# Revised Backbone-Virtual-Bond-Angle Potentials to Treat the L- and D-Amino Acid Residues in the Coarse-Grained United Residue (UNRES) Force Field

Adam K. Sieradzan,*,[†] Andrei Niadzvedtski,[†] Harold A. Scheraga,[‡] and Adam Liwo[†]

[†]Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland
[‡]Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301, United States

ABSTRACT: Continuing our effort to introduce D-amino-acid residues in the united residue (UNRES) force field developed in our laboratory, in this work the $C^\alpha \cdots C^\alpha \cdots C^\alpha$ backbone-virtual-bond-valence-angle (θ) potentials for systems containing D-amino-acid residues have been developed. The potentials were determined by integrating the combined energy surfaces of all possible triplets of terminally blocked glycine, alanine, and proline obtained with ab initio molecular quantum mechanics at the MP2/6-31G(d,p) level to calculate the corresponding potentials of mean force (PMFs). Subsequently, analytical expressions were fitted to the PMFs to give the virtual-bond-valence potentials to be used in UNRES. Alanine represented all types of amino-acid residues except glycine and proline. The blocking groups were either the N-acetyl and N',N'-dimethyl or N-acetyl and pyrrolidyl group, depending on whether the residue next in sequence was an alanine-type or a proline residue. A total of 126 potentials (63 symmetry-unrelated potentials for each set of terminally blocking groups) were determined. Together with the torsional, double-torsional, and side-chain-rotamer potentials for polypeptide chains containing D-amino-acid residues determined in our earlier work (Sieradzan et al. *J. Chem. Theory Comput.*, **2012**, 8, 4746), the new virtual-bond-angle (θ) potentials now constitute the complete set of physics-based potentials with which to run coarse-grained simulations of systems containing D-amino-acid residues. The ability of the extended UNRES force field to reproduce thermodynamics of polypeptide systems with D-amino-acid residues was tested by comparing the experimentally measured and the calculated free energies of helix formation of model KLALKLALxxLKLALKLA peptides, where x denotes any D- or L- amino-acid residue. The obtained results demonstrate that the UNRES force field with the new potentials reproduce the changes of free energies of helix formation upon D-substitution but overestimate the free energies of helix formation. To test the ability of UNRES with the new potentials to reproduce the structures of polypeptides with D-amino-acid residues, an ab initio replica-exchange folding simulation of thurincin H from *Bacillus thuringiensis*, which has D-amino-acid residues in the sequence, was carried out. UNRES was able to locate the native α-helical hairpin structure as the dominant structure even though no native sulfide−carbon bonds were present in the simulation.
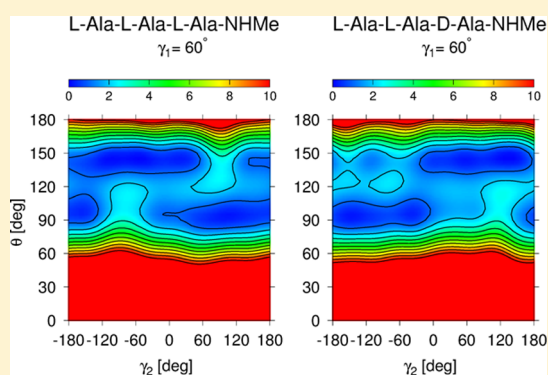
1. INTRODUCTION

D-amino acids are currently of great interest in drug design.[1−3] D-amino acid substitution can increase protein resistance to proteolysis[4] and can enhance the thermodynamic stability[5] of proteins. This idea has been used in designing CD4 analogues as potential HIV binding domain inhibitors.[1] Moreover, D-amino acids can increase specificity of, e.g., D-tyrosine conjugated with naproxen.[2]

Despite increases in computing power, all-atom molecular dynamic simulations are still limited to either small system size[6] or short time scale.[7] Therefore, coarse-grained models of proteins continue to play a great role in protein simulations,[8,9] including computer-aided drug design.[10] The coarse-grained treatment speeds up the time scale of simulations by 3−4 orders of magnitude compared to all-atom treatment with explicit solvent.[11,12] However, most of the coarse-grained force fields that are capable of predicting the structures of proteins are

knowledge-based and, naturally, treat L-amino-acid residues rather than D-amino acid residues, because of an insufficient number of D-amino-acid residues in protein structural databases to derive the respective statistical potentials. By contrast, the united residue (UNRES) model and force field of polypeptide chains,[13−25] which we have been developing for over two decades in our laboratory, is physics-based and can easily be extended to include D-amino-acid residues, because it does not depend on structural databases. The capability of UNRES to predict protein structures is being continuously assessed, since 1998, in the consecutive community wide experiments of the Critical Assessment of Techniques for Protein Structure Prediction (CASP) with very good results.[26−31] In particular, in the recent

CASP10 exercise, UNRES, as one of only two methods, predicted correct domain packing of target T0663.[31]

In our previous work,[32] we introduced a general treatment of D-amino-acid residues in UNRES. We demonstrated that the potentials that depend on single amino-acid residue types as far as local interactions are concerned, namely, the side-chain rotamer and third-order backbone-local–backbone-electrostatic correlation potentials can be obtained from the respective alanine-type L-residue potentials by applying the inversion operation. Conversely, we found that only approximate symmetry relations exist between the local potentials that depend on pairs (the torsional potentials) or triplets (double-torsional and virtual-bond-angle-potentials) of consecutive amino-acid residues. In our previous work,[32] we determined the physics-based torsional and double-torsional potentials that involve L- and D-amino-acid residues and applied approximate symmetry relations to estimate the virtual-bond-angle potentials. With these modifications, the UNRES model enabled us to simulate folding pathways of the BBAT1 protein[33] and stability of gramicidin D.[32] The determination of the exact virtual-bond-angle potentials for polypeptides and proteins containing D-amino-acid residues had then been postponed to further work because of substantial effort required to accomplish this task.

In the present work, using the ab initio potential-energy surfaces of terminally blocked glycine, alanine, and proline determined in our previous work[32] to calculate the respective potentials of mean force, we determine the physics-based virtual-bond-angle potentials for backbone sections that include D-amino-acid residues. This work concludes the extension of UNRES to polypeptides and proteins with D-amino-acid residues, which have not been determined in our earlier work.[32] We use the harmonic-extrapolation method developed in our first paper on the determination of virtual-bond-angle potentials,[20] where we treated only L-amino-acid residues. Moreover, because of limited computational resources available at that time, in that work, energy minimization was carried out with the RHF scheme with single-point MP2 correction, and the N,N'-dimethyl group was used to mimic the link of a residue to the proline residue at the C-terminus. Therefore, in this work, we also revised the virtual-bond-valence angle potentials for systems composed of L-amino-acid residues that were determined in ref 20. We demonstrate the performance of the modified UNRES force field with the examples of thermodynamics of helix formation in model peptides and folding of thurincin H, a natural peptide that contains D-amino-acid residues.

2. METHODS

2.1. UNRES Model of the Polypeptide Chain. In the UNRES model, a polypeptide chain is represented as a sequence of α-carbon (C^{α}) atoms with attached united side chains (SC's) and united peptide groups (p's) positioned halfway between two consecutive C^{α}s. Only the united side chains and united peptide groups act as interaction sites, while the C^{α} atoms assist only in the definition of geometry (Figure 1). The effective energy function is expressed by eq 1. This effective energy function originates from the restricted free energy (RFE) or the potential of mean force (PMF) of the chain constrained to a given coarse-grained conformation along with the surrounding solvent.[16,17,21] The PMF is then expanded into factors,[17] which are the cluster-cumulant functions introduced by Kubo.[34] Factors of order m are PMF components that arise from m groups of atomic-detailed interactions, each of which comprises interactions between the atoms that belong to a given coarse-grained site or to two coarse-
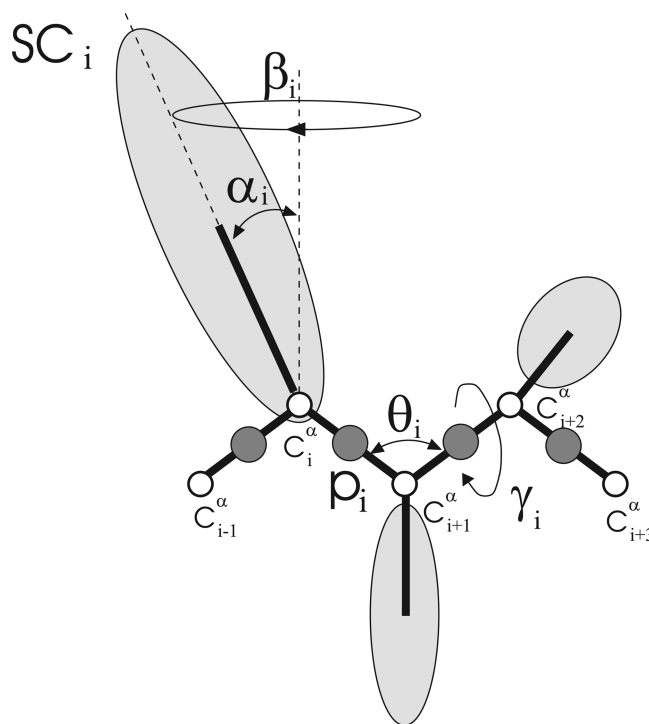
Figure 1. UNRES model of polypeptide chains. The interaction sites are peptide-group centers (p), and side-chain centers (SC) attached to the corresponding α-carbons with different $C^{\alpha} \cdots SC$ bond lengths, d_{SC}. The peptide groups are represented as dark gray circles and the side chains are represented as light gray ellipsoids of different size. The α-carbon atoms are represented by small open circles. The geometry of the chain can be described either by the virtual-bond vectors $\mathbf{dC}_i$ (from C^{α}_i to C^{α}_{i+1}), $i = 1,2,...,n-1$, and $\mathbf{dX}_i$ (from C^{α}_i to SC_i), $i = 2,...,n-1$, represented by thick lines, where n is the number of residues, or in terms of virtual-bond lengths, backbone virtual-bond angles θ_i, $i = 1,2,...,n-2$, backbone virtual-bond-dihedral angles γ_i, $i = 1,2,...,n-3$, and the angles α_i and β_i, $i = 2,3,...,n-1$ that describe the location of a side chain with respect to the coordinate frame defined by C^{α}_{i-1}, C^{α}_i, and C^{α}_i, C^{α}_{i+1}.

grained sites. For example, the side-chain–side-chain inter-actions are factors of order 1, whereas the torsional potentials are factors of order 2.

$$
\begin{aligned}
U = {}& w_{SC} \sum_{i<j} U_{SC_i SC_j} + w_{SCp} \sum_{i \neq j} U_{SC_i p_j} + w_{pp}^{VDW} \sum_{i<j-1} U_{p_i p_j}^{VDW} \\
& + w_{pp}^{el} f_2(T) \sum_{i<j-1} U_{p_i p_j}^{el} + w_{tor} f_2(T) \sum_i U_{tor}(\gamma_i) \\
& + w_{tord} f_3(T) \sum_i U_{tord}(\gamma_i, \gamma_{i+1}) + w_b \sum_i U_b(\theta_i) \\
& + w_{rot} \sum_i U_{rot}(\alpha_{SC_i}, \beta_{SC_i}) + w_{bond} \sum_i U_{bond}(d_i) \\
& + w_{corr}^{(3)} f_3(T) U_{corr}^{(3)} + w_{corr}^{(4)} f_4(T) U_{corr}^{(4)} + w_{turn}^{(3)} f_3(T) U_{turn}^{(3)} \\
& + w_{turn}^{(4)} f_4(T) U_{turn}^{(4)} + w_{SC-corr} f_2(T) \\
& \times \sum_{n=1}^{3} \sum_i U_{SC-corr}(\tau_i^{(n)})
\end{aligned}
\tag{1}
$$

where the U's are energy terms, θ_i is the backbone virtual-bond angle, γ_i is the backbone virtual-bond-dihedral angle, α_i and β_i are the angles defining the location of the center of the united side chain of residue i (Figure 1), and d_i is the length of the ith virtual bond, which is either a $C^{\alpha} \cdots C^{\alpha}$ virtual bond or $C^{\alpha} \cdots SC$ virtual

bond; $\tau^{(1)}$ is the $SC_i \cdots C_i^\alpha \cdots C_{i+1}^\alpha \cdots C_{i+2}^\alpha$ torsional angle, $\tau^{(2)}$ is the $C_{i-1}^\alpha \cdots C_i^\alpha \cdots C_{i+1}^\alpha \cdots SC_{i+1}$ torsional angle, and $\tau^{(3)}$ is the $SC_i \cdots C_i^\alpha \cdots C_{i+1}^\alpha \cdots SC_{i+1}$ torsional angle.[35] Each energy term is multiplied by an appropriate weight, w_x (where x stands for the interactions contained in the respective energy term); these quantities are determined by force-field calibration[19,36] (see the end of this section for more information). The terms corresponding to factors of order higher than 1 in the cluster-cumulant expansion of the PMF are additionally multiplied by the respective temperature factors that were introduced in our earlier work[19] and which reflect the temperature dependence of the first generalized-cumulant term in those factors, as discussed in refs 19 and 37. The factors f_n are defined by eq 2.

$$f_n(T) = \frac{\ln(e + 1/e)}{\ln\{\exp[(T/T_o)^{n-1}] + \exp[-(T/T_o)^{n-1}]\}} \quad (2)$$

where $T_o = 300$ K.

The term $U_{SC_iSC_j}$ represents the mean free energy of the hydrophobic (hydrophilic) interactions between the side chains, which implicitly contains the contributions from the interactions of the side chain with the solvent. The term $U_{SC_ip_j}$ denotes the excluded-volume potential of the side-chain−peptide-group interactions. The peptide-group interaction potential is split into two parts: the Lennard-Jones interaction energy between peptide-group centers ($U_{p_ip_j}^{VDW}$) and the average electrostatic energy between peptide-group dipoles ($U_{p_ip_j}^{el}$); the second of these terms accounts for the tendency to form backbone hydrogen bonds between peptide groups p_i and p_j. The terms U_{tor}, U_{tord}, U_b, U_{rot}, and U_{bond} are the virtual-bond-dihedral angle torsional terms, virtual-bond dihedral angle double-torsional terms, virtual-bond angle bending terms, side-chain rotamer, and virtual-bond-deformation terms; these terms account for the local properties of the polypeptide chain. For U_{tor}, U_{tord}, U_b, three types of residues are distinguished: alanine, glycine, and proline where alanine describes all amino acid types except glycine and proline. The terms $U_{corr}^{(m)}$ represent correlation or multibody contributions from the coupling between backbone-local and backbone-electrostatic interactions, and the terms $U_{turn}^{(m)}$ are correlation contributions involving m consecutive peptide groups; they are, therefore, termed turn contributions. The multibody terms are indispensable for reproduction of regular α-helical and β-sheet structures.[16,17,38] $U_{SC-corr}$ are correlation potentials between local conformations of side-chain and backbone. These potentials significantly improve the description of loop structures.[35]

The energy-term weights are determined by force-field calibration to reproduce the structure and folding thermodynamics of selected training proteins.[19,36] In this work, we used the energy-term weights from the force field calibrated in our earlier work[36] with the tryptophan cage (PDB code: 1L2Y)[39] and the tryptophan zipper 2 (PDB code: 1LE1).[40] It should be noted that the force field of our previous work[36] was calibrated with peptides that contain only L-amino-acid residues. The hierarchical optimization procedure was used[19,36] whose goal is to achieve a free-energy landscape in which the free energy decreases with increasing the native likeness below the folding-trasition temperature and free energy increases with increasing the native likeness above the folding-transition temperature.

2.2. Determination of the New U_b Potentials. The procedure developed in our earlier work[20] was used to determine new virtual $C^\alpha \cdots C^\alpha \cdots C^\alpha$ valence angle θ (Figure 2) bending
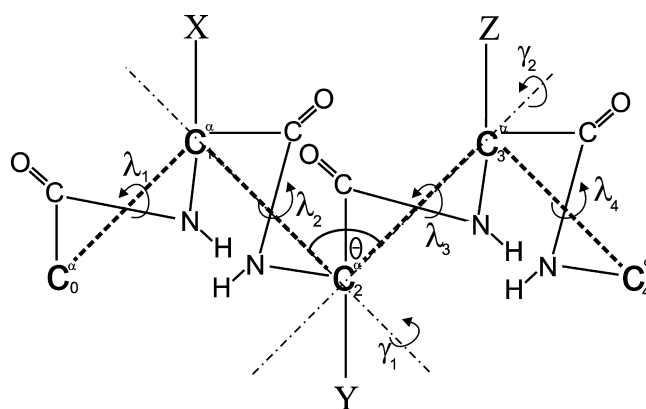
Figure 2. Illustration of the model terminally blocked tripeptides constructed to compute the integrals of eqs 3. Each X, Y, and Z denotes side-chains of L-Ala, D-Ala, Gly, L-Pro, or D-Pro.

potentials (U_b). In that procedure, the respective potentials of mean force (PMFs) were first calculated from nonadiabatic energy maps of terminally blocked amino-acid residues, and then a three-dimensional Fourier series was fitted to them. Energy maps had been computed[20] in the angles of rotation $\lambda^{(1)}$ and $\lambda^{(2)}$ about the virtual-bond $C^\alpha \cdots C^\alpha$ axes; these angles have been defined in ref 41 (Figure 3). Energy was minimized with respect to all degrees of freedom except for $\lambda^{(1)}$ and $\lambda^{(2)}$. Therefore, energy maps are minimized in the θ angle.
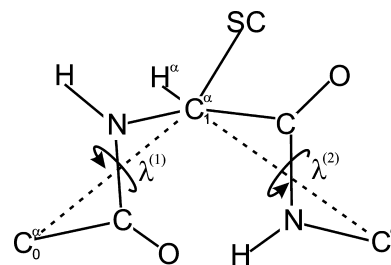
Figure 3. Definition of the dihedral angles $\lambda^{(1)}$ and $\lambda^{(2)}$ for rotation of a peptide unit about the $C^\alpha \cdots C^\alpha$ virtual bonds of a peptide unit.[41]

To calculate the PMF on a θ grid, harmonic extrapolation has to be used,[20] as expressed by eq 3 (which has the same form as eq 15 in ref 20); it should be noted that eq 3 includes a harmonic-entropy contribution.

$$U_{b,XYZ}(\theta, \gamma_1, \gamma_2) = -\beta^{-1}\ln\{2^{-4}\pi^{-4+N/2}V_{y_2'}^{-1}$$

$$\times \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}(\det \mathbf{H}_{x_2'x_2'}^* \det \mathbf{H}_{y_2'y_2'}^* \det \mathbf{H}_{z_2'z_2'}^*)^{-1/2}$$

$$\times \exp\{-\beta[e_X^* + e_Y^* + e_Z^* + \frac{1}{2}\left(\mathbf{H}_{\theta^*\theta^*} - \frac{1}{4}\mathbf{H}_{\theta^*y_2'}^*\mathbf{H}_{y_2'y_2'}^{*-1}\mathbf{H}_{\theta^*y_2'}^{*T}\right)$$

$$\Delta\theta^{*2}]\}d\lambda_1 d\lambda_2 d\lambda_3 d\lambda_4\} + \beta^{-1}\ln\{(2\pi)^{-4}$$

$$\times \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}(\det \mathbf{H}_{x_2'x_2'}^* \det \mathbf{H}_{y_2'y_2'}^* \det \mathbf{H}_{z_2'z_2'}^*)^{-1/2}$$

$$\times \exp\{-\beta[e_X^* + e_Y^* + e_Z^*]\}d\lambda_1 d\lambda_2 d\lambda_3 d\lambda_4\} \quad (3)$$

where $U_{b,XYZ}$ denotes the bending potentials of mean force, X, Y, and Z denote L-alanine L-proline, D-alanine, D-proline, or glycine, where alanine represents all amino acids except proline or glycine, $\beta = (RT)^{-1}$, T is the absolute temperature, R is the universal gas constant, $\mathbf{x}_2'$, $\mathbf{y}_2'$, and $\mathbf{z}_2'$ are the vectors of variables other than θ, $\lambda^{(1)}$, and $\lambda^{(2)}$ of residues X, Y, and Z, respectively. $V_{y_2'}$ is the volume of the space spanned by the secondary variables

pertaining to the second residue (Y), contained in the $\mathbf{y}'_2$ vector, $\mathbf{H}$ denotes the Hessian matrix (matrix of the second derivatives of energy in geometrical variables); the terms with $\mathbf{H}$ contain the harmonic-entropy contributions, and the asterisks indicate the values corresponding to the points on the nonadiabatic energy maps. $\lambda_1 - \lambda_4$, shown in Figure 3, were determined from eqs 4a–f:

$$\lambda_X^{(1)} = \lambda_1 \tag{4a}$$

$$\lambda_X^{(2)} = \gamma_1 - \pi - \lambda_2 \tag{4b}$$

$$\lambda_Y^{(1)} = \lambda_2 \tag{4c}$$

$$\lambda_Y^{(2)} = \gamma_2 - \pi - \lambda_3 \tag{4d}$$

$$\lambda_Z^{(1)} = \lambda_3 \tag{4e}$$

$$\lambda_Z^{(2)} = \lambda_4 \tag{4f}$$

where the angles $\lambda^{(1)}$ and $\lambda^{(2)}$ are shown in Figure 3.

The second multiple integral in eq 3 corresponds to a sum of torsional and double-torsional potentials and must be subtracted from the PMF due to local interactions in tripeptide systems sketched in Figure 2, consistent with the definition of the PMF factors in UNRES,[17] because the torsional and double-torsional potentials constitute separate UNRES energy terms. These energy terms were determined for systems that include D-amino-acid residues in our earlier work,[32] and here in eq 3.

After determination of $U_{b,XYZ}$, a three-dimensional Fourier series (eq 5) was used to fit the free energy surfaces. As variables, γ_1, γ_2, and $\theta/2$ were used. The use of $\theta/2$ instead of θ is justified by the fact that the squares of interatomic distances within a peptide unit are naturally expressed in terms of powers of $\cos(\theta/2)$ and $\sin(\theta/2)$ as shown in our previous work.[20]

$$
\begin{aligned}
U_b(\theta, \gamma_1, \gamma_2) = a_0 &+ \sum_{i=1}^{10} a_i \sin(i\theta/2) \\
&+ \sum_{i=1}^{4} \sum_{j=1}^{6} \sin(i\theta/2)[b_{ij}\cos(j\gamma_1) + c_{ij}\sin(j\gamma_1) \\
&+ d_{ij}\cos(j\gamma_2) + e_{ij}\sin(j\gamma_2)] + \sum_{i=1}^{4}\sum_{j=2}^{6}\sum_{k=1}^{m-1} \sin(i\theta/2) \\
&\times \{ f_{ijk}\cos[j\gamma_1 + (k-j)\gamma_2] + f_{ijk}\cos[j\gamma_1 - (k-j)\gamma_2] \\
&+ g_{ijk}\sin[m\gamma_1 + (k-j)\gamma_2] + g_{ijk}\sin[j\gamma_1 - (k-j)\gamma_2] \}
\end{aligned}
\tag{5}
$$

In our previous work,[20] we computed the energy maps of terminally blocked amino-acid residues by using ab initio molecular quantum mechanics with the 6-31G(d,p) basis set. Because less powerful resources were available at that time, the energy was minimized in the Restricted Hartree–Fock (RHF) scheme and, subsequently, the single-point Møller–Plesset (MP2) energy correction was computed for the energy-minimized conformation. Previously, the limited resources enabled us to calculate Hessian matrices only in the RHF scheme. Moreover, interfacing a residue to the following proline residue was modeled by blocking it with an N',N'-dimethyl group. In this work, all energy maps were recalculated by minimizing the energy in the MP2/6-31G(d,p) scheme; additionally, we used the C-terminal pyrrolidine group and not

the N',N'-dimethylamine blocking group for residues preceding proline. Consequently, all diabatic energy maps of *trans-N*-acetyl-alanyl-N'-methylamide, *trans-N*-acetyl-alanyl-N'-pyrrolidyla-mide, *trans-N*-acetyl-glycyl-N'-methylamide, *trans-N*-acetyl-glyc-yl-N'-pyrrolidylamide *trans-N*-acetyl-prolyl-N'-methylamide, and *trans-N*-acetyl-prolyl-N'-pyrrolidylamide were computed. The angles $\lambda^{(1)}$ and $\lambda^{(2)}$ (Figure 3) were used as grid variables; both varied from $-180°$ to $+180°$ with a $15°$ step. The energy was minimized with respect to all other degrees of freedom. The Hessians (second-derivative matrices or force-constant matrices) were also calculated in the MP2 scheme at every grid point to determine the harmonic-entropy contribution and harmonic approximation to the PMF value (eq 3).

2.3. Thermodynamic Stability of α-Helix Formation in KLALKLALxxLKLALKLA Peptides. In our previous work[32,33] and also in this work (section 3.3), we assessed the ability of the modified UNRES force field to reproduce the structure of proteins with D-amino-acid residues. However, it is equally important to assess if the force field can distinguish different chirality by energy. To assess whether the new potentials can reproduce the change of the energetics of local interactions upon D-substitution, we studied a set of 39 peptides KLALKLALxxL-KLALKLA where xx denote two consecutive natural amino-acid residues.[42] All coded L-amino-acid residues and their D-amino-acid counterparts as well as glycine were considered. For these peptides, the free-energy differences related to the formation of α-helical structure were determined from molar ellipticities at $\lambda = 222$ nm and reversed-phase high pressure liquid chromatog-raphy.[42] The experimental values of free energies could thus be compared with the results of our simulations. To carry out simulations, multiplexed replica-exchange molecular dynamics (MREMD), as implemented in version 3.2 of UNRES (available at www.unres.pl),[12,43] was used. The Berendsen thermostat[44] was used to control temperature. The equations of motions were integrated by using the variable time step (VTS) algorithm[11] as implemented in the UNRES program. Each simulation was performed at the following 16 temperatures with 2 copies at each temperature: 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, and 400 K. Each simulation lasted about 50 ns (approximately 10 million steps, with a step length of 4.89 fs). The weighted histogram analysis method (WHAM)[19,45] was used to evaluate the thermodynamic quantities and conforma-tion-dependent ensemble averages, as implemented in our earlier work.[19]

The Gibbs free-energy changes ($\Delta\Delta G$) at $T = 300$ K were calculated from the fractions of amino-acid residue in the α-helical state at this temperature, averaged (using WHAM) over conformations. For each conformation, the fraction of residues in the α-helical state was evaluated using the following two methods. In the first approach, the coarse-grained UNRES structure was converted to all-atom representation with the use of our energy-based method, based on the optimal alignment of peptide-group dipoles to reconstruct the backbone[46] and optimal side-chain packing to reconstruct all-atom side chains from the positions of side-chain centroids.[47] A residue was considered to be in the α-helical state if its ϕ and ψ angles were in the A region of the Ramachandran map, as defined by Zimmerman et al.[48] ($\phi = -75° \pm 40°$ and $\psi = -40° \pm 35°$). In the second approach, a residue was considered to be in the α-helical state if the peptide group preceding it formed a hydrogen-bonding contact with the third succeeding peptide group; the presence of a hydrogen-bonding contact was assessed based on the mean-field energy of interactions, which depends on the distance between the centers
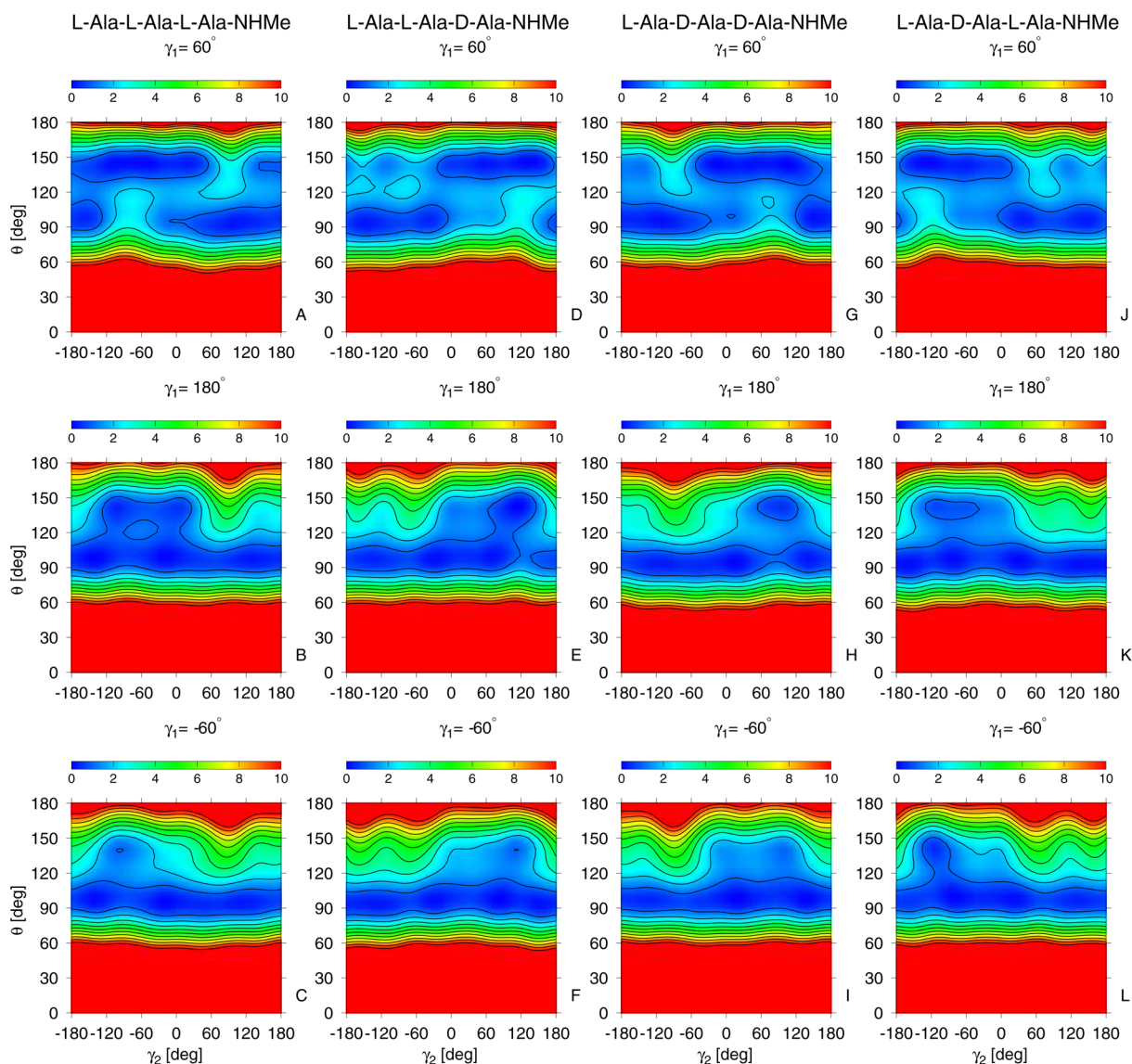
Figure 4. Sample contour plots of the valence bond bending potentials $U_{L-Ala-(D,L)-Ala-(D,L)-Ala-NHMe}(\theta, \gamma^{(1)}, \gamma^{(2)})$ as functions of θ and $\gamma^{(2)}$ angles for alanine-type tripeptides with NHMe terminal groups, where (D,L)-Ala indicates a D- or an L-Ala residue, for three selected values of the $\gamma^{(1)}$ angle =60°, 180°, −60°. γ_1 is always fixed and its value is printed at the top of each panel; γ_2 is a variable. Finally, the energy scale (kcal/mol) is on the top of the figure.

of the two peptide groups; the details of this method are described in ref 49. The free energies of α-helical-structure formation, $\Delta G^{hel}(x)$, were calculated from eq 6. For comparison with experimental values,[42] we calculated the Gibbs free energies of α-helix formation expressed relative to glycine, $\Delta\Delta G^{hel}(Gly \rightarrow x)$, where x = D-x or L-x, and the differences between the free energies of the D- and L-amino acid substituted host−guest peptides, $\Delta\Delta G^{hel}(L-x \rightarrow D-x)$. These quantities are expressed by eqs 7 and 8, respectively.

$$\Delta G^{hel}(x) = -RT\ln \frac{f_{hel}(x)}{1 - f_{hel}(x)} \tag{6}$$

$$\Delta\Delta G^{hel}(Gly \rightarrow x) = \Delta G^{hel}(x) - \Delta G^{hel}(Gly) \tag{7}$$

$$\Delta\Delta G^{hel}(D-x \rightarrow L-x) = \Delta G^{hel}(D-x) - \Delta G^{hel}(L-x) \tag{8}$$

where R is the universal gas constant, T is the absolute temperature, and $f_{hel}(x)$ is the ensemble-averaged fraction of α-helical structures in the ensemble for the host−guest peptide

containing a pair of specific residues x. The temperature T was set at 298 K, as in the experiment.[42]

2.4. Simulation of Folding Pathway of Thurincin H. To evaluate the predictive power of the UNRES force field extended to treat D-amino-acid residues, a small peptide thurincin H[50] (PDB code: 2LBZ), which contains D-allo-threonine residues, was selected. The sequence of this peptide is L-Asp-L-Trp-L-Thr-L-Cys-L-Trp-L-Ser-L-Cys-L-Leu-L-Val-L-Cys-L-Ala-L-Ala-L-Cys-L-Ser-L-Val-L-Glu-L-Leu-L-Leu-D-Asn-L-Leu-L-Val-D-allo-Thr-L-Ala-L-Ala-D-allo-Thr-Gly-L-Ala-D-Ser-L-Thr-L-Ala-L-Ser. Thurincin H contains the following four sulfur−C$^\alpha$ bonds S(Cys^4)−C$^\alpha$(D-Ser^{28}), S(Cys^7)−C$^\alpha$(D-allo-Thr^{25}), S(Cys^{10})−C$^\alpha$(D-allo-Thr^{22}), and S(Cys^{13})−C$^\alpha$(D-Asn^{19}), where the sulfur is bonded to the C$^\beta$ atom of Cys and to the C$^\alpha$ atom of the D-residue. Such sulfide bonds are rare in natural peptides and proteins, as opposed to the frequently occurring disulfide bonds, which makes this target even more interesting to study. The simulation was started from the all-extended conformation, with all virtual-bond-dihedral angles γ set at 180°. The UNRES version and technical details were those described in section 2.3 except that
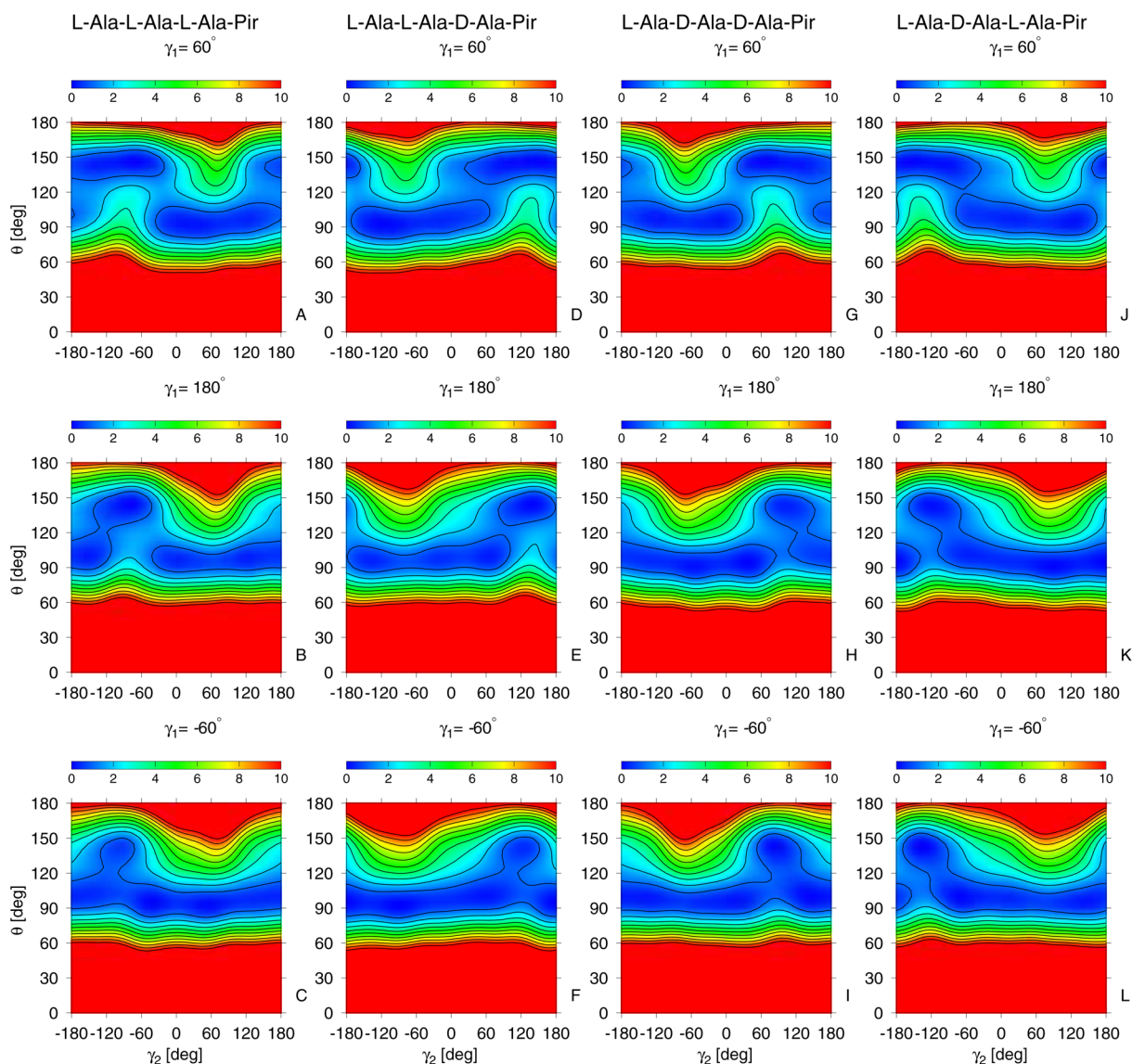
Figure 5. Same as Figure 4, but for (D,L)-Ala-Pir.

the duration of the simulation was 300 ns (about 60 million steps) to ensure the convergence of the heat-capacity curve.

Because the difference between allo-threonine and threonine is too small at the coarse-grained level to warrant the development of separate side-chain-interaction parameters for allo-threonine for UNRES, D-threonine was used in the simulation instead of D-allo-threonine. Although thurincin H contains four sulfur–C^{α} bonds, no restraints were imposed on the S–C^{α} distances in the simulation and, consequently, no information about the structure of the peptide was input. 2000 last snapshots from each trajectory (snapshot frequency being 20 000 steps) were analyzed as described in our earlier work.[19] First, WHAM was run to compute the heat-capacity curve and the weights of conformations at the selected temperature, T_a, (10 K below the heat-capacity maximum; in this study T_a = 300 K), then cluster analysis with the use of Ward's minimum variance method[51] was carried out at T_a to partition the set of conformations into families and the families were ranked according to decreasing fraction in the ensemble.

3. RESULTS AND DISCUSSION

3.1. Virtual-Bond-Angle Potentials. A total of 125 virtual-valence-angle-bending potentials, of which 63 are not related by symmetry to each other, were determined for systems with each of the C-terminally blocking group (NHMe or Pir). These numbers are obtained as follows. In each position of the XYZ triplet, Gly, L-Ala, D-Ala, L-Pro or D-Pro can appear, which provides 5 possibilities. This gives 5 × 5 × 5 = 125 possible PMF surfaces. Of these, one corresponds to the Gly-Gly-Gly triplet, the other 124 involve at least one nonglycine residue. Those 124 consist of symmetry-related pairs obtained by flipping the chirality of all residues involved and, consequently, the corresponding pairs of PMF surfaces are related by the change of the signs of both the γ_1 and γ_2 angles. Thus, there are 124/2 + 1 = 63 symmetry-unrelated PMF surfaces.

Maps of the virtual-bond-angle potentials in the γ_2 and θ angles (see Figure 2 for the definition of the θ, γ_1, and γ_2 angles) are displayed in Figures 4A−L and 5A−L, respectively, for the following 8 selected systems: L-Ala-L-Ala-L-Ala-NHMe (Figure 4A−C), L-Ala-L-Ala-D-Ala-NHMe (Figure 4D−F), L-Ala-D-Ala-D-Ala-NHMe (Figure 4G−I), L-Ala-D-Ala-L-Ala-NHMe (Figure

Table 1. Calculated and Experimental Gibbs Free Energy Differences of α-Helical Structure Formation in the KLALKLALxxLKLALKLA Host−Guest Peptides for the D- $[\Delta\Delta G^{hel}(\text{Gly} \rightarrow \text{D-x})]$ and the L-Amino-Acid Residues $[\Delta\Delta G^{hel}(\text{Gly} \rightarrow \text{L-x})]$ Expressed Relative to Glycine, And the Free-Energy Differences of α-Helical Structure Formation between the D- and L-Amino-Acid Substituted Host−Guest Peptides $[\Delta\Delta G^{hel}(\text{L-x} \rightarrow \text{D-x})]$ at 300 K

residue	$\Delta\Delta G^{hel}(\text{Gly} \rightarrow \text{D-x})^a$			$\Delta\Delta G^{hel}(\text{Gly} \rightarrow \text{L-x})^a$		$\Delta\Delta G^{hel}(\text{L-x} \rightarrow \text{D-x})^b$			
name	contacts^c	all-atom^d	exp^e	contacts^c	all-atom^d	contacts^c	all-atom^d	exp^e	exp^f
Pro	1.44	0.45	6.04	0.48	−0.25	0.96	0.70		−0.6
Lys	−0.69	−0.95	3.44	−3.33	−3.30	2.64	2.35		4.1
Arg	−1.01	−1.21	2.87	−3.48	−3.39	2.48	2.17		4.2
His	−1.00	−1.24	−0.63	−3.55	−3.42	2.54	2.18	1.76	1.6
Asp	−0.90	−1.12	−0.32	−3.49	−3.28	2.58	2.16	0.66	1.8
Glu	−1.12	−1.19	1.48	−3.48	−3.41	2.36	2.23		2.6
Asn	−0.83	−1.19	1.76	−3.36	−3.34	2.54	2.15		3.2
Gln	−1.00	−1.20	1.61	−3.50	−3.43	2.49	2.23		2.2
Ser	−0.99	−1.23	1.97	−3.31	−3.42	2.31	2.19		3.4
Thr	−0.89	−1.21	4.74	−3.65	−3.40	2.76	2.18		6.9
Ala	−1.00	−1.14	4.38	−3.30	−3.38	2.30	2.23		4.2
Tyr	−1.37	−1.39	5.68	−3.75	−3.44	2.37	2.06	4.52	3.8
Trp	−1.02	−1.19	3.84	−3.58	−3.47	2.56	2.28	4.69	4.1
Val	−0.77	−0.82	4.86	−3.61	−3.36	2.84	2.54	5.95	7.1
Leu	−0.99	−1.13	2.77	−3.61	−3.47	2.62	2.33		3.0
Ile	0.06	−0.43	5.11	−4.01	−3.46	4.07	3.03	5.84	6.0
Phe	−1.08	−1.16	4.82	−3.84	−3.48	2.76	2.32		3.1
Cys	0.08	−0.69	1.58	−1.50	−3.42	1.58	2.72		1.0
Met	−0.89	−1.18	2.95	−3.73	−3.55	2.84	2.37		2.8

^{a}See eq 7 for definition. ^{b}See eq 8 for definition. ^{c}Calculated based on peptide-group contacts; see section 2.3. ^{d}Calculated based on the backbone ϕ and ψ angles obtained by conversion of united-residue polypeptide chains to all-atom chains by using our energy-based procedure;[46,47] see section 2.3. ^{e}Based on molar ellipticity; Tables 1 and 2 of ref 42. ^{f}Based on chromatographic capacity factors determined by RP-HPLC (the bar heights in Figure 5 of ref 42).

4J−L), L-Ala-L-Ala-L-Ala-Pir (Figure 5A−C), L-Ala-L-Ala-D-Ala-Pir (Figure 5D−F), L-Ala-D-Ala-D-Ala-Pir (Figure 5G−I), and L-Ala-D-Ala-L-Ala-Pir (Figure 5J−L), and for the following selected values of γ_1: $\gamma_1 = 60°$ (Figures 4A,D,G,J and 5A,D,G,J), $\gamma_1 = 180°$ (Figures 4B,E,H,K and 5B,E,H,K) and $\gamma_1 = 60°$ (Figures 4C,F,I,L and 5C,F,I,L). The angle γ_2 and not γ_1 was chosen as the second variable because the U's depend on this angle more strongly than on γ_1, as already concluded by Levitt[52] based on a statistical analysis of protein structures. Figure 4 shows all symmetry-unrelated PMF surfaces for alanine-type residues. As already mentioned in this section, the PMF surfaces for the other alanine-type triplets can be obtained from those by inversion in the γ_1 and γ_2 angles. It can be seen that two broad regions of minima (blue color) appear in the PMF surfaces of all 8 model systems shown in Figures 4A−L and 5A−L. The first region occurs around $\theta = 90°$, a value characteristic of α-helical or turn structures, and the second one is centered around $\theta = 140°$, a value characteristic of extended or β-sheet structures. The sizes of the $\theta = 90°$ region are always comparable regardless of the values of γ_1, whereas the $\theta = 140°$ region is the largest for $\gamma_1 = 60°$ (or $\gamma_1 = -60°$ for the symmetry-related surfaces). For $\gamma_1 = 180°$, the two regions nearly merge and there is only a small free-energy barrier between them. For $\gamma_1 = 60°$, there is a noticeable barrier of ≈2 kcal/mol. For example, for a potential of mean force for $U_{\text{L-Ala−L-Ala−L-Ala}}$ (Figure 4A−C and 5A−C), when $\gamma_1 = 180°$ and $\theta = 140°$, the free-energy difference between $\gamma_2 = 0°$ and $\gamma_2 = 75°$ is ≈8 kcal/mol.

Because of a weak dependence of U on γ_1 and a strong dependence on γ_2, $U_{\text{L-Ala−L-Ala−L-Ala}}$ (Figures 4A−C and 5A−C) and $U_{\text{L-Ala−D-Ala−L-Ala}}$ (Figures 4J−L and 5J−L) are similar to each other whereas $U_{\text{L-Ala−L-Ala−L-Ala}}$ (Figures 4A−C and 5A−C) is almost a mirror image of $U_{\text{L-Ala−L-Ala−D-Ala}}$ (Figures 4D−F and 5D−F) and $U_{\text{L-Ala−D-Ala−D-Ala}}$ (Figures 4G−I and 5G−I). This

observation is universal for each amino acid type for all γ_1, γ_2. However, this pseudosymmetry is not a full symmetry and does not follow the symmetry eq 9 as:

$$U_{\text{b};\overline{XYZ}}(\gamma_1, \gamma_2, \theta) = U_{\text{b};XYZ}(-\gamma_1, -\gamma_2, \theta) \tag{9}$$

where X, Y, and Z denote amino-acid-residue types of a given chirality (L or D), and a bar over a symbol denotes the change of enantiomer chirality; this operation does not affect glycine.

Comparison of the PMF surfaces corresponding to systems with the NHMe C-terminal blocking groups (panels A−L of Figure 4) with their counterparts with the Pir C-terminal blocking groups (panels A−L of Figure 5) reveals that the potentials are similar to each other; however, the regions of minima are narrower for the Pir C-terminal blocking group. Therefore, if the residue next in sequence is a proline residue, this restricts the range of the available θ angle slightly. This tendency is the same as for the torsional and double-torsional potentials.[32] It is, however, less noticeable because the virtual-bond-angle potentials are potentials corresponding to higher-order terms of the cluster-cumulant expansion of the PMF.

3.2. Thermodynamic Stability of α-Helix Formation in KLALKLALxxLKLALKLA Peptides. The calculated and experimental[42] free-energy differences of α-helix formation of the host−guest peptides with general sequence KLALKLALxxL-KLALKLA (where x denotes a natural L-amino-acid residue or its D-counterpart), expressed relative to glycine (eq 7), as well as the free-energy differences of helix formation in the L- and D-amino-acid-substituted peptides (eq 8), are summarized in Table 1. It can be seen from Table 1 that the two methods used to estimate the fraction of α-helical structure described in section 2.3 resulted in similar $\Delta\Delta G^{hel}(\text{Gly} \rightarrow x)$ and $\Delta\Delta G^{hel}(\text{L-x} \rightarrow \text{D-x})$, in the respective "contact" and "all-atom" columns, which suggests that
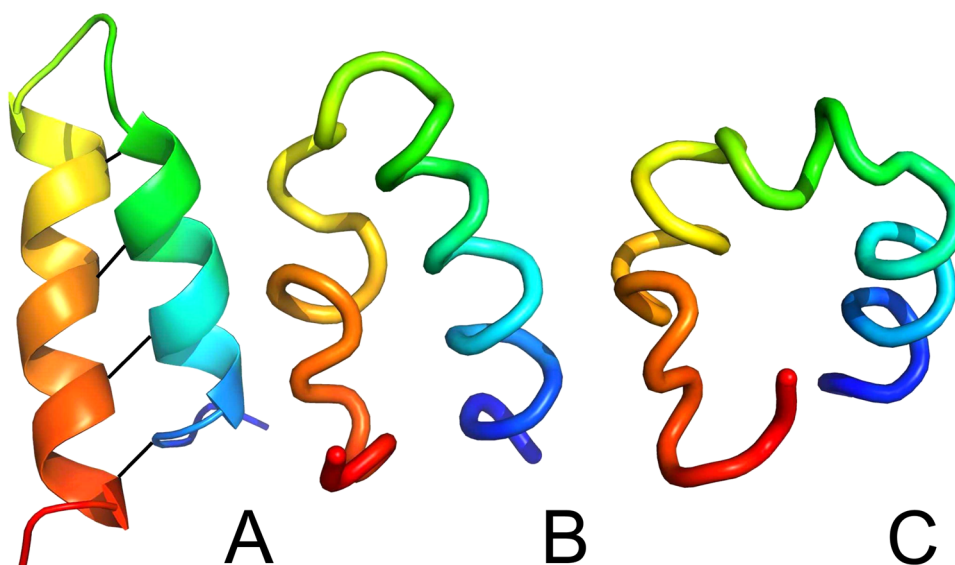
Figure 6. (A) Model 1 from the ensemble of NMR structures of thurincin H (PDB code: 2LBZ);[50] black lines mark the sulfide (S—C$^\alpha$) bonds, (B) the calculated structure of thurincin H with the lowest C$^\alpha$RMSD (3.86 Å) obtained from the MREMD simulation, and (C) average structure of the top cluster of the calculated structures of thurincin H at 310 K; C$^\alpha$-RMSD from the NMR Model 1 structure is 5.72 Å.

the calculated values are not critically sensitive to the method. For example, for lysine, the $\Delta\Delta G^{hel}$(Gly → D-Lys) determined by the contact-based method is −0.69 kcal/mol and that determined based on the backbone φ and ψ angles is −0.95 kcal/mol. The value of $\Delta\Delta G^{hel}$(Gly → L-Lys) determined by the contact-based method is −3.33 kcal/mol and that determined based on the backbone ϕ and ψ angles is −3.30 kcal/mol. These values give $\Delta\Delta G^{hel}$(D-Lys → L-Lys) = 2.64 and 2.35 kcal/mol, as determined by the contact-based method and from the backbone ϕ and ψ angles, respectively. These differences are even smaller than those arising from use of different experimental methods to estimate $\Delta\Delta G^{hel}$ (ref 42).

As can be seen from Table 1, the calculated values of $\Delta\Delta G^{hel}$(L-x → D-x) > 0, indicating that inserting a pair of D-residues into the host sequence destabilizes the α-helical structure compared to the insertion of a pair of the corresponding L-amino-acid residues, consistent with the experimental data except for proline.[42] For proline, the experimental value of $\Delta\Delta G^{hel}$(L-Pro → D-Pro) is slightly negative (−0.6 kcal/mol[42]), which means that insertion of two consecutive L-proline residues destabilizes α-helical structure slightly more than the insertion of two consecutive D-prolines. Conversely, the calculated values of $\Delta\Delta G^{hel}$(L-Pro → D-Pro) are slightly positive, independent of the method (0.96 kcal/mol for the contact-based method and 0.70 kcal/mol for the method based on the backbone ϕ and ψ angles, respectively). This observation does not indicate any gross disagreement between the UNRES and the experimental results but suggests that proline affects α-helix stability to a similar extent independent of chirality. Thus, although the extended UNRES force field was not specifically recalibrated to reproduce the thermodynamical properties of D-substituted polypeptides and proteins, it reproduces the helix-destabilizing trend of D-amino acid substitution.

It can be seen from Table 1 that the calculated values of $\Delta\Delta G^{hel}$(Gly → x) are mostly negative, which indicates that UNRES predicts that glycine destabilizes α-helix formation more than most D-amino acids, in contrast to experiment. Two or more glycine residues in a row act as a clear helix breaker when inserted

into clearly helix-forming L-amino-acid-residue sequences, such as poly(hydroxybutylglutamine)[53] but the experimental data of ref 42 demonstrate that a pair of D-residues except D-His and D-Asp is a stronger helix breaker than the Gly-Gly sequence (because the experimental values of $\Delta\Delta G^{hel}$(Gly → x) are negative only for D-His and D-Asp; see column 3 of Table 1). This observation indicates that parametrization of UNRES must be enhanced to include more thermodynamic data, as is planned for further work.

3.3. Simulations of the Folding of Thurincin H. The experimental Model 1 structure of thurincin H is shown in Figure 6A. Its structure is an α-helical hairpin (Figure 6A), which is stabilized by four S—C$^\alpha$ bonds, which were not present in the simulations (see section 2.4) because this would impose distance restraints, while we intended to assess whether the extended UNRES can reproduce the topology of the fold in unrestricted conformational search. Nevertheless, the UNRES force field was able to fold this peptide (Figure 6). The lowest-root-mean-square deviation over the C$^\alpha$ atoms (C$^\alpha$-RMSD between structures in Figure 6A,B) obtained after equilibration at T = 310 K was 3.86 Å. This value is comparable to the value of 3.04 Å of the mean RMSD of the structures from the experimental conformational ensemble of thurincin H determined by NMR (PDB code: 2LBZ) from the average structure of the ensemble. Clustering results reveal that the best folded structure belongs to the most probable cluster (28%) at T = 310 K. The average RMSD of the top cluster at 310 K is 5.72 Å (Figure 6C). The results suggest that thurincin H is able to fold without formation of the S—C$^\alpha$ bonds but, on the other hand, that the formation of these bonds can provide additional stability of the structure. When the S—C$^\alpha$ bonds are included, the lowest RMSD drops to 2.40 Å (Niadzvedtski, A; Sieradzan, A. K., unpublished). Because the resulting structure is not very different from the lowest-RMSD structure shown in Figure 6B, it is not shown.

4. SUMMARY

In this work, we determined the physics-based C$^\alpha$ ··· C$^\alpha$ ··· C$^\alpha$ backbone-virtual-bond-angle potentials for systems that include D-amino-acid residues, and we also revised the backbone-virtual-

bond-angle potentials for L-amino-acid residues that were determined in our earlier work[20] at a less advanced level of theory and are used with the present version of UNRES. Together with the virtual-bond-torsional, virtual-bond-double-torsional, side-chain-local, and backbone-electrostatic-correlation potentials determined in our recent work,[32] the new virtual-bond-angle potentials now enable us to simulate peptides and proteins, containing D-amino-acid residues, with the UNRES force field. To derive the virtual-bond-angle potentials, we used the energy surfaces of terminally blocked glycine, L-alanine, and L-proline calculated in our previous work[32] at the MP2/6-31G(d,p) ab initio level.

Because the change of the chirality of all residues is equivalent to the inversion of the sign of the virtual-bond dihedral angles γ_1 and γ_2, there are only 63 independent virtual-bond-angle bending potentials for each of the two types (NHMe or pyrrolidine) of C-terminal blocking groups. For L-amino-acid residues and the NHMe C-terminal blocking group, the potentials are very similar to the less accurate potentials determined in our previous work.[20] In that work,[20] the potential-energy surfaces used to determine the PMF surfaces were calculated by energy minimization with the RHF scheme with single-point MP2 correction; the Hessians were also calculated in the RHF/6-31G(d,p) scheme. In this work, the potential-energy surfaces were obtained by full minimizations of the MP2/6-31G(d,p) energy and the Hessians were also calculated with the MP2/6-31G(d,p) scheme. Moreover, in our previous work,[20] the N',N'-dimethylamine group and not the pyrrolidine group was used to represent a succeeding proline residue.

The extension of UNRES to D-amino-acid residues, which includes the torsional, double-torsional, local, and correlation potentials determined in our previous work[32] and the virtual-bond-angle (θ) potentials determined in this work, were tested with regard to reproducing the change in the stability of the α-helical structure of the KLALKLALxxLKLALKLA peptides upon substitution of D-amino-acid residues. The force field was able to reproduce the experimentally measured[42] decrease of ΔG values of α-helical-structure formation upon D-substitution of their L-counterparts in a qualitative manner. However, to reproduce the thermodynamics more quantitatively, enhancement of terms corresponding to local interactions is required, presumably deriving potentials that are more specific to sequence. Recently[35] we started to introduce sequence-specific torsional potentials into UNRES.

We also tested the enhanced UNRES force field by studying the simulated folding of thurincin H, which has D-amino-acid residues. The UNRES force field was able to fold the thurincin H peptide to native-like conformations in unrestricted MREMD simulations, without the S—C$^\alpha$ covalent links that occur in the experimental structure. This suggests that thurincin H first folds into an appropriate structure even without the presence of the S—C$^\alpha$ bonds and then its conformational stability is increased by the formation of these bonds. The influence of S—C$^\alpha$ bond formation on folding and stability of thurincin H will be investigated.

■ AUTHOR INFORMATION

Corresponding Author

*A. Sieradzan. Phone: +48-58-523-5350. Fax: +48-58-523-5472. E-mail: adasko@sun1.chem.univ.gda.pl.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Meier, J.; Kassler, K.; Sticht, H.; Eichler, J. *Beilstein J. Org. Chem.* **2012**, *8*, 1858−1866.

(2) Li, J.; Kuang, Y.; Gao, Y.; Du, X.; Shi, J.; Xu, B. *J. Am. Chem. Soc.* **2013**, *135*, 542−545.

(3) Siegrist, M. S.; Whiteside, S.; Jewett, J. C.; Aditham, A.; Cava, F.; Bertozzi, C. R. *ACS Chem. Biol.* **2013**, *8*, 500−505.

(4) Tugyi, R.; Uray, K.; Iván, D.; Fellinger, E.; Perkins, A.; Hudecz, F. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 413−418.

(5) Ali, M. H.; Peisach, E.; Allen, K. N.; Imperiali, B. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 12183−12188.

(6) Lindorff-Larsen, K.; Trbovic, N.; Maragakis, P.; Piana, S.; Shaw, D. E. *J. Am. Chem. Soc.* **2012**, *134*, 3787−3791.

(7) Sanbonmatsu, K. Y.; Joseph, S.; Tung, C.-S. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 15854−15859.

(8) Kolinski, A.; Skolnick, J. *Polymer* **2004**, *45*, 511−524.

(9) Maisuradze, G. G.; Senet, P.; Czaplewski, C.; Liwo, A.; Scheraga, H. A. *J. Phys. Chem. A* **2010**, *114*, 4471−4485.

(10) Horwacik, I.; Kurciński, M.; Bzowska, M.; Kowalczyk, A. K.; Czaplicki, D.; Koliński, A.; Rokita, H. *Int. J. Mol. Med.* **2011**, *28*, 47−57.

(11) Khalili, M.; Liwo, A.; Jagielska, A.; Scheraga, H. A. *J. Phys. Chem. B* **2005**, *109*, 13798−13810.

(12) Liwo, A.; Khalili, M.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 2362−2367.

(13) Liwo, A.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Scheraga, H. A. *Protein Sci.* **1993**, *2*, 1715−1731.

(14) Liwo, A.; Ołdziej, S.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Scheraga, H. A. *J. Comput. Chem.* **1997**, *18*, 849−873.

(15) Liwo, A.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Ołdziej, S.; Scheraga, H. A. *J. Comput. Chem.* **1997**, *18*, 874−887.

(16) Liwo, A.; Kaźmierkiewicz, R.; Czaplewski, C.; Groth, M.; Ołdziej, S.; Wawak, R. J.; Rackovsky, S.; Pincus, M. R.; Scheraga, H. A. *J. Comput. Chem.* **1998**, *19*, 259−276.

(17) Liwo, A.; Czaplewski, C.; Pillardy, J.; Scheraga, H. A. *J. Chem. Phys.* **2001**, *115*, 2323−2347.

(18) Liwo, A.; Ołdziej, S.; Czaplewski, C.; Kozłowska, U.; Scheraga, H. A. *J. Phys. Chem. B* **2004**, *108*, 9421−9438.

(19) Liwo, A.; Khalili, M.; Czaplewski, C.; Kalinowski, S.; Ołdziej, S.; Wachucik, K.; Scheraga, H. A. *J. Phys. Chem. B* **2007**, *111*, 260−285.

(20) Kozłowska, U.; Liwo, A.; Scheraga, H. A. *J. Phys.: Condens. Matter* **2007**, *19*, 285203-1−28503.15.

(21) Liwo, A.; Czaplewski, C.; Ołdziej, S.; Rojas, A. V.; Kaźmierkiewicz, R.; Makowski, M.; Murarka, R. K.; Scheraga, H. A. Simulation of protein structure and dynamics with the coarse-grained UNRES force field. In *Coarse-Graining of Condensed Phase and Biomolecular Systems*; Voth, G., Ed.; CRC Press, Taylor & Francis Group: Boca Raton, FL, 2008; pp 1391−1411.

(22) Kozłowska, U.; Maisuradze, G. G.; Liwo, A.; Scheraga, H. A. *J. Comput. Chem.* **2010**, *31*, 1154−1167.

(23) Makowski, M.; Liwo, A.; Sobolewski, E.; Scheraga, H. A. *J. Phys. Chem. B* **2011**, *115*, 6119−6129.

(24) Makowski, M.; Liwo, A.; Scheraga, H. A. *J. Phys. Chem. B* **2011**, *115*, 6130−6137.

(25) Sieradzan, A. K.; Scheraga, H. A.; Liwo, A. *J. Chem. Theory Comput.* **2012**, *8*, 1334−1343.

(26) Liwo, A.; Lee, J.; Ripoll, D. R.; Pillardy, J.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 5482−5485.

(27) Lee, J.; Liwo, A.; Ripoll, D. R.; Pillardy, J.; Saunders, J. A.; Gibson, K. D.; Scheraga, H. A. *Int. J. Quantum Chem.* **2000**, *77*, 90−117.

(28) Pillardy, J.; Czaplewski, C.; Liwo, A.; Lee, J.; Ripoll, D. R.; Kaźmierkiewicz, R.; Oldziej, S.; Wedemeyer, W. J.; Gibson, K. D.; Arnautova, Y. A.; Saunders, J.; Ye, Y.-J.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 2329−2333.

(29) Ołdziej, S.; et al. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 7547−7552.

(30) Liwo, A.; He, Y.; Scheraga, H. A. *Phys. Chem. Chem. Phys.* **2011**, *13*, 16890−16901.

(31) He, Y.; Mozolewska, M. A.; Krupa, P.; Sieradzan, A. K.; Wirecki, T. K.; Liwo, A.; Kachlishvili, K.; Rackovsky, S.; Jagieła, D.; Ślusarz, R.; Czaplewski, C. R.; Ołdziej, S.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 14936−14941.

(32) Sieradzan, A. K.; Hansmann, U. H. E.; Scheraga, H. A.; Liwo, A. *J. Chem. Theory Comput.* **2012**, *8*, 4746−4757.

(33) Sieradzan, A. K.; Liwo, A.; Hansmann, U. H. E. *J. Chem. Theory Comput.* **2012**, *8*, 3416−3422.

(34) Kubo, R. *J. Phys. Soc. Jpn.* **1962**, *17*, 1100−1120.

(35) Krupa, P.; Sieradzan, A. K.; Rackovsky, S.; Baranowski, M.; Ołdziej, S.; Scheraga, H. A.; Liwo, A.; Czaplewski, C. *J. Chem. Theory Comput.* **2014**, *9*, 4620−4632.

(36) He, Y.; Xiao, Y.; Liwo, A.; Scheraga, H. A. *J. Comput. Chem.* **2009**, *30*, 2127−2135.

(37) Shen, H.; Liwo, A.; Scheraga, H. A. *J. Phys. Chem. B* **2009**, *113*, 8738−8744.

(38) Kolinski, A.; Skolnick, J. *J. Chem. Phys.* **1992**, *97*, 9412−9426.

(39) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. *Nat. Struct. Biol.* **2002**, *9*, 425−430.

(40) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 5578−5583.

(41) Nishikawa, K.; Momany, F. A.; Scheraga, H. A. *Macromolecules* **1974**, *7*, 797−806.

(42) Krause, E.; Bienert, M.; Schmieder, P.; Wenschuh, H. *J. Am. Chem. Soc.* **2000**, *122*, 4865−4870.

(43) Czaplewski, C.; Kalinowski, S.; Liwo, A.; Scheraga, H. A. *J. Chem. Theory Comput.* **2009**, *5*, 627−640.

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

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

(46) Kaźmierkiewicz, R.; Liwo, A.; Scheraga, H. A. *J. Comput. Chem.* **2002**, *23*, 715−723.

(47) Kaźmierkiewicz, R.; Liwo, A.; Scheraga, H. A. *Biophys. Chem.* **2003**, *100*, 261−280; *Erratum: Biophys. Chem.* **2003**, *106*, 91.

(48) Zimmerman, S. S.; Pottle, M. S.; Némethy, G.; Scheraga, H. A. *Macromolecules* **1977**, *10*, 1−9.

(49) Ołdziej, S.; Liwo, A.; Czaplewski, C.; Pillardy, J.; Scheraga, H. A. *J. Phys. Chem. B* **2004**, *108*, 16934−16949.

(50) Sit, C. S.; vanBelkum, M. J.; McKay, R. T.; Worobo, R. W.; Vederas, J. C. *Angew. Chem., Int. Ed.* **2011**, *50*, 8718−8721.

(51) Murtagh, F. *Multidimensional clustering algorithms*; Physica-Verlag: Vienna, 1985.

(52) Levitt, M. *J. Mol. Biol.* **1976**, *104*, 59−107.

(53) Ananthanarayanan, V. S.; Andreatta, R. H.; Poland, D.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 417−424.