JCTC Journal of Chemical Theory and Computation

Computer Simulations of Peptides from the p53 DNA Binding Domain

Mey Khalili[†,‡] and David J. Wales*[,§]

Center for Cancer Research Nanobiology Program, National Cancer Institute, Frederick, Maryland 21702, MITRE Corporation, 7515 Colshire Drive, McLean, Virginia 22102, and Department of Chemistry, Lensfield Road, Cambridge CB2 1EW, United Kingdom

Received December 5, 2008

Abstract: We have studied the dynamics and thermodynamics of two of the four evolutionarily conserved segments from the p53 DNA binding domain, using molecular dynamics and replica exchange simulations. These two regions contain well-defined elements of secondary structure (a β hairpin for region II and an α helix for region V) and bind to DNA in the intact protein. They are also mutational hot spots. The goal of our study was to determine the stability and folding propensity of these peptides in isolation. We used three force fields and solvent models (CHARMM19 with EEF1, CHARMM27 with GBMV, GROMOS96 with SPC). The predicted stability, folding propensity, and secondary structures depend upon the potential. Secondary structure predictors identify helical propensity for region II, in agreement with one of the force fields (CHARMM/GBMV). However, the other two potentials favor β structure for this peptide, although the conformations may differ from the crystal. For region V secondary structure predictions are unclear. Only one force field (CHARMM/GBMV) produces low-lying free energy minima that retain some of the α helical structure from the crystal structure. The other two potentials appear to favor β structure for this peptide.

1. Introduction

p53 has been one of the most extensively studied proteins since its discovery in 1977.[1−5] The p53 gene is mutated in a large number of human cancers.[6−9] Furthermore, in many tumor cells that have the wild-type p53 protein, activity is hindered by the overexpression of regulators or viral oncogenes.[10−17] The active p53 protein is a noncovalent tetramer of four 393 residue monomers,[18−20] which serves as a transcription factor involved in regulating the cell cycle, repairing damaged DNA, and initiating cell death through apoptosis.[8] Each monomer contains two folded domains, namely a core DNA-binding domain (DBD, around residues 94−297)[18,21] and a tetramerization domain (around residues 323−360).[20,22] The remainder of the sequence appears to be intrinsically disordered,[8,9,20] including the N-terminal transactivation domain (NTD, residues 1−67)[20,23−25] and the C-terminal negative regulatory domain (residues 360−393).[22] This intrinsic disorder may reflect the multifunctional nature of p53 and its "promiscuous" binding to a variety of other proteins.[9,26] In the present work our focus is on two peptides from the DBD, which binds to specific elements of the p53 gene promoters.[27,28] This domain has been crystallized,[18,29,30] and its NMR structure has been determined with and without[26] its consensus DNA sequence.

The tumor suppressor activity of p53 arises from its ability to act as a transcription factor and induce the expression of a number of proteins that are involved in DNA repair or the inhibition of cell proliferation and apoptosis.[4,8,31] In normal cells p53 is closely regulated and has a high turnover (half-life between 5 and 40 min),[32] due to the inhibitory negative feedback mechanism of the oncoprotein mdm2 (murine double minute clone 2, human equivalent hdm2).[33] mdm2 binds to p53 in the NTD and thereby inhibits transcriptional

* Corresponding author e-mail: dw34@cam.ac.uk.
[†] National Cancer Institute.
[‡] MITRE Corporation.
[§] Department of Chemistry.

10.1021/ct8005387 CCC: $40.75
Published on Web 04/10/2009

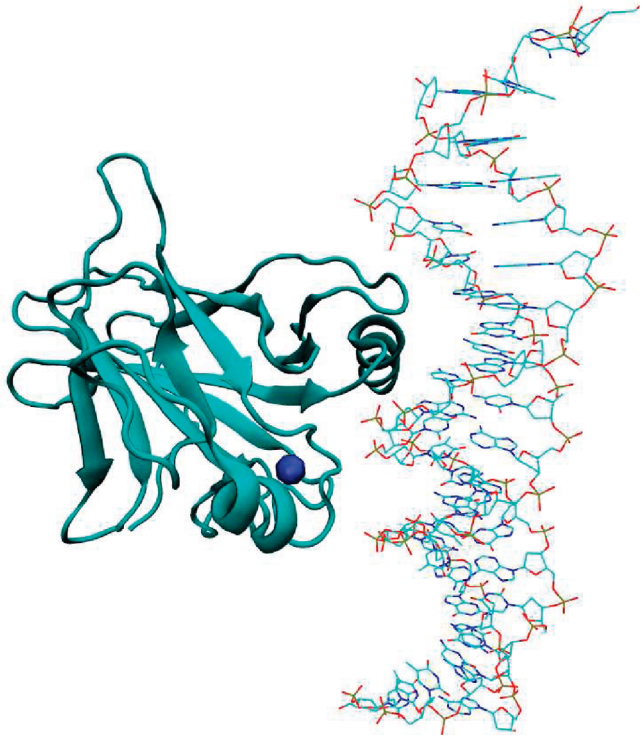

Figure 1. Monomer of the p53 DNA binding domain bound to its consensus DNA-binding site. The zinc cation is shown in dark blue.

activity;[34,35] a second binding site has also been identified for hdm2 in the DBD.[36] High levels of mdm2 may promote polyubiquitination of p53, marking the protein for degradation through the proteasome pathway.[37,38]

Given the crucial role that p53 plays in cell cycle regulation, and the large number of cancers associated with the malfunction of this protein, it is not surprising that there are many ongoing research efforts to rescue or mimic its function in cancer cells.[8,9,17,39] In the most direct application of gene therapy, tumor suppressor genes are expressed in the cells where these genes are defective. For TP53 (the gene encoding the p53 protein), encouraging results have been obtained for some classes of tumor.[40,41] Gene therapy using adenovirus mediated wild-type p53 gene transfer has achieved some success.[42−45] However, this approach does not work for tumors in which the p53 protein is functional, but its activity is hindered by the overexpression of its regulators. For example, some cancers are caused by excess levels of mdm2 or hdm2.[46,47] For these classes of tumor, research has focused on small peptides that bind to the NTD and prevent hdm2 from binding to p53.[17,48]

The crystal structure of the DNA-binding domain of human p53 (PDB code: 1TUP) was determined in a complex with consensus DNA by X-ray crystallography in 1994.[18] It consists of an immunoglobulin-like β sandwich composed of two antiparallel β sheets packed face-to-face across a hydrophobic core[18] (Figure 1). This domain contains four of the five regions that are conserved across species[18,49] (Figure 2), which are numbered II−V, following Cho et al.[18] Regions III (residues 171−181), IV (residues 234−255), and V (residues 270−289) bind to DNA directly. Region II (residues 117−142) binds to DNA through a small segment

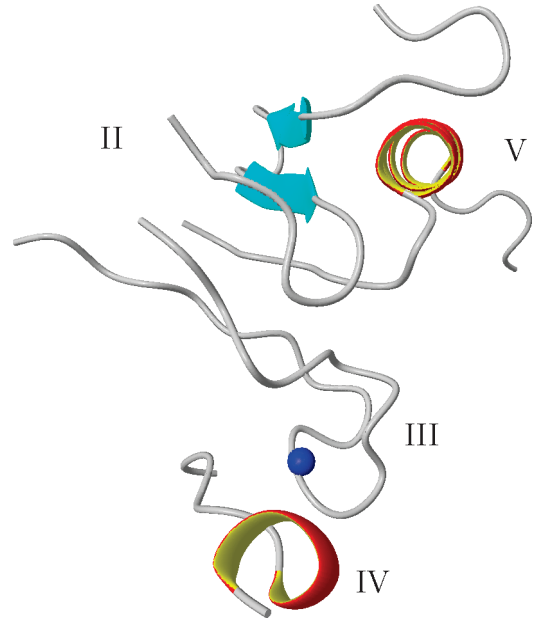

Figure 2. Conserved sites in the DNA-binding domain of p53.[18] Together these sites constitute four of the five conserved regions in a p53 monomer.

(residues 117−123) and corresponds to a loop and hairpin denoted as L1 (residues 112−124) and S2−S2′ (residues 124−142).[18] Region V corresponds to the end of a strand denoted as S10 (residues 271−274) and the α helix H2 (residues 278−286 in the 1994 crystal structure).[18] Regions II and V therefore contain well-defined secondary structures, and form a characteristic loop-sheet-helix motif,[18] where conserved residues make specific contacts with the major groove of DNA.[9] The DBD binds specifically to a palindromic double-stranded DNA promoter site, containing two decameric motifs of the form 5′-PuPuPuC(A/T)(A/T)GPyPyPy-3′ separated by up to 13 base pairs, where Pu is A/G and Py is T/C.[27]

Around 50% of human cancers involve mutations of the p53 gene.[8,9,39] Ninety-five percent of these correspond to the DBD, and 75% are single missence mutations.[39] Examples include[39] R175H (in region III, which disrupts the zinc binding domain), C242S (replaces a zinc ligand in IV), G245S (IV), R249S (IV), M237I (IV), R273H (V), and R282W (V). Arg-248, which makes a minor groove contact from a loop denoted L3, is another mutational hot spot.[18] In general, p53 mutations can be divided into three categories:[9,50] (i) DNA-contact mutations that have little effect on the folding and stability of the DBD (e.g., R273H); (ii) mutations that cause local distortions, mainly in the proximity of the DNA-binding site (e.g., R249S, destabilization $\sim$ 2 kcal/mol); (iii) mutations that cause global unfolding (in the β sandwich), which are destabilized by more than 3 kcal/mol. Bullock and Fersht note that mutations in the DNA-contact regions are oncogenic in any of these three categories, while tumorigenic mutations in the β sandwich fall into category iii.[39] For example, zinc-binding-site mutations in the DBD may have a distorted structure that fails to bind DNA.[39,50] One current research goal is to find small molecules that can rescue the function of destabilized p53 mutants by a mass action effect, namely binding strongly to the native state,

but only weakly to the denatured protein.[8,9,39] This strategy will be particularly challenging for DNA-contact mutants and mutations that perturb the zinc-binding region.[9] Alternatively, a superstable quadruple mutant of the p53 DBD has been designed, where an increase in structural rigidity results in increased thermostability and restoration of activity in several oncogenic mutants.[51,52] This superstable mutant has been used to demonstrate that hot-spot mutations have the same effect upon stability in the isolated core domain and full-length protein.[53]

During the present study, the structure of the free p53 core domain solved with NMR spectroscopy was published by Fersht and co-workers,[21] who characterize it as a marginally stable protein at body temperature, making p53 particularly susceptible to inactivation by destabilizing mutations.[9] Compared to the crystal structure solved in 1994, the α helix that belongs to region V is more than one turn longer in the NMR structure. Secondary structure predictors indicate that this longer version of region V might form a helix in isolation (Table 3). The calculations for region V in the present work all refer to residues 272−289 identified in the 1994 crystal structure.[18] Future work will consider a region V peptide that includes all the residues characterized as helical in the latest NMR structure.[21]

Our aim in this paper is to study the foldability and stability of peptides corresponding to regions II and V as isolated fragments. In native p53 the H2 α helix and the S2−S2′ β hairpin are packed together, and our motivation is partly to see whether this packing is required to stabilize the two elements of secondary structure.

It is well known that the force fields used to study biological molecules are not perfect. There have been a number of studies linking the choice of force field to differences in structural propensities in biological molecules. For example, Eisenmenger and Hansmann performed thermodynamic studies of Met-enkephalin and found subtle differences in the energy landscapes obtained for the ECEPP/2 and ECEPP/3 force fields.[54] Lwin and Luo studied a β hairpin with several alternative AMBER potentials, and with implicit and explicit solvents. They found significant differences between the implicit and explicit solvent simulations.[55] Yoda et al.[56] examined the secondary structure propensities of six protein force fields and also found clear differences. The alternative structures and energy landscapes that we have characterized for the two p53 peptides are discussed in the Results section.

2. Methods

2.1. Explicit Solvent MD Simulations. Explicit solvent simulations were performed with the GROMACS software.[57,58] We used the GROMOS96 53a6[59] force field, which has been parametrized to reproduce the enthalpies of hydration and polar solvation for a range of compounds. Each fragment was solvated by adding a cubic box of 216 SPC flexible water molecules[60] up to 9 nm from the periphery of the molecule. The volume of the box was chosen so that the density was approximately 100 g/L. Chloride and sodium ions were added to neutralize each system, and an initial minimization was performed by constraining the fragment and running molecular dynamics for the water molecules. Each system was then simulated with periodic boundary conditions. The temperature was held constant by the Berendsen thermostat,[61] and the Lincs algorithm[62] was used to constrain the hydrogens. The particle mesh Ewald method was employed[63,64] for the electrostatics and the van der Waals terms with a cutoff of 14 Å. The cutoff distance for the short-range neighbor list was set to 9 Å. The cutoff distance for the Lennard-Jones short-range neighbor list was set to 14 Å. Snapshots were collected every 500 ps with an integration time step of 2 fs. The neighbor list was updated every five steps. For the high temperature simulations at 340 K, the coupling constant to the Berendsen barostat was increased from 0.5 to 0.9 to allow for larger fluctuations in the pressure.

2.2. Implicit Solvent MD Simulations. Implicit solvent simulations were performed with the CHARMM package.[65,66] To provide additional comparisons, we considered two different implicit solvent models, namely EEF1[67] and GB-MV.[68] Implicit water solvation with EEF1 is incorporated into the CHARMM19 polar hydrogen function.[69] In this force field, the effective energy for a given conformation of the protein is the free energy of the system consisting of the macromolecule and the solvent averaged over all the solvent degrees of freedom at a given temperature. The total free energy of the protein-solvent system is the sum of the average effective energy and the configuration entropy. The ability of EEF1 to represent solvent effects was demonstrated by Lazaridis and Karplus, who were able to reproduce conformations for proteins that are comparable to those in explicit solvent simulations.[67] EEF1 has also been able to produce results similar to those of explicit solvent simulations in folding/unfolding studies.[70] Furthermore, the electrostatic energy obtained with EEF1 correlates very well with the screened Coulomb potential-implicit solvent model and reasonably well with the Poisson−Boltzmann energies.[71]

To increase the efficiency of the molecular dynamics (MD) simulations with the EEF1 solvation model, we used the multiple time step (MTS) algorithm with Langevin dynamics.[72] In this method, which is based on Trotter factorization, force linearization is combined with force splitting techniques. Unlike the conventional multiple time step algorithm,[73] this method merges the slow and fast motions via extrapolation rather than impulses, which results in a more significant time step increase. The combination of Langevin dynamics and the EEF1 implicit solvent enabled us to obtain relatively long simulation times. The runs were performed at 300 K with the MTS 4 24 integration scheme. The time steps were 0.5 fs for the fast motions, 2 fs for the intermediate motions, and 48 fs for the slow motions, yielding a 48 fs time step overall. The cutoff for short-range forces was 6.0 Å.

The generalized Born (GB) implicit solvent model is one of the most successful methods for approximating solvation energies. The most important, and time-consuming, part of this theory is the calculation of Born radii, which define the spherically averaged distance of each atom to the solvent boundary. Still and co-workers[74] introduced an expression for the Born energy, which often provides good agreement

Table 1. Secondary Structure Prediction for Region II with Different Servers[a]

	GTAKSVTCTYSPALNKMFCQLAKTCP
Jpred[88]	−−−−−EEE−−−HHHHHHHHH−−−−−−
PSIPRED[89]	CCCCEEEEECCHHHHHHHHHCCCCC
GoreIV[90]	CCCCCEEEECCCCCCCCCCCEEEEEC
HNN[91]	CCCCEEEEECCHHHHHHHHHHHCCC
nnpredict[92]	−−−−−EEE−−−−−−HHHHHHH−−−−−
porter[93]	CCCCCCCEEEECCCCEEEECCCCCEC
SOPMA[94]	TCCHHEEECCCHHHHHHHHHHTCCT
SWISS-MODEL[95]	CCCEEEEEEECCCHHHHHHHHHCCC
SCRATCH[96]	CCCCEEEECCCHHHHHHHHHHCCCC

[a] The top of the table contains the sequence. The first column lists different secondary structure servers used for predictions. The letters in the rows in the second column correspond to the propensity of each residue in the sequence: C, coil; E, extended; H, helical; T, turn. Most servers indicate that region II has helical propensity.

with the Poisson energy. The advantage of the Still equation is that, unlike the Poisson equation, it is an analytic function of atomic positions and can be used in algorithms that require force calculations. For GB simulations we used the CHARMM27 force field, which includes the ϕ and ψ grid correction CMAP surface term.[75] We used the GBMV implicit solvent model[68] with the γ surface tension factor set to 0.015 kcal/mol Å^2 to model the hydrophobic effect.[76] A Nose-Hoover thermostat[77,78] was used to keep the temperature constant. The cutoff for nonbonded list generation was set to 20 Å, the cutoff for nonbonded interactions was set to 18 Å, and the onset of switching for nonbonded interactions was set to 16 Å.

2.3. Replica Exchange Molecular Dynamics Simulations. The thermodynamic properties of the fragments were studied using replica exchange molecular dynamics (REMD).[79−81] Replica exchange simulations were performed with the multiscale modeling tools for structural biology (MMTSB)[82] toolkit. We used the same two force fields that were employed for the implicit solvent simulations, namely CHARMM19 with the EEF1 solvation model and CHARMM27 with the GBMV force field. The number of replicas was set to 32 with temperature values selected exponentially between 200 and 900 K, which produces more replicas around the lower bound. We selected the temperature range to permit efficient sampling of the energy landscape, while allowing for more replicas around the temperature range of interest (298 K). We set the number of replicas to 32 to ensure significant overlap between the energy histograms of adjacent replicas.

The length of the MD simulations in each replica cycle was 20 ps, which should lie above the upper bound for the relaxation time of the water molecules.[83] The replica cycles were repeated until convergence was achieved, as judged from the behavior of the heat capacity curve and the free energy surfaces. The PDB structures were energy minimized and equilibrated at 300 K for 1 ns. They were then used as the initial structure for each replica. The averages for different observables, and the free energies, were calculated using the weighted histogram analysis method.[84]

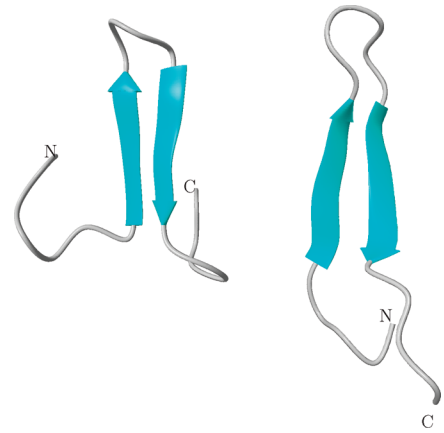

Figure 3. Crystal structure (left) and final structure (right) of region II after 45 ns of MD simulations with CHARMM19/EEF1. At the end of the run the hairpin has elongated by two residues on each side.

3. Results

3.1. Region II: Dynamics. The fastest methods for predicting the secondary structure propensity of a sequence are knowledge-based, homology modeling algorithms.[85] These methods compare a given sequence to a database of sequences for which the structure is known and calculate the statistical propensity of a sequence to form a particular secondary structure. The structure propensity of region II predicted with a number of secondary structure predictors is shown in Table 1. The consensus is that the fragment corresponding to region II is likely to form a helix.

To compare the knowledge-based results with the physics-based methods, we performed MD simulations using CHARMM19/EEF1 and the MTS algorithm at 300 K, to determine whether the hairpin remains folded on a reasonable time scale. The simulation was started from the minimized crystal structure of region II (residues 117−142, Figure 2), and continued for 45 ns.

We chose the hairpin section of region II to calculate the root-mean-square deviation (rmsd), because the rest of the chain is essentially structureless, and fluctuates significantly during the simulation. Examining the rmsd (not shown), it is evident that the hairpin is stable with an rmsd around 2.2 Å during the 45 ns simulation, and does not unfold. Figure 3 (right) shows the conformation at the end of the run, along with the crystal structure (left). During the simulation the hairpin grows from residues 123−136 to span residues 121−139. The results of MD simulations with CHARMM19/EEF1 contrast with the secondary structure predictors (Table 1), which classify region II as helical. The disagreement could be due to the fact that the minimum on the free energy landscape containing the crystal structure is deep enough to prevent the trajectory from escaping at 300 K within 45 ns of simulation. To test this hypothesis, we considered higher temperature MD simulations with CHARMM27/GBMV. We employed an alternative force field at this stage both to provide a comparison with CHARMM19/EEF1 and because the EEF1 parameters are actually fitted specifically for physiological temperatures.[67] As before, these simulations were started from the crystal structure following a local minimization.

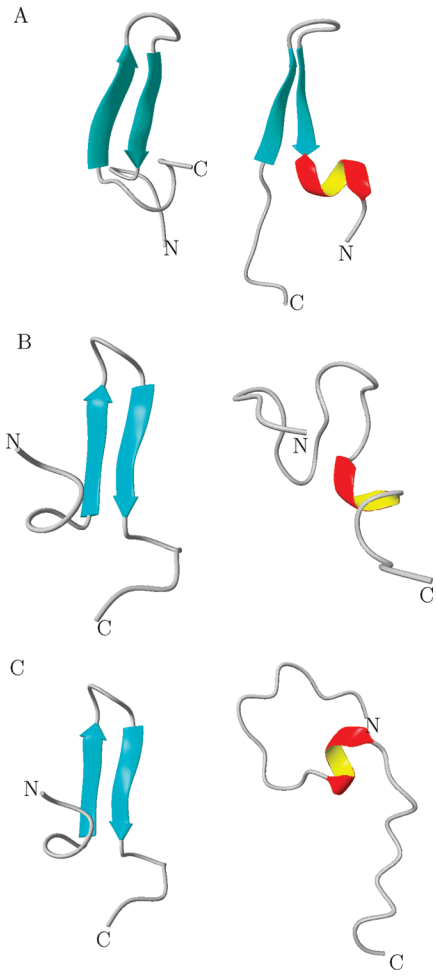

Figure 4. Initial structures (left) and final structures (right) of region II with CHARMM27/GBMV after MD runs at (A) 350, (B) 400, and (C) 450 K. A helical turn appears at 350 K. At 400 and 450 K the hairpin is completely lost, but the helix turn persists, and shifts through the sequence. Since the hairpin is expected to be thermodynamically more stable than a helix,[96,97] this result might indicate some preference for helical conformations.

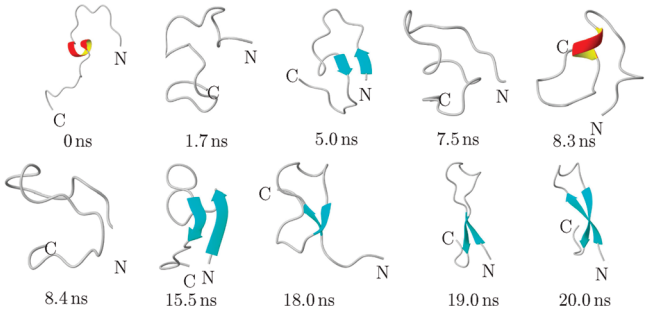

Figure 5. Snapshots of the folding pathway of region II with GROMOS96/SPC at 340 K. There is a clear competition between the formation of a hairpin and a helix during the first 9 ns of the simulation.

With CHARMM27/GBMV, region II is stable on the MD time scale at 350 K (rmsd never increases above 3.5 Å, data not shown), but unfolds at higher temperatures (rmsd higher than 6.0 Å, data not shown). The final structure of region II at the end of the 350 K simulations includes the original hairpin (Figure 4). However, a helical turn appears at the N-terminus. At 400 and 450 K the hairpin is lost in about 1.5 and 1 ns, respectively. However, the helix turn persists, and it shifts its position from the N-terminus to the C-terminus of the unfolded sequence (Figure 4B,C, right). The persistence of the helical turns could reflect a structural preference, or a bias in the force field toward helical conformations. Either way, the behavior of region II with CHARMM27/GBMV is more consistent with the secondary structure predictors, which classify this region as helical (Table 1).

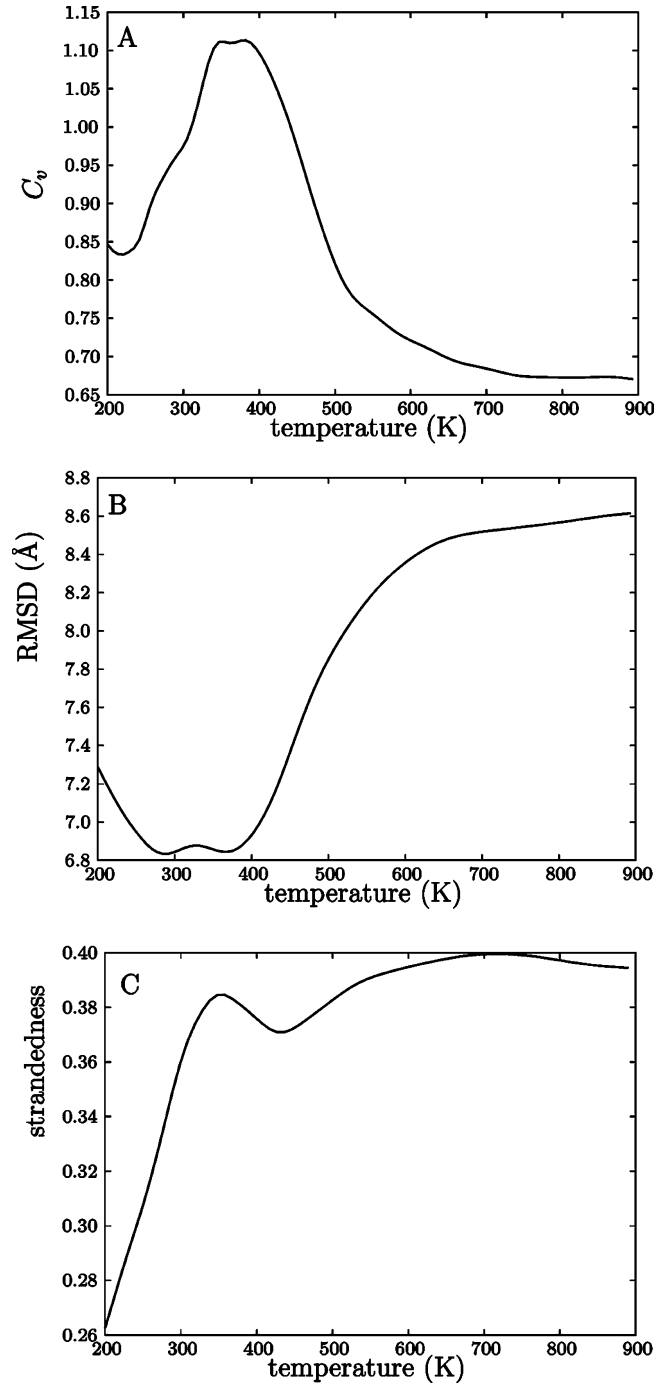

Figure 6. Observables calculated in REMD simulations of region II with CHARMM19/EEF1. (A) Heat capacity [kcal/(mol K)], (B) rmsd from the crystal structure, and (C) strandedness.

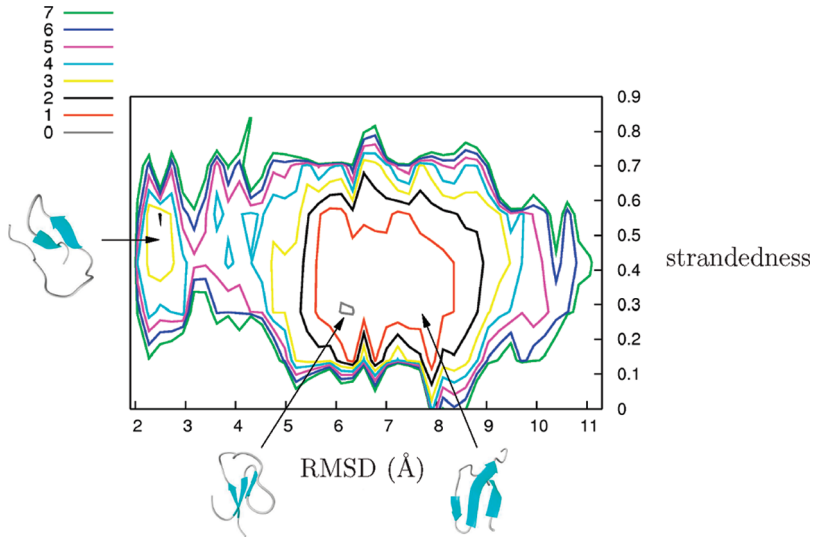

Figure 7. Two-dimensional free energy surface (kcal/mol) of region II as a function of rmsd and strandedness obtained using CHARMM19/EEF1 at 298 K.

Next we tried to fold region II from an unfolded conformation. We used the final structure obtained at 450 K (Figure 4C) as the starting conformation. This choice was made to help the helix nucleate and elongate, if indeed the helical conformation is favorable for this sequence. We used the GROMOS96 force field and explicit SPC water. We also chose a folding temperature of 340 K, to enhance sampling and speed up the folding. We note that this choice might bias the sequence toward β conformations. The simulations were performed for 21 ns. The rmsd of the hairpin from the crystal structure starts around 7.5 Å and drops to about 5.7 Å at about 8 ns, fluctuating around this value thereafter. Figure 5 shows snapshots of region II along its folding pathway. Initially, there is competition between the formation of helix and hairpin. The helical turn disappears at about 2 ns, and a hairpin forms, but does not persist. The helix reappears after 8.3 ns, but disappears again quickly. The same hairpin appears briefly and elongates, but the structure readjusts so that the hairpin corresponding to the crystal structure nucleates at about 18 ns and elongates. The final hairpin is slightly shifted (the turn is shifted by two residues in the C-terminus direction) from the crystal structure. Hence, with the GROMOS96 force field, the sequence appears to favor a hairpin, in contrast to the secondary structure predictions, but in agreement with the crystal structure.

3.2. Region II: Thermodynamics. To study the thermodynamics of region II, we performed REMD simulations with both the CHARMM19/EEF1 and CHARMM27/GBMV force fields. With CHARMM19/EEF1, the system reached equilibrium after 2.36 μs combined time (3685 replica cycles for each of 32 temperatures, each consisting of 10 000 MD steps with 0.002 ps time step). Figure 6 shows different observables at the end of the simulations. The heat capacity curve (Figure 6A) is broad, suggesting the absence of a sharp transition to a single minimum below the folding temperature. Figure 6B shows the average rmsd of the hairpin from the crystal structure at different temperatures. At all temperatures the rmsd is more than 6.8 Å from the crystal conformation. It dips around the physiological temperature range, which indicates that the force field produces more native-like conformations here. Before the system "melts", it adopts conformations that are about 7 Å from the crystal structure. Visual inspection of the low-energy conformations indicated that almost all of them were β structures. We therefore examined a different observable to measure the propensity of the sequence to form extended conformations.

Calculating the extent of β conformations present in MD snapshots is more difficult than for α helices. In particular, changes that are relatively minor upon visual inspection, such as a slight shift in the hairpin formation, can produce large rmsd values. β strands also involve well-separated portions of a chain, making a distance measure more difficult. We used the ϕ and ψ angles to measure the propensity of the chain to form extended structures, following Yoda et al.,[56] who measured these two angles and compared them with the values of the Ramachandran map corresponding to extended structures. We considered a residue to be in the β (or extended) state if the backbone dihedral angles (ϕ and ψ) were in the ranges $-130° \pm 50°$ and $135° \pm 45°$, respectively. We then counted the number of residues with dihedral angles in this range, and divided by the total number of residues in the chain. We call this measure "strandedness", following Yoda et al. Usually a strandedness of 0.4 or higher means that most of the chain is extended, and not surprisingly, strandedness increases with temperature. The strandedness of the crystal conformation is 0.24 (Figure 3, left), while that of the final structure obtained from the MD simulations with CHARMM19/EEF1 is 0.44 (Figure 3, right). Figure 6C shows that the strandedness of region II increases with temperature and reaches a maximum of 0.38, after which it decreases before increasing again as the chain unfolds. This maximum near the physiological temperature is another indication that the force field favors extended conformations around the physiological temperature.

We also measured helicity, i.e., the propensity of the chain to form a helix, using the number of backbone carbonyl, i, and nitrogen, $i + 4$, atoms separated by less than 4 Å. There are 22 possible contacts in a sequence of 26 residues, and

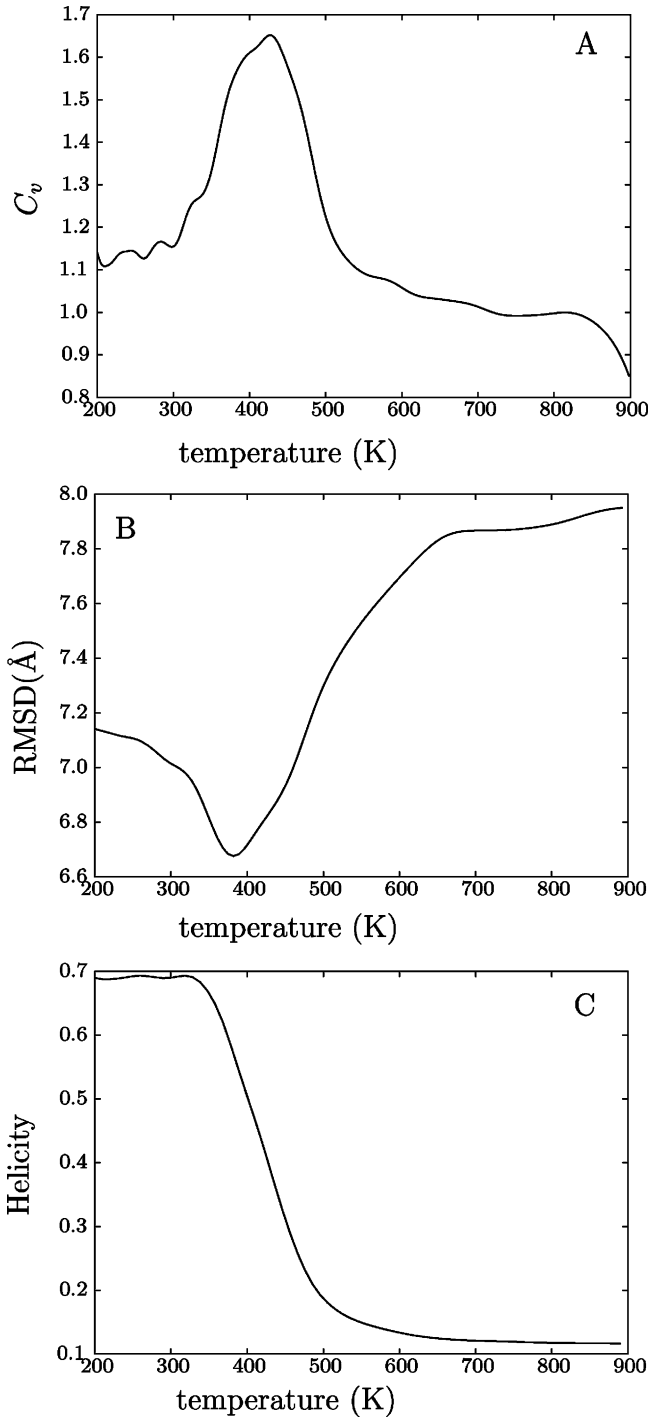

Figure 8. Calculated observables from REMD simulations of region II with CHARMM27/GBMV. (A) Heat capacity [kcal/(mol K)], (B) rmsd from the crystal structure, and (C) helicity. The sequence forms helical conformations with this force field rather than the native β strand.

helicity is defined as the percentage of them that actually form. A plot of average helicity vs temperature (not shown) reveals values below a helicity of 0.045 (corresponding to one helical hydrogen bond) throughout the temperature range considered. Hence the helical content of this sequence is negligible for CHARMM19/EEF1.

Figure 7 shows the two-dimensional free energy surface as a function of rmsd and strandedness at 298 K. The global free energy minimum corresponds to structures that are about

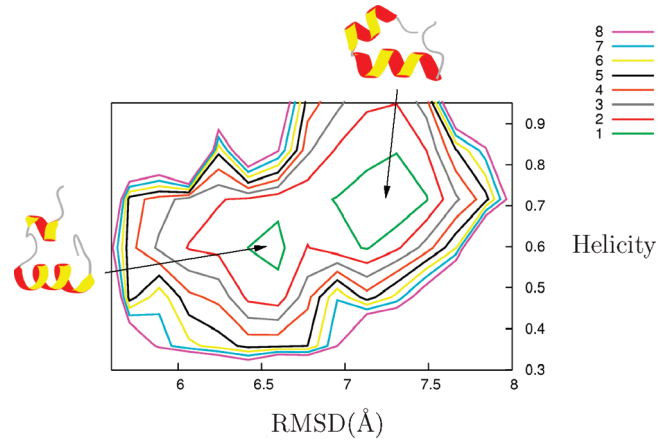

Figure 9. Two-dimensional free energy surface (kcal/mol) for region II with CHARMM27/GBMV at 298 K.

Table 2. Secondary Structure Predictions for Region V (from the 1994 Crystal Structure) with Different Secondary Structure Prediction Servers[a]

	VRVCACPGRDRRTEEENL
Jpred	sequence too short
PSIPRED	sequence too short
GorelV	CCCCCCCCCCCCCEEEEC
HNN	CEEECCCCCCCCCCCCCC
nnpredict	−−−−−HHHHH−−−−−−−−
porter	CEECCCCCCCCCCHHCCC
SOPMA	EEEECCTTCCCCCHHHHH
SWISS-MODEL	CEEEECCCCCCCCCCCCC
SCRATCH	CEEEECCCCCCCCCCCCC

[a] There is no strong consensus on the structure of this sequence.

Table 3. Secondary Structure Predictions for Region V (from the New NMR Structure) with Different Secondary Structure Prediction Servers[a]

	VRVCACPGRDRRTEEENLRKKGEPHH
Jpred	−EE−−−−−−−−−−−−−HHHH−−−−−−
PSIPRED	CCEECCCCCCCCCHHHHHHHCCCCCC
GorelV	CCCCCCCCCCCCHHHHHHHHCCCEEC
HNN	CEEECCCCCCCCCHHHHHHHCCCCCC
nnpredict	−−−−−−−−−−−−−HHHHH−−−−−−−−
porter	CEECCCCCCCCCHHHHHHHHCCCCCC
SOPMA	EEEECCTTCCCCCHHHHHHHTCCCCC
SWISS-MODEL	CEEEECCCCCCCCCHHHCCCCCCCCC
SCRATCH	CEEEECCCCCCCCCHHHHHHHCCCCCC

[a] Helical propensity is generally predicted.

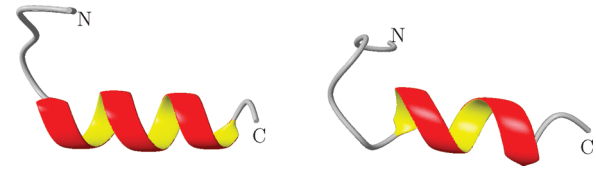

Figure 10. Crystal structure of region V (left) and final structure (right) after 70 ns of MD simulation with CHARMM19/EEF1. The helix loses hydrogen bonds from both ends, while retaining the central part of the helix.

6 Å from the crystal structure and have a strandedness of 0.3. However, this minimum belongs to a broad region of low free energy spanning rmsd values between 6 and 8 Å and strandedness between 0.2 and 0.5. Therefore, at room

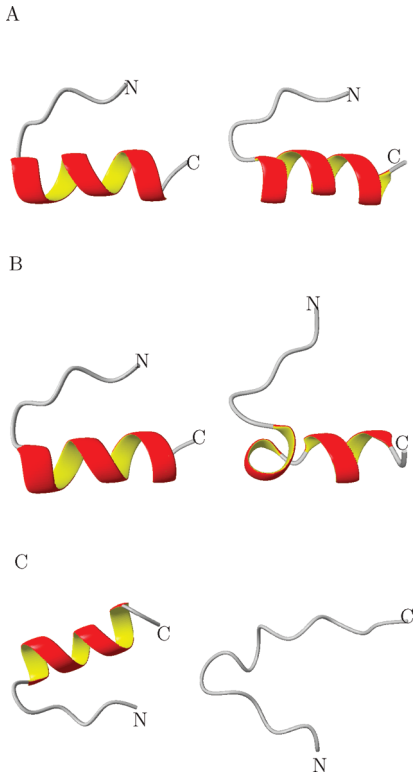

Figure 11. Initial structures (left) and final structures (right) of region V with CHARMM27/GBMV at (A) 350, (B) 400, and (C) 450 K. As for region II, region V remains close to its crystal conformation at 350 K, but changes structure at higher temperatures.

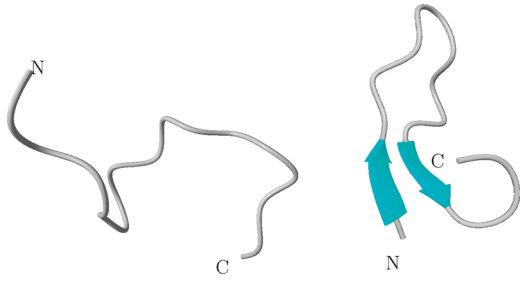

Figure 12. Initial unfolded structure (left) and conformation at the end of MD simulations (right) for region V with the GROMOS96/SPC force field after 20 ns. The sequence adopts a hairpin, which persists to the end of the run.

temperature, the sequence does not have a single well-defined structure. There is also a free energy minimum corresponding to a narrow range of rmsd around 2 Å and a relatively wide range of strandedness. This feature corresponds to the crystal structure, and lies 2 kcal/mol higher than the global minimum, separated by a barrier of about 3 kcal/mol. As expected from the CHARMM19/EEF1 MD simulations, this minimum is deep enough to trap the system at room temperature on short time scales. Although this free energy minimum contains conformations that are comparable to the crystal, it also contains non-native β structures.

To check whether changing the temperature would make the free energy minimum corresponding to the crystal structure more favorable, we recalculated the free energy profile at 310, 390, 390, and 450 K, respectively (data not

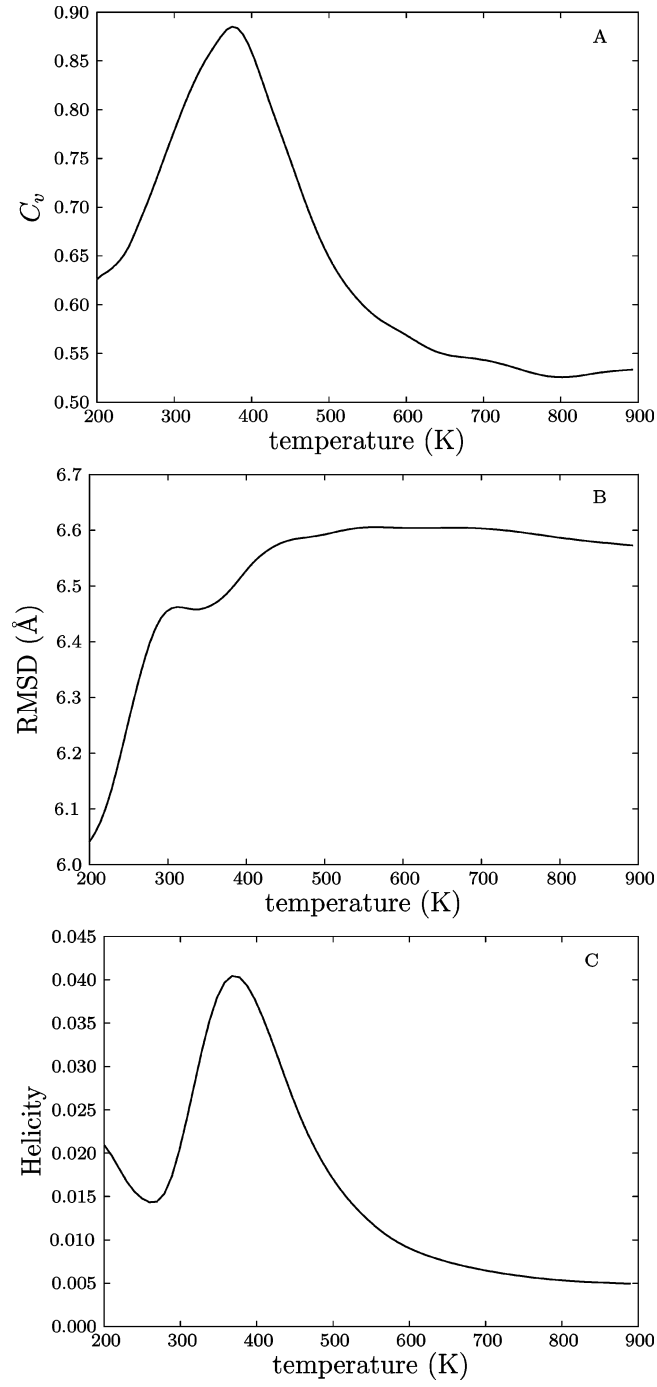

Figure 13. Calculated observables for REMD simulations of region V with CHARMM19/EEF1. (A) Heat capacity [kcal/(mol K)], (B) rmsd, and (C) helicity.

shown). At temperatures above 350 K this minimum becomes progressively shallower, indicating marginal stability. However, the global minimum for 298 K persists and broadens at higher temperature, as expected, indicating that the sequence prefers to be in the more extended conformations at all temperatures with this potential.

Next we repeated the REMD simulations with the CHARMM27/GBMV force field, which gives very different results. The system reached equilibrium after 27.52 ns combined time (43 000 replica cycles for each 32 temperatures, each consisting of 10 000 MD steps with a 0.002 ps time step). Figure 8 shows different observables from this

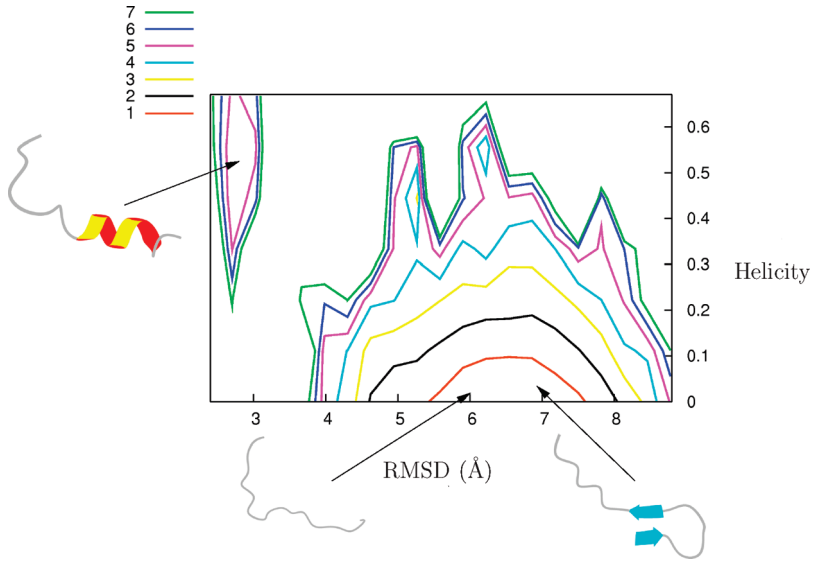

Figure 14. Two-dimensional free energy surface (kcal/mol) of region V for CHARMM19/EEF1 at 298 K, projected onto rmsd and helicity.

simulation. As for Figure 6A, the heat capacity is broad, again indicating no well-defined transition to a single minimum below the transition temperature. At all temperatures the rmsd from the crystal structure (Figure 8B) is more than 6.8 Å, and the average helicity is about 0.7 at temperatures below 400 K, which corresponds to 15 helical hydrogen bonds (Figure 8C). A corresponding two-dimensional free energy surface is shown in Figure 9. There are two minima of equal depth, separated by a small barrier. It is clear that both minima contain conformations with high helicity. It is also clear that they can easily interconvert, since the barrier of 1 kcal/mol is too small to confine the structures in either one of the minima at physiological temperatures. Sequence II in isolation therefore prefers to form helices with this force field, in agreement with secondary structure predictors, and in contrast to the crystal structure and the other two potentials considered. Our MD runs at lower temperatures were clearly not long enough to achieve equilibrium, although the run at 350 K did produce one helical turn, as discussed above. The broad shape of the heat capacity curve and the broad minima on the free energy surface are due to the fact that the N-terminus of the sequence forms many conformations with different amounts of helicity.

3.3. Region V: Dynamics. As mentioned in the Introduction, region V considered here (residues 272−289) is based on the crystal structure obtained in 1994, and does not include the additional helical turn beyond residue 289 reported for the latest NMR structure.[21] Apart from containing many charged residues, which help it to bind directly to DNA, this region is also notable for having a proline nearer the N-terminal part of the sequence (P278). This proline forces the N-terminus to adopt a kinked position relative to the rest of the sequence. In contrast to region II, there is a lack of consensus among secondary structure predictors for residues 272−289 of region V, although most of them predict the sequence to be unstructured in isolation (Table 2). However, when we consider residues 272−297, which include the complete helix from the NMR structure, the homology modeling methods generally predict region V to be helical (Table 3).

We performed 70 ns MD simulations for the 18-residue peptide from region V with CHARMM19/EEF1 using the MTS algorithm. The rmsd fluctuates around 2.5 Å, with relatively minor variations until 20 ns (data not shown). From this point, it increases by 1.5 Å and the fluctuations grow significantly, because the N-terminal and C-terminal parts of the helix unfold somewhat and the ends move more freely. The helicity decreases from unity to about 0.6 in the first 20 ns of the run, and then stabilizes (data not shown). This value corresponds to the loss of approximately three helical hydrogen bonds. Beyond this point the rmsd increases significantly, indicating that first the three helical hydrogen bonds are lost, and then the chain becomes more labile. Figure 10 shows the crystal structure (left), and the structure at the end of the 70 ns MD simulations (right). The hydrogen bonds lost belong to both ends of the helix, while the core remains intact. As mentioned above, due to the position of P278, the orientation of the N-terminus relative to the rest of the peptide rarely changes.

High temperature MD simulations with CHARMM27/GBMV were performed for region V at 350, 400, and 450 K, for intervals of 10, 9, and 9 ns, respectively. As for region II, region V is stable on this time scale at 350 K with the helix intact, but unfolds at higher temperatures. For the 400 K run the helix loses two helical hydrogen bonds by the end of the trajectory; the final structure is a distorted helix with a kink in the middle (Figure 11 B). Helicity is completely lost in the 450 K simulations; the final structure has only one helical hydrogen bond, and lacks any other helical features (Figure 11C).

To see if region V folds to its crystal structure, we considered MD runs using explicit solvent simulations with the GROMOS96 force field and explicit SPC water. We used a temperature of 340 K and ran the simulations for 20 ns to see if we would observe any helix nucleation. The rmsd

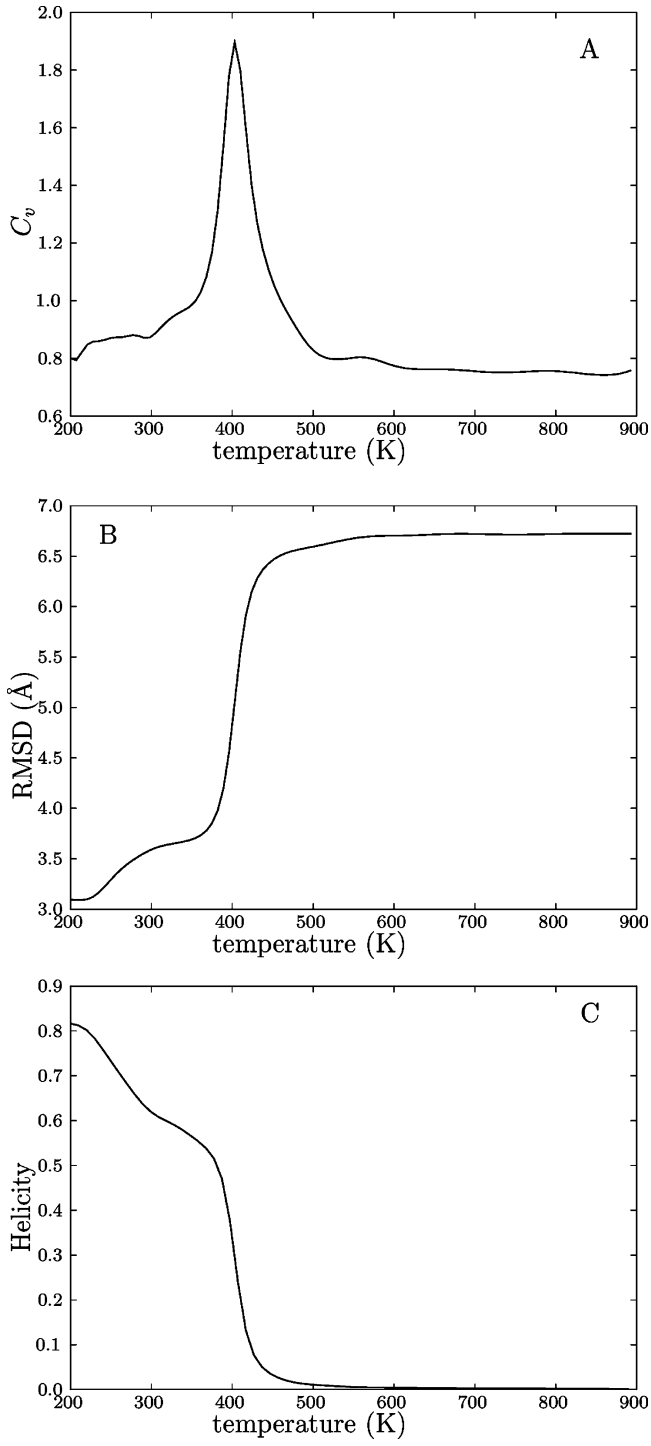

Figure 15. Calculated observables from REMD simulations of region V with CHARMM27/GBMV. (A) Heat capacity [kcal/(mol K)], (B) rmsd of the helical region, and (C) helicity.

during the simulation fluctuates around 6.5 Å, and the helicity does not increase above 0.1 (data not shown), indicating that there is no helical hydrogen-bond formation. Instead we observe a degree of hydrophobic collapse to compact conformations, but because the sequence is highly charged it prefers to maximize its contact with water. This peptide is stable in the hairpin configuration of Figure 12 and remains in this conformation until the end of the simulation for GROMOS96/SPC.

3.4. Region V: Thermodynamics. Figure 13 shows various properties calculated in replica exchange MD simulations for region V with CHARMM19/EEF1. The simulations were performed for 1.5 μs (2417 replica cycles for each of the 32 temperature windows, each consisting of 10 000 MD steps with a time step of 0.002 ps). Figure 13A shows the heat capacity, which is broad with a small peak, similar to that of region II with this force field, indicating the absence of a sharp transition. Figure 13B shows the average rmsd of the helical segment (residues 277−289) as a function of temperature. At all temperatures, the average rmsd is about 6 Å from the crystal structure, indicating that helical conformations are not favored at any temperature with this force field. This observation is further confirmed by the helicity, which is less than 0.1 at all temperatures, indicating that no helical hydrogen bonds are present.

We calculated the free energy surface at 298 K (Figure 14) as a function of rmsd and helicity, revealing two distinct minima. The higher minimum is relatively narrow, and corresponds to structures about 2.8 Å from the crystal structure with helicities of 0.4−0.6, in good agreement with the values obtained in the MD simulations for this potential. This minimum corresponds to a state where the secondary structure observed in the crystal is partially retained. However, the global free energy minimum corresponds to structures that lie 5.5−7.5 Å from the crystal structure, with helicities below 0.1. This minimum is 4 kcal/mol lower in energy than the one associated with the crystal structure, and corresponds to extended and unfolded structures, in agreement with the unfolding observed in MD simulations for this potential. Our MD trajectories were not long enough to escape completely from the higher minimum.

We repeated the REMD simulations with CHARMM27/GBMV starting from the crystal structure. The simulations were performed for 5.0 μs (7805 replica cycles for each of 32 temperature windows, 10 000 MD steps for each with a time step of 0.002 ps). Calculated observables are shown in Figure 15. Compared to the CHARMM19/EEF1 force field for this region (Figure 13), and both force fields for region II, the heat capacity curve is sharp (Figure 15A), indicating a more well-defined transition. There is a clear signature for this transition in both the rmsd (Figure 15B) and helicity (Figure 15C), indicating melting as the temperature increases. The rmsd is higher (by 1 Å) and the helicity is lower (by about two hydrogen bonds) at 350 K for the REMD simulations with CHARMM27/GBMV compared to the MD runs with the same force field, indicating that the MD results are consistent, but did not reach equilibrium.

The free energy surface projected onto rmsd and helicity obtained with CHARMM27/GBMV has two comparable free energy minima, separated by a small barrier of 1 kcal/mol (Figure 16). One minimum spans conformations that are close to the crystal structure (rmsd about 1.5−3.3 Å), with helicity between 0.6 and 0.8. The other minimum consists of partially unfolded conformations that are between 4.5 and 5.5 Å from the crystal structure and have a helicity of around 0.4. This value still corresponds to four helical hydrogen bonds, which is significant for a small helix, indicating that there is still appreciable helical character. Inspection of

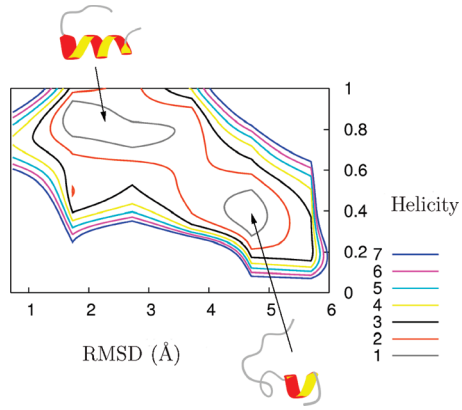

Figure 16. Free energy surface (kcal/mol) of region V with CHARMM27/GBMV as a function of rmsd and helicity at 298 K.

conformations from this minimum confirms that they have several helical turns. Hence the CHARMM27/GBMV force field supports an α-helical structure for both the region II and V peptides considered in the present work.

In summary, both CHARMM/EEF1 and CHARMM/GBMV support free energy minima corresponding to helical conformations close to the crystal structure for region V. However, for CHARMM/EEF1 the corresponding state is predicted to be unstable relative to extended structures, while for CHARMM/GBMV an equilibrium with partially unfolded helical conformations would be expected. The GROMOS96/SPC simulations did not produce any helical structure.

4. Conclusions

We have studied two evolutionarily conserved peptide fragments from the p53 DNA binding domain, both containing well-defined secondary structure in the intact protein. We used a number of force fields and solvent models, which produce somewhat different predictions for the structure propensity and stability. The region II peptide folds to its near-native β hairpin with the GROMOS96 force field and SPC explicit solvent. It is also stable in its hairpin conformation in MD runs with CHARMM19/EEF1 at 300 K, but forms a helical turn at the N-terminus for CHARMM27/GBMV at 350 K, which persists at higher temperatures where the hairpin is lost.

REMD simulations for region II with CHARMM19/EEF1 produce a free energy minimum containing a wide range of β conformations that are different from the crystal structure, perhaps suggesting a propensity for aggregation. With CHARMM27/GBMV the most stable conformations are helical in equilibrium for physiological temperatures, indicating that the MD trajectories were trapped in higher energy configurations. All the low-lying conformations for this potential have a well-defined helix in the middle of the sequence, with a varying degree of helicity at the N-terminus. The predictions of helical structure from CHARMM27/GBMV and homology modeling, which contrast with the β hairpin seen in the crystal structure and with the other potentials, suggest that the secondary structure preference of region II in the complete protein may be significantly affected by the environment.

For region V defined by residues 272−289 the consensus of homology models is that this peptide would lack secondary structure, although two of the models classify it as helical. In molecular dynamics simulations with CHARMM19/EEF1 and higher temperature runs for CHARMM27/GBMV the helix loses some of its helical structure. Simulations for region V with GROMOS96/SPC produced a non-native hairpin.

REMD simulations of region V for CHARMM19/EEF1 indicate that a variety of β-like conformations are present at physiological temperatures, as for region II with this potential. In contrast, for CHARMM27/GBMV the free energy minima below the folding transition temperature correspond to a range of helical structures that exhibit various degrees of unfolding compared to the crystal structure. Although there is a high energy minimum corresponding to the crystal structure for CHARMM19/EEF1, only the CHARMM27/GBMV combination predicts that the secondary structure observed in the crystal structure might be partially retained in the isolated peptide. The MD simulation in explicit water for GROMOS96/SPC resulted in collapse to a hairpin structure. These results again suggest that the environment in the complete protein may help to stabilize the secondary structure observed in the crystal. However, a longer sequence including additional residues associated with the extended helix observed in the NMR structure[21] would probably exhibit more helical character in isolation.

We should not necessarily expect small peptide fragments of larger proteins to exhibit stable secondary structure, and the present study reaffirms previous work[56,86] that reports a lack of consensus among different bimolecular force fields for such systems. It is well known that all force fields have some inherent bias. For example, CHARMM19/EEF1 may be less reliable at higher temperatures, because the experimental values used for specific heat capacities and enthalpies to parametrize this force field correspond to 300 K.[67] In this force field the partial charges are adjusted[67] so that the total charges of all residues are zero, and it is unclear what effect this will have on the electrostatics of sequences such as region V, which involve significant charges.

The generalized Born model has been reported to overstabilize secondary structure,[87] which may contribute to the formation of two distinct minima on the free energy surfaces of regions II and V separated by small barriers. It has also been reported that the GROMOS96 force field favors β hairpins,[56] which could explain why region II folds to a β hairpin comparable to the crystal structure, while homology models classify it as helical, and why region V also collapses to a β hairpin. We therefore conclude that the simulation results with different force fields should be considered as suggestions of how these systems might behave, which could be useful in interpreting future experiments. Finally, we note that intrinsically disordered regions[26] of proteins such as p53 may provide an even more stringent test of different force fields in computer simulations.

References

(1) Lane, D. P.; Crawford, L. V. *Nature (London)* **1979**, *278*, 261–263.

(2) Oren, M.; Rotter, V. *Cell. Mol. Life Sci.* **1999**, *55*, 9–11.

(3) May, P.; May, E. *Oncogene* **1999**, *18*, 7621–7636.

(4) Vogelstein, B.; Lane, D.; Levine, A. J. *Nature (London)* **2000**, *408*, 307–310.

(5) *25 Years of p53 Research*; Hainaut, P., Wiman, K. G., Eds.; Springer: New York, 2005.

(6) Linzer, D. I.; Levine, A. J. *Cell* **1979**, *17*, 43–52.

(7) Soussi, T.; Dehouche, K.; Beroud, C. *Hum. Mutat.* **2000**, *15*, 105–113.

(8) Römer, L.; Klein, C.; Dehner, A.; Kessler, H.; Buchner, J. *Angew. Chem., Int. Ed.* **2006**, *45*, 6440–6460.

(9) Joeger, A. C.; Fersht, A. R. *Oncogene* **2007**, *26*, 2226–2242.

(10) Freedman, D. A.; Wu, L.; Levine, A. J. *Cell. Mol. Life Sci.* **1999**, *55*, 96–107.

(11) Sherr, C. J. *Genes Dev.* **1998**, *12*, 2984–2991.

(12) Dasika, G. K.; Lin, S. C.; Zhao, S.; Sung, P.; Tomkinson, A.; Lee, E. Y. *Oncogene* **1999**, *18*, 7883–7899.

(13) Bartek, J.; Lukas, J. *Curr. Opin. Cell. Biol.* **2001**, *13*, 738–747.

(14) Salomoni, P.; Pandolfi, P. P. *Cell* **2002**, *108*, 165–170.

(15) Colombo, E.; Marine, J. C.; Danovi, D.; Falini, B.; Pelicci, P. G. *Nat. Cell Biol.* **2002**, *4*, 529–533.

(16) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Lecine, A. J.; Pavletich, N. P. *Science* **1996**, *274*, 948–953.

(17) Hupp, T. R.; Lane, D. P.; Ball, K. L. *Biochem. J.* **2000**, *352*, 1–17.

(18) Cho, Y.; Gorina, S.; Jeffrey, P. D.; Pavletich, N. P. *Science* **1994**, *265*, 346–355.

(19) Weinberg, R. L.; Veprintsev, D. B.; Fersht, A. R. *J. Mol. Biol.* **2004**, *341*, 1145–1159.

(20) Tidow, H.; Melero, R.; Mylonas, E.; Freund, S. M. V.; Grossmann, J. G.; Carazo, J. M.; Svergun, D. I.; Valle, M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 12324–12329.

(21) Nadillas, J. M. P. C.; Tidow, H.; Freund, S. M. V.; Rutherford, T. J.; Ang, H. C.; Fersht, A. F. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 2109–2114.

(22) Rustandi, R. R.; Baldisseri, D. M.; Weber, D. J. *Nat. Struct. Biol.* **2000**, *7*, 570–574.

(23) Clore, G. M.; Ernst, J.; Clubb, R.; Omichinski, J. G.; Kennedy, W. M. P.; Sakaguchi, K. *Nat. Struct. Biol.* **1995**, *2*, 321–333.

(24) Lee, W.; Harvey, T. S.; Yin, Y.; Yau, P.; Litchfield, D.; Arrowsmith, C. H. *Nat. Struct. Biol.* **1994**, *1*, 877–890.

(25) Dawson, R.; Muller, L.; Dehner, A.; Klein, C.; Kessler, H.; Buchner, J. *J. Mol. Biol.* **2003**, *332*, 1131–1141.

(26) Wells, M.; Tidow, H.; Rutherford, T. J.; Markwick, P.; Jensen, M. R.; Mylonas, E.; Svergun, D. I.; Blackledge, M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 5762–5767.

(27) el Deiry, W. S.; Kern, S. E.; Pietenpol, J. A.; Kinzler, K. W.; Vogelstein, B. *Nat. Genet.* **1992**, *1*, 45–49.

(28) Kern, S. E.; Kinzler, K. W.; Bruskin, A.; Jarosz, D.; Friedman, P.; Prives, C.; Vogelstein, B. *Science* **1991**, *252*, 1708–1711.

(29) Kitayner, M.; Rozenberg, H.; Kessler, N.; Rabinovich, D.; Shaulov, L.; Haran, T. E.; Shakked, Z. *Mol. Cell* **2006**, *22*, 741–753.

(30) Ho, W. C.; Fitzgerald, M. X.; Marmorstein, R. *J. Biol. Chem.* **2006**, *281*, 20494–20502.

(31) Sionov, R. V.; Haupt, Y. *Oncogene* **1999**, *18*, 6145–6157.

(32) Reihsaus, E.; Kohler, M.; Kraiss, S.; Oren, M.; Montenarh, M. *Oncogene* **1990**, *5*, 137–145.

(33) Barak, Y.; Juven, T.; Haffner, R.; Oren, M. *EMBO J.* **1993**, *12*, 461–468.

(34) Momand, J.; Zambetti, G. P.; Olson, D. C.; George, D.; Levine, A. J. *Cell* **1992**, *69*, 1237–1245.

(35) Uesugi, M.; Verdine, G. L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14801–14806.

(36) Yu, G. W.; Rudiger, S.; Veprintsev, D.; Freund, S.; Fernandez-Fernandez, M. R.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1227–1232.

(37) Honda, R.; Tanaka, H.; Yasuda, H. *FEBS Lett.* **1997**, *420*, 25–27.

(38) Li, M.; Brooks, C. L.; Wu-Baer, F.; Chen, D.; Baer, R.; Gu, W. *Science* **2003**, *302*, 1972–1975.

(39) Bullock, A. N.; Fersht, A. R. *Nat. Cancer Rev.* **2001**, *1*, 68–76.

(40) Roth, J. A.; Nguyen, D.; Lawrence, D. D.; Kemp, B. L.; Carrasco, C. H.; Ferson, D. Z.; Hong, W. K.; Komaki, R.; Lee, J. J.; Nesbitt, J. C.; Pisters, K. M. W.; Putnam, J. B.; Schea, R.; Shin, D. M.; Walsh, G. L.; Dolormente, M. M.; Han, C. I.; Martin, F. D.; Yen, N.; Xu, K.; Stephens, L. C.; McDonnell, T. J.; Mukhopadhyay, T.; Cai, D. *Nat. Med.* **1996**, *2*, 985–991.

(41) Seth, P.; Katayose, D.; Li, Z. W.; Kim, M.; Wersto, R.; Shanmugam, N.; Ohri, E.; Mudahar, B.; Rakkar, A. N. S.; Kadoli, P.; Cowan, K. *Cancer Gene Ther.* **1997**, *4*, 383–390.

(42) Kawabe, S.; Munshi, A.; Zumstein, L. A.; Wilson, D. R.; Roth, J. A.; Meyn, R. E. *Int. J. Radiat. Biol.* **2001**, *77*, 185–194.

(43) Kataoka, M.; Schumacher, G.; Cristiano, R. J.; Atkinson, E. N.; Roth, J. A.; Mukhopadhyay, T. *Cancer Res.* **1998**, *58*, 4761–4765.

(44) Nielsen, L. L.; Dell, J.; Maxwell, E.; Armstrong, L.; Maneval, D.; Catino, J. J. *Cancer Gene Ther.* **1997**, *4*, 129–138.

(45) Spitz, F. R.; Nguyen, D.; Skibber, J. M.; Cusack, J.; Roth, J. A.; Cristiano, R. J. *Anticancer Res.* **1996**, *16*, 3415–3422.

(46) Kubbutat, M. H. G.; Jones, S. N.; Vousden, K. H. *Nature (London)* **1997**, *387*, 299–303.

(47) Haupt, Y.; Maya, R.; Kazaz, A.; Oren, M. *Nature (London)* **1997**, *387*, 296–299.

(48) Wasylyk, C.; Salvi, R.; Argentini, M.; Dureuil, C.; Delumeau, I.; Abecassis, J.; Debussche, L.; Wasylyk, B. *Oncogene* **1999**, *18*, 1921–1934.

(49) Pavletich, N. P.; Chambers, K. A.; Pabo, C. O. *Genes Dev.* **1993**, *7*, 2556–2564.

(50) Bullock, A. N.; Henckel, J.; Fersht, A. R. *Oncogene* **2000**, *19*, 2000.

(51) Nikolova, P. V.; Henckel, J.; Lane, D. P.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14675–14680.

(52) Joerger, A. C.; Allen, M. D.; Fersht, A. R. *J. Biol. Chem.* **2004**, *279*, 1291–1296.

(53) Ang, H. C.; Joerger, A. C.; Mayer, S.; Fersht, A. R. *J. Biol. Chem.* **2006**, *281*, 21934–21941.

(54) Eisenmenger, F.; Hansmann, U. H. E. *J. Phys. Chem. B* **1997**, *101*, 3304–3310.

(55) Lwin, T. Z.; Luo, R. *Protein Sci.* **2006**, *15*, 2642.

(56) Yoda, T.; Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **2004**, *386*, 460–467.

(57) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43–56.

(58) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.

(59) Oostenbrink, C.; Villa, A.; Mark, A. E.; Gunsteren, W. F. V. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

(60) Berendsen, H. J.; Postma, J. P.; van Gunsteren, W. F.; Hermans, J. Interaction models for water in relation to protein hydration. In *Intermolecular Forces*; D. Reidel Publishing Co.: Dordrecht, 1981; pp 331−342.

(61) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(62) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

(63) Darden, T.; York, D.; Pedersn, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(64) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. *J. Chem. Phys.* **1995**, *103*, 8577–8592.

(65) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.

(66) Brooks, C. L.; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. Charmm: The energy function and its parameterization with an overview of the program. In *The Encyclopedia of Computational Chemistry*; John Wiley and Sons: Chichester, 1998; pp 271−277.

(67) Lazaridis, T.; Karplus, M. *Proteins: Struct., Funct., Genet.* **1999**, *35*, 133–152.

(68) Lee, M. S.; Salsbury, F. R.; Brooks, C. L. *J. Chem. Phys.* **2002**, *116*, 10606–10614.

(69) Neria, E.; Fischer, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 1902–1921.

(70) Inuzuka, Y.; Lazaridis, T. *Proteins* **2000**, *41*, 21–32.

(71) Hassan, S. A.; Mehler, E. L. *Proteins* **2002**, *47*, 45–61.

(72) Barth, E.; Schlick, T. *J. Chem. Phys.* **1998**, *109*, 1617–1632.

(73) Tuckerman, M.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990–2001.

(74) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.

(75) MacKerell, A. D.; Feig, M.; Brooks, C. L. *J. Am. Chem. Soc.* **2004**, *126*, 698–699.

(76) Karanicolas, J.; Brooks, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3432–3437.

(77) Nosé, S. *Mol. Phys.* **1984**, *52*, 255–268.

(78) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.

(79) Geyer, C. J. *Proceedings of the 23rd symposium on the interface*; Computing Science and Statistics Interface Foundation of North America: Fairfax Station, VA, 1991.

(80) Hansmann, U. H. E.; Okamoto, Y. *Physica A* **1994**, *212*, 415–437.

(81) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604.

(82) Feig, M.; Karanicolas, J.; Brooks, C. L. *J. Mol. Graphics Modell.* **2004**, *22*, 377–395.

(83) Sasai, M.; Ohmine, I.; Ramaswamy, R. *J. Chem. Phys.* **1991**, *96*, 3045–3053.

(84) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(85) Wallner, B.; Elofsson, A. *Protein Sci.* **2005**, *14*, 1315–1327.

(86) Bursulaya, B.; Brooks, C. L. *J. Phys. Chem. B.* **2002**, *104*, 12378–12383.

(87) Cuff, J. A.; Clamp, M. E.; Siddiqui, A. S.; Finlay, M.; Barton, G. J. *Bioinformatics* **1998**, *14*, 892–893.

(88) McGuffin, L. J.; Bryson, K.; Jones, D. T. *Bioinformatics* **2000**, *16*, 404–405.

(89) Garnier, J.; Gibrat, J. F.; Robson, B. *Methods Enzymol.* **1996**, *266*, 540–553.

(90) Guermeur, Y. Ph.D. Thesis, Université Paris 6, Paris, France, 1997.

(91) Kneller, D. G.; Cohen, F. E.; Langridge, R. *J. Mol. Biol.* **1990**, *214*, 171–182.

(92) Pollastri, G.; McLysaght, A. *Bioinformatics* **2004**, *21*, 1719–1720.

(93) Geourjon, C.; Delage, G. *Comput. Appl. Biosci.* **1995**, *11*, 681–684.

(94) Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. *Bioinformatics* **2006**, *22*, 195–201.

(95) Pollastri, G.; Przybylski, D.; Rost, B.; Baldi, P. *Proteins* **2002**, *47*, 228–235.

(96) Wang, P.; Matsukawa, S.; Kameda, T.; Kurosu, H.; Ando, I. *J. Mol. Struct.* **1997**, *435*, 229–234.

(97) Imamura, H.; Chen, J. Z. Y. *Proteins: Struct., Funct., Bioinf.* **2007**, *67*, 459–468.

CT8005387