

Determination of Effective Potentials for the Stretching of C^α...C^α Virtual Bonds in Polypeptide Chains for Coarse-Grained Simulations of Proteins from *ab Initio* Energy Surfaces of N-Methylacetamide and N-Acetylpyrrolidine

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

[†]Faculty of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland

[‡]Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301, United States

ABSTRACT: The potentials of mean force (PMFs) for the deformation of the C^α...C^α virtual bonds in polypeptide chains were determined from the diabatic energy surfaces of N-methylacetamide (modeling regular peptide groups) and N-acetylpyrrolidine (modeling the peptide groups preceding proline), calculated at the Møller–Plesset (MP2) *ab initio* level of theory with the 6-31G(d,p) basis set. The energy surfaces were expressed in the C^α...C^α virtual-bond length (d) and the H–N–C^α...C' improper dihedral angle ($α$) that describes the pyramidalicity of the amide nitrogen, or in the C^α–C'(O)–N–C^α dihedral angle ($ω$) and the angle $α$. For each grid point, the potential energy was minimized with respect to all remaining degrees of freedom. The PMFs obtained from the ($d, α$) energy surfaces produced realistic free-energy barriers to the trans–cis transition (10 and 13 kcal/mol for the regular and proline peptide groups, respectively, compared to 12.6–13.9 kcal/mol and 17.3–19.6 kcal/mol determined experimentally for glycylglycine and N-acetylprolines, respectively), while those obtained from the ($ω, α$) energy maps produced either low-quality PMF curves when direct Boltzmann summation was implemented to compute the PMFs or too-flat curves with too-low free-energy barriers to the trans–cis transition if harmonic extrapolation was used to estimate the contributions to the partition function. An analytical bimodal logarithmic-Gaussian expression was fitted to the PMFs, and the potentials were implemented in the UNRES force field. Test Langevin-dynamics simulations were carried out for the Gly–Gly and Gly–Pro dipeptides, which showed a 10⁶-fold increase of the simulated rate of the trans–cis isomerization with respect to that measured experimentally; effectively the same result was obtained with the analytical Kramers theory of reaction rate applied to the UNRES representation of the peptide groups. Application of Kramers' theory to the computation of the rate constants from the all-atom *ab initio* energy surfaces of the model compounds studied resulted in isomerization rates close to the experimental values, which demonstrates that the increase of the isomerization rate in UNRES simulations results solely from averaging out the secondary degrees of freedom.

1. INTRODUCTION

Great progress in extending the time and size scales of all-atom simulations has recently been achieved by use of world-distributed computing (the FOLDING@HOME project);¹ development of very efficient load-balanced parallel codes such as GROMACS,² NAMD,³ and DESMOND,⁴ with implementation of all-atom molecular dynamics (MD) programs on graphical processor units (GPUs);⁵ and the construction of dedicated machines.⁶ The latter advancement has recently enabled millisecond-scale simulations of small proteins to be carried out.^{7,8} Nevertheless, all-atom *ab initio* MD simulations with explicit treatment of water are still restricted to small proteins, and therefore, folding simulations of larger proteins require a coarse-grained approach. Such an approach is used in the UNRES (UNited RESidue) model developed in our laboratory,^{9–22} in which each amino acid residue is reduced to two interacting sites, namely the peptide group and the side-chain group. This reduction of polypeptide-chain representation results in a 3 order of magnitude extension of the UNRES time scale compared to the experimental time scale.^{23,24} Recently,²⁵ we parallelized the energy and energy-gradient evaluation in UNRES, extending its application to simulations of large proteins in days of wall-clock time,

provided that massively parallel resources are available. By contrast to many coarse-grained force fields, UNRES is based upon physical principles, i.e., a cluster-cumulant²⁶ decomposition of the Restricted Energy Function (REF) or Potential of Mean Force (PMF) of polypeptide chains in water^{11,12,15} and by calibrating the resulting effective energy function based on structural and thermodynamic data pertaining to the folding of small proteins^{13–16,18,20} instead of statistical analyses of protein databases.

One of the slowest processes in protein-folding pathways is often the trans–cis isomerization of peptide bonds; this process often plays a key role in protein folding.^{27,28} Moreover, the cis configuration of a peptide group often has a major effect on the structure of a protein. One example, in which trans–cis isomerization is crucial, is ribonuclease A,^{27,28} in which trans–cis isomerization of three proline residues (Pro⁹³, Pro¹¹⁴, Pro¹¹⁷) determines the folding rate. Another example is the gene-3-protein in which N-terminal-domain Pro¹⁶¹ is in the cis form.²⁹ The loop containing this residue is flexible and, together

Received: November 23, 2011

Published: February 24, 2012



with adjacent parts of the chain, forms a folding nucleus which determines the folding rate.²⁹

The aim of this work was to extend the applicability of the UNRES force field to model trans–cis isomerization of peptide groups. This aim was achieved by calculating the potentials of mean force (PMFs) of virtual C^α…C^α bonds as functions of virtual-bond length and fitting analytical expressions to these potentials, which were subsequently implemented in UNRES.

2. METHODS

2.1. UNRES Model of Polypeptide Chains. In the UNRES force field,^{9–22} a polypeptide chain is represented by a sequence of α-carbon atoms with united side chains attached to them and peptide groups positioned halfway between two consecutive α carbons (Figure 1). The effective energy function is defined as the free energy of the chain constrained to a given coarse-grained conformation plus the surrounding solvent [this free energy is termed a restricted free energy (RFE) or a potential of mean force (PMF)^{11,12,19}]. The energy of the virtual-bond chain is expressed by eq 1.

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

where 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 united side-chain center of residue i (Figure 1), and d_i is the length of the i th virtual bond, which is either a C^α…C^α virtual bond or a C^α…SC virtual bond. Each term is multiplied by an appropriate weight, w_x , and the terms corresponding to factors of order higher than 1 are additionally multiplied by the respective temperature factors which were introduced in our recent work¹⁸ and which reflect the dependence of the first generalized-cumulant term in those factors on temperature, as discussed in refs 18 and 30. The factors f_n are defined by eq 2.

$$f_n(T) = \frac{\ln[\exp(1) + \exp(-1)]}{\ln\{\exp[(T/T_0)^{n-1}] + \exp[-(T/T_0)^{n-1}]\}} \quad (2)$$

where $T_0 = 300$ K.

The term $U_{SC,SC}$ 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 $U_{SC,p}$ denotes the excluded-volume potential of the side-chain–peptide-group interactions. The peptide-group interaction potential is split into two parts: the

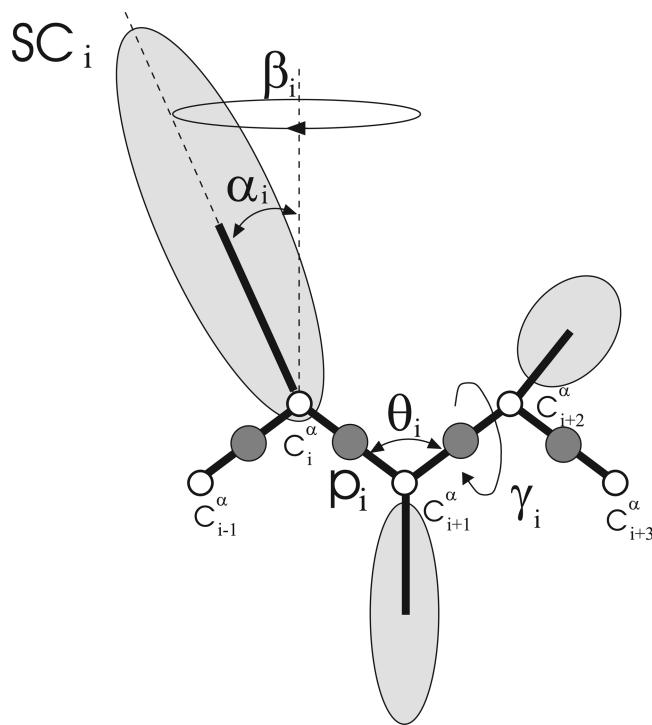


Figure 1. The UNRES model of polypeptide chains. The interaction sites are peptide-group centers (p) and side-chain centers (SC) attached to the corresponding α carbons with different C^α…SC bond lengths, d_{SC} . The peptide groups are represented as gray circles, and the side chains are represented as gray ellipsoids of different sizes. The α -carbon atoms are represented by small open circles. The geometry of the chain can be described either by the virtual-bond vectors $d\mathbf{C}_i$ (from C^α to C^α_{i+1}), $i = 1, 2, \dots, n - 1$, and $d\mathbf{X}_i$ (from C^α to SC_i), $i = 2, \dots, n - 1$, represented by thick lines, where n is the number of residues, or in terms of virtual-bond lengths, backbone virtual-bond angles θ_i , $i = 1, 2, \dots, n - 2$, backbone virtual-bond-dihedral angles γ_i , $i = 1, 2, \dots, n - 3$, and the angles α_i and β_i , $i = 2, 3, \dots, n - 1$, 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+1}.

Lennard-Jones interaction energy between peptide-group centers ($U_{p_i p_j}^{VDW}$) and the average electrostatic energy between peptide-group dipoles ($U_{p_i p_j}^{el}$); 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, virtual-bond angle bending, side-chain rotamer, and virtual-bond-deformation terms; these terms account for the local properties of the polypeptide chain. In the current version of UNRES, the contributions to U_{bond} from backbone virtual-bond stretching are unimodal and correspond to the distortion of the trans configuration of peptide bonds. 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 fifth and sixth order correlation terms present in earlier versions of UNRES^{12–15} were later¹⁶ found to be unnecessary for the force field to perform well; consequently, the version of UNRES developed after the work published in ref 16, including this work, does not contain these higher-order terms. The multibody terms are indispensable for the reproduction of regular α -helical and β -sheet structures.^{11,12,31}

The energy-term weights are determined by force-field calibration to reproduce the structure and folding thermodynamics of selected training proteins.^{18,20}

2.2. Langevin Dynamics with UNRES. The canonical molecular dynamics implementation of UNRES has been described in our earlier work.^{23,24,32} Briefly, the C^α...C^α and C^α...SC virtual-bond vectors (Figure 1) are gathered into a vector of generalized coordinates $\mathbf{q} = (\mathbf{dC}_o, \mathbf{dC}_1, \dots, \mathbf{dC}_n, \mathbf{dX}_1, \mathbf{dX}_2, \dots, \mathbf{dX}_n)^T$. The vector \mathbf{dC}_o specifies the position of the first C^α atom of the chain, \mathbf{dC}_i specifies the C_i^α...C_{i+1}^α virtual-bond vector, while \mathbf{dX}_i specifies the C_i^α...SC_i virtual-bond vector. The vectors $\dot{\mathbf{q}}$ and $\ddot{\mathbf{q}}$ denote generalized velocities and generalized accelerations, respectively. The virtual bonds are represented as elastic rods with mass distribution that scales with the length of a rod.³² The Cartesian coordinates of the interacting sites $\mathbf{x} = (\mathbf{r}_{p_1}, \mathbf{r}_{p_2}, \dots, \mathbf{r}_{p_{n-1}}, \mathbf{r}_{SC_1}, \mathbf{r}_{SC_2}, \dots, \mathbf{r}_{SC_n})^T$ are related to the generalized coordinates by a linear transformation $\mathbf{x} = \mathbf{A}\mathbf{q}$, where \mathbf{A} is a constant matrix with $a_{i(k),j} = 0$ [$i(k)$ being a Cartesian coordinate of site k] if the coordinates up to j correspond to virtual-bond vectors of the part of the chain to the right of site k , $a_{i(k),j} = 1$ if the coordinates correspond to the virtual-bond vectors to the left of site k or to a C^α...SC virtual bond containing the side chain with index k , and $a_{i(k),j} = 1/2$ if the coordinates correspond to the virtual-bond vector containing the peptide group with index $i(k)$. The same relationship holds between the time derivatives of \mathbf{x} and $\dot{\mathbf{q}}$.

In matrix notation, the complete equations of motion for Langevin dynamics with the UNRES force field can be written as eq 3.²³

$$(\mathbf{A}^T \mathbf{M} \mathbf{A} + \mathbf{H})\ddot{\mathbf{q}} = -\nabla_{\mathbf{q}} U(\mathbf{q}) - \mathbf{A}^T \boldsymbol{\Gamma} \mathbf{A} \dot{\mathbf{q}} + \mathbf{A}^T \mathbf{f}^{\text{rand}} \quad (3)$$

where \mathbf{M} is the diagonal matrix of masses of the sites (united peptide groups and united side chains) such that m_{ii} is the mass of the site corresponding to the i th generalized coordinate, \mathbf{H} (a diagonal matrix) is the part of the inertia matrix corresponding to the internal stretching motion of the virtual bonds with $h_{ii} = (1/12)m_p$ (m_p being the mass of a peptide group) for peptide groups and $h_{ii} = (1/3)m_{SC(j(i))}$ ($m_{SC(j(i))}$ being the mass of the side chain corresponding to the i th generalized coordinates) for side chains, $\boldsymbol{\Gamma}$ is the diagonal friction tensor (represented by the friction matrix) acting on the interacting sites such that γ_{ii} is the coefficient of the site corresponding to the i th coordinate, \mathbf{f}^{rand} is the vector of random forces acting on interacting sites, U is the UNRES potential energy defined by eq 1, and $\nabla_{\mathbf{q}}$ denotes the gradient in \mathbf{q} . The random forces (\mathbf{f}^{rand}) are calculated from eq 4.^{33–35}

$$f_i^{\text{rand}} = \sqrt{\frac{2\gamma_i RT}{\delta t}} N(0, 1) \quad (4)$$

where f_i^{rand} is the i th component of the random force vector \mathbf{f}^{rand} , γ_i is the friction coefficient associated with the i th coordinate of the interaction sites, R is the universal gas constant, T is the absolute temperature, δt is the integration time step, and $N(0,1)$ is the normal distribution with zero mean and unit variance. Together, the random and friction forces constitute a thermostat that maintains the average temperature at the preset value.

2.3. Determination of New Backbone U_{bond} Potentials.

As mentioned in section 2.1, in the UNRES force field used prior to this work, it was assumed that all peptide groups have trans geometry, and consequently, the backbone-virtual-bond potential was represented by a simple quadratic function with an equilibrium virtual-bond length of 3.8 Å. In this section, we introduce a new potential function for backbone-virtual-bond deformation which accounts for the presence of both trans and cis configurations of peptide groups and for transitions between these configurations.

N-methylacetamide and N-acetylpyrrolidine were used as models of the regular and proline-type peptide groups, respectively. To determine the respective PMFs, the potential-energy surfaces of these systems were calculated by using the MP2/G6-31(d,p) *ab initio* method. The energy was evaluated on a two-dimensional grid on which one coordinate was either the C^α–C’–N–C^α dihedral angle (ω) for rotation about the C’–N bond ranging from -180° to $+180^\circ$ with a 10° step and the improper H–N–C^α...C’ or C^δ–N–C^α...C’ angle (α ; for N-methylacetamide or N-acetylpyrrolidine, respectively), which defines the pyramidicity of the amide nitrogen, ranging from 90° to 270° with a 10° step, or the virtual C^α...C^α bond length (d) ranging from 2 to 5 Å with a 0.01 Å step and the angle α . The energy of the system becomes very high when $\alpha < 90^\circ$ or $\alpha > 270^\circ$. We assigned α to the range from 0° to 360° , instead of the range from -180° to $+180^\circ$, in order to have a contiguous range of α in which the energy is computed; if α ranged from -180° to $+180^\circ$ as does the dihedral angle ω , this range would consist of two parts, one from -180° to -90° and the other one from 90° to 180° .

The energy was minimized with respect to all degrees of freedom (including the bond lengths and bond angles) except for the grid variables. Energy Hessians (second derivative matrices or force-constant matrices) were calculated at each grid point. The (ω, α) grid seems to be a natural choice, because the first angle is a reaction coordinate for trans–cis isomerization, and the second one defines the pyramidicity of the amide nitrogen, which arises from distortion of the peptide bond from planarity and, consequently, a decrease of the C’–N bond order. However, use of ω as a reaction coordinate does not cover small and large C^α...C^α virtual-bond lengths; hence, we also used $d_{\text{C}^{\alpha}\dots\text{C}^{\alpha}}$ as a variable. The PMFs were initially calculated from the (ω, α) surfaces by direct Boltzmann summation to compute histograms in d or by using the harmonic-extrapolation method developed in our earlier work.^{17,21} The PMFs obtained by direct Boltzmann summation are expressed by eq 5.

$$F(d_i, T) = -\beta^{-1} \ln \sum_{\omega} \sum_{\alpha} \exp[-\beta e^*(\omega, \alpha)] \quad d_{i-1} \leq d(\omega, \alpha) < d_i \quad (5)$$

where $F(d, T)$ is the PMF for the virtual-bond length d , at absolute temperature T , $\beta = 1/RT$, R is the universal gas constant, $e^*(\omega, \alpha)$ is the energy at the (ω, α) grid point, and the superscript “*” denotes the value optimized in all degrees of freedom except ω and α . The interval length in evaluating the histograms was 0.05 Å.

With the harmonic extrapolation,^{17,21} each point of the grid contributed to every value of $F(d, T)$ (eq 6), as given by eq 6.

$$\begin{aligned} F(d, T) \approx -\beta^{-1} \ln \sum_{\omega} \sum_{\alpha} [\det \mathbf{H}^*(\omega, \alpha)]^{-1/2} \\ \times \exp \left\{ -\beta \left[e^*(\omega, \alpha) + \frac{1}{2} [H_{dd}^*(\omega, \alpha) \right. \right. \\ \left. \left. - \frac{1}{4} \mathbf{H}_{d\mathbf{x}}^*(\omega, \alpha) \mathbf{H}_{\mathbf{x}\mathbf{x}}^*(\omega, \alpha) \mathbf{H}_{d\mathbf{x}}^{*\mathrm{T}}(\omega, \alpha) \right] \right. \\ \left. \times (d - d^*(\omega, \alpha))^2 \right\} \end{aligned} \quad (6)$$

where $H_{dd}^*(\omega, \alpha) = \partial^2 e^*(\omega, \alpha) / \partial d^2$, $\mathbf{H}_{d\mathbf{x}}^*(\omega, \alpha)$ is the part of the energy Hessian corresponding to the virtual-bond distance and nongrid variables (gathered in vector \mathbf{x}), and $\mathbf{H}_{\mathbf{x}\mathbf{x}}^*(\omega, \alpha)$ is the energy Hessian in the nongrid variables. These quantities are calculated at the (ω, α) grid point, corresponding to the energy that is minimized in all nongrid variables.

For the (d, α) grid, the PMFs were calculated by direct Boltzmann summation. Two sets of potentials were computed, the first one without (eq 7) and the second one with (eq 8) estimation of the contribution to the partition function from the nongrid degrees of freedom (mainly the entropy contribution); this contribution was calculated by using a harmonic approximation.

$$F(d, T) = -\beta^{-1} \ln \sum_{\alpha} \exp[-\beta e^*(d, \alpha)] \quad (7)$$

$$\begin{aligned} F(d, T) = -\beta^{-1} \ln \sum_{\alpha} [\det \mathbf{H}^*(d, \alpha)]^{-1/2} \\ \times \exp[-\beta e^*(d, \alpha)] \end{aligned} \quad (8)$$

To obtain the effective potentials for $C^\alpha \cdots C^\alpha$ virtual-bond deformation, we fitted the analytical expression given by eq 9 to the calculated PMFs; we used the PMFs calculated from eq 8 (i.e., from (d, α) maps with the harmonic contribution from the nongrid variables).

$$\begin{aligned} U(d, T) = -\beta^{-1} \ln \left\{ \exp \left[-\beta \left(\frac{1}{2} k_{cis} (d - d_{cis}^\circ)^2 \right. \right. \right. \\ \left. \left. + V_{cis}^\circ + S_{cis}(T - T_0) \right) \right] \\ \left. + \exp \left[-\beta \left(\frac{1}{2} k_{trans} (d - d_{trans}^\circ)^2 \right. \right. \right. \\ \left. \left. + V_{trans}^\circ + S_{trans}(T - T_0) \right) \right] \right\} \end{aligned} \quad (9)$$

where $U(d, T)$ is the effective energy term for virtual $C^\alpha \cdots C^\alpha$ bond deformation in the UNRES force field and k_{cis} (k_{trans}) is the force constant corresponding to the cis (the trans) form, V_{cis}° (V_{trans}°) is the free-energy value of the cis (trans) form at $T_0 = 298$ K, cis is the reference for N-acetylpyrrolidine, trans is for N-methylacetamide, S_{cis} (S_{trans}) is the first derivative of the free energy of an unstrained cis (trans) peptide bond with respect to temperature, and d_{cis}° (d_{trans}°) is the equilibrium value of the virtual $C^\alpha \cdots C^\alpha$ bond length for the cis (trans) form. This expression

behaves as a harmonic potential in the neighborhood of the cis or trans configuration, respectively, and includes the dependence of the effective potential on temperature. Marquardt's method³⁶ was used as a fitting procedure. The PMF curves were calculated and fitted on a grid from $T = 270$ K to $T = 350$ K with a 5 K step size.

2.4. Simulation of the Kinetics of cis–trans Isomerization with the New Backbone U_{bond} Potentials. After implementation of eq 9 in the UNRES force field, four sets of 512 independent canonical Langevin-dynamics simulations each were run for the Gly–Gly and four sets for the Gly–Pro dipeptide, respectively. The sets were run at $T = 300, 320, 340$, and 360 K, respectively. Each trajectory consisted of 10^9 steps with a 4.89 fs time-step length. To speed up simulations, the friction was set to 10^{-2} of the value of the friction of water, which is 0.8902 cP at $T = 298$ K,³⁷ as in our earlier work.²³ All simulations were started from the trans conformation ($d_{C^\alpha \cdots C^\alpha} = 3.8$ Å). For each series of simulations, the fraction of the cis conformation averaged over all 512 trajectories was calculated as a function of time, and a first-order kinetic equation was fitted (eq 10) to the simulation data.

$$x_{cis} = \frac{k_1}{k_1 + k_{-1}} \{1 - \exp[-(k_1 + k_{-1})t]\} \quad (10)$$

where x_{cis} is the mole fraction of the cis conformation and k_1 and k_{-1} are the reaction rate constants of conversion from the trans form to the cis form and from the cis form to the trans form, respectively, and t denotes time.

The forward k_1 and the backward k_{-1} reaction-rate constants determined from simulations were compared with those computed from Kramers' theory of reaction rates (eq 11).³⁸

$$\begin{aligned} k = \frac{\Omega}{\eta_{\text{water}} \frac{6\pi r}{m_{\text{red}}}} \sqrt{2\pi k_B T} \\ \times \{ \sqrt{m_{\text{red}}} \int_{\zeta_1}^{\zeta_2} \exp\{\beta[U(\zeta) - U(\zeta_1)]\} d\zeta \}^{-1} \end{aligned} \quad (11)$$

with

$$\Omega = \sqrt{\frac{\kappa}{m_{\text{red}}}} \quad (12)$$

where Ω is the frequency of the harmonic oscillator for the trans form (forward reaction) or for the cis form (backward reaction); $\eta_{\text{water}} = 0.803$ cP (8.03×10^{-4} kg s⁻¹) is the viscosity of water; r and m_{red} are the radius of a hydrated peptide group and the reduced mass of a peptide group, respectively (it should be noted that the reduced masses cancel out when eq 12 is plugged into eq 11), $r = 3.8$ Å for regular peptide groups and $r = 5.6$ Å for proline-type peptide groups; κ is the force constant obtained from eq 9, $\kappa = k_{trans}$ for the forward reaction and $\kappa = k_{cis}$ for the backward reaction, respectively; ζ is the reaction coordinate, $\zeta_1 = \zeta_{trans}$ and $\zeta_2 = \zeta_{cis}$ for the forward reaction and $\zeta_1 = \zeta_{cis}$ and $\zeta_2 = \zeta_{trans}$ for the backward reaction, respectively; and k_B is the Boltzmann constant. The same values of η_{water} , κ , and r were used for both UNRES and the all-atom model. It should be noted that k does not depend on the reduced mass explicitly because the powers of the reduced mass cancel out in eq 11.

Equation 11 is valid if the substrate can attain thermal equilibrium before the reaction occurs to any remarkable extent (high viscosity limit).³⁸ This assumption certainly holds for the

trans–cis isomerization of peptide groups because the trans–cis conversion is very slow.^{39–41} It is also valid for the simulations carried out in this work because the test molecules resided in a given (trans or cis) state before they converted to the other state (see section 3.3).

For the UNRES model of peptide groups, ζ was defined as the virtual $C^\alpha\cdots C^\alpha$ bond length. For the all-atom representation of peptide groups, ζ was defined as the Cartesian distance traveled by the methyl-carbon atom of the acetyl group upon variation of ω (the dihedral angle of rotation about the $C'\text{--}N$ bond) and by the amide-hydrogen atom or the C^δ atom of the pyrrolidine ring upon variation of α , as given by eq 13.

$$\zeta_i = \sum_{k=1}^{i-1} \sqrt{(l_1 \sin \theta_1 \Delta \omega_k)^2 + (l_2 \sin \theta_2 \Delta \alpha_k)^2} \quad (13)$$

where l_1 and θ_1 are the length of the $N\text{--}C^\alpha$ bond and the $C'\text{--}N\text{--}C^\alpha$ bond angle of the acetyl group and l_2 and θ_2 are the length of the $N\text{--}X$ and the $C'\text{--}N\text{--}X$ bond angle, where $X = H$ for N-methylacetamide and $X = C^\delta$ for N-acetylpyrrolidine, respectively. $\Delta \omega_k$ and $\Delta \alpha_k$ are the changes of the ω and α angles, respectively, from point k to point $k+1$ of the reaction path.

3. RESULTS AND DISCUSSION

3.1. Potential Energy Surfaces of N-Methylacetamide and N-Acetylpyrrolidine. The potential-energy surfaces of N-methylacetamide and N-acetylpyrrolidine, constructed on the (ω, α) and on the (d, α) grids, are shown in Figures 2A and B and 3A and B, respectively, while the potential-energy profiles along the gentlest ascent/descent paths determined from the potential-energy maps in ω and α , together with the substrate, product, and transition-structure geometries, are shown in Figure 4A and B, respectively.

It can be seen that the potential-energy surfaces for both N-acetylpyrrolidine and N-methylacetamide have two distinct sets of minima (Figure 2A and 2B). For N-methylacetamide, two equivalent lowest-energy minima [at $(\omega = 180^\circ, \alpha = 170^\circ)$ and $(\omega = 180^\circ, \alpha = 190^\circ)$, respectively] correspond to the trans conformation and two equivalent higher-energy minima [at $(\omega = 10^\circ, \alpha = 160^\circ)$ and $(\omega = -10^\circ, \alpha = 200^\circ)$, respectively] correspond to the cis conformation (Figure 2A), while for N-acetylpyrrolidine there is a single cis and a single trans minimum; both have the same energy, and occur at $(\omega = 180^\circ, \alpha = 180^\circ)$ and $(\omega = 0^\circ, \alpha = 180^\circ)$, respectively (Figure 2B). In the (d, α) potential-energy surfaces, the cis-form minima correspond to the $C^\alpha\cdots C^\alpha$ virtual-bond lengths $d = 2.91 \text{ \AA}$ and $d = 2.92 \text{ \AA}$ for N-methylacetamide and N-acetylpyrrolidine, respectively, while the trans-form minima have $d = 3.80 \text{ \AA}$ and $d = 3.81 \text{ \AA}$ for N-methylacetamide and N-acetylpyrrolidine, respectively (Figure 3). The trans minimum also is bifurcated for N-methylacetamide (Figure 3A). The potential-energy barriers to the transition from the trans to the cis conformation are 16.6 and 16.9 kcal/mol for N-methylacetamide and N-acetylpyrrolidine, respectively. The potential-energy barrier obtained for N-acetylpyrrolidine is comparable to the values from 17.6 to 19.3 kcal/mol calculated by Kang and co-workers^{42,43} for N-acetyl-N'methylprolineamide by the MP2/6-31G(d) approach (for terminally blocked proline, these values also depend on the pucker state of the pyrrolidine ring⁴³).

For N-methylacetamide, the trans form exhibits almost no pyramidalicity of the nitrogen atom and has a planar amide bond, $(\omega = 180^\circ, \alpha = 170^\circ)$ or $(\omega = 180^\circ, \alpha = 190^\circ)$. Conversely, the

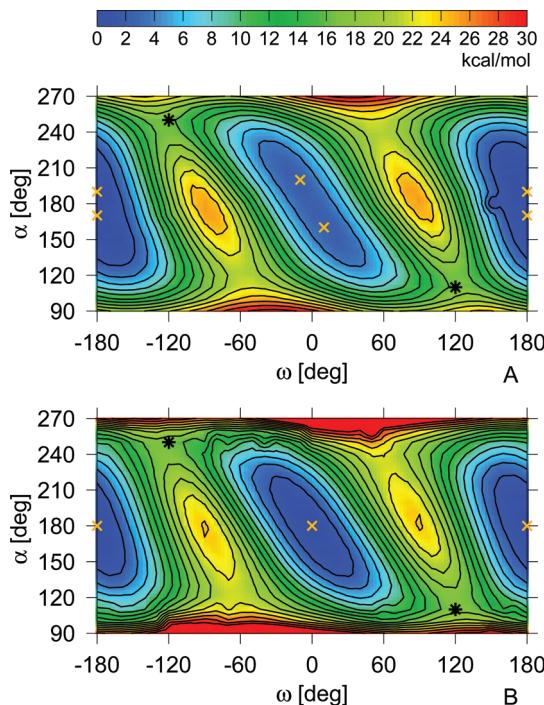


Figure 2. Contour plot of the MP2/6-31G(d,p) potential-energy surface of (A) N-methylacetamide and (B) N-acetylpyrrolidine in the ω angle of rotation about the $C'\text{--}N$ bond (the $C^\alpha\cdots C'\text{--}N\text{--}C^\alpha$ dihedral angle) and the improper-dihedral angle α ($H\text{--}N\text{--}C^\alpha\cdots C'$) that describes the pyramidalicity of the amide nitrogen. The potential-energy scale is shown in the upper panel. For N-methylacetamide (part A), two equivalent trans minima occur at $(\omega = 180^\circ, \alpha = 170^\circ)$ and $(\omega = 180^\circ, \alpha = 190^\circ)$, respectively, two equivalent cis minima occur at $(\omega = 10^\circ, \alpha = 160^\circ)$ and $(\omega = -10^\circ, \alpha = 200^\circ)$, respectively, and two equivalent saddle points (corresponding to the transition state) occur at $(\omega = 120^\circ, \alpha = 110^\circ)$ and $(\omega = -120^\circ, \alpha = 250^\circ)$, respectively. For N-acetylpyrrolidine (part B), the trans and the cis minima occur at $(\omega = 180^\circ, \alpha = 180^\circ)$ and $(\omega = 0^\circ, \alpha = 180^\circ)$, respectively, and the transition states occur at $(\omega = 120^\circ, \alpha = 110^\circ)$ and $(\omega = -120^\circ, \alpha = 250^\circ)$, respectively. For both systems, the potential-energy surfaces exhibit inverse symmetry, and equivalent critical points (minima and transition states) can be obtained by the following transformation: $\omega^{*'} = -\omega^*$, $\alpha^{*'} = 360^\circ - \alpha^*$, where an asterisk marks a critical point. The positions of the minima are marked with orange X's, and the positions of the saddle points are marked with black asterisks.

cis conformation is not flat at the energy minimum and exhibits noticeable pyramidalicity of the nitrogen atom, $(\omega = 10^\circ, \alpha = 160^\circ)$ or $(\omega = -10^\circ, \alpha = 200^\circ)$. Both systems exhibit high pyramidalicity of the nitrogen atom at the transition state, $(\omega = 120^\circ, \alpha = 110^\circ)$ and $(\omega = -120^\circ, \alpha = 250^\circ)$. The structures at the minima, and transition states on the potential-energy surfaces of the two compounds are illustrated in Figure 4A and B for N-methylacetamide and N-acetylpyrrolidine, respectively.

3.2. Potentials of Mean Force for $C^\alpha\cdots C^\alpha$ Virtual Bond Deformation. The PMFs for the virtual $C^\alpha\cdots C^\alpha$ bond stretching calculated by using the four approaches discussed in section 2.3 at $T = 298 \text{ K}$ are compared in Figure 5A and B for N-methylacetamide and N-acetylpyrrolidine, respectively. For the PMFs computed from the (ω, α) energy maps by direct Boltzmann summation (eq 5), the difference in free energy between the cis and the trans form equals 2 kcal/mol for N-methylacetamide, while the two forms of N-acetylpyrrolidine have nearly the same free energy. The free-energy barriers are 14 and 15 kcal/mol, respectively. These numbers are close to the respective

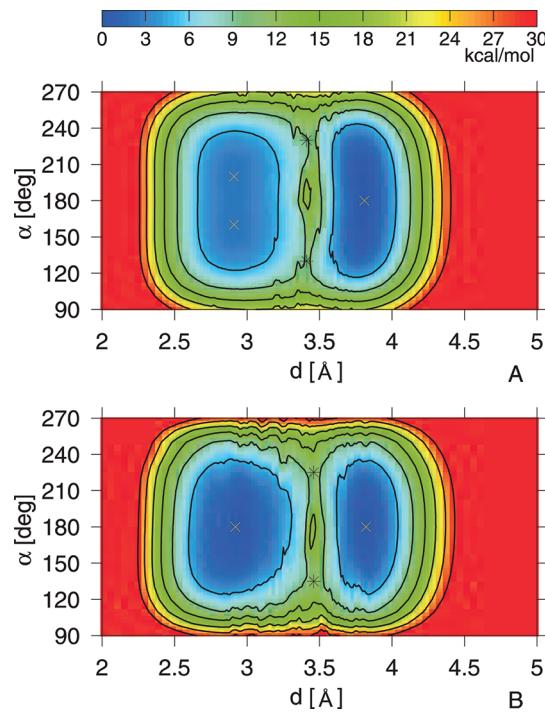


Figure 3. Contour plot of the MP2/6-31G(d,p) potential-energy surface of (A) N-methylacetamide and (B) N-acetylpyrrolidine in the virtual $C^{\alpha}\cdots C^{\alpha}$ bond length distance (d) and improper-dihedral angle (α ($H-N-C^{\alpha}\cdots C'$) that describes the pyramidalicity of the amide nitrogen. The potential-energy scale is shown in the upper panel. For N-methylacetamide (part A) the trans minimum occur at ($d = 3.80 \text{ \AA}$ and $\alpha = 180^\circ$) and two equivalent cis minima occur ($d = 2.91 \text{ \AA}$, $\alpha = 160^\circ$ and $(d = 2.91 \text{ \AA}, \alpha = 200^\circ)$, and two equivalent saddle points (corresponding to the transition state) occur at ($d = 3.41 \text{ \AA}, \alpha = 120^\circ$) and ($d = 3.41 \text{ \AA}, \alpha = 200^\circ$), respectively. For N-acetylpyrrolidine (part B), the trans and the cis minima occur at ($d = 3.81 \text{ \AA}, \alpha = 180^\circ$) and ($d = 2.92 \text{ \AA}, \alpha = 180^\circ$) respectively, and the transition states occur at ($d = 3.46 \text{ \AA}, \alpha = 130^\circ$) and ($d = 3.46 \text{ \AA}, \alpha = 230^\circ$). For both systems, the potential-energy surfaces exhibit reflection symmetry about the axis. $\alpha = 180^\circ$ and equivalent critical points (minima and transition states) can be obtained by the following transformation: $\alpha^{*'} = 360^\circ - \alpha^*$, where an asterisk marks a critical point. The positions of the minima are marked with orange X's, and the positions of the saddle points are marked with black asterisks.

experimental values. The free-energy differences between the trans and the cis conformations determined by NMR are 3 kcal/mol for glycylglycine,⁴¹ 0.1 to 0.7 kcal/mol for substituted N-acetylpyrrolidine derivatives^{39,40} (as calculated on the basis of the respective equilibrium constants reported in ref 40), and 1.48 and 1.34 kcal/mol for N-acetylglycylproline methyl ester in water and toluene, respectively.⁴⁴ The measured free-energy barrier to the trans–cis transition is 12.6 to 13.9 kcal/mol for glycylglycine⁴¹ and 17.3 to 19.6 kcal/mol for substituted N-acetylproline derivatives.^{39,40} It can be seen from Figure 5, though, that the PMF curves (red) are ragged and cover a limited $C^{\alpha}\cdots C^{\alpha}$ distance range. These features arise from the fact that the dihedral angle ω and not the $C^{\alpha}\cdots C^{\alpha}$ distance d (which is the coarse-grained variable) was the reaction coordinate of the trans–cis transition and that the energy was minimized in the nongrid variables when constructing the energy maps, which limited the lower and the upper limits of d to those corresponding to relaxed cis and relaxed trans geometry. Consequently, the corresponding PMF curve (red) cannot be a basis for the derivation of the coarse-grained backbone U_{bond} potentials.

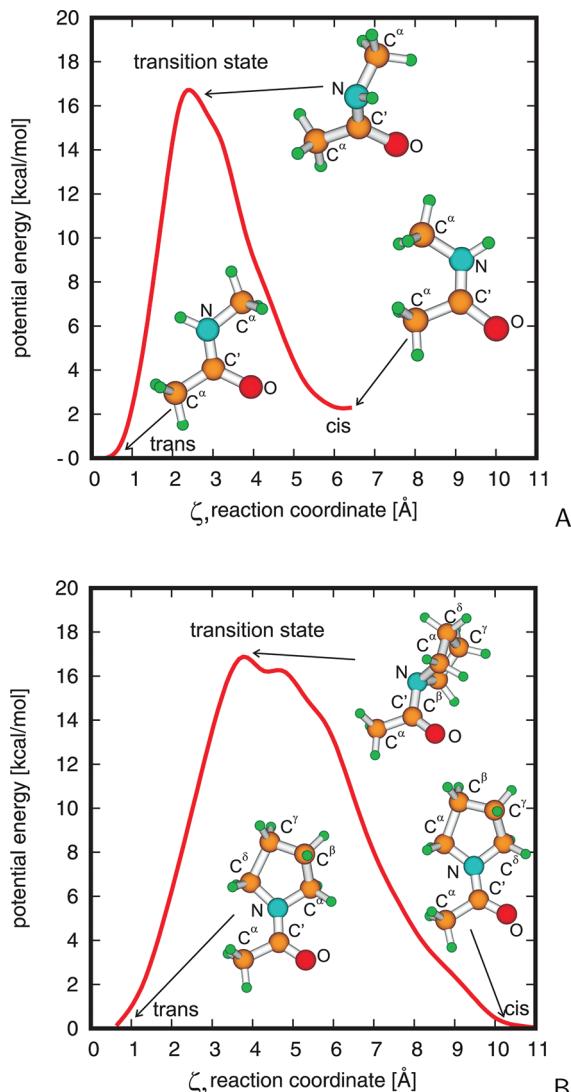


Figure 4. Potential-energy profiles of (A) N-methylacetamide and (B) N-acetylpyrrolidine constructed as gentlest ascent/gentlest descent paths passing from the trans to the cis minimum through the transition state. The substrate (trans), product (cis), and transition structures are inserted into the graphs. The reaction coordinate ζ is obtained by summing up the Cartesian distances traveled by the C atom of a methyl group and the amide hydrogen atom (for N-methylacetamide) or the C^{δ} atom of the pyrrolidine ring for N-acetylpyrrolidine upon changing ω and α from one to the next point of the reaction path closest to the grid points (eq 13).

Computing the PMFs from the (ω, α) energy maps by taking into account the contributions from every grid point through the use of harmonic extrapolation (eq 6)¹⁷ results in smooth PMF curves (green) covering the entire range of d (Figure 5). However the free-energy barriers are far too small (5 and 7 kcal/mol for N-methylacetamide and N-acetylpyrrolidine, respectively) compared with the experimental values 12.6–13.9 kcal/mol for glycylglycine⁴¹ and 17.3–19.6 kcal/mol for substituted N-acetylproline derivatives,^{39,40} respectively. Therefore, the harmonic extrapolation, which was applied with success in our earlier work on the determination of virtual-bond valence¹⁷ and side-chain-rotamer potentials,^{21,22} seems to overestimate the contributions from outside a given grid point for the systems considered in this study, in which the target

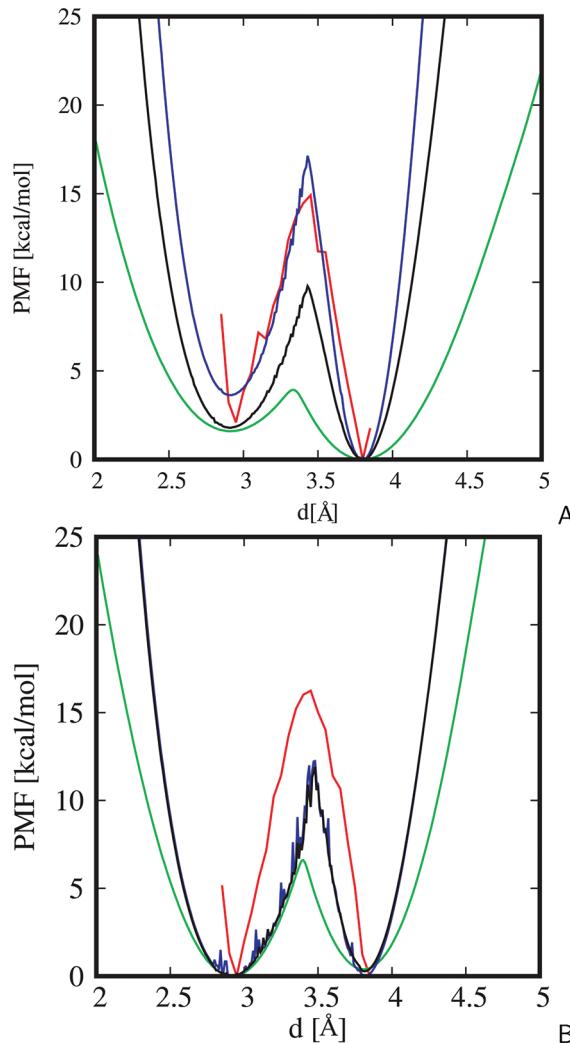


Figure 5. Potentials of mean force for (A) N-methylacetamide and (B) N-acetylpyrrolidine at $T = 298\text{ K}$ calculated from the (ω,α) or (d,α) *ab initio* MP2/6-31G(d,p) potential-energy surfaces of these compounds by using different methods. Red lines: direct Boltzmann summation over the (ω,α) potential-energy maps (eq 5). Green lines: summation over (ω,α) energy maps with harmonic extrapolation of the energy from all grid points to a given point (eq 6).¹⁷ Blue lines: direct Boltzmann summation over the (d,α) potential-energy maps (eq 7). Black lines: direct Boltzmann summation over the (d,α) potential-energy maps taking into account the entropy contribution in harmonic approximation (eq 8).

reaction coordinate, the $\text{C}^{\alpha}\cdots\text{C}^{\alpha}$, is a “harder” variable (i.e., a variable whose variation involves large energy changes), as in this study.

For the PMFs calculated from the energy surfaces constructed on the (d,α) grid (blue), the difference between the values of the cis and trans forms is 2 kcal/mol for N-methylacetamide and 0 kcal/mol for N-acetylpyrrolidine, respectively, whereas the free energy barrier equals 17 and 13 kcal/mol, respectively. After inclusion of the entropic contributions calculated from a harmonic approximation (black), the difference between the PMF at the minima becomes 2 kcal/mol and -0.3 kcal/mol for N-methylacetamide and N-acetylpyrrolidine, respectively, while the free energy barrier equals 10 and 13 kcal/mol, respectively. Lowering of the free-energy barrier is caused by the high entropy of the transition state for N-methylacetamide. The values of the free-energy barriers to the trans–cis transitions obtained for

N-methylacetamide and N-acetylpyrrolidine are in good agreement with the experimental values determined for glycylglycine, which range from 12.6 to 13.9 kcal/mol,⁴¹ and for N-acylprolines, which range from 17.3 to 19.6 kcal/mol.^{39,40} Consequently, the PMFs calculated from the (d,α) energy surfaces with inclusion of the harmonic contribution to the entropy were used to derive the backbone U_{bond} potentials for UNRES (eq 9). The curves obtained by fitting eq 9 to the PMFs calculated from eq 8 are compared with the PMFs in Figure 6A and B for N-methylacetamide and

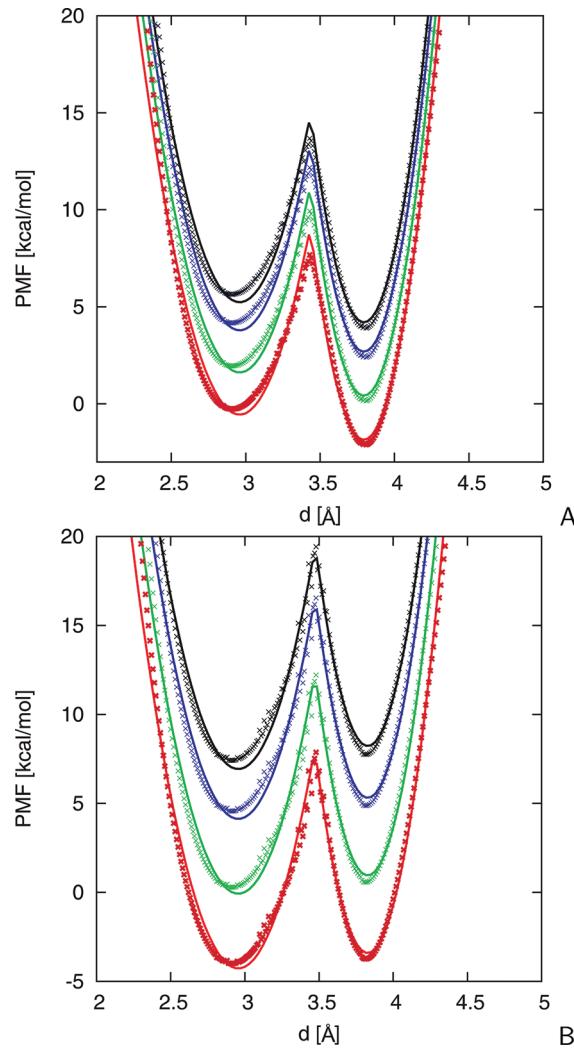


Figure 6. Plots of the potentials of mean force for (A) N-methylacetamide and (B) N-acetylpyrrolidine as functions of the $\text{C}^{\alpha}\cdots\text{C}^{\alpha}$ virtual-bond length calculated from the (d,α) potential-energy maps by using eq 8 (crosses) and fits of analytical approximation of the PMF as given by eq 9 (lines). Red, $T = 270\text{ K}$; green, $T = 300\text{ K}$; blue, $T = 330\text{ K}$; black, $T = 350\text{ K}$. The parameters of eq 8 resulting from fitting are collected in Table 1.

N-acetylpyrrolidine, respectively, while the parameters of the new UNRES potentials for $\text{C}^{\alpha}\cdots\text{C}^{\alpha}$ virtual-bond deformation are summarized in Table 1.

3.3. Kinetics of the trans–cis Transition. The fractions of the conformations with cis peptide bonds at various temperatures obtained in UNRES Langevin-dynamics simulations with the new potentials are shown in Figure 7A and B for N-methylacetamide and N-acetylpyrrolidine, respectively. It can be seen that the transition from the trans to the cis form is

Table 1. Parameters of the Peptide Bond Potentials Determined by Fitting eq 9 to the PMFs Calculated from eq 8

parameter	N-methyl-acetamide	N-acetyl-pyrrolidine
V_{cis}^o [kcal/mol]	1.471	-0.352
V_{trans}^o [kcal/mol]	0.265	0.665
k_{cis} [kcal/(mol × Å ²)]	87	94
k_{trans} [kcal/(mol × Å ²)]	169	181
S_{cis} [kcal/(mol × K)]	0.0724	0.140
S_{trans} [kcal/(mol × K)]	0.0756	0.145
d_{cis}^o [Å]	2.96	2.96
d_{trans}^o [Å]	3.79	3.83

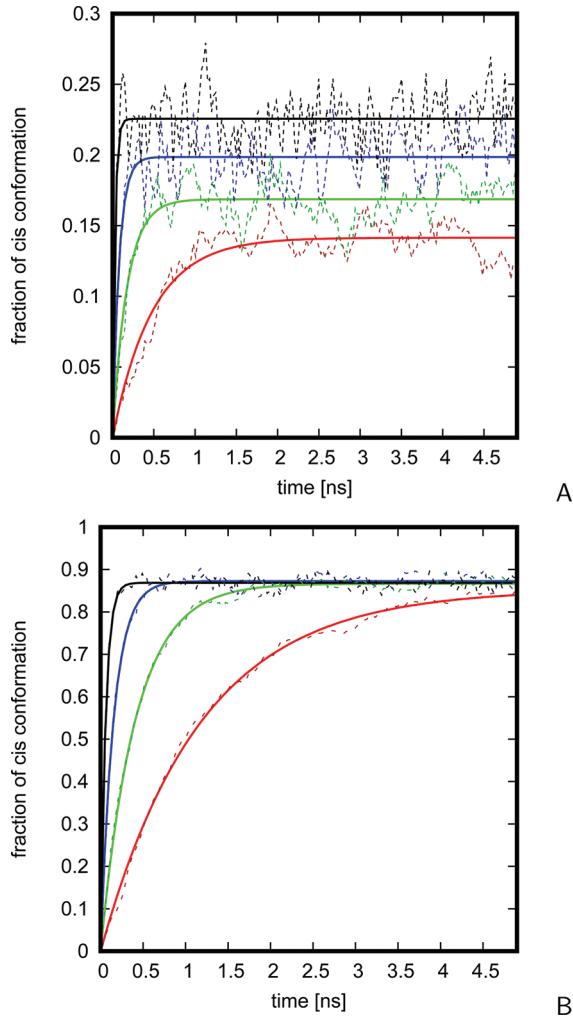


Figure 7. Fraction of the cis configuration of (A) the Gly–Gly system and (B) the Gly–Pro system calculated by UNRES Langevin dynamics simulation (ragged dotted lines) and fits of the first-order kinetic equation (eq 10) to simulation data. Red, $T = 300$ K; green, $T = 320$ K; blue, $T = 340$ K; black, $T = 360$ K.

faster for the proline-type peptide bond because of the greater relative stability of the cis form with respect to that of the trans form; however, the equilibrium is attained faster for the regular type of peptide bond. The trans form dominates for the regular and the cis form dominates for the proline-type peptide bond (Figure 7A and B). After scaling down by the factor of 100 to account for the fact that the friction of water was scaled down by this factor in UNRES Langevin dynamics simulations

(section 2.4), the forward (from trans to cis) rate constants calculated by fitting eq 10 to the respective simulation data at room temperature are on the order of 10^7 s⁻¹ greater at $T = 300$ K, compared to experimental values which are on the order of 10 s⁻¹ at room temperature.^{39–41}

The Arrhenius plots of the trans–cis and cis–trans rate constants (scaled down by the factor of 100 as stated above) are shown in Figure 8A and B for the Gly–Gly and the Pro–Gly

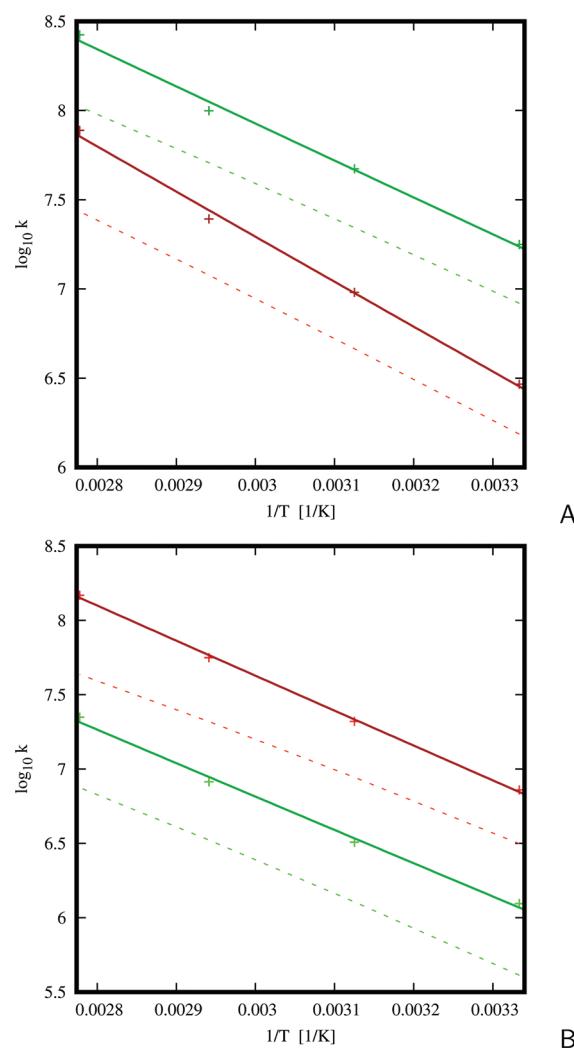


Figure 8. Arrhenius plots of the rate constants of forward and backward reaction of the trans–cis isomerization of peptide bonds (red, forward reaction; green, backward reaction) for (A) Gly–Gly and (B) Pro–Gly systems. Crosses, simulation; solid lines, Kramers’ equation fit to values obtained from simulations; dotted lines, Arrhenius-like fit to values obtained from simulations. The fitting equation are $\log k_1 = 13.46 - 1/2 \log T - 1910/T$, $\log k_{-1} = 12.75 - 1/2 \log T - 2359/T$ for Gly–Gly and $\log k_1 = 13.29 - 1/2 \log T - 2187/T$, $\log k_{-1} = 12.14 - 1/2 \log T - 2080/T$ for Gly–Pro.

peptides, respectively. It can be seen that the logarithms of the simulated rate constants obey a nearly linear relationship with inverse temperature and fit a linear Arrhenius-like relationship very well (solid lines in Figure 8A and B). We adopted here the temperature dependence of the pre-exponential term from Kramers’ equation (see the legend to Figure 8 for the respective equations). In Figure 8A and B, we also present the plots of the logarithms of the rate constants calculated from the

PMF curves (crosses in Figure 6A and B) by using Kramers' theory (eq 11). It can be seen that these rate constants correlate very well with those determined from simulations (Figure 8) except that the latter are about 2.5 times greater. This difference is understandable in view of the fact that Kramers' theory involves a harmonic approximation to the energy in the neighborhood of the substrate and the product of the reaction. Another reason could be a possible overestimation of the effective Stokes radii of the regular and proline peptide groups and a possible overestimation of the friction coefficient of water because its macroscopic value is used in eq 11.

The rate constants calculated by using Kramers' equation are on the same order of magnitude as those determined from UNRES Langevin dynamics simulations (Figure 8). Therefore, in order to determine whether our *ab initio* calculations of the parent potential-energy surfaces translate to the kinetics of the trans–cis isomerization comparable to experimental results, we used Kramers' equation to estimate the rate constants from the reaction energy profiles shown in Figure 4A and B for N-methylacetamide and N-acetylpyrrolidine, respectively. (While UNRES simulations could be used, all-atom MD simulations could not be used to determine the rates of these processes. This is because the experimental trans–cis isomerization rates are on the order of several reciprocal seconds.^{39–41}) We obtained $k_1 = 6.2 \text{ s}^{-1}$ and $k_{-1} = 337 \text{ s}^{-1}$ for N-methylacetamide and $k_1 = 3.1 \text{ s}^{-1}$ and $k_{-1} = 2.3 \text{ s}^{-1}$ for N-acetylpyrrolidine, respectively, at $T = 300 \text{ K}$. The value of k_1 for N-methylacetamide is comparable to the experimental k_1 value for zwitterionic glycylglycine,⁴¹ which is 13 s^{-1} , while the total $k_1 + k_{-1}$ rate constant for N-acetylpyrrolidine is within the range of trans–cis isomerization rate constants determined by NMR for N-acetylprolines,⁴⁰ which range from 0.05 s^{-1} to 100 s^{-1} (as calculated from the Arrhenius equation parameters provided in ref 40). Therefore, the fact that the rate constants are by about 6 orders of magnitude greater in UNRES simulations results from the reduction of the number of the degrees of freedom in the UNRES representation of polypeptide chains.

4. CONCLUSIONS

By calculating the energy surfaces of N-methylacetamide and N-acetylpyrrolidine, which represent the regular and the proline peptide bonds, respectively, with the use of *ab initio* molecular quantum mechanics at the MP2/6-31G(d,p) level, converting them into the potentials of mean force, and fitting analytical expressions to the PMFs, we have determined the effective potentials for the deformation of the C^a…C^a virtual bonds as functions of virtual-bond distance. The potentials are bimodal (Figure 6), with one minimum (at the C^a…C^a distance of $d \approx 3.8 \text{ \AA}$) corresponding to the trans and the other one (at $d \approx 2.8 \text{ \AA}$) to the cis configuration of a peptide bond (as calculated from standard bond lengths and bond angles in proteins⁴⁵). Moreover, the potentials include temperature dependence, which affects the difference between the free energy of the trans and the cis forms. The calculated difference between the PMF minima corresponding to the cis and the trans forms of N-methylacetamide (2 kcal/mol) agrees well with the experimental free-energy difference between the cis and the trans forms of glycylglycine (3 kcal/mol).⁴¹ The cis and trans forms of N-acetylpyrrolidine have nearly the same PMF values; however, the trans form is likely to be favored in the crowded environment of a polypeptide chain. (The free energies of the cis conformations of N-acetylpyrrolidines are only 0.1 to 0.7 kcal/mol higher than those of the trans conformations,⁴⁰ but the free energy of the cis form of N-N-acetylglycylproline methyl ester is almost 1.5 kcal/mol higher

than that of the trans form.⁴⁴) The free-energy barriers to the trans–cis transition of N-methylacetamide (10 kcal/mol) and N-acetylpyrrolidine (13 kcal/mol) agree with the experimental free-energy barriers measured for glycylglycine (12.6–13.9 kcal/mol)⁴¹ and for N-acetylprolines (17.3–19.6 kcal/mol).^{39,40}

On the basis of the results of UNRES Langevin dynamics simulations and Kramers' theory of reaction rate,³⁸ it can be concluded that the reduction of polypeptide chain representation in UNRES results in a 6 orders of magnitude greater rate of trans–cis transition of the peptide bonds compared to experimental data, which makes coarse-grained simulations of protein folding, including systems in which the cis–trans isomerization is a crucial step, viable.

It should be noted that, to be used in UNRES, the virtual C^a…C^a bond potentials must be accompanied by the local [backbone-valence (U_b), backbone-dihedral angle (U_{tor} and U_{tord}), and side-chain-rotamer (U_{rot})] potentials involving cis peptide bonds, as well as peptide-group interaction (U_{pp}) and correlation (U_{corr}) potentials. Work on the determination of these potentials is currently underway in our laboratory.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +48585235430. E-mail: adam@chem.univ.gda.pl.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by grants from the Polish Ministry of Science and Higher Education (538-8372-0510-1), the National Institutes of Health (GM-14312), and the National Science Foundation (MCB10-19767). This research was supported by an allocation of advanced computing resources provided by the National Science Foundation (<http://www.nics.tennessee.edu/>) and by the National Science Foundation through TeraGrid resources provided by the Pittsburgh Supercomputing Center. Computational resources were also provided by (a) the Informatics Center of the Metropolitan Academic Network (IC MAN) in Gdańsk, (b) our 624-processor Beowulf cluster at the Baker Laboratory of Chemistry, Cornell University, and (c) our 184-processor Beowulf cluster at the Faculty of Chemistry, University of Gdańsk.

■ REFERENCES

- (1) Pande, V. S.; Baker, I.; Chapman, J.; Elmer, S.; Kalig, S.; Larson, S. M.; Rhee, Y. M.; Shirts, M. R.; Snow, C. D.; Sorin, E. J.; Zagrovic, B. *Biopolymers* **2003**, *68*, 91–109.
- (2) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (3) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (4) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. *ACM/IEEE SC 2006 Conference (SC.06)*; New York, NY, USA, 2006; pp 43–43.
- (5) Friedrichs, M. S.; Eastman, P.; Vaidyanathan, V.; Houston, M.; Legrand, S.; Beberg, A. L.; Ensign, D. L.; Bruns, C. M.; Pande, V. S. *J. Comput. Chem.* **2009**, *30*, 864–872.
- (6) Shaw, D. E.; et al. *Commun. ACM* **2008**, *51*, 91–97.
- (7) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. *Science* **2010**, *330*, 341–346.

- (8) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. *Science* **2011**, *334*, 517–520.
- (9) Liwo, A.; Oldziej, S.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Scheraga, H. A. *J. Comput. Chem.* **1997**, *18*, 849–873.
- (10) Liwo, A.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Oldziej, S.; Scheraga, H. A. *J. Comput. Chem.* **1997**, *18*, 874–887.
- (11) Liwo, A.; Kaźmierkiewicz, R.; Czaplewski, C.; Groth, M.; Oldziej, S.; Wawak, R. J.; Rackovsky, S.; Pincus, M. R.; Scheraga, H. A. *J. Comput. Chem.* **1998**, *19*, 259–276.
- (12) Liwo, A.; Czaplewski, C.; Pillardy, J.; Scheraga, H. A. *J. Chem. Phys.* **2001**, *115*, 2323–2347.
- (13) Liwo, A.; Arlukowicz, P.; Czaplewski, C.; Oldziej, S.; Pillardy, J.; Scheraga, H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1937–1942.
- (14) Oldziej, S.; Kozłowska, U.; Liwo, A.; Scheraga, H. A. *J. Phys. Chem. A* **2003**, *107*, 8035–8046.
- (15) Liwo, A.; Oldziej, S.; Czaplewski, C.; Kozłowska, U.; Scheraga, H. A. *J. Phys. Chem. B* **2004**, *108*, 9421–9438.
- (16) Oldziej, S.; Łagiewka, J.; Liwo, A.; Czaplewski, C.; Chinchio, M.; Nania, M.; Scheraga, H. A. *J. Phys. Chem. B* **2004**, *108*, 16950–16959.
- (17) Kozłowska, U.; Liwo, A.; Scheraga, H. A. *J. Phys.: Condens. Matter* **2007**, *19*, 285203.
- (18) Liwo, A.; Khalili, M.; Czaplewski, C.; Kalinowski, S.; Oldziej, S.; Wachucik, K.; Scheraga, H. *J. Phys. Chem. B* **2007**, *111*, 260–285.
- (19) Liwo, A.; Czaplewski, C.; Oldziej, S.; Rojas, A. V.; Kaźmierkiewicz, R.; Makowski, M.; Murarka, R. K.; Scheraga, H. A. In *Coarse-Graining of Condensed Phase and Biomolecular Systems*; Voth, G., Ed.; CRC Press: Boca Raton, FL, 2008; Chapter 8, pp 1391–1411.
- (20) He, Y.; Xiao, Y.; Liwo, A.; Scheraga, H. A. *J. Comput. Chem.* **2009**, *30*, 2127–2135.
- (21) Kozłowska, U.; Liwo, A.; Scheraga, H. A. *J. Comput. Chem.* **2010**, *31*, 1143–1153.
- (22) Kozłowska, U.; Maisuradze, G. G.; Liwo, A.; Scheraga, H. A. *J. Comput. Chem.* **2010**, *31*, 1154–1167.
- (23) Khalili, M.; Liwo, A.; Jagielska, A.; Scheraga, H. *J. Phys. Chem. B* **2005**, *109*, 13798–13810.
- (24) Liwo, A.; Khalili, M.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2362–2367.
- (25) Liwo, A.; Oldziej, S.; Czaplewski, C.; Kleinerman, D. S.; Blood, P.; Scheraga, H. A. *J. Chem. Theory Comput.* **2010**, *6*, 583–595.
- (26) Kubo, R. *J. Phys. Soc. Jpn.* **1962**, *17*, 1100–1120.
- (27) Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953–4963.
- (28) Wedemeyer, W. J.; Welker, E.; Scheraga, H. A. *Biochemistry* **2002**, *41*, 14637–14644.
- (29) Jakob, R. P.; Zierer, B. K.; U., U. W.; Hofman, S. D.; Lorentz, S. H.; Balbach, J.; Bobbek, H.; Schmid, F. X. *J. Mol. Biol.* **2010**, *399*, 331–346.
- (30) Shen, H.; Liwo, A.; Scheraga, H. A. *J. Phys. Chem. B* **2009**, *113*, 8738–8744.
- (31) Kolinski, A.; Skolnick, J. *J. Chem. Phys.* **1992**, *97*, 9412–9426.
- (32) Khalili, M.; Liwo, A.; Rakowski, F.; Grochowski, P.; Scheraga, H. *J. Phys. Chem. B* **2005**, *109*, 13785–13797.
- (33) de Gennes, P.-G. *Scaling Concepts in Polymer Physics*; Cornell University Press: Ithaca, NY, 1979; Chapter VI.
- (34) Veitshans, T.; Klimov, D.; Thirumalai, D. *Fold. Des.* **1996**, *2*, 1–22.
- (35) Cieplak, M.; Hoang, T. X.; Robbins, M. O. *Proteins: Struct., Funct., Genet.* **2002**, *49*, 104–113.
- (36) Marquardt, D. W. *J. Soc. Indust. Appl. Math.* **1963**, *11*, 431–441.
- (37) Eicher, L. D.; Zwolinski, B. *J. Phys. Chem.* **1971**, *75*, 2016–2024.
- (38) Kramers, H. A. *Physica* **1940**, *7*, 284–304.
- (39) Maia, H. L.; Orrell, K. G.; Rydon, H. N. *J. Chem. Soc. D* **1971**, 1209.
- (40) Maia, H. L.; Orrell, K. G.; Rydon, H. N. *J. Chem. Soc. D* **1976**, 761–763.
- (41) Holtz, J. S. W.; Li, P.; Asher, S. A. *J. Am. Chem. Soc.* **1999**, *121*, 3762–3766.
- (42) Jhon, J. S.; Kang, Y. K. *J. Phys. Chem. A* **1999**, *103*, 5436–5439.
- (43) Kang, Y. K.; Choi, H. Y. *Biophys. Chem.* **2006**, *111*, 135–142.
- (44) Eberhardt, E. S.; Loh, S. N.; Raines, R. T. *Tetrahedron Lett.* **1993**, *34*, 3055–3056.
- (45) Engh, R. A.; Huber, R. *Acta Crystallogr., Sect. A* **1991**, *47*, 392–400.