J|A|C|S
ARTICLES
Published on Web 03/22/2007

Exploring Multiple Timescale Motions in Protein GB3 Using Accelerated Molecular Dynamics and NMR Spectroscopy

Phineus R. L. Markwick,* Guillaume Bouvignies, and Martin Blackledge*

Contribution from the Institute de Biologie Structurale Jean-Pierre Ebel, CNRS, CEA, UJF, UMR 5075, 41 rue Jules Horowitz, F-38027 Grenoble, Cedex, France.

Received December 7, 2006; E-mail: phineus@pasteur.fr; martin.blackledge@ibs.fr

Abstract: Biological function relies on the complex spectrum of conformational dynamics occurring in biomolecules. We have combined Accelerated Molecular Dynamics (AMD) with experimental results derived from NMR to probe multiple time-scale motions in the third IgG-binding domain of Protein G (GB3). AMD is shown to accurately reproduce the amplitude and distribution of slow motional modes characterized using residual dipolar couplings, reporting on dynamics up to the millisecond timescale. In agreement with experiment, larger amplitude slower motions are localized in the β-strand/loop motif spanning residues 14−24 and in loop 42−44. Principal component analysis shows these fluctuations participating in the primary mode, substantiating the existence of a correlated motion traversing the β-sheet that culminates in maximum excursions at the active site of the molecule. Fast dynamics were simulated using extensive standard MD simulations and compared to order parameters extracted from ^{15}N relaxation. Notably 60 2-ns fully solvated MD simulations exploring the different conformational substates sampled from AMD resulted in better reproduction of order parameters compared to the same number of simulations starting from the relaxed crystal structure. This illustrates the inherent dependence of protein dynamics on local conformational topology. The results provide rare insight into the complex hierarchy of dynamics present in GB3 and allow us to develop a model of the conformational landscape native to the protein, appearing as a steep sided potential well whose flat bottom comprises multiple similar but discrete conformational substates.

Introduction

Proteins require substantial conformational flexibility in order to correctly fold and to perform their native function.[1,2] Nuclear magnetic resonance (NMR) is exquisitely sensitive to protein dynamics on a vast range of timescales and therefore provides access to motional parameters governing the function of biological systems.[3] In particular spin relaxation measurements provide precise information about local dynamics on picosecond to nanosecond timescales.[4−6] Dynamics occurring on slower timescales are however of particular interest because many biologically important processes, such as enzymatic catalysis,[7] signal transduction,[8] ligand binding, and allosteric regulation[9] are expected to occur in this range.[10] Relaxation dispersion experiments have successfully extended the range of timescales to identify conformational exchange between states in proteins,[11] relating microsecond molecular motion to enzymatic function,[12] and resolving thermodynamic, and even structural, descriptions of very weakly populated excited states.[13,14]

Residual dipolar couplings (RDCs) have emerged as a powerful addition to these techniques.[15,16] The sensitivity of RDCs to the relative orientation of a pair of coupled spins has led to their routine use in protein structure determination.[17,18] However it is in terms of molecular dynamics that a second, equally important aspect of RDCs has recently become apparent.[19,20] RDCs represent averages over all orientations of the dipolar interaction sampled up to the timescale defined by the inverse of the measured dipolar coupling (up to the millisecond range) and thereby hold great promise for the description of slower motion in proteins. This sensitivity of RDCs to local conformational fluctuation has recently been exploited to examine the extent and nature of local dynamic behavior present in proteins over this time range.[21−28]

(1) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. *Science* **1991**, *254*, 1598−1603.
(2) Benkovic, S. J.; Hammes-Schiffer, S. *Science* **2003**, *301*, 1196−202.
(3) Mittermaier, A.; Kay, L. E. *Science* **2006**, *312*, 224−228.
(4) Kay, L. E.; Torchia, D. A.; Bax, A. *Biochemistry* **1989**, *28*, 8972−8979.
(5) Palmer, A. G. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 129−155.
(6) Bruschweiler, R. *Curr. Opin. Struct. Biol.* **2003**, *13*, 175−183.
(7) Tousignant, A.; Pelletier, J. N. *Chem. Biol.* **2004**, *11*, 1037−1042.
(8) Rousseau, F.; Schymkowitz, J. *Curr. Opin. Struct. Biol.* **2005**, *15*, 23−30.
(9) Kern, D.; Zuiderweg, E. R. P. *Curr. Opin. Struct. Biol.* **2003**, *13*, 748−757.
(10) Eisenmesser, E. Z.; Bosco, D. A.; Akke, M.; Kern, D. *Science* **2002**, *295*, 1520−1523.
(11) Mulder, F. A.; Mittermaier, A.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *Nat. Struct. Biol.* **2001**, *8*, 932−935.
(12) Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D. M.; Wolf-Watz, M.; Bosco, D. A.; Skalicky, J. J.; Kay, L. E.; Kern, D. *Nature* **2005**, *438*, 117−121.
(13) 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.
(14) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313*, 1638−1642.
(15) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1697−1697.
(16) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Nat. Struct. Biol.* **1997**, *4*, 292−297.
(17) Bax, A. *Protein Sci.* **2003**, *12*, 1−16.
(18) Prestegard, J. H.; Bougault, C. M.; Kishore, A. I. *Chem. Rev.* **2004**, *104*, 3519−3540.
(19) Prestegard, J. H.; al-Hashimi, H. M.; Tolman, J. R. *Q. Rev. Biophys.* **2000**, *33*, 371−424.
(20) Blackledge, M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *46*, 23−61.

10.1021/ja0687668 CCC: $37.00

We have recently investigated the presence of slow motions in protein GB3, a small globular protein of 56 amino acids, by interpreting an extensive set of RDCs using the three-dimensional Gaussian Axial Fluctuation model (3D-GAF),[29] a description of local peptide plane dynamics defined in terms of diffusive motions around three orthogonal axes attached to the peptide plane. A heterogeneous distribution of slower motions was found in the protein, including a network of correlated motions in the β-sheet of the protein, propagated across the sheet via the hydrogen-bonding network.[30] The directionality of motions found in the interaction site of GB3 was proposed to coincide with the conformational adjustment required for molecular recognition.[31] This study delivered a site-specific anisotropic motional description of each peptide plane in the protein, providing an estimate of the direction and extent of motions present on the peptide backbone. The aim of the current work is to combine state-of-the-art Molecular Dynamics techniques, designed to extend the timescale available to molecular simulation into the millisecond range, with the results derived from experimental NMR, to develop a molecular description of multiple timescale motions present in GB3.

MD simulation enables the exploration of conformational energy landscapes accessible to proteins[32,33] and therefore represents a powerful tool for understanding and interpreting experimental NMR data in terms of molecular motion.[6,34−37] While classical MD simulation can be employed to predict fast stochastic motions responsible for spin relaxation, more sophisticated methods are required to access the slower motions probed by RDCs. Considerable progress has been made toward the development of methods to sample the complete phase space of biomolecules, allowing functionally important motions to be sampled more efficiently and rapidly. In general, these methods can be divided into two groups: The first involves the identification of transition pathways between known initial and final states, such as two well-defined conformational states of a protein. Such methods include targeted molecular dynamics,[38] transition path sampling,[39] and essential molecular dynamic sampling.[40] The second contains those methods that efficiently sample low-energy molecular conformations, allowing the rapid identification of thermodynamically dominant regions on the potential energy surface. These methods include conformational flooding,[41] replica exchange MD,[42] umbrella sampling,[43] high-temperature MD, leap dynamics,[44] and a number of related techniques recently reviewed by Elber.[45] A key aspect common to many of these methods is to enhance the escape rate from one low-energy conformational state to another. One way of achieving this is to lower the free energy barriers separating the low-energy states on the rugged potential energy landscape, thus accelerating the occurrence of slower dynamic events by application of a biasing potential. This is the basis of an efficient biased-potential method, known as accelerated molecular dynamics (AMD).[46]

AMD has recently been successfully used to simulate peptidyl cis−trans isomerization within the flexible Gly rich flaps of HIV-1 protease on the microsecond to millisecond timescale.[47] In our case the aim was to use AMD to characterize the slower motions present in GB3 and to compare the results with motional parameters extracted from NMR measurements.

Results and Discussion

Accelerated MD Enhances Extent of Backbone Fluctuations. Accelerated Molecular Dynamics simulations extend the ability of MD to sample backbone conformations, producing a time-averaged distribution of conformational states whose relative populations depend on the parameters E_b and α, used to accelerate the molecular motion (see Methods).[46] These parameters therefore need to be adjusted to elicit the required amount of dynamics. We note that the parametrization used here was found to be robust and that more aggressive acceleration resulted in similar results to those shown (vide infra), with higher amplitude dynamics occurring in the more flexible parts of the molecule, until eventual unfolding occurred.

Figure 1 compares the backbone root-mean-square deviation (rmsd) fluctuation over a standard 10 ns MD simulation of GB3, to 1 of the 12 AMD simulations. Note that only slightly larger fluctuations with respect to the average structure are observed for the AMD simulation (0.73 ± 0.16)Å compared to (0.54 ± 0.09)Å for the standard MD. The average backbone coordinates of the AMD simulation compare very closely with those of the X-ray and RDC-refined X-ray structures (0.54 Å and 0.38 Å, respectively), indicating that although the accelerated MD trajectories sample broader conformational space, they appear to be distributed about a mean conformation that resembles the time- and ensemble-averaged experimentally determined structures.

Slow Molecular Motions from AMD. Orientational fluctuations of internuclear vectors sampled by AMD can be compared to experimentally determined order parameters ($S^2_{RDC,NH}$ and $S^2_{RDC,CC}$) extracted from a recent analysis of an extensive set

(21) Meiler, J.; Prompers, J. J.; Peti, W.; Griesinger, C.; Bruschweiler, R. *J. Am. Chem. Soc.* **2001**, *123*, 6098−6107.
(22) Peti, W.; Meiler, J.; Bruschweiler, R.; Griesinger, C. *J. Am. Chem. Soc.* **2002**, *124*, 5822−5833.
(23) Briggman, K. B.; Tolman, J. R. *J. Am. Chem. Soc.* **2003**, *125*, 10164−10165.
(24) Ulmer, T. S.; Ramirez, B. E.; Delaglio, F.; Bax, A. *J. Am. Chem. Soc.* **2003**, *125*, 9179−9191.
(25) Clore, G. M.; Schwieters, C. D. *J. Am. Chem. Soc.* **2004**, *126*, 2923−2938.
(26) Bernado, P.; Blackledge, M. *J. Am. Chem. Soc.* **2004**, *126*, 4907−4920.
(27) Bernado, P.; Blackledge, M. *J. Am. Chem. Soc.* **2004**, *126*, 7760−7761.
(28) Bouvignies, G.; Bernado, P.; Blackledge, M. *J. Magn. Reson.* **2005**, *173*, 323−338.
(29) Lienin, S. F.; Bremi, T.; Brutscher, B.; Bruschweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1998**, *120*, 9870−9879.
(30) Bouvignies, G.; Bernado, P.; Meier, S.; Cho, K.; Grzesiek, S.; Bruschweiler, R.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13885−13890.
(31) Derrick, J. P.; Wigley, D. B. *J. Mol. Biol.* **1994**, *243*, 906−918.
(32) Karplus, M.; McCammon, J. A. *Nat. Struct. Biol.* **2002**, *9*, 646−652.
(33) Karplus, M.; Kuriyan, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6679−6685.
(34) Levy, R. M.; Karplus, M.; Wolynes, P. G. *J. Am. Chem. Soc.* **1981**, *103*, 5998−6011.
(35) Bruschweiler, R.; Roux, B.; Blackledge, M.; Griesinger, C.; Karplus, M.; Ernst, R. R. *J. Am. Chem. Soc.* **1992**, *114*, 2289−2302.
(36) Case, D. A. *Acc. Chem. Res.* **2002**, *35*, 325−331.
(37) Chandrasekhar, I.; Clore, G. M.; Szabo, A.; Gronenborn, A. M.; Brooks, B. R. *J. Mol. Biol.* **1992**, *226*, 239−250.
(38) Schlitter, J.; Engels, M.; Kruger, P. *J. Mol. Graph.* **1994**, *12*, 84−89.
(39) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.
(40) Amadei, A.; Linssen, A. B.; Berendsen, H. J. *Proteins* **1993**, *17*, 412−425.
(41) Muller, E. M.; de Meijere, A.; Grubmuller, H. *J. Chem. Phys.* **2002**, *116*, 897−905.
(42) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *J. Chem. Phys.* **2003**, *118*, 6664−6675.
(43) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(44) Kleinjung, J.; Bayley, P.; Fraternali, F. *FEBS Lett.* **2000**, *470*, 257−262.
(45) Elber, R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 151−156.
(46) Hamelberg, D.; Mongan, J.; McCammon, J. A. *J. Chem. Phys.* **2004**, *120*, 11919−11929.
(47) Hamelberg, D.; McCammon, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 13778−13779.

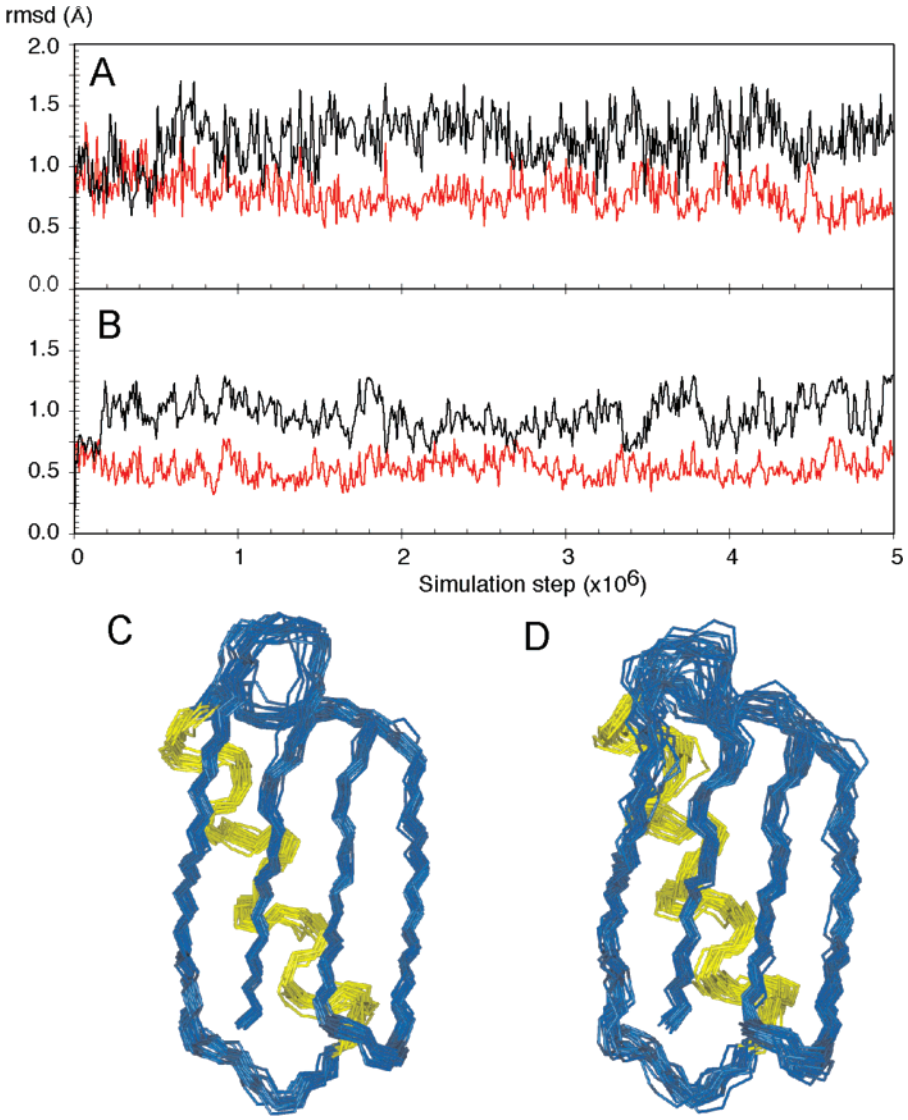

Figure 1. Backbone root-mean-square deviation (rmsd) fluctuation (A) over 1 of the 12 AMD simulations (5×10^6 steps with a time step of 2 fs) (B) compared to a standard 10 ns MD simulation. In both cases the rmsd relative to first (black) and average (red) structures are shown after removing overall translational and rotational motion. (C) Representative ensemble of MD sampling. (D) Representative ensemble of the AMD sampling.

of experimental dipolar couplings,[24,48] using the 3D-GAF description of peptide plane motion and relating to motions up to the millisecond timescale. Although the level of acceleration was tuned by adjusting the two parameters E_b and α to match the lowest experimentally determined values of $S^2_{RDC,NH}$, the correspondence seen throughout the chain between experimentally determined and simulated order parameters is quite striking (Figure 2). Higher amplitude slow motions between residues 15 and 20, comprising the first loop (I) and strand $\beta 2$, that were parametrized from the 3D GAF analysis are also present in the AMD ensemble. The fine structure found among sequential $S^2_{RDC,NH}$ and $S^2_{RDC,CC}$ values along the protein sequence are remarkably well reproduced from simulation, for example, residues 23−30 and 41−50. This information is also reported in terms of a colored ribbon representation in Figure 3a. For the three points G14, E26, and D52 where RDC-derived order parameters were found to be significantly higher (>0.05) than the relaxation-derived order parameters in the previous study, and are therefore considered nonphysical, the relaxation-derived parameter ($S^2_{LS,NH}$) is shown. These results clearly show the high correlation between simulated motional modes detected from the AMD and conformational disorder detected from experimental RDCs in GB3.

Fast Molecular Motions Sampled from Combined MD and AMD. NMR spin relaxation reports on a complex average of motions. While relaxation-active phenomena are limited to a time range up to and including the overall rotational diffusion of the protein (around 5ns for GB3), these motions are averaged over the ensemble of conformations sampled up to the millisecond range (the chemical shift coalescence limit). Sampling of fast motions in different substates is challenging using classical MD simulations, whose conformational range is likely to depend on the initial conformation. Although long timescale simulations, up to the microsecond range, can in principle extend the range of sampling, they still remain rare.[49] Because the RDC-derived order parameters shown here are well reproduced by

(48) Meier, S.; Haussinger, D.; Jensen, P.; Rogowski, M.; Grzesiek, S. *J. Am. Chem. Soc.* **2003**, *125*, 44−45.

(49) Nederveen, A. J.; Bonvin, A. M. J. J. *J. Chem. Theory Comp.* **2005**, *1*, 363−374.

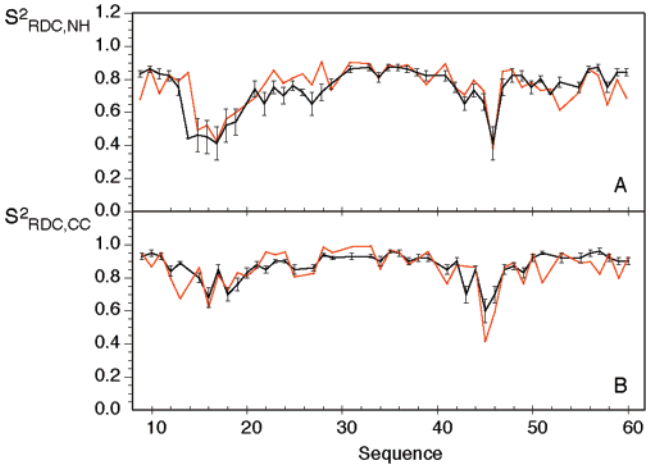

Figure 2. Slow motional order parameters extracted from RDCs and simulation. These order parameters report on fluctuations occurring up to the millisecond time range. (A) Comparison of order parameters ($S^2_{RDC,NH}$) calculated from the AMD ensemble (black) compared to values determined using 3D-GAF analysis of experimental RDCs (red). Uncertainty in the 3D-GAF order parameters was estimated at 0.05. For the three residues 14, 26, 52 for which the RDC-derived order parameters were found to be larger than the relaxation-derived values (a nonphysical result), the latter values were given. (B) $S^2_{RDC,CC}$ extracted from AMD ensemble (black) compared to values determined using 3D-GAF analysis of experimental RDCs (red). Uncertainty in the 3D-GAF order parameters was estimated at 0.03.

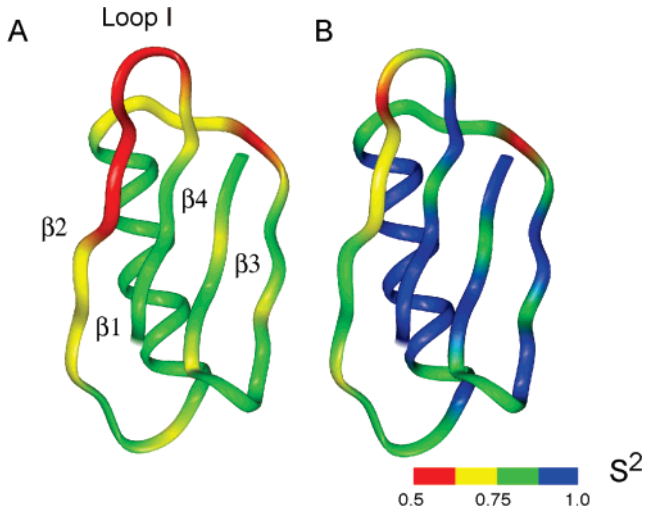

Figure 3. Fast and slow motion in GB3. (A) Ribbon diagram showing the distribution of $S^2_{RDC,NH}$ motional order parameters along the peptide chain of GB3 as determined from accelerated molecular dynamics. The color code refers to the $S^2_{RDC,NH}$ value. (B) Ribbon diagram showing the distribution of $S^2_{LS,NH}$ fast motional order parameters along the peptide chain of GB3 as determined from AMD. The color code refers to the $S^2_{LS,NH}$ value.

AMD, we can assume that appropriate conformational disorder representing time and ensemble averages up to the chemical shift limit is present in the ensemble. The increased statistical sampling of different substates explored by the AMD therefore allows us to study fast motions occurring within the diverse conformational states and provides rare insight into the complex relationship between slow structural fluctuations and rapid dynamics.

$^{15}N-^{1}H$ autorelaxation-derived order parameters ($S^2_{LS,NH}$) have been calculated from 60 separate 2-ns MD simulations, using starting coordinates extracted from different conformational states from the AMD simulation, and subsequently

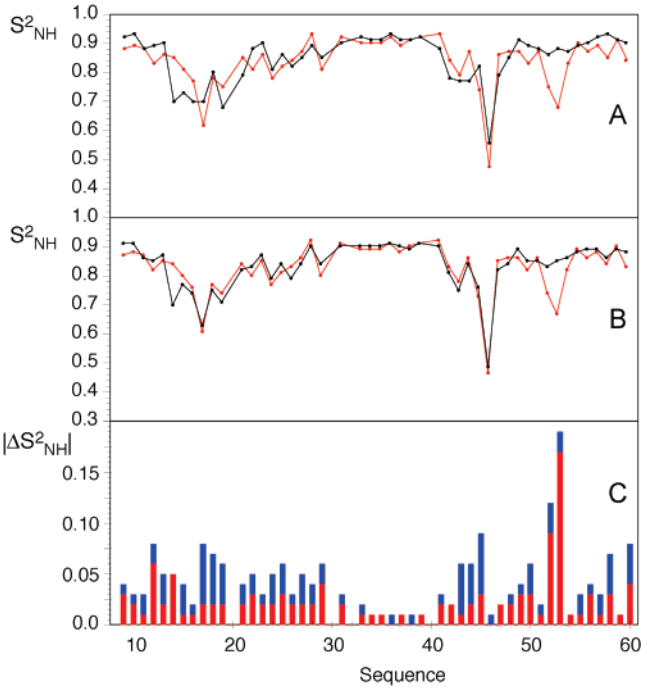

Figure 4. Lipari−Szabo order parameters[61] extracted from ^{15}N relaxation and simulation. (A) $S^2_{LS,NH}$ extracted from $^{15}N-^{1}H$ autorelaxation (red) compared to average values calculated from 60 2-ns MD simulations in explicit solvent starting from the relaxed high-resolution crystal structure (black). (B) $S^2_{LS,NH}$ extracted from $^{15}N-^{1}H$ autorelaxation (red) compared to average values calculated from 60 2-ns MD simulations in explicit solvent starting from structures extracted from the AMD ensemble (black). (C) Absolute values of differences between experimental and simulated $S^2_{LS,NH}$: (Blue) differences between $S^2_{LS,NH}$ values shown in part A and (red) differences between $S^2_{LS,NH}$ values shown in part B.

averaged. These averages are compared to experimental values[50] in Figure 4. Rare differences between experimental and simulated results are found for residues D52 and E53 in the final loop between strands β3 and β4, for which the larger amplitude motions found experimentally are not generated. We note that slow motions for these two sites are also not well reproduced by the AMD simulation.

In order to pinpoint the origin of this remarkable reproduction we have compared the results with $S^2_{LS,NH}$ predicted from 60 classical MD simulations from the same relaxed initial structure with different random seed generators (scenarios I and II in Methods). A clear improvement over the simulated values found in the classical MD case is visible, indicating that the increased phase space sampled by the AMD is responsible for the improved agreement with experiment. This is supported by analysis of each of the 60 2-ns trajectories (Table 1 in the Supporting Information). Remarkably the individual AMD-derived trajectory that best fits the experimental data has a mean $|\Delta S^2|$ per residue of 0.037 (the average value over the 60 trajectories is 0.042), and the averaged $|\Delta S^2|$ per residue from all 60 random seed generators, shown in Figure 3A, is 0.039, whereas the averaged $|\Delta S^2|$ per residue from the 60 AMD simulations, shown in Figure 3B, is 0.024. We note that the major contribution to the improvement is essentially localized in the loop regions of the molecule, where longer timescale dynamics are observed.

(50) Hall, J. B.; Fushman, D. *J. Biomol. NMR* **2003**, *27*, 261−275.

The transformation of experimental relaxation rates into dynamic parameters involves a number of assumptions concerning the nature of the relevant physical interactions.[36] In particular the $^{15}N-^{1}H$ internuclear distance and local chemical shift anisotropies are both assumed known and uniform along the protein backbone, while local motions are considered sufficiently independent of overall tumbling to warrant mathematical decoupling of the contributions to the spectral density functions. In light of these potential sources of error and the already significant experimental uncertainty, the similarity between experimental and simulated order parameters is remarkable. Although the prediction of order parameters from MD is in general accurate, the reason for problems encountered in reproducing local detail remains a subject of debate. One possible explanation is that the analytical force fields employed in contemporary classical MD simulations have not been sufficiently accurate. Indeed recent improvements in force fields have resulted in better reproduction of experimental data.[51] It has been suggested that inconsistencies between experimental and simulated order parameters also result from poor statistical mechanical sampling of the different conformational states of the protein that contribute to the relaxation measurement. Our results support this proposition, providing an example where more complete sampling of the ensemble leads to global improvement in the accuracy of fast dynamics. This observation provides evidence that rapid motional modes are dictated by the specific topography of local substates in the conformational landscape. More generally the close reproduction of both RDC- and relaxation-derived order parameters suggests that the substates sampled in the AMD simulations are realistic, encoding a more accurate description of a broad spectrum of dynamics in proteins. Examination of the 60 simulations deriving from AMD substates reveals no identifiable correlation between the rmsd over the individual trajectories or the rmsd relative to the X-ray structure and the improvement in S^2 (Table 1 Supporting Information).

Collective Motions in Protein G. The collective nature of the slower motions present in GB3 has been analyzed using quasi-harmonic principal component analysis of the AMD ensemble.[52] The results are illustrated in Figure 5, where the first motional mode is depicted. The regions of most collective motion again coincide closely with regions showing the greatest difference between $S^2_{RDC,NH}$ and $S^2_{LS,NH}$. The principal mode concerns a correlated motion relating the C-terminal region of the α-helix continuing into the loop III 41−45 and loop I between strands β1 and β2. This mode also describes a collective motion in the β-sheet, the outer strand (β2) of which constitutes the active site of the molecule. This dynamic mode was shown using 3D-GAF analysis to be essentially a γ-motion (about the $^{\alpha}C-^{\alpha}C$ axis) corresponding to the expected conformational sampling required for formation of a hydrogen-bonded complex with the physiological partner protein Fab.[31] Recent extended (100−200 ns) MD simulations of GB1 also identified motion in this strand and proposed that the movement was correlated with fluctuations throughout the molecule.[53] AMD simulation therefore provides additional evidence, initially detected from the 3D-GAF analysis, that RDCs are identifying slow correlated motion. The first mode alone accounts for 30% of the covariance and therefore contributes significantly to the slower timescale dynamics.

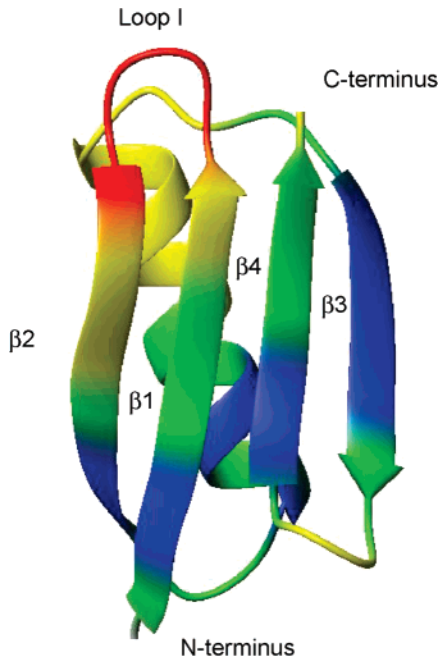

Figure 5. Quasi-harmonic principle component analysis of AMD ensemble. Residue fluctuation for the lowest eigenmode in GB3. The colors represent the percentage of the total amplitude of the eigenmode. Blue < 1%, 1% < green < 3%, 3% < yellow < 5%, red > 5%.

The description of the collective motion across the β-sheet present in the AMD simulation differs in detail with that parametrized using the 3D-GAF analysis.[30] The alternation of β and γ 3D-GAF motions found in the central part of the β-sheet and observable in the experimental $S^2_{RDC,NH}$ values for residues 9, 11, 58, and 60 are of smaller amplitude or undetected from simulation. Analysis of trans-hydrogen bond scalar couplings indicated these fluctuations to be correlated across the β-sheet and mediated via the hydrogen-bonding network. We note that the imposed directionality inherent to the 3D-GAF approach may influence the effective amplitudes in the case of collective motions and that the representation of hydrogen bonding interactions available from classical mechanical force fields may in turn limit their ability to characterize dynamic effects occurring across hydrogen bonds.

The slow dynamic modes in proteins may be linked to motions that underlie the unfolding mechanism. The principal collective mode shown in Figure 5 includes one of the two regions of the molecule that have previously been identified as melting hot spots on the basis of temperature-dependent RDC measurements.[54] Interestingly we also find that more aggressively accelerating the dynamic events using AMD (by raising E_b and α, see Methods) to the point of unfolding reveals this same motional mode as precursor of the complete unfolding process.

Describing the Conformational Landscape Native to GB3. In light of these observations we are able to propose a schematic description of the conformational landscape native to GB3

(51) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712−725.
(52) Case, D. A. *Curr. Opin. Struct. Biol.* **1994**, *4*, 285−290.
(53) Lange, O. F.; Grubmuller, H.; de Groot, B. L. *Angew. Chem., Int. Ed.* **2005**, *44*, 3394−3399.
(54) Ding, K. Y.; Louis, J. M.; Gronenborn, A. M. *J. Mol. Biol.* **2004**, *335*, 1299−1307.

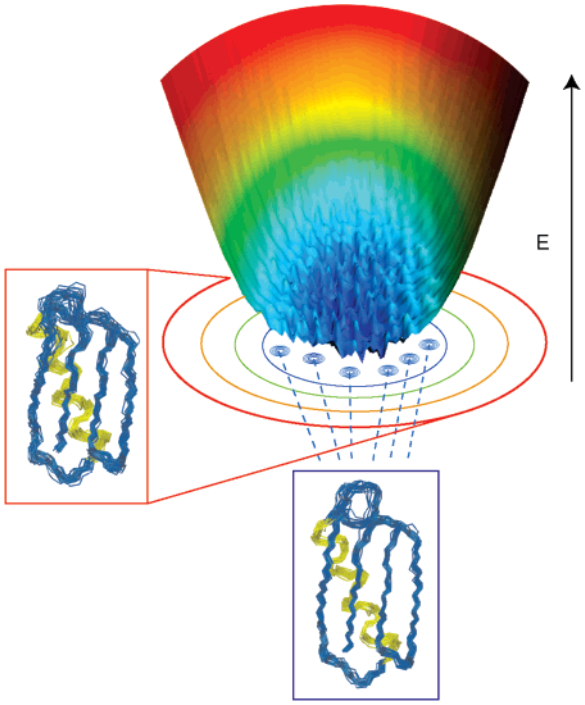

Figure 6. Schematic representation of the potential energy landscape associated with GB3. The conformation appears to reside in a steep sided well, whose shallow base comprises numerous low-energy conformational states. The conformational ensemble populating the whole base is indicated in the red box (AMD sampling), while the local sampling of the different coformational substates is represented in the blue box (MD sampling).

(shown schematically in Figure 6). It appears that the protein lies in a deep and narrow potential energy well with a relatively shallow or weakly harmonic base (backbone rms "diameter" of about 1.1−1.3 Å). The base of the well in fact comprises a large number of local potential energy minima, similar to Frauenfelder's description,[1] separated by relatively small energy barriers. These conformational microstates are sufficiently well-populated to contribute significantly to an accurate description of both fast and slower motions. We have no reliable information concerning the timescale of the motions detected using AMD, except that they appear in the window between the relaxation-imposed limit of the overall correlation time and the characteristic timescale of the RDC averaging (from nanosconds to milliseconds). Nevertheless the relative ease with which the barriers between local conformational substates are overcome in the AMD calculation indicates that these slower motions occur predominantly in the fast limit of the window, that is, the submicrosecond range. Note that vastly increased acceleration is needed to drive the system out of the potential energy well, a process that results in unfolding. These observations would be consistent with the highly stable nature of GB3.

Conclusions

Recent advances in experimental NMR techniques sensitive to slow structural fluctuations have been accompanied by parallel developments in MD methodologies aiming to extend the range of attainable timescales. Here we have combined state-of-the-art experimental and simulation techniques for exploring long timescale dynamics to probe a broad spectrum of backbone motions in GB3. We have shown that AMD reproduces the main features and many of the details of slow motions recently identified from extensive RDC measurements combined with a 3D-GAF analysis of peptide plane motions. The close reproduction of site-to-site variation and relative amplitudes in both loop and secondary structural elements underlines the power of AMD as a complementary tool for studying slow motions observed by experimental NMR.

One of the more remarkable aspects of this study is the improvement in the simulation of ^{15}N relaxation derived order parameters when additional conformational substates accessed during the AMD are included in the fast motion analysis. This is verified using extensive MD simulations in different conformational substates sampled by AMD, compared to an equal number of simulations starting from the relaxed crystal structure. This demonstration that local dynamics are dependent on small differences in the topography of conformational substates illustrates the complex coupling between structure and dynamics and has important implications for the accurate reproduction of motional modes from molecular simulation.

The description of the dynamic properties of protein GB3 over multiple timescales delivers crucial information about the conformational energy landscape native to the molecule in solution, revealing a steep-sided conformational energy potential, whose stippled base comprises multiple local minima. We expect the combination of AMD with NMR-based detection of slower motions to contribute significantly to our understanding of functional dynamics in proteins.

Methods

Accelerated Molecular Dynamics. The details of accelerated MD have been discussed previously in the literature,[46,47] and we merely provide a brief summary here. Following Voter's hyperdynamics scheme,[55] to provoke accelerated dynamic events a reference or "boost" energy, E_b is defined lying above the minimum of the potential energy surface. At each step in the simulation, if the potential energy $V(r)$ lies below this boost energy, a continuous non-negative bias potential is added to the actual potential. If the potential energy is greater than the boost energy, it remains unaltered. This essentially raises the low-energy valleys, decreasing the magnitude of energy barriers and accelerating the exchange between low energy conformational states, but retaining the essential details of the landscape. The extent to which the potential energy surface is modified depends on the energy difference between the boost energy and the actual potential. Explicitly, the modified potential, $V^*(r)$, is defined as

$$V^*(r) = V(r) \tag{1}$$

if the potential energy is greater than or equal to the boost energy and

$$V^*(r) = V(r) + \Delta V(r) \tag{2}$$

when the potential energy is less than the boost energy. The energy modification is give by

$$\Delta V(r) = \frac{(E_b - V(r))^2}{\alpha + (E_b - V(r))} \tag{3}$$

The extent of acceleration is determined by the choice of the boost energy and the acceleration parameter, α. During the course of the MD simulation, if the potential energy is modified, the forces on the atoms are recalculated for the modified potential, and the use of the biasing potential as defined above ensures that the derivative of the modified potential will not be discontinuous at points where $V(r) = E_b$.

The increased conformational sampling induced by the AMD approach depends on the parameters used to accelerate the molecular

(55) Voter, A. F. *Phys. Rev. Lett.* **1997**, *78*, 3908−3911.

motion in the following way: If the boost energy E_b is too large and the acceleration parameter is too small, the modified potential energy surface becomes iso-energetic, resulting in a random walk through phase space, causing the system to spend a large proportion of time sampling energetically unfavorable conformational space (and eventually unfolding). If E_b is chosen to be too small the acceleration is negligible and the conformational sampling resembles that found by classical methods.

Calibration of Boost Energy and Acceleration Parameter. In the present work, the protocol used for defining E_b and α was to initially choose a reasonable value of $\alpha = 60$ kcal·mol^{-1}. We consider that the maximum energy barrier which we wish to traverse is of the order of 6 kcal·mol^{-1}, and so α is approximately 10 times this barrier height. We then performed a series of short (2.5×10^5 step) AMD simulations, systematically increasing the boost energy. For each simulation, backbone rmsd and N−H order parameters were calculated (vide infra) and compared to the results for the standard MD simulations. At some particular value of E_b, we observed an increase in the backbone rmsd and a decrease in the N−H order parameters S^2_{NH}. This value is then used for production. In this way, we ensure that the boost energy is not too large and that we are sampling the low potential energy basins reasonably well. In order to confirm this, we repeated the procedure using a larger value of $\alpha = 100$ kcal·mol^{-1}. When we compared the two AMD simulations, we saw an increase in the dynamic fluctuations in the same regions. Finally E_b and α were varied by a small amount to adjust the rate of acceleration in order to acquire minimal $S^2_{RDC,NH}$ that were comparable in magnitude to their experimental counterparts.

Extraction of Canonical Average Observables. One of the favorable characteristics of this method is that it yields a canonical average of an observable, so that thermodynamic and other equilibrium properties can be accurately determined. The corrected canonical ensemble average of the system is obtained by reweighting each point in the configuration space on the modified potential by the strength of the Boltzmann factor of the bias energy $\exp^{\beta\Delta V[r(ti)]}$, at that particular point.

Simulation Details. The X-ray structure of the third IG domain (PDB code 1IGD)[31] was placed in a periodically repeating box with 10 000 water molecules and two Na^+ counterions. Initially 12 standard classical MD simulations were performed. In each case the system was brought to thermodynamic equilibrium at 300 K, 1 bar of pressure using a weak coupling thermostat with a different random seed generator, before performing a standard 1-ns MD simulation. These 12 simulations provided the starting point for the biased potential AMD simulations performed over 5×10^6 steps. The acceleration was applied across all dihedral angles in the solute. The boost energy was set to 300 kcal·mol^{-1} above the dihedral angle energy (estimated from the average dihedral angle energy from the initial unbiased 1-ns MD simulation). The α parameter was set to 60 kcal·mol^{-1} (vide supra).

In all simulations performed, bonds involving protons were constrained using the SHAKE algorithm[56] and a time step of 2 fs was employed. In the accelerated MD simulations, structures were saved for analysis every 100 steps. All simulations were performed under periodic boundary conditions using weak coupling temperature and pressure conditions. Electrostatic interactions were treated using the Particle Mesh Ewald (PME)[57] method with a direct space sum limit of 10 Å. The Cornell et al. 1999 force field (ff99) was used for solute residues,[58] and the TIP3P water force-field[59] was employed for solvent molecules, while for the solvated MD calculations the newly developed ff99SB force field was used for solute residues. All simulations were performed using a modified in-house version of the AMBER8 code.[60]

Calculation of the Order Parameters. The internal dynamics present in the different simulations of GB3 were assessed by calculating order parameters ($S^2_{LS,NH}$) relevant for Lipari−Szabo type analysis[61] of ^{15}N autorelaxation data (vide infra) and for RDC-derived order parameters ($S^2_{RDC,NH}$ and $S^2_{RDC,CC}$). After superposition of the structures collected across the MD trajectories onto the heavy backbone atoms of all residues 6−61 of the average structure, the order parameters were calculated from[37]

$$S^2 = \frac{1}{2}[3\sum_{i=1}^{3}\sum_{j=1}^{3}\langle \mu_i \mu_j \rangle^2 - 1] \quad (4)$$

where μ_i are the Cartesian coordinates of the normalized internuclear vector of interest. For each AMD simulation, the temporal weighting correction was performed before calculation of the average structure and superposition. Resulting order parameters were then averaged over multiple AMD and compared to experimentally determined order parameters.

We have predicted $S^2_{LS,NH}$ values using classical MD simulations from starting coordinates extracted from different conformational states encountered along the AMD simulation. Two sets of simulation conditions were compared: (I) In the first case, 60 2-ns trajectories were acquired from the same initial structure (third IG domain pdb code 1IGD), in each case using a different random seed generator for the weak-coupling thermostat. (II) A series of 60 2-ns trajectories were simulated from conformational states accessed during the 12 AMD simulations. In both cases resulting $S^2_{LS,NH}$ order parameters from the 60 trajectories were averaged.

In total more than 240 ns of fully solvated classical MD and 120 ns of Accelerated MD were performed. All calculations and analyses were carried out on 16 processors of an HP SC45 (800 Alpha 1.25GHz processors with a quadrics network).

Quasi-Harmonic Analysis. For quasi-harmonic analysis, the accelerated MD trajectories were concatenated and the structures were fitted to the average coordinates (rms fit of the $^{\alpha}C$ atoms from residues 6−60). The mass-weighted covariance matrix was then calculated and diagionalized to obtain quasi-harmonic eigenmodes and eigenvalues as described previously.[52]

Acknowledgment. The authors would like to acknowledge useful discussions with Donald Hamelberg. G.B. and P.R.L.M. receive grants from the CEA. This work was supported by the EU through EU-NMR JRA3 and French Research Ministry through ANR NT05-4_42781. The authors would like to thank the Centre de Calcul Recherche et Technologie (CCRT) (CEA, Bruyéres-le-Châtel).

Supporting Information Available: Table summarizing the rms deviations, potential energies, and $|\Delta S^2|$ for the 60 AMD simulations. Reference 60 in full. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0687668

(56) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(57) Cheatham, T. E.; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193−4194.
(58) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 2309−2309.
(59) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(60) Case et al. *AMBER 8*; University of California: San Francisco, CA, 2004.
(61) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.