX-HD binds DNA cooperatively with Hox transcription factors and plays an important role in anterior−posterior patterning in vertebrate development.[17] DNA binding causes the C-terminal extension of PBX-HD to fold into a fourth α helix that packs against the rest of the homeodomain.[1,2] PBX-HD interacts with Hox in the presence but not the absence of DNA.[18] Truncation of the C-terminal extension decreases the DNA affinity by a factor of 5[19] and reduces PBX-HD/Hox interactions.[20−22] Interestingly, the C-terminal extension does not directly contact the DNA. Thus, folding and docking of the C-terminal extension is involved in a network of allosteric communication that links the binding sites for Hox and DNA and modulates the function of PBX.

We recently used ^{15}N CPMG relaxation dispersion experiments to characterize millisecond time scale motions of PBX-HD.[23] They showed that the protein transiently adopts a conformation similar to the DNA bound conformation,[23] even in the absence of DNA. This points to a Monod−Wyman−Changeux (MWC)-like allosteric mechanism, in which folding and docking of the C-terminal extension stabilizes a form of PBX-HD with higher affinity for both DNA and Hox proteins. In the present study, the combination of NMR techniques described above clearly indicates that at least two different dynamical processes occur in PBX-HD on the millisecond time scale. Notably, the second dynamical process is almost completely obscured when CPMG data are analyzed by standard methods. The results are consistent with PBX-HD exchanging among three conformational states, a ground state in which the C-terminal region is largely unstructured, the previously identified four-helix form that is poised to bind ligands, and an additional state in which the C-terminal extension is likely misfolded against the rest of the protein. Misfolding is a fundamental part of the protein folding energy landscape. These results suggest that it can also impinge on more local conformational transitions as well.

■ MATERIALS AND METHODS

Sample Preparation. We studied a version of PBX-HD running from residues 233−313 of the murine pre-B cell leukemia transcription factor.[19] PBX-HD is a member of the TALE (three amino acid loop extension) class of homeo-domains,[24] thus named because the helix 1−2 loop contains a three-residue insertion not present in most homeodomains. We used a consensus numbering scheme running from 1 to 78 in which the three residues of the TALE are designated 23a,b,c. In this convention, the consensus homeodomain sequence is comprised of residues 1−59, while residues 60−78 include the C-terminal extension.

Uniformly ^{15}N- and ^{2}H-labeled protein was expressed in *Escherichia coli* BL21 (DE3) cells harboring the Rosetta plasmid (Novagen), in 95% ^{2}H$_2$O-minimal media with ^{1}H-glucose.[25] We estimate that the final level of deuteration at nonexchangeable positions was about 75%. The protein was purified as described previously[23] using SP-sepharose ion-exchange and size-exclusion chromatography. The NMR sample contained 400 μM [^{15}N,^{2}H]-enriched protein, in a buffer containing 10% ^{2}H$_2$O, 1 mM EDTA, 10 μM DSS, and 20 mM 2-(4-morpholino)ethane-sulfonic acid (MES) at pH 6.0.

Oligodeoxynucleotides with the sequences 5′-GCGCATGA-TTGCCC-3′ and 5′-GGGCAATCATGCGC-3′ were purchased from Alpha DNA (Montreal, Canada) and purified by Q-sepharose ion-exchange chromatography. The DNA duplex was annealed by slow cooling from 95 to 25 °C in a buffer containing 20 mM sodium phosphate, 150 mM NaCl, and 1 mM EDTA at pH 7.5. A molar equivalent of ^{15}N-enriched PBX-HD was added, and the complex was dialyzed into a buffer containing 10% ^{2}H$_2$O, 1 mM EDTA, 10 μM DSS, and 20 mM sodium phosphate at pH 6.0.

NMR Spectroscopy. NMR experiments were performed at 18.8 T and 25 °C on a Varian INOVA spectrometer equipped with a high-sensitivity cold probe. Additional ^{15}N CPMG experiments were performed on a Varian INOVA spectrometer at 11.7 T.

^{15}N R_1 relaxation rates[26] were obtained with relaxation delays between 10 and 63 ms. {^{1}H}−^{15}N ssNOEs[26] were determined from peak intensities in spectra obtained with/without 5 s proton presaturation, with a total delay of 12 s between scans (including presaturation). The total acquisition times were 12 and 68 h for the nitrogen R_1 and {^{1}H}−^{15}N ssNOE experiments, respectively. The pulse schemes of Hansen et al[13] were used to measure four ^{1}H/^{15}N relaxation rates, $R_{1\rho}(2H'_zN_z)$, $R_{1\rho}(2H_zN'_z)$, $R_{1\rho_2}(2H'_zN'_z)$, and $R_1(2H_zN_z)$, where the primes indicate spin-locked nuclei. Relaxation delays between 1 and 50 ms were used for $R_{1\rho}(2H'_zN_z)$, $R_{1\rho}(2H_zN'_z)$, and $R_{1\rho_2}(2H'_zN'_z)$ experiments, while delays between 5 and 330 ms were used for $R_1(2H_zN_z)$. ^{1}H and ^{15}N continuous wave (CW) spin-lock field strengths of 12.3 and 2.0 kHz were employed, respectively. Total acquisition times were 9.25 h for $R_1(2H_zN_z)$ (1.16 h/2D data set) and 9 h for $R_{1\rho}(2H'_zN_z)$, $R_{1\rho}(2H_zN'_z)$, and $R_{1\rho}^2(2H'_zN'_z)$ (1.12 h/2D data set).

^{1}H-decoupled ^{15}N CPMG experiments[14] with power compensation schemes[4] were performed with a 40 ms relaxation delay, ν_{CPMG} ranging from 25 to 1000 Hz, and ^{1}H decoupling fields of 13−15 kHz. Transverse relaxation rates, R_2, and their associated uncertainties, σ_{R2}, were calculated as described previously.[4]

Spin Relaxation Analysis. NMR spectra were processed, and peak intensities were determined using the NmrPipe/NmrDraw software package.[27] Spin relaxation rates were determined by nonlinear least-squares fitting of peak intensities to the exponential decay function

$$I(T_{\text{relax}}) = I_0 \exp(-RT_{\text{relax}}) \tag{1}$$

where T_{relax} is the delay time, I_0 is the initial peak intensity, and R is the relaxation rate. Uncertainties in relaxation rates were estimated from the deviations between experimental peak intensities and their back-calculated values.[28]

Exchange-free transverse relaxation rates, R_{dd}, were obtained from the linear combination of the four relaxation rates according to the expressions[13]

$$R_{\Sigma,\rho} = 0.5[R_{1\rho}(2H'_z N_z) + R_{1\rho}(2H_z N'_z) - R_{1\rho^2}(2H'_z N'_z) - R_1(2H_z N_z)] \tag{2}$$

and

$$R_{dd} \approx \frac{R_{\Sigma,\rho}}{\sin^2 \theta_H \sin^2 \theta_N} \tag{3}$$

θ_X is the angle between the effective magnetic field in the rotating frame and the z-axis, $\tan \theta_X = \omega_{SL,X}/\Omega_X$, where $\omega_{SL,X}$ is the spin-locked field strength and Ω_X is the offset of the nuclear resonance frequency from the RF carrier for either ^{15}N or ^{1}H.

Values of the spectral density function, $J(\omega)$, were calculated at $\omega = 0$, ω_N, and $0.87\omega_H$, where ω_N and ω_H are the ^{15}N and ^{1}H Larmor frequencies, respectively. In all cases, the high-frequency approximation, in which $J(0.921\omega_H)$ and $J(0.955\omega_H)$ are both equivalent to $J(0.870\omega_H)$, was used, according to the following set of equations[3,4]

$$\text{NOE} = 1 + \left(\frac{d^2}{4}\right)\left(\frac{\gamma_H}{\gamma_N}\right)[5J(0.870\omega_H)]R_1 \tag{4}$$

$$R_1 = \left(\frac{d^2}{4}\right)[3J(\omega_N) + 7J(0.870\omega_H)] + c^2 J(\omega_N) \tag{5}$$

$$R_{dd} = \left(\frac{d^2}{8}\right)[4J(0) - 3J(\omega_N) + 0.845J(0.870\omega_H)] \tag{6}$$

where $d = [\mu_0 h\gamma_H\gamma_N/8\pi^2]\langle r_{NH}^{-3}\rangle$, $c = (\omega_N/\sqrt{3})(\sigma_\parallel - \sigma_\perp)$, μ_0 is the permeability of free space, γ_H and γ_N are the respective gyromagnetic ratios, ω_H and ω_N are the respective nuclear Larmor frequencies of ^{15}N and ^{1}H, h is Planck's constant, r_{NH} is the length of the amide bond, and $\sigma_\parallel$ and $\sigma_\perp$ are the parallel and perpendicular components of the axially symmetric ^{15}N chemical shift tensor.[29]

Exchange-free ^{15}N $R_2{}^0$ relaxation rates were back-calculated using the values of the spectral density function extracted above, according to the expression

$$R_2{}^0 = \left(\frac{d^2}{8}\right)[4J(0) + 3J(\omega_N) + 13J(0.870\omega_H)] + \left(\frac{c^2}{6}\right)[4J(0) + 3J(\omega_N)] \tag{7}$$

where c and d have the same meaning as in eqs 4−6. The largest source of error in this calculation is deviations in the ^{15}N CSA tensor from its mean values. However, assuming an uncertainty in the magnitude of the CSA, Δc, of ± 13 ppm,[30] the contribution to the uncertainty in back-calculated $R_2{}^0$ values is only about 0.3 s^{-1}.

CPMG Two-State Data Analysis. ^{15}N CPMG relaxation dispersion profiles were analyzed according to a two-state exchange model

$$A \underset{k_{BA}}{\overset{k_{AB}}{\rightleftharpoons}} B \tag{8}$$

In the fast exchange limit, relaxation dispersion profiles are given by the expression[31,32]

$$R_2 = R_2{}^0 + \frac{\Phi_{ex}}{k_{ex}}\left(1 - \frac{4\nu_{\text{CPMG}}}{k_{ex}}\tanh\left(\frac{k_{ex}}{4\nu_{\text{CPMG}}}\right)\right) \tag{9}$$

p_A and p_B are the relative populations of states A and B, $k_{ex} = k_{AB} + k_{BA}$, $\Phi_{ex} = p_A p_B \Delta\omega^2$, $\Delta\omega$ is difference between ^{15}N resonance frequencies in the two states (rad s^{-1}), $\nu_{\text{CPMG}} = 1/(2\tau_{cp})$, where τ_{cp} is the delay between successive refocusing pulses in the CPMG pulse train, and $R_2{}^0$ is the population-weighted average intrinsic (broadening-free) transverse relaxation rate, $R_2{}^0 = p_A R_{2A}{}^0 + p_B R_{2B}{}^0$. Data obtained at 11.7 and 18.8 T were analyzed together for each residue in all analyses.

A global two-state CPMG data analysis was performed as described previously.[5,6] Dispersion profiles for all residues were simultaneously fitted to yield a single exchange rate ($k_{ex}{}^{global}$), while Φ_{ex} and intrinsic relaxation rates at 11.7 and 18.8 T were obtained on a per-residue basis. This was accomplished by fixing k_{ex} at values between 100 and 1000 s^{-1} in increments of 10 s^{-1} and adjusting all other parameters to minimize the χ^2 function

$$\chi^2(k_{ex}) = \sum \frac{(R_2{}^{exp} - R_2{}^{calc}(k_{ex}))^2}{\sigma_{R2}{}^2} \tag{10}$$

$R_2{}^{calc}(k_{ex})$ was computed using eq 9 and the sum runs over all R_2 data points for a single residue at both magnetic field strengths. Aggregate χ^2 versus k_{ex} profiles were then calculated according to

$$\chi_{\text{total}}{}^2(k_{ex}) = \sum \chi^2(k_{ex}) \tag{11}$$

where the sum runs over data for all selected residues. The global value of k_{ex} was then selected such that

$$\chi_{\text{total}}{}^2(k_{ex}{}^{global}) = \min(\chi_{\text{total}}{}^2(k_{ex})) \tag{12}$$

Uncertainties in $k_{ex}{}^{global}$ values were estimated using a Monte Carlo approach in which randomized $\chi_{\text{total}}{}^2(k_{ex})$ profiles were calculated by summing $\chi^2(k_{ex})$ profiles for subsets of residues selected randomly according to a bootstrap procedure.[33] The standard deviations of the k_{ex} estimates thus obtained were taken as the uncertainties in global k_{ex} values.

CPMG Three-State Data Analysis. NMR relaxation dispersion data were fit to a linear three-state kinetic model

$$C \underset{k_{-2}}{\overset{k_2}{\leftrightarrows}} A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} B \tag{13}$$

In the fast exchange limit, the corresponding relaxation dispersion profiles are given by the expressions[8]

$$R_2 = R_2{}^0 + \frac{\varphi_{\text{fast}}}{\kappa_{\text{fast}}}\left(1 - \frac{4\nu_{\text{CPMG}}}{\kappa_{\text{fast}}}\tanh\left[\frac{\kappa_{\text{fast}}}{4\nu_{\text{CPMG}}}\right]\right)$$
$$+ \frac{\varphi_{\text{slow}}}{\kappa_{\text{slow}}}\left(1 - \frac{4\nu_{\text{CPMG}}}{\kappa_{\text{slow}}}\tanh\left[\frac{\kappa_{\text{slow}}}{4\nu_{\text{CPMG}}}\right]\right)$$

$$\varphi_{\text{fast}} = \frac{(-\kappa_{\text{slow}}\alpha_1 + \alpha_2)}{Z} \qquad \varphi_{\text{slow}} = \frac{(\kappa_{\text{fast}}\alpha_1 - \alpha_2)}{Z}$$

$$\kappa_{\text{fast}} = \frac{(k_{\text{ex}} + Z)}{2} \qquad \kappa_{\text{slow}} = \frac{(k_{\text{ex}} - Z)}{2}$$

$$k_{\text{ex}} = k_1 + k_{-1} + k_2 + k_{-2}$$

$$Z = (k_{\text{ex}}{}^2 - 4B)^{1/2}$$

$$B = k_{-1}k_{-2} + k_1k_{-2} + k_{-1}k_2$$

$$\alpha_1 = p_A p_B \Delta\omega_{\text{AB}}{}^2 + p_B p_C \Delta\omega_{\text{BC}}{}^2 + p_A p_C \Delta\omega_{\text{AC}}{}^2$$

$$\alpha_2 = p_A(k_1\Delta\omega_{\text{AB}}{}^2 + k_2\Delta\omega_{\text{AC}}{}^2)$$

$$p_A = \frac{k_{-1}k_{-2}}{B}$$

$$p_B = \frac{k_1 k_{-2}}{B}$$

$$p_C = \frac{k_{-1}k_2}{B}$$

(14)

p_A, p_B, and p_C are the relative populations of states A, B, and C, and $\Delta\omega_{\text{XY}}$ is the difference between ^{15}N resonance frequencies in states X and Y. Global exchange rates, κ_{fast} and κ_{slow}, and per-residue φ_{fast} and φ_{slow} parameters were calculated by performing a grid search in κ_{fast} and κ_{slow} for all residues with dispersion magnitudes greater than 1 s^{-1}. Data for one residue, Q64, were not in the fast time scale regime and were not included in any subsequent analysis. κ_{fast} and κ_{slow} were arrayed over a range of 5000–75 000 and 100–10 000 s^{-1}, respectively. The values of $R_2{}^0$ for 18.8 T data were fixed to the spectral-density-derived values obtained using eq 7. Values of φ_{fast}, φ_{slow}, and $R_2{}^0$(11.7 T) were adjusted on a per-residue basis to minimize the χ^2 function

$$\chi^2(\kappa_{\text{fast}}, \kappa_{\text{slow}}) = \sum \frac{(R_2{}^{\text{exp}} - R_2{}^{\text{calc}}(\kappa_{\text{fast}}, \kappa_{\text{slow}}))^2}{\sigma_{R2}{}^2} \quad (15)$$

$R_2{}^{\text{calc}}$ was computed using eq 14, and the sum runs over all R_2 data points for a single residue at both magnetic field strengths. Aggregate χ^2 surfaces were then calculated according to

$$\chi_{\text{total}}{}^2(\kappa_{\text{fast}}, \kappa_{\text{slow}}) = \sum \chi^2(\kappa_{\text{fast}}, \kappa_{\text{slow}}) \quad (16)$$

where the sum runs over data for all selected residues. Globally optimized values for κ_{fast} and κ_{slow} were then chosen such that

$$\chi_{\text{total}}{}^2(\kappa_{\text{fast}}{}^{\text{global}}, \kappa_{\text{slow}}{}^{\text{global}}) = \min(\chi_{\text{total}}{}^2(\kappa_{\text{fast}}, \kappa_{\text{slow}})) \quad (17)$$

Uncertainties in global κ_{fast} and κ_{slow} values were estimated using a Monte Carlo approach in which randomized $\chi_{\text{total}}{}^2(\kappa_{\text{fast}}, \kappa_{\text{slow}})$ surfaces were calculated by summing $\chi^2(\kappa_{\text{fast}}, \kappa_{\text{slow}})$ profiles for subsets of residues selected randomly according to a bootstrap procedure.[33] The standard deviations

of the κ_{fast} and κ_{slow} estimates thus obtained were taken as uncertainties in the respective values.

Values of φ_{fast} and φ_{slow} were then obtained for all residues by minimizing eq 15, fixing κ_{fast} and κ_{slow} to their globally optimized values. Uncertainties in φ_{fast} and φ_{slow} were obtained by repeating the analysis using the Monte Carlo simulations described above. The standard deviations of the resultant sets of φ_{fast} and φ_{slow} estimates were taken as the uncertainties in their respective values. One-dimensional χ^2 versus φ_{fast} profiles were calculated with (i) fixed and (ii) varied $R_2{}^0$(18.8 T) values as follows: Equation 15 was minimized with κ_{fast} and κ_{slow} fixed equal to 9500 and 2400 s^{-1}, respectively, and φ_{fast} was incremented between 0 and 0.5 ppm^2 while (i) fixing $R_2{}^0$(18.8 T) = LOSMQ-$R_2{}^0$ and optimizing φ_{slow} and $R_2{}^0$(11.7 T) and (ii) optimizing φ_{slow}, $R_2{}^0$(11.7 T), and $R_2{}^0$(18.8 T).

■ RESULTS AND DISCUSSION

Here, we present a two-step approach for extracting quantitative kinetic, thermodynamic, and chemical shift parameters from CPMG data for systems undergoing rapid three-state conformational exchange.

In the first step, we use spin relaxation experiments together with spectral density mapping[34] to estimate ^{15}N $R_2{}^0$ values that are free from dynamical broadening contributions. These back-calculated $R_2{}^0$ values are subsequently employed as constraints in the analysis of ^{15}N CPMG relaxation dispersion data with a three-state kinetic model.

In the second step, we use a novel grid search selection procedure to extract the full kinetic description of exchange (k_1, k_{-1}, k_2, and k_{-2} in eq 13) from the CPMG dynamics parameters. As inputs for this grid search selection, we generated criteria based on an observed correlation between dynamical parameters extracted in step 1 (φ_{slow} in eq 14) and the changes in PBX-HD ^{15}N chemical shifts produced by binding to DNA.

Spin Relaxation Experiments. We performed steady-state {^{1}H}–^{15}N ssNOE and ^{15}N R_1 measurements[26] together with a recently developed suite of NMR experiments that measure the relaxation rates of four different forms of ^{1}H/^{15}N magnetization,[13] ^{1}H/^{15}N longitudinal two-spin order, spin-locked antiphase ^{15}N single-quantum coherence, spin-locked antiphase ^{1}H single-quantum coherence, and spin-locked ^{1}H/^{15}N multiple-quantum coherence. In what follows, these will be referred to as LOSMQ (longitudinal order, single and multiple quantum) experiments. These four rates can be used to compute ^{15}N transverse relaxation rates, R_{dd}, that are solely dependent on ^{1}H/^{15}N dipole–dipole interactions and contain no contribution from conformational exchange broadening. Note that in general, R_{dd} is similar but not equal to R_2, as is evident from a comparison of eqs 6 and 7. In particular, R_2, but not R_{dd}, contains contributions from chemical shift anisotropy (CSA) and conformational exchange.

Spectral Density Mapping. The spectral density function, $J(\omega)$, is related to the "amount" of motion affecting a nucleus at a frequency ω. The spin relaxation data described above (ssNOE, R_1, R_{dd}) were used to compute $J(0)$, $J(\omega_N)$, and $J(0.87\omega_H)$ for each backbone amide group using a reduced spectral density mapping approach.[34] These parameters give information on fast motions in PBX-HD and are free of exchange broadening contributions. In general, higher flexibility on the nanosecond to picosecond time scales leads to elevated values of $J(\omega_N)$ and $J(0.87\omega_H)$ and reduced values of $J(0)$. Spectral density values are plotted in Figure 1. Small $J(0)$ and
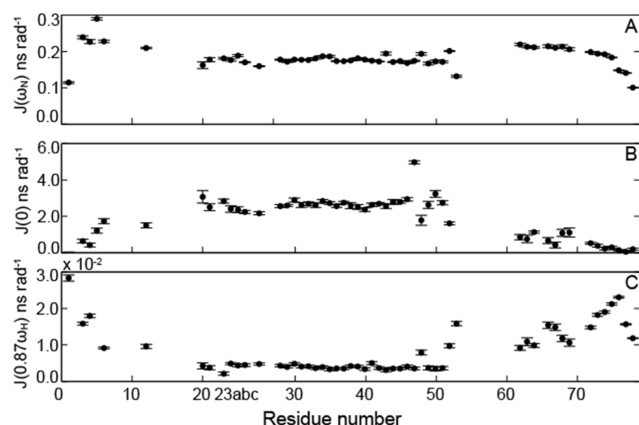
Figure 1. Values of the spectral density function, $J(\omega)$, for backbone amide groups of PBX-HD plotted as a function of residue number at frequencies ω = (A) $\omega_N = 5.1 \times 10^8$ rad s^{-1}, (B) 0, and (C) $0.87\omega_H = 4.4 \times 10^9$ rad s^{-1}, where ω_H and ω_N are the Larmor frequencies of ^{1}H and ^{15}N at 18.8 T, respectively.

large $J(0.87\omega_H)$ values, indicative of flexibility, are observed for the C-terminal extension (residues G61−A78) and in the N-terminal region (residues A1−K10). Elevated $J(0.87\omega_H)$ in residues K52 and R53 in the C-terminus of helix 3 may indicate some fraying of this helix as well. These findings are largely consistent with those of previous studies. The C-terminal extension is largely unstructured, as indicated by C^α and C^β secondary chemical shifts,[16,35] although it does experience transient excursions to a more ordered form.[23] The flexible N-terminal region is a conserved feature of homeodomains that inserts into the minor groove when in complex with DNA.[15] Backbone relaxation studies of other homeodomains have also revealed elevated dynamics in the C-terminal portion of helix 3.[36,37]

CPMG Experiments. ^{15}N CPMG relaxation dispersion experiments[14] were performed on PBX-HD at 25 °C at two spectrometer field strengths, 11.7 and 18.8 T.[38] The experiments yielded ^{15}N R_2 relaxation rates that varied as a function of the repetition rate of refocusing pulses (ν_{CPMG}) applied during the CPMG delay period (Figure 2). The magnitudes of the resultant relaxation dispersion profiles varied as approximately the square of the spectrometer field, indicating that exchange is fast on the NMR chemical shift time scale.[39] The data were initially analyzed using a two-state model for fast time scale exchange (eq 9). These fits yielded, on a per-residue basis, the exchange rate k_{ex} and the parameter Φ_{ex}, which depends on populations and chemical shift differences, $\Phi_{ex} = p_A p_B \Delta\omega_{AB}^2$. In addition, the asymptotic transverse relaxation rates, R_2^0, for $\nu_{CPMG} = \infty$ at 11.7 and 18.8 T were extracted. A two-state global analysis[5,6] was then performed, yielding a single value of k_{ex} as well as separate values of Φ_{ex}, R_2^0(11.7 T), and R_2^0(18.8 T) for each residue. The global analysis provided excellent agreement with the experimental data, and group-fit residual χ^2 values are only slightly (1.3-fold) higher than those obtained from individual fits. Back-calculated R_2 rates are in good agreement with experimental values (Figure 1). The extracted value of k_{ex}, 2600 ± 200 s^{-1}, agrees well with the value reported previously for ordering of the C-terminal region of PBX-HD, 2300 ± 200 s^{-1}.[23] A small discrepancy between the two numbers is expected as the current study employs a partially deuterated protein sample, while the previous study used fully protonated molecules. Deuteration modulates the strength of van der Waal's interactions and consequently can alter the populations and exchange rates of interconverting states.[9,10] Despite the good agreement obtained with the two-state model, we find that that the exchange process involves more than two states, as discussed below.

Evidence for Multisite Exchange from Cross-Validation of R_2^0. For systems in the fast exchange regime, relaxation dispersion profiles asymptotically approach exchange-free,
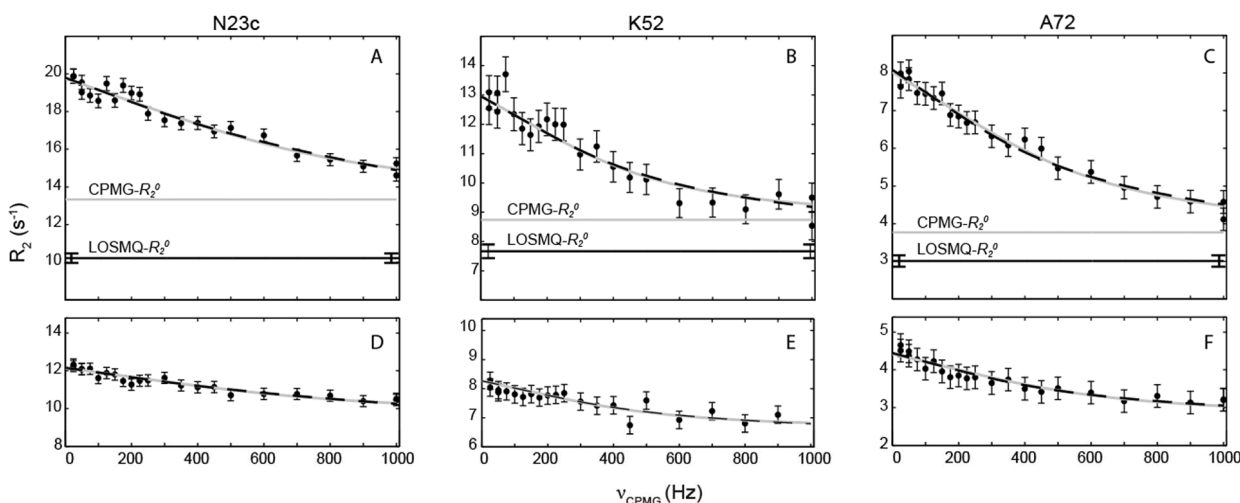
Figure 2. ^{15}N CPMG relaxation dispersion NMR data for three representative residues located in the Hox binding site (N23c), DNA binding site (K52), and C-terminal extension (A72) of PBX-HD measured at (A−C) 18.8 and (D−F) 11.7 T. The three residues of the TALE insert are designated 23a,b,c in the consensus numbering scheme. R_2 is the ^{15}N transverse relaxation rate, and $\nu_{CPMG} = 1/(2\tau)$, where τ is the delay between successive refocusing pulses in the CPMG pulse train. The horizontal gray lines (CPMG-R_2^0) correspond to the asymptotic values of R_2 ($\nu_{CPMG} = \infty$), extracted as adjustable parameters from global fits of CPMG data using a two-state exchange model, eq 8. The horizontal black lines (LOSMQ-R_2^0) correspond to exchange-free ^{15}N R_2 values obtained from LOSMQ relaxation data using eq 7. Gray curves correspond to global fits of CPMG data using a two-state model, eq 8. Black dashed curves correspond to global fits using a linear three-state exchange model, eq 13, with the asymptotic R_2^0(18.8 T) fixed equal to LOSMQ-R_2^0.

population-weighted average transverse relaxation rates (R_2^0) with increasing ν_{CPMG}.[8,31] In typical CPMG analyses, as above, values of R_2^0 are varied as adjustable parameters in the fits. However, independent estimates of R_2^0 can be calculated using the experimental spectral densities $J(0)$, $J(\omega_N)$, and $J(0.87\omega_H)$ according to eq 7. In what follows, these independent estimates of asymptotic transverse relaxation rates will be referred to as LOSMQ-R_2^0 because they are closely related to R_{dd} values and depend on the LOSMQ suite of experiments. LOSMQ-R_2^0 values are free of conformational exchange contributions. In contrast, R_2^0 estimates extracted from CPMG analyses can still contain contributions from microsecond time scale motions that remain incompletely quenched at the highest ν_{CPMG} field strengths used. Thus, when CPMG-derived R_2^0 values lie above those obtained by LOSMQ, the implication is that the molecule undergoes additional dynamics that are not included in the model used to fit the CPMG data. In principle, LOSMQ-R_2^0 values should not exceed those extracted from CPMG dispersion profiles, although, in practice, experimental errors and uncertainties in predicting CSA contributions to relaxation rates can lead to small discrepancies in this direction.

LOSMQ-R_2^0 rates were calculated for PBX-HD using 18.8 T spin relaxation data. As expected, almost all LOSMQ-R_2^0 values lie at or below CPMG-R_2^0, with only 3 out of 32 residues showing LOSMQ-R_2^0 marginally ($\sim$0.4 s^{-1}) greater than CPMG-R_2^0. For many residues in PBX-HD, the LOSMQ-R_2^0 values are significantly less than those extracted from the two-state CPMG analysis, as illustrated in Figure 2. Notably, many of the affected residues lie in the C-terminal extension, in helix 3, and in the helix 1−2 loop, which we previously found to undergo concerted motions implicated in allosteric interactions with DNA and Hox homeodomains.[23] This implies that these residues experience both millisecond time scale dynamics and more rapid microsecond time scale motions that are incompletely quenched at the maximum CPMG pulse repetition rates employed in this study (ν_{CPMG} = 1000 Hz).

An important caveat is that this comparison is only quantitative when CPMG pulse sequences that measure in-phase ^{15}N transverse relaxation rates, such as rczz[40] or ^{1}H-decoupled[14] schemes are used. In contrast, relaxation rates obtained using traditional relaxation-compensated CPMG schemes[41] are a mixture of in-phase and antiphase ^{15}N transverse relaxation rates and cannot be compared directly to LOSMQ-derived values. Also note that alternative methods can be used to obtain independent R_2^0 estimates, such as the ^{1}H−^{15}N dipole/^{15}N CSA relaxation interference approach developed by Palmer and co-workers.[42]

Three-State Analysis of CPMG Data. In order to better understand the additional microsecond time scale dynamics uncovered by the R_2^0 cross-validation, we analyzed the CPMG data using a global model with three exchanging states, C ↔ A ↔ B (eq 13), that contains microscopic rate constants k_1, k_{-1}, k_2, and k_{-2} that correspond to the A → B, B → A, A → C, and C → A transitions, respectively. A key feature of this analysis is that CPMG-R_2^0(18.8 T) asymptotic relaxation rates were fixed equal to the LOSMQ-R_2^0 values obtained from the spectral density analysis. As discussed below, this constraint was necessary for extracting the full set of three-state exchange parameters. While it is possible that this linear three-site model is an oversimplification of the true exchange process, the fact that it provides excellent agreement with the experimental results implies that more complicated models with any

additional adjustable parameters would be incompletely defined by these data.

When all exchange processes are in the fast regime, as is the case here, relaxation dispersion profiles produced by the linear three-state model are mathematically equivalent to the sum of two two-state dispersion curves, with the parameters κ_{fast}, κ_{slow} and φ_{fast}, φ_{slow} in eq 14 taking the place of k_{ex} and Φ_{ex} in eq 9.[8,43] Note that the macroscopic κ and φ parameters are fairly complex combinations of all microscopic rate constants, $k_{1,-1,2,-2}$, and thus do not necessarily refer to physically distinct processes, as discussed below. We performed a simultaneous analysis of the CPMG data set, such that all residues share the same global values of κ_{fast} and κ_{slow}. The parameters φ_{fast} and φ_{slow} depend on the individual chemical shift differences between the states and are thus distinct for each residue. The parameter values were obtained by performing a grid search of κ_{fast} and κ_{slow}, fixing R_2^0(18.8 T) to the LOSMQ-derived value, and optimizing φ_{fast}, φ_{slow}, and R_2^0(11.7 T) on a per-residue basis at each point on the grid. The resultant χ^2 surface is plotted in Figure 3. A distinct minimum was obtained at κ_{fast} = 9500 ± 2300 s^{-1} and κ_{slow} = 2400 ± 200 s^{-1}. The dispersion profiles back-calculated using these values are in excellent agreement with the experimental data (Figure 2). Thus a linear
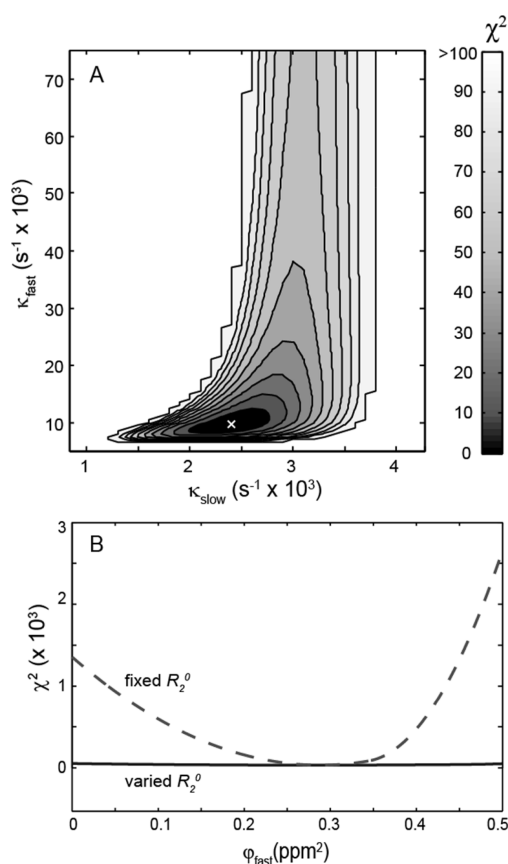

Figure 3. (A) Residual χ^2 surface for global fits of a linear three-state exchange model (eq 14) to CPMG data, employing a two-dimensional grid search in κ_{fast} and κ_{slow}. The surface was normalized relative to the minimum χ^2 value (indicated by the white x), which was assigned a value of 0. (B) χ^2 profile for fits of a linear three-state exchange model to CPMG data for N23c employing a one-dimensional search in φ_{fast} with fixed values of κ_{fast} = 9500 s^{-1} and κ_{slow} = 2400 s^{-1} and R_2^0(18.8 T) either fixed equal to LOSMQ-R_2^0 (dashed curve) or varied as an adjustable parameter in the fit (solid black curve).

three-state exchange model accounts for all CPMG data and LOSMQ-R_2^0 values, while a two-state exchange model does not.

In addition to the kinetic parameters κ_{slow} and κ_{fast}, the global analysis yields two exchange parameters, φ_{fast} and φ_{slow}, for each residue. These are functions of chemical shift differences among the three states and thus contain structural information on the exchange process. We compared the CPMG-derived values of φ_{fast} and φ_{slow} with the squared differences in the ^{15}N chemical shift between the free and DNA-bound forms of PBX-HD, $\Delta\delta^2$, determined from HSQC spectra of the two forms of the protein (Figure 4A,B). For residues in the C-terminal extension
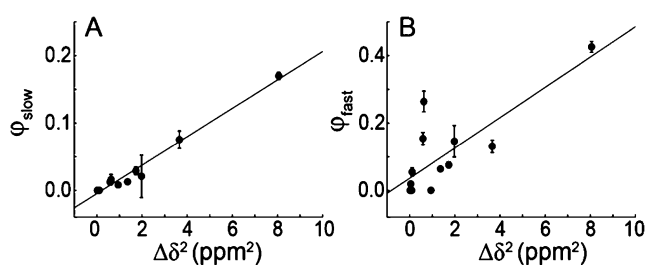
Figure 4. Dynamical parameters extracted from an analysis of CPMG data using a linear three-state exchange model, (A) φ_{slow} and (B) φ_{fast}, plotted as a function of squared changes in the ^{15}N chemical shift produced by DNA binding, $\Delta\delta^2$, for residues in the C-terminal extension of PBX-HD (G61−A78).

(residues 60−78), the level of agreement between φ_{slow} and $\Delta\delta^2$ is quite high, with a correlation coefficient of $r = 0.99$ and a slope of $m_\varphi = 0.0211$. In contrast, the correlation between φ_{fast} and $\Delta\delta^2$ is much lower for these residues, with a correlation coefficient of $r = 0.58$. This is consistent with our previous finding that conformational exchange in PBX-HD involves a weakly populated state that has similar chemical shifts to those of the DNA-bound form of the protein.[23] In the following section, we demonstrate that, based on this correlation, it is possible to relate the phenomenological κ and φ parameters to the microscopic rate (k) and chemical shift ($\Delta\omega$) parameters that describe the actual physical exchange process.

We investigated whether the three-state analysis depends on fixing R_2^0(18.8 T) to the LOSMQ-derived values. We repeated the grid search analysis, allowing R_2^0(18.8 T) to vary as an adjustable parameter in the fits. Interestingly, we obtained a very similar χ^2 minimum at $\kappa_{slow} = 1600 \pm 300$ s^{-1} and $\kappa_{fast} = 6000 \pm 1600$ s^{-1}. In other words, extraction of κ parameters does not rely on fixing R_2^0(18.8 T) to the LOSMQ-derived R_2^0 value, for this data set. However, the situation for φ_{fast} is quite different, as illustrated by the dependence of χ^2 on φ_{fast} in Figure 3B. When R_2^0(18.8 T) is fixed equal to the LOSMQ-derived value, a clear χ^2 minimum is obtained, indicating that the value of the parameter is well-defined. In contrast, when the value of R_2^0(18.8 T) is allowed to vary, the χ^2 profile is flat, indicating that the value of φ_{fast} is essentially undefined. In other words, the LOSMQ-R_2^0 constraint is required for the extraction of the full set of microscopic rate constants needed in the dynamical analysis below.

Elucidation of the Physical Exchange Process. CPMG and spin relaxation data for PBX-HD can be quantitatively explained by millisecond−microsecond time scale interconversion among three states. However the κ and φ parameters extracted in this analysis do not, in themselves, yield much physical insight into the nature of the exchange process. In what

follows, we draw a distinction between the CPMG-derived macroscopic model, which is described by the parameters κ_{slow}, κ_{fast}, φ_{slow}, and φ_{fast}, and what we will refer to as "microscopic models", which are described by the rate constants k_1, k_{-1}, k_2, k_{-2}, (and hence p_A, p_B, and p_C), as well as the chemical shift differences $\Delta\omega_{AB}$, $\Delta\omega_{AC}$, and $\Delta\omega_{BC}$. In general, a microscopic model is of more value as it provides an explicit description of the conformational dynamics. In our usage, a "microscopic model" refers to a unique set of values for the k, p, and $\Delta\omega$ parameters. The challenge lies in the fact that a great many different microscopic models can be consistent with the same macroscopic exchange parameters. Here, we show that with the correlation observed between CPMG-derived φ_{slow} values and $\Delta\delta^2$, the chemical shift changes associated with DNA binding can be used to differentiate among the many possible microscopic models and define the physical exchange process with a high degree of precision.

We tackled this problem by first identifying sets of k_1, k_{-1}, k_2, and k_{-2} values that are compatible with the experimentally determined values of κ_{fast} and κ_{slow}. We then applied restrictions based on the φ_{slow} versus $\Delta\delta^2$ correlation in order to narrow down the possibilities to a single set of microscopic models. We found it convenient to frame this procedure in terms of a two-parameter optimization of k_2 and k_{-2}. If values of k_2 and k_{-2} are selected and κ_{fast} and κ_{slow} are experimentally determined, the values of k_1 and k_{-1} and the populations of the states are uniquely specified by the expressions

$$k_{-1} = \frac{B + k_{-2}(k_2 + k_{-2} - k_{ex})}{k_2}$$

$$k_1 = k_{ex} - k_{-1} - k_2 - k_{-2}$$

$$k_{ex} = \kappa_{fast} + \kappa_{slow} = k_1 + k_{-1} + k_2 + k_{-2}$$

$$B = \frac{1}{4}(k_{ex}^2 - Z^2) = k_{-1}k_{-2} + k_1k_{-2} + k_{-1}k_2$$

$$Z = \kappa_{fast} - \kappa_{slow} = \sqrt{k_{ex}^2 - 4B}$$

$$p_A = \frac{k_{-1}k_{-2}}{B}$$

$$p_B = \frac{k_1k_{-2}}{B}$$

$$p_C = \frac{k_{-1}k_2}{B} \tag{18}$$

which follow directly from eq 14. Furthermore, the requirement that all rate constants take positive values means that many (k_2,k_{-2}) values are incompatible with a given pair of κ_{fast} and κ_{slow}. The set of all compatible (k_2,k_{-2}) values, and hence all possible microscopic models, associated with a κ_{fast}, κ_{slow} pair can be visualized as a bounded region in k_2k_{-2} space, as illustrated in Figure 5. The area associated with our experimental κ_{fast}, κ_{slow} values is quite large; it includes a wide variety of different physical descriptions, in some cases with populations distributed roughly equally among the states and in other cases with populations that are skewed toward A, B, or C. Below, we use the correlation between φ_{slow} and $\Delta\delta^2$ to substantially reduce the solution space for possible microscopic models.
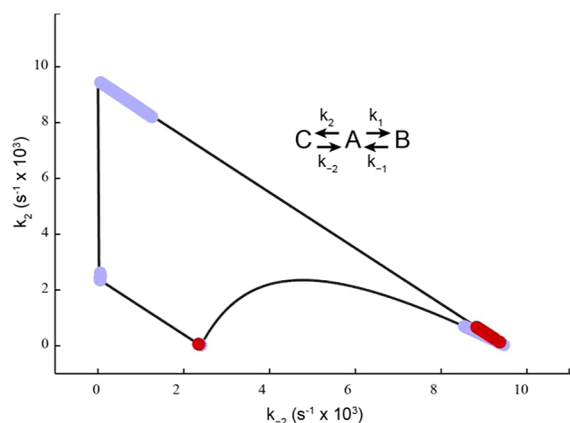
Figure 5. Transition rate constants consistent with the three-state CPMG-derived values of $\kappa_{fast} = 9500$ s^{-1} and $\kappa_{slow} = 2400$ s^{-1}, mapped as a function of k_2 and k_{-2}. The values of k_1 and k_{-1} are uniquely determined by those of κ_{fast}, κ_{slow}, k_2 and k_{-2}, according to eq 18. In the absence of additional information, the entire bounded region represents possible microscopic exchange models. Light blue points indicate the solutions consistent with constraints based on the correlation between φ_{slow} and $\Delta\delta^2$. Dark red points indicate that the solutions are additionally consistent with reasonable limits for differences in chemical shift between exchanging states, as described in the text.

In order to proceed, we recast the expressions for φ_{fast} and φ_{slow} explicitly in terms of the chemical shift differences between the states ($\Delta\omega_{AB}$, $\Delta\omega_{AC}$, $\Delta\omega_{BC}$) as follows

$$\varphi_{fast} = \lambda_{f,AB}\Delta\omega_{AB}^2 + \lambda_{f,AC}\Delta\omega_{AC}^2 + \lambda_{f,BC}\Delta\omega_{BC}^2$$

$$\lambda_{f,AB} = \frac{(p_A k_1 - \kappa_{slow}p_A p_B)}{Z}$$

$$\lambda_{f,AC} = \frac{(p_A k_2 - \kappa_{slow}p_A p_C)}{Z}$$

$$\lambda_{f,AB} = \frac{(-\kappa_{slow}p_B p_C)}{Z}$$

$$\varphi_{slow} = \lambda_{s,AB}\Delta\omega_{AB}^2 + \lambda_{s,AC}\Delta\omega_{AC}^2 + \lambda_{s,BC}\Delta\omega_{BC}^2$$

$$\lambda_{s,AB} = \frac{(\kappa_{fast}p_A p_B - p_A k_1)}{Z}$$

$$\lambda_{s,AC} = \frac{(\kappa_{fast}p_A p_C - p_A k_2)}{Z}$$

$$\lambda_{s,BC} = \frac{(\kappa_{fast}p_B p_C)}{Z} \tag{19}$$

The φ parameters are linear combinations of the squared chemical shift differences between the three states, with coefficients, λ, that depend on the exchange rates and populations. Any microscopic model, that is, any point in $k_2 k_{-2}$ space, is associated with a unique set of λ values. This connection provides the basis for narrowing down the $k_2 k_{-2}$ solution space based on the extrinsic chemical shift information.

The linear relationship observed between φ_{slow} and the DNA-induced changes in chemical shifts, $\Delta\delta^2$, implies that both parameters largely correlate with the same chemical shift difference parameter, either $\Delta\omega_{AB}^2$, $\Delta\omega_{AC}^2$, or $\Delta\omega_{BC}^2$. It follows that the chemical shifts observed for residues in the C-

terminal extension of the DNA-bound form of the protein correspond to those of either conformational state A, B, or C in the absence of DNA. For the sake of the discussion below, we will suppose that B is the DNA-bound-like state, although similar arguments can be equally well applied if either A or C is bound-like (see below). In this example, the chemical shifts in the bound state are then given by $\delta_{bound} = \omega_B$, while those in the free state are given by $\delta_{free} = p_A\omega_A + p_B\omega_B + p_C\omega_C$. Accordingly, $\Delta\delta^2$ is given by the expression

$$\Delta\delta^2 = p_A^2\Delta\omega_{AB}^2 + p_C^2\Delta\omega_{BC}^2 - 2p_A p_C\Delta\omega_{AB}\Delta\omega_{BC} \tag{20}$$

In order for $\Delta\delta^2$ to correlate with a chemical shift parameter, either $p_A \gg p_C$ and $\Delta\delta^2 \approx p_A^2\Delta\omega_{AB}^2$ or $p_C \gg p_A$ and $\Delta\delta^2 \approx p_C^2\Delta\omega_{BC}^2$. In order for φ_{slow} to correlate with a chemical shift parameter, either $\varphi_{slow} \approx \lambda_{s,AB}\Delta\omega_{AB}^2$ and $(\lambda_{s,AC},\lambda_{s,BC}) \approx 0$, or $\varphi_{slow} \approx \lambda_{s,BC}\Delta\omega_{BC}^2$ and $(\lambda_{s,AB},\lambda_{s,AC}) \approx 0$, or $\varphi_{slow} \approx \lambda_{s,AC}\Delta\omega_{AC}^2$ and $(\lambda_{s,AB},\lambda_{s,BC}) \approx 0$. Note that in this example, only the first two cases could yield φ_{slow} versus $\Delta\delta^2$ correlations. In addition, the slope of the φ_{slow} versus $\Delta\delta^2$ plot (m_φ) provides an additional constraint for identifying physical solutions consistent with the CPMG data. The slope is approximately equal to $\lambda_{s,AB}/p_A^2$ when $(\Delta\delta^2,\varphi_{slow}) \propto \Delta\omega_{AB}^2$, or it is approximately equal to $\lambda_{s,AB}/p_C^2$ when $(\Delta\delta^2,\varphi_{slow}) \propto \Delta\omega_{BC}^2$.

Following this logic, we performed a grid search in $k_2 k_{-2}$ space, evaluating whether each point was consistent with the φ_{slow} versus $\Delta\delta^2$ correlation. For each (k_2,k_{-2}) pair, we calculated the corresponding λ and p values and categorized the solution as consistent or inconsistent, according to the criteria above. Out of 1.4×10^8 points in the grid search, 3.1×10^7 were associated with all positive rate constants; only 7.0×10^3 of these were considered to be consistent with the φ_{slow} versus $\Delta\delta^2$ correlation and are plotted in Figure 5.

Although this procedure significantly reduced the available region in $k_2 k_{-2}$ space, a number of qualitatively different microscopic models remained after this analysis. In order to discriminate among these possibilities, we applied another set of constraints, based on realistic magnitudes for the $\Delta\omega$ parameters. We estimated that it is quite unlikely that any $\Delta\omega$ exceeds 5 ppm (particularly because $\Delta\delta$ values are all less than 3 ppm). Values of both φ_{slow} and φ_{fast} are on the order of 0.1 ppm^2. Thus the nonzero λ_s parameters should be greater than about 0.004. For example, if $(\Delta\delta^2,\varphi_{slow}) \propto \Delta\omega_{AB}^2$, then $(\lambda_{s,AC},\lambda_{s,BC}) \approx 0$, $\lambda_{s,AB} \approx \varphi_{slow}/\Delta\omega_{AB}^2$, and $\lambda_{s,AB}$ should be larger than about 0.004. Similarly, in order to account for the relatively large values of φ_{fast} that we observe, we applied the condition that $(\lambda_{f,AB} + \lambda_{f,AC} + \lambda_{f,BC}) > 0.004$. Combining these arguments with those above yields the six criteria listed in Table 1. All points in $k_2 k_{-2}$ space that meet any one of these criteria are considered to be consistent with the φ_{slow} versus $\Delta\delta^2$ correlation and physically reasonable. The number of accepted microscopic models identified in the grid search is 388. Notably, all of these models are qualitatively the same (Table 2). In every case, the distribution of molecules is heavily skewed toward the middle state, A, with a population of more than 90%. State A undergoes exchange on the order of 2000 s^{-1} with the DNA-bound-like state, which is populated to about 2%, and ultrafast exchange on the order of 9000 s^{-1} with a state populated to about 5%. Both symmetry-related microscopic models are extracted from the analysis (class I and class II in Table 2), in which either C or B is the DNA-bound-like state. The variability of the extracted microscopic models is quite low, as reflected in the small standard deviations reported in Table

10324

dx.doi.org/10.1021/jp306127m | J. Phys. Chem. B 2012, 116, 10317–10329

Table 1. Criteria for Grid Search Selection

i. $(\omega_A \approx \delta_{\text{bound}})\ (p_B \ll p_A)^a$ and $((\lambda_{s,AB}, \lambda_{s,BC}) \approx 0)^b$ and $(\lambda_{s,AC} \approx p_C^2 \times m_\varphi)^c$ and $(\lambda_{s,AC} > 0.004)$ and $(\lambda_{f,AB} + \lambda_{f,AC} + \lambda_{f,BC}) > 0.004)$

ii. $(\omega_A \approx \delta_{\text{bound}})\ (p_C \ll p_A)^a$ and $((\lambda_{s,AC}, \lambda_{s,BC}) \approx 0)^b$ and $(\lambda_{s,AB} \approx p_B^2 \times m_\varphi)^c$ and $(\lambda_{s,AB} > 0.004)$ and $(\lambda_{f,AB} + \lambda_{f,AC} + \lambda_{f,BC}) > 0.004)$

iii. $(\omega_B \approx \delta_{\text{bound}})\ (p_A \ll p_B)^a$ and $((\lambda_{s,AB}, \lambda_{s,AC}) \approx 0)^b$ and $(\lambda_{s,BC} \approx p_C^2 \times m_\varphi)^c$ and $(\lambda_{s,BC} > 0.004)$ and $(\lambda_{f,AB} + \lambda_{f,AC} + \lambda_{f,BC}) > 0.004)$

iv. $(\omega_B \approx \delta_{\text{bound}})\ (p_C \ll p_A)^a$ and $((\lambda_{s,AC}, \lambda_{s,BC}) \approx 0)^b$ and $(\lambda_{s,AB} \approx p_A^2 \times m_\varphi)^c$ and $(\lambda_{s,AB} > 0.004)$ and $(\lambda_{f,AB} + \lambda_{f,AC} + \lambda_{f,BC}) > 0.004)$

v. $(\omega_C \approx \delta_{\text{bound}})\ (p_A \ll p_C)^a$ and $((\lambda_{s,AB}, \lambda_{s,AC}) \approx 0)^b$ and $(\lambda_{s,BC} \approx p_B^2 \times m_\varphi)^c$ and $(\lambda_{s,BC} > 0.004)$ and $(\lambda_{f,AB} + \lambda_{f,AC} + \lambda_{f,BC}) > 0.004)$

vi $(\omega_C \approx \delta_{\text{bound}})\ (p_B \ll p_C)^a$ and $((\lambda_{s,AB}, \lambda_{s,BC}) \approx 0)^b$ and $(\lambda_{s,AC} \approx p_A^2 \times m_\varphi)^c$ and $(\lambda_{s,AC} > 0.004)$ and $(\lambda_{f,AB} + \lambda_{f,AC} + \lambda_{f,BC}) > 0.004)$

^{a}Considered true if the smaller value of p is less than 0.1. ^{b}Considered true if $-2 \times \Delta_m < \lambda < 2 \times \Delta_m$, where Δ_m is the standard error of m_φ; $2 \times \Delta_m = 0.1 m_\varphi = 0.0021$. ^{c}Considered true if $p^2 \times m_\varphi - 2 \times \Delta_m < \lambda < p^2 \times m_\varphi + 2 \times \Delta_m$.

**Table 2. Microscopic Exchange Model Parameters**

	class I^a		class II^b	
	mean	stdev	mean	stdev
p_A	0.93	0.02	0.94	0.2
p_B	0.022	0.001	0.04	0.02
p_C	0.047	0.015	0.022	0.001
k_1	55	4	400	200
k_{-1}	2348	3	9100	200
k_2	450	150	55	4
k_{-2}	9040	150	2348	3
$\lambda_{f,AB}$	0.00035	0.0001	0.04	0.02
$\lambda_{f,AC}$	0.04	0.01	0.0003	0.0001
$\lambda_{f,BC}$	−0.00035	0.0001	−0.0003	0.0001
$\lambda_{s,AB}$	0.020	0.001	−0.0011	0.0005
$\lambda_{s,AC}$	−0.0013	0.0004	0.020	0.001
$\lambda_{s,BC}$	0.0014	0.0004	0.0012	0.0005

^{a}Microscopic solutions in which state B resembles the DNA-bound form of the protein ($\omega_B \approx \delta_{\text{bound}}$) ^{b}Microscopic solutions in which state C resembles the DNA-bound form of the protein ($\omega_C \approx \delta_{\text{bound}}$)

2. Thus, the criteria above are sufficient to produce a precise description of the physical exchange process. These solutions represent special cases of three-state exchange in which the macroscopic κ and φ parameters correspond directly to the microscopic exchange processes. For class I processes, B is the DNA-bound-like state, $\kappa_{\text{fast}} \approx k_2 + k_{-2}$, $\varphi_{\text{fast}} \approx \lambda_{f,AC}\Delta\omega_{AC}^2$, $\kappa_{\text{slow}} \approx k_1 + k_{-1}$, and $\varphi_{\text{slow}} \approx \lambda_{f,AB}\Delta\omega_{AB}^2$. For class II solutions, the situation is essentially the same, with states B and C reversed. It must be emphasized that this simple physical interpretation of the κ and φ parameters is not true in general and, in this case, can only be made based on the results of the grid search selection procedure.

PBX-HD Undergoing Local Folding and Misfolding. We have found that PBX-HD experiences two separate exchange processes. A ground state that is populated to 93.2 ± 1.6% interconverts directly with a state populated to 2.2 ± 0.1% on a time scale of $\kappa_{\text{slow}} = 2400 \pm 200\ \text{s}^{-1}$. The squared chemical shift differences between the two states are proportional to φ_{slow} and are shown in Figure 6a. For residues in the C-terminal region of the protein, φ_{slow} correlates strongly with the squared chemical shift differences, $\Delta\delta^2$, between the free and DNA-bound forms of the protein. The rest of the homeodomain also senses this exchange process, but φ_{slow} does not correlate with $\Delta\delta^2$ for these residues. We therefore conclude that the $\kappa_{\text{slow}}, \varphi_{\text{slow}}$ process corresponds to transient
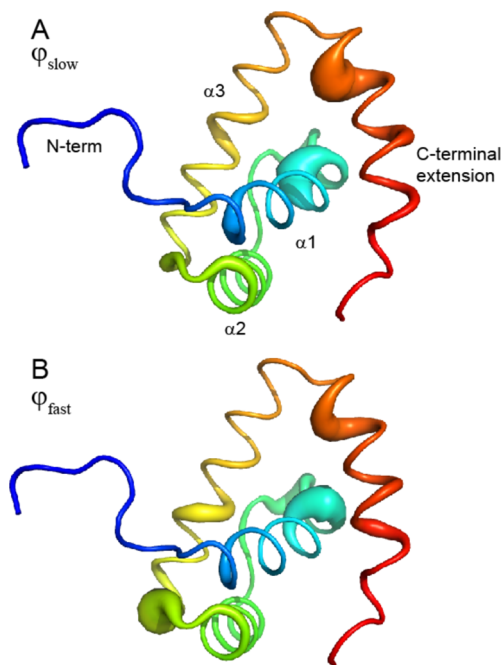
Figure 6. Ribbon diagrams of PBX-HD in the DNA-bound conformation (DNA not shown). The ribbon radius is proportional to CPMG-derived values of (A) φ_{slow}, which are proportional to squared chemical shift differences between the ground state of PBX-HD and the four-helix excited state, and (B) φ_{fast}, which are proportional to squared chemical shift differences between the ground state and the misfolded excited state. Ribbons are colored blue to red from the N- to C-termini.

excursions to a state that partially resembles the DNA-bound form of the protein, in which the C-terminal region of the protein is folded into a fourth α helix, as was concluded previously.[23] In what follows, we will refer to this as the "four-helix" state.

The ground state also exchanges directly with a state populated to 4.6 ± 1.6% on a time scale of $\kappa_{\text{fast}} = 9500 \pm 2300\ \text{s}^{-1}$. The squared chemical shift differences between the two states are proportional to φ_{fast}. Large values of φ_{fast} are obtained both in the in the C-terminal extension and in the rest of the homeodomain. This implies that the exchange process involves a concerted conformational change involving the entire molecule. Many residues with large φ_{slow} also exhibit large φ_{fast} values, suggesting that there are some similarities between the two excited states. However, there are some significant differences between the two sets of values as well, providing confidence that the fast and slow processes are distinct. For example, in the helix 2−3 loop, T41 senses the fast process ($\varphi_{\text{fast}} = 0.743$) but not the slower one ($\varphi_{\text{slow}} = 0.053$). Overall, the correlation between φ_{slow} and φ_{fast} is quite low ($r = 0.50$). Thus, the two excited states are related but not identical. Notably, the grid search procedure indicates that the major state lies between the two minor states. Therefore, the more rapid exchange process does not represent the formation of an on-pathway local folding intermediate because in this case, the minor state would be the central one. In addition, this third form of the protein does not lie on the far side of the four-helix state from the ground state. We therefore conclude that the $\kappa_{\text{fast}}, \varphi_{\text{fast}}$ process involves concerted conformational excursions to a form that is a kinetic dead-end, in the sense that the

protein must return to the ground state before it is able to adopt the four-helix state that is poised to bind DNA.

There are several potential explanations for the nature of this third state. It could correspond to the fully unfolded form of the protein. However, many residues in the C-terminal region produce large values of φ_{fast}. This part of the protein is already unstructured in the ground state[16] and would not be expected to strongly sense the unfolding of the rest of the homeodomain. Several residues in the rest of the homeodomain with large chemical shift differences between the folded and unfolded states do not exhibit large φ_{fast} values.[3] Furthermore, the extent of unfolding at 25 °C, as determined by DSC, is much lower than the population of 5% observed for this state.[23] Thus, it is unlikely that the fast dynamic process corresponds to protein unfolding. Alternatively, the fast process could correspond to self-association/dissociation dynamics on the microsecond time scale. In this case, dilution of the sample would be expected to lessen this contribution to transverse relaxation; residues with large φ_{fast} should exhibit a drop in R_2, while those with small φ_{fast} should exhibit little or no change in R_2. In fact, dilution of the sample by a factor of 2 produces very little effect. There is a uniform decrease in $R_2(\nu_{CPMG} = 1000)$ of about 0.40 s^{-1}, consistent with a small decrease in the effective rotational correlation time, as has been observed previously.[23] However, residues with large values of φ_{fast} are not preferentially affected. Upon dilution, the 10 residues with the largest values of φ_{fast} exhibit an average decrease in R_2 of 0.37 ± 0.1 s^{-1}, while those with the smallest values of φ_{fast} exhibit an average decrease in R_2 of 0.52 ± 0.1 s^{-1}. Therefore, it is unlikely that the $\kappa_{fast},\varphi_{fast}$ process is related to self-association. Instead, it is most probable that this process reflects transient interactions between the C-terminal extension and the rest of homeodomain that differ structurally, thermodynamically, and kinetically from those in the four-helix excited state or the DNA-bound state. These transitions are off-pathway from the local folding process and thus appear to be an example of local misfolding. Misfolding has long been recognized as a kinetic trap in the process of global protein folding.[44] The involvement of misfolding in local conformational transitions has been less well explored.

PBX-HD Allosteric Mechanism. Our previous investigation of PBX-HD dynamics employed CPMG relaxation dispersion methods but did not include LOSMQ-derived R_2^0 constraints.[23] Consequently, only the slower transition between the ground state and the four-helix excited state was detected. On the basis of these results, we proposed a MWC-type[23] allosteric model in which the weakly populated four-helix state has greater affinity for both DNA and the Hox homeodomain. Interacting with either ligand would shift the population from the ground state to the four-helix state, thereby increasing the tendency of PBX-HD to bind the second partner.[23] Here, we have found that the free form of the protein can adopt at least three different conformational states, the previously identified three-helix ground state and four-helix excited state, as well as an additional excited state in which the C-terminal extension is locally misfolded. The existence of locally misfolded proteins does not rule out our prior thermodynamic model; allostery can still be largely driven by a shift in population from three-helix to four-helix forms of the protein. However, there are implications for how allosteric structural transitions are encoded by the protein energy landscape. In most instances of classical allosteric theory, such as the MWC[45] or Koshland–Nemethy–Filmer (KNF)[46] models, protein molecules exist in two conformations, often referred to as the "tense" and "relaxed" states. This implies a tightly controlled energy landscape with a small number energy minima corresponding to discrete functional states. More recently, allostery has been recast in terms of folding and binding energy funnels.[47,48] These descriptions invoke the breadth of the protein conformational ensemble: "proteins are inherently dynamic and sample a vast ensemble of conformations",[47] that represent "the repertoire of molecules available for the binding event."[48] The presence of ligands alters the energy landscape, resulting in a redistribution of the populations of the conformational substates toward those with the highest binding affinities. This new view de-emphasizes the importance of distinct structural isomers and focuses on the statistical nature of the protein conformational ensemble.[49] Our results for PBX-HD are reminiscent of both the old and new paradigms. The free protein exchanges among three conformational states separated by significant energy barriers. One of the states resembles the ligand-bound form. Thus, classical theory involving transitions among discrete states provides a useful explanation for allostery in this molecule. Yet, the existence of the "locally misfolded" excited state raises important questions that lie beyond this description: To what extent are local folding and misfolding in PBX-HD driven by the same interactions, for example hydrophobic collapse? How are folding and misfolding of the C-terminal extension tied to conformational transitions in the rest of the homeodomain? These questions are best addressed within the framework of folding and binding funnels, in accordance with the new ensemble view of allostery.

Three-State Dynamical Parameters from CPMG Data. There have been a number of prior studies in which linear three-state models were used to analyze NMR CPMG relaxation dispersion data.[7−9,50] In most of these cases, exchange rates were in the intermediate time scale regime. On these time scales, the challenges and the required analytical approaches are different from those in the present work. In many examples, the existence of multisite exchange could be deduced from a clear failure of global two-state fitting to account for CPMG data from all residues.[12,50] The largest challenge in these cases was in the fitting of NMR relaxation dispersion profiles. Under intermediate exchange conditions, the data are governed by four global parameters, k_1, k_{-1}, k_2, and k_{-2}, and two residue-specific parameters, $\Delta\omega_{AB}$ and $\Delta\omega_{AC}$, which may take either positive or negative values ($\Delta\omega_{BC} = \Delta\omega_{AB} + \Delta\omega_{AC}$). As a consequence of this large number of adjustable parameters, the residual χ^2 surfaces for fitting dispersion profiles contain numerous local minima.[12] Large CPMG data sets are typically required in these studies. These have included measurements acquired at multiple temperatures,[50] pressures,[51] and ligand concentrations,[10] involving different types of nuclei,[52] and in some cases, they were derived from relaxation dispersion experiments directed at double-quantum,[53] zero-quantum,[53] multiple-quantum,[54] as well as single-quantum coherences.[13,14] However, once the fitting optimization was completed for these examples, the physical picture of exchange was well-defined as the microscopic rate constants and chemical shift differences were yielded directly by the analysis.

The situation is different when both exchange processes are in the fast regime. In the current example, a global two-state exchange model provides good agreement with all PBX-HD dispersion profiles, as long as asymptotic R_2^0 relaxation rates are adjusted in the fits. The existence of multistate exchange only becomes clearly evident when CPMG data are compared with

LOSMQ-derived R_2^0 values. Independent R_2^0 estimates have been used previously to extract three-state exchange parameters.[8] In particular, Palmer and co-workers combined $^1H/^{15}N$ η_{xy} CSA/dipolar cross-correlation-derived R_2^0 estimates[42] with rczz CPMG data.[40] In the current work, R_2^0 estimates are obtained from the LOSMQ suite of relaxation experiments and are combined with ^1H-decoupled CPMG data. These two approaches are complementary as the LOSMQ experiments can be applied in cases where the cross-correlation experiments are problematic.[13] In addition, a previous study combined LOSMQ-derived R_2^0 rates with CPMG data. However, in that example, a formal two-state model was used to describe the slower process, while the second process was assumed to be in the extreme fast limit with contributions to relaxation dispersion profiles that were invariant with ν_{CPMG}.[55] In the present case, the contribution from the faster process shows a sloped dispersion profile (Supporting Information Figure S1), and a full three-state fit must be used.

Under fast exchange conditions, dispersion profiles are governed by two global parameters, κ_{fast} and κ_{slow}, and two residue-specific parameters, φ_{fast} and φ_{slow}, which are both positive. The smaller number of parameters makes global fitting of fast exchange data more straightforward than analyses of intermediate exchange data. We have shown here that the global χ^2 optimum can be robustly determined using a grid search in κ_{fast} and κ_{slow}. However, the parameters yielded by fast exchange analyses do not translate directly into a unique physical description of the exchange process. In general, a given set of κ_{fast}, κ_{slow}, φ_{fast}, φ_{slow} can map onto a wide variety of different microscopic physical models. This problem has been previously addressed by altering the temperature to effectively "freeze" out one of the dynamical processes;[8] however, this method requires the collection of large variable-temperature data sets and would not be expected to work for all proteins.

Here, we present a complementary approach that allows the somewhat ambiguous parameters extracted from fits of fast exchange CPMG data to be rigorously interpreted in terms of microscopic physical models. In the first step, CPMG data are analyzed to yield the parameters (κ_{fast}, κ_{slow}, φ_{fast}, φ_{slow}). In the second step, we use a grid search selection procedure that exploits additional information to narrow down the set of possible physical descriptions of exchange (k_1, k_{-1}, k_2, k_{-2}, $\Delta\omega_{AB}^2$, $\Delta\omega_{AC}^2$, $\Delta\omega_{BC}^2$) that are consistent with the experimentally determined parameters. One advantage of this stepwise approach is that is reduces the problem to a two-dimensional search in (κ_{fast}, κ_{slow}) space, followed by another in (k_2, k_{-2}) space. Note that this approach does not seek to use additional information to improve the fits of raw relaxation dispersion data. Rather, it exploits this information to convert the parameters extracted from the fits (κ_{fast}, κ_{slow}, φ_{fast}, φ_{slow}) into a specific physical model. Accordingly, the additional information is not needed in the first step of fitting the dispersion profiles, and no adjustment of these fits is required after the second step of identifying a microscopic model.

In the current work, we used the observed correlation between φ_{slow} and the squared changes in the chemical shift produced by DNA binding, $\Delta\delta$, to generate selection criteria for the grid search. These were combined with reasonable limits on the maximum values of $\Delta\omega$ (i.e., minimum values of λ). In this case, the criteria were sufficient to identify essentially a single solution, in which the ground state of PBX-HD exchanges directly with two weakly populated excited states. Under these conditions, the macroscopic (κ_{fast}, φ_{fast}) and

(κ_{slow}, φ_{slow}) parameters correspond closely to the two physical exchange processes, as has been noted previously.[43]

In addition to the ones used here, a number of other selection criteria can be envisaged. One potential source is population data obtained from three-state fits of ITC binding data[56] or DSC folding data.[57] Criteria can also be generated if the chemical shifts of one of the states or chemical shift differences between any two states are known (see the Supporting Information). Chemical shifts of an exchanging unfolded state could be obtained from theoretical random-coil chemical shifts[58,59] or NMR melt experiments,[3] while those of an exchanging bound state could be obtained from titration experiments. Further criteria could be derived from additional kinetic data, for instance, those from stopped-flow measurements. The grid search selection procedure offers the versatility of combining criteria from a wide variety of sources. Its ability to deliver a unique solution will depend on the number and type of selection criteria available. We anticipate that even in cases where more than one possible solution remains, the grid search selection procedure will still significantly narrow down the options and improve our understanding of exchange processes.

■ CONCLUSIONS

We have developed a two-step approach for obtaining detailed descriptions of multistate conformational exchange from NMR relaxation data. In the first step, we use a recently developed suite of LOSMQ NMR relaxation experiments to calculate exchange-free ^{15}N R_2 transverse relaxation rates (LOSMQ-R_2^0). These values are combined with CPMG relaxation dispersion data, both as a cross-validation tool to detect multisite exchange and as a constraint to improve multistate fits of CPMG data. The second step of our approach involves a novel grid search selection strategy for elucidating physical exchange processes that are fast on the NMR time scale. In the fast regime, analyses of CPMG data yield macroscopic parameters that cannot be directly related to physical exchange rates and populations, a priori.[8] Our procedure exploits additional information to permit the extraction of microscopic exchange parameters from macroscopic CPMG-derived parameters.

We applied this approach to the PBX homeodomain, which we had previously shown to undergo millisecond time scale dynamics that are linked to allostery. Using the LOSMQ-R_2^0 cross-validation, we found that PBX-HD experiences motions on the microsecond time scale that were not obvious from the CPMG data alone. A global three-state analysis of CPMG data using the LOSMQ-R_2^0 values as constraints yielded well-defined exchange parameters. Exploiting a strong correlation between the extracted CPMG parameters and changes in chemical shifts produced by binding DNA, we elucidated a detailed physical model of three-state exchange in PBX-HD. We found that the major three-helix form of the protein interconverts directly with two excited states. One of these corresponds to a form of the protein that resembles the DNA-bound state, in which the C-terminal extension folds and packs against the rest of the homeodomain, forming a fourth α-helix. In the second excited state, the C-terminal extension transiently forms interactions with the rest of the homeodomain that differ from those formed in the four-helix form of the protein, that is, it is misfolded. Misfolding has long been recognized as a feature of the protein folding energy landscape. Our observation of local misfolding in PBX-HD supports the idea that the energy

landscapes of folding, binding, and allostery contain similar features.[48]

ASSOCIATED CONTENT

ⓈSupporting Information

Decomposition of a three-state CPMG relaxation dispersion curve into fast and slow components. Description of population- and chemical-shift-based criteria for the grid search selection procedure. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: anthony.mittermaier@mcgill.ca.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Mittermaier, A. K.; Kay, L. E. *Trends Biochem. Sci.* **2009**, *34*, 601−611.

(2) Palmer, A. G.; Kroenke, C. D.; Loria, J. P. *Methods Enzymol.* **2001**, *339*, 204−238.

(3) Farber, P.; Darmawan, H.; Sprules, T.; Mittermaier, A. *J. Am. Chem. Soc.* **2010**, *132*, 6214−6222.

(4) Demers, J. P.; Mittermaier, A. *J. Am. Chem. Soc.* **2009**, *131*, 4355−4367.

(5) Bruschweiler, S.; Schanda, P.; Kloiber, K.; Brutscher, B.; Kontaxis, G.; Konrat, R.; Tollinger, M. *J. Am. Chem. Soc.* **2009**, *131*, 3063−3068.

(6) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313*, 1638−1642.

(7) Korzhnev, D. M.; Kay, L. E. *Acc. Chem. Res.* **2008**, *41*, 442−451.

(8) Grey, M. J.; Wang, C.; Palmer, A. G., III. *J. Am. Chem. Soc.* **2003**, *125*, 14324−14335.

(9) Sugase, K.; Dyson, H. J.; Wright, P. E. *Nature* **2007**, *447*, 1021−1025.

(10) Tolkatchev, D.; Xu, P.; Ni, F. *J. Am. Chem. Soc.* **2003**, *125*, 12432−12442.

(11) Korzhnev, D. M.; Bezsonova, I.; Lee, S.; Chalikian, T. V.; Kay, L. E. *J. Mol. Biol.* **2009**, *386*, 391−405.

(12) Neudecker, P.; Korzhnev, D. M.; Kay, L. E. *J. Biomol. NMR* **2006**, *34*, 129−135.

(13) Hansen, D. F.; Yang, D.; Feng, H.; Zhou, Z.; Wiesner, S.; Bai, Y.; Kay, L. E. *J. Am. Chem. Soc.* **2007**, *129*, 11468−11479.

(14) Hansen, D. F.; Vallurupalli, P.; Kay, L. E. *J. Phys. Chem. B* **2008**, *112*, 5898−5904.

(15) Gehring, W. J.; Affolter, M.; Burglin, T. *Annu. Rev. Biochem.* **1994**, *63*, 487−526.

(16) Sprules, T.; Green, N.; Featherstone, M.; Gehring, K. *Biochemistry* **2000**, *39*, 9943−9950.

(17) Chang, C. P.; Shen, W. F.; Rozenfeld, S.; Lawrence, H. J.; Largman, C.; Cleary, M. L. *Genes Dev.* **1995**, *9*, 663−674.

(18) Sprules, T.; Green, N.; Featherstone, M.; Gehring, K. *J. Biol. Chem.* **2003**, *278*, 1053−1058.

(19) Green, N. C.; Rambaldi, I.; Teakles, J.; Featherstone, M. S. *J. Biol. Chem.* **1998**, *273*, 13273−13279.

(20) Knoepfler, P. S.; Kamps, M. P. *Mol. Cell. Biol.* **1995**, *15*, 5811−5819.

(21) Lu, Q.; Kamps, M. P. *Mol. Cell. Biol.* **1996**, *16*, 1632−1640.

(22) Peltenburg, L. T. C.; Murre, C. *Development* **1997**, *124*, 1089−1098.

(23) Farber, P. J.; Mittermaier, A. *J. Mol. Biol.* **2011**, *405*, 819−830.

(24) Burglin, T. R. *Nucleic Acids Res.* **1997**, *25*, 4173−4180.

(25) Marley, J.; Lu, M.; Bracken, C. *J. Biomol. NMR* **2001**, *20*, 71−75.

(26) Farrow, N. A.; Muhandiram, R.; Singer, A. U.; Pascal, S. M.; Kay, C. M.; Gish, G.; Shoelson, S. E.; Pawson, T.; Forman-Kay, J. D.; Kay, L. E. *Biochemistry* **1994**, *33*, 5984−6003.

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

(28) Mittermaier, A.; Kay, L. E. *Protein Sci.* **2004**, *13*, 1088−1099.

(29) Cornilescu, G.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 10143−10154.

(30) Kroenke, C. D.; Rance, M.; Palmer, A. G. *J. Am. Chem. Soc.* **1999**, *121*, 10119−10125.

(31) Luz, Z.; Meiboom, S. *J. Chem. Phys.* **1963**, *39*, 366−&.

(32) Allerhand, A.; Gutowsky, H. S.; Jonas, J.; Meinzer, R. A. *J. Am. Chem. Soc.* **1966**, *88*, 3185−3193.

(33) Efron, B.; Tibshirani, R. *Stat. Sci.* **1986**, *1*, 54−77.

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

(35) Jabet, C.; Gitti, R.; Summers, M. F.; Wolberger, C. *J. Mol. Biol.* **1999**, *291*, 521−530.

(36) Fausti, S.; Weiler, S.; Cuniberti, C.; Hwang, K. J.; No, K. T.; Gruschus, J. M.; Perico, A.; Nirenberg, M.; Ferretti, J. A. *Biochemistry* **2001**, *40*, 12004−12012.

(37) Gumral, D.; Nadalin, L.; Corazza, A.; Fogolari, F.; Damante, G.; Viglino, P.; Esposito, G. *FEBS J* **2008**, *275*, 435−448.

(38) Kovrigin, E. L.; Kempf, J. G.; Grey, M. J.; Loria, J. P. *J. Magn. Reson.* **2006**, *180*, 93−104.

(39) Millet, O.; Loria, J. P.; Kroenke, C. D.; Pons, M.; Palmer, A. G. *J. Am. Chem. Soc.* **2000**, *122*, 2867−2877.

(40) Wang, C.; Grey, M. J.; Palmer, A. G., III. *J. Biomol. NMR* **2001**, *21*, 361−366.

(41) Loria, J. P.; Rance, M.; Palmer, A. G. *J. Am. Chem. Soc.* **1999**, *121*, 2331−2332.

(42) Kroenke, C. D.; Loria, J. P.; Lee, L. K.; Rance, M.; Palmer, A. G. *J. Am. Chem. Soc.* **1998**, *120*, 7905−7915.

(43) O'Connell, N. E.; Grey, M. J.; Tang, Y.; Kosuri, P.; Miloushev, V. Z.; Raleigh, D. P.; Palmer, A. G., III. *J. Biomol. NMR* **2009**, *45*, 85−98.

(44) Dobson, C. M. *Nature* **2003**, *426*, 884−890.

(45) Monod, J.; Wyman, J.; Changeux, J. P. *J. Mol. Biol.* **1965**, *12*, 88−118.

(46) Koshland, D. E., Jr.; Nemethy, G.; Filmer, D. *Biochemistry* **1966**, *5*, 365−385.

(47) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol.* **2009**, *5*, 789−796.

(48) Kumar, S.; Ma, B.; Tsai, C. J.; Sinha, N.; Nussinov, R. *Protein Sci.* **2000**, *9*, 10−19.

(49) Gunasekaran, K.; Ma, B.; Nussinov, R. *Proteins: Struct., Funct., Bioinf.* **2004**, *57*, 433−443.

(50) Korzhnev, D. M.; Salvatella, X.; Vendruscolo, M.; Di Nardo, A. A.; Davidson, A. R.; Dobson, C. M.; Kay, L. E. *Nature* **2004**, *430*, 586−590.

(51) Korzhnev, D. M.; Bezsonova, I.; Evanics, F.; Taulier, N.; Zhou, Z.; Bai, Y.; Chalikian, T. V.; Prosser, R. S.; Kay, L. E. *J. Am. Chem. Soc.* **2006**, *128*, 5262−5269.

(52) Mittermaier, A.; Korzhnev, D. M.; Kay, L. E. *Biochemistry* **2005**, *44*, 15430−15436.

(53) Orekhov, V. Y.; Korzhnev, D. M.; Kay, L. E. *J. Am. Chem. Soc.* **2004**, *126*, 1886−1891.

(54) Korzhnev, D. M.; Kloiber, K.; Kay, L. E. *J. Am. Chem. Soc.* **2004**, *126*, 7320−7329.

(55) Hansen, D. F.; Feng, H. Q.; Zhou, Z.; Bai, Y. W.; Kay, L. E. *J. Am. Chem. Soc.* **2009**, *131*, 16257−16265.

(56) Cliff, M. J.; Williams, M. A.; Brooke-Smith, J.; Barford, D.; Ladbury, J. E. *J. Mol. Biol.* **2005**, *346*, 717−732.

(57) Carra, J. H.; Anderson, E. A.; Privalov, P. L. *Biochemistry* **1994**, *33*, 10842−10850.

(58) Schwarzinger, S.; Kroon, G. J. A.; Foss, T. R.; Chung, J.; Wright, P. E.; Dyson, H. J. *J. Am. Chem. Soc.* **2001**, *123*, 2970−2978.

(59) Schwarzinger, S.; Kroon, G. J. A.; Foss, T. R.; Wright, P. E.; Dyson, H. J. *J. Biomol. NMR* **2000**, *18*, 43−48.