Free Energies for Folding and Refolding of Four Types of β Turns: Simulation of the Role of D/L Chirality

Yibing Yan,[†] Bruce W. Erickson,*[,†] and Alexander Tropsha*[,‡]

Contribution from the Department of Chemistry and the Laboratory for Molecular Modeling, School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Received January 31, 1994[⊗]

Abstract: Rational protein design would benefit from quantitative estimates of the free-energy differences (ΔG) among β turn conformers. We have simulated a chirally representative set of nine loop dipeptides of the form CH_3CO-L1-L2-$NHCH_3$, where L1 and L2 were Gly, L-Ala, or D-Ala. The ΔG value for each glycine (type-II) and inverse-glycine (type-II′) β-turn conformer (relative to the Gly-Gly conformers) was estimated in an explicit water environment by the slow-growth method, in which an α-hydrogen atom of a Gly residue is replaced by a methyl group to give L-Ala or D-Ala. The ΔG values ranged from 0.9 to 6.5 kcal/mol. The molecular symmetry of this model system allowed estimation of the ΔG values for II → II′ refolding, which were spread from −4.6 to 4.6 kcal/mol. The relative free-energy change ($\Delta\Delta G_f$) for folding of each loop dipeptide from its extended (ε) conformer into its type-II or type-II′ β-turn conformer (relative to folding of the Gly-Gly loop dipeptide) was calculated from a thermodynamic cycle. The $\Delta\Delta G_f$ values for ε → II or ε → II′ folding ranged from −2.2 to 2.5 kcal/mol, and those for ε → I or ε → I′ folding spanned from −2.1 to 0.9 kcal/mol. The contribution of an L-Ala or D-Ala residue to the $\Delta\Delta G_f$ for folding of the loop dipeptides into four β-turn conformations (I, I′, II, II′) ranged from −1.1 to 1.7 kcal/mol. The $\Delta\Delta G_f$ values for folding of Gly-Gly, Gly-L-Ala, and L-Ala-Gly into each of these β-turn conformations correlated with the relative occurrences of these β-turn conformations in natural proteins, suggesting that formation of a β turn during protein folding is mainly guided by local interactions. During rational design of a synthetic protein, placing D-Ala at L1 alone should favor the type-II′ β turn, D-Ala at L2 alone should favor the type-II β turn, and D-Ala at both L1 and L2 should favor the type-I′ β turn. Using Gly at L1 should favor either a type-I′ or -II′ β turn and Gly at L2 should favor either a type-I or -II β turn.

Rational design of well-folded proteins requires better understanding of the local folding of secondary structures. Reverse β turns are among the simplest elements of secondary structure and comprise about 25% of all residues in proteins.[1] Four major types of β turns are the common (type-I), inverse-common (type-I′), glycine (type-II), and inverse-glycine (type-II′) turns.[2] They mainly occur between two antiparallel β strands and serve to reverse the direction of the peptide chain of these strands. Traditionally these turns are described by a tetrapeptide segment whose residues are designated i, $i+1$, $i+2$, and $i+3$ from N to C terminus.[3,4] The terminal residues i and $i+3$ usually have β-strand main chain conformations but with their peptide bonds oriented antiparallel to one another. The central residues $i+1$ and $i+2$ (also called[5] the loop residues L1 and L2) reverse the direction of the peptide chain and have characteristic main chain dihedral angles for each type of β turn.

Reverse turns help to maintain the compact structure of a globular protein and may also help to initiate protein folding, as suggested by recent theoretical studies.[6-8] For example, the importance of local folding of β turns became evident during the engineering of betabellin, a 64-residue β protein.[9-15] Each 32-residue chain of betabellin is designed to fold into a β sheet comprised of four antiparallel β strands joined by three β turns. These β turns should form properly during folding of the β sheet and maintain their integrity upon the dimerization of two β sheets. Successful folding of this designed, nongenetic protein requires placing within the sequence of each betabellin chain three pairs of residues with high propensity to form the appropriate β turns.

* Authors to whom all correspondence should be addressed.
† Department of Chemistry.
‡ School of Pharmacy.
⊗ Abstract published in *Advance ACS Abstracts*, July 1, 1995.

(1) Kabsch, W.; Sanders, C. *Biopolymers* **1983**, 2577-2637.
(2) Richardson, J. S.; Richardson, D. C. In *Prediction of Protein Structure and the Principles of Protein Conformation*; Fasman, G. D., Ed.; Plenum: New York, 1989; pp 1-98.
(3) Venkatachalam, C. M. *Biopolymers* **1968**, *6*, 1425-1436.
(4) Richardson, J. S. *Adv. Protein Chem.* **1981**, *34*, 167-339.
(5) Sibanda, B. L.; Thornton, J. M. *Nature* **1985**, *316*, 170-174.
(6) Scholnik, J.; Kolinski, A. *J. Mol. Biol.* **1991**, *221*, 499-531.
(7) Soman, K. V.; Karini, A.; Case, D. A. *Biopolymers* **1991**, *31*, 1351-1361.
(8) Dill, K. A.; Fiebig, K. M.; Chan, H. S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1942-1946.
(9) Erickson, B. W.; Daniels, S. B.; Reddy, P. A.; Unson, C. G.; Richardson, J. S.; Richardson, D. C. In *Computer Graphics and Molecular Modeling*; Fletterick, R., Zoller, M., Ed.; Cold Spring Harbor Laboratory Press: Plainview, NY, 1986; pp 53-57.
(10) Richardson, J. S.; Richardson, D. C. In *Protein Engineering*; Oxender, D. L., Fox, C. F., Eds.; Alan R. Liss: New York, 1987; pp 149-163.
(11) Erickson, B. W.; Daniels, S. B.; Reddy, P. A.; Higgins, M. L.; Richardson, J. S.; Richardson, D. C. *ICSU Short Rep.* **1988**, *8*, 4-5.
(12) Richardson, J. S.; Richardson, D. C. In *Protein Folding*; Gierasch, L. M., King, J., Eds.; AAAS: Washington, DC, 1988; pp 5-17.
(13) McClain, R. D.; Daniels, S. B.; Williams, R. W.; Pardi, A.; Hecht, M.; Richardson, J. S.; Richardson, D. C.; Erickson, B. W. In *Peptides: Chemistry, Structure and Biology*; Rivier, J. E., Marshall, G. R., Eds.; ESCOM Science Publishers: Leiden, The Netherlands, 1990; pp 682-684.
(14) McClain, R. D.; Yan, Y.; Williams, R. W.; Donlan, M. E.; Erickson, B. W. In *Peptides: Chemistry, Structure and Biology;* Smith, J. A., Rivier, J. E., Eds.; ESCOM Science Publishers: Leiden, The Netherlands, 1991; pp 364-365.
(15) Yan, Y.; Erickson, B. W. In *Peptides: Chemistry, Structure, and Biology;* Hodges, R. S., Smith, J. A., Eds.; ESCOM Science Publishers: Leiden, The Netherlands, 1994; pp 1027-1029.

0002-7863/95/1517-7592$09.00/0

Analysis of β-turn models[3] and β turns in proteins[16] as well as molecular mechanics calculations of β-turn models[17] have identified several preferred main chain conformations for the loop residues L1 and L2. Statistical analysis of known protein structures indicated that a variety of loop dipeptides are found in specific β-turn conformations but certain dipeptides are favored.[18] Structural analysis of β turns[3,16] and experimental studies[19,20] have shown that the presence of a D-amino acid at either or both loop positions favors the formation of different types of β turns, which can be correlated with the presence of the achiral Gly residue at these loop positions in these types of β turns in natural proteins. For example, the use of two D-amino acid residues at each turn improved the folding stability and water solubility of recent betabellins.[13-15] What is needed is a quantitative measure of this role of chirality in stabilizing the four major types of β turns. Current experimental approaches for analysis of sequence−stability relationships are in general not applicable to an isolated β turn because short polypeptide sequences usually do not form one particular conformation in solution.[21] Molecular dynamics simulations in explicit solvent provide an appealing alternative to experimentation.[22-25] The accuracy of these methods in reproducing experimental data has been shown in several cases,[24] including accurate calculations of the relative α-helical propensity of certain amino acids[26] and correct estimates of the relative binding constants of several ligands for their receptors.[27,28]

Recently we evaluated quantitatively the role of chirality in stabilizing type-I and type-I′ β turns by free-energy simulation.[29,30] As β-turn models, we used a set of nine blocked dipeptides of the form CH_3CO-L1-L2-$NHCH_3$, where the loop residues L1 and L2 were Gly, L-Ala, or D-Ala. In this paper, we expand this work to include the type-II and type-II′ β turns. For each model dipeptide, we calculated the relative free energy of folding from the extended conformation into either the type-II or the type-II′ β turn and the free energy of refolding from one β-turn conformation into the other. As expected, we found that the chirality of the loop residues plays a major role in determining the preferred β-turn conformation. The L-Ala-D-Ala dipeptide forms the most stable type-II β turn and D-Ala-L-Ala dipeptide forms the most stable type-II′ β turn in aqueous solution. Comparison of the results of our calculations with conformational trends of native proteins supports the idea that the folding of β turns occurs largely due to local interactions.

Methods

Peptide Models. The classical β-turn model[3,17,31,32] is the loop dipeptide CH_3CO-L1-L2-$NHCH_3$ where L1(i + 1) and L2(i + 2) are the first and the second loop residues, respectively.[5] This acetylated dipeptide methylamide can be perceived as a tetrapeptide lacking the N-terminal amino group and the C-terminal carboxyl group. This loop dipeptide is the shortest peptide that has the four main chain dihedral angles (ϕ_{L1}, ψ_{L1}, ϕ_{L2}, ψ_{L2}) needed to define the different β-turn types. We used a chirally representative set of nine loop dipeptides that contains all possible combinations of achiral glycine (Gly, G) and the simplest chiral amino acids, L-alanine (L-Ala, A) and D-alanine (D-Ala, a). These loop dipeptides and their two-letter codes are L-Ala-L-Ala (AA), L-Ala-Gly (AG), L-Ala-D-Ala (Aa), Gly-L-Ala (GA), Gly-Gly (GG), Gly-D-Ala (Ga), D-Ala-L-Ala (aA), D-Ala-Gly (aG), and D-Ala-D-Ala (aa). A particular conformation is indicated by appending one of the following symbols to the two-letter code of the loop dipeptide: I for the type-I (common) β turn, I′ for the type-I′ (inverse-common) β turn, II for the type-II (glycine) β turn, II′ for the type-II′ (inverse-glycine) β turn, and ε for the extended main chain conformation (ϕ_{L1}, ψ_{L1}, ϕ_{L2}, and ψ_{L2} each 180°). For instance GGε denotes the Gly-Gly loop dipeptide with a fully extended peptide chain and aAII′ denotes the D-Ala-L-Ala loop dipeptide in the type-II′ β-turn conformation.

Molecular Dynamics Simulation. Molecular dynamics simulation was carried out with the CEDAR program,[33] which used a force field having the same nonbonded parameters[34] as those of the GROMOS program[35] but a different description of geometry and geometric deformations.[36] It employed the SHAKE algorithm[37] in order to maintain bond lengths constant and a time step of 2 fs. Peptide molecules were surrounded with SPC water[35] in rectangular boxes and periodic boundary conditions were used. The sizes of the boxes in each dimension were equal to the size of the peptide in this dimension plus twice the non-bonded cutoff value; for instance, the GAII peptide was solvated with 437 water molecules in a box with dimensions 24 × 24 × 23.5 Å. Mean temperature and pressure were maintained constant (300 K and 1 atm, respectively) by small adjustments at each time step of the kinetic energy and the dimensions of the periodic box.[35] An all-atom representation and a 8-Å cutoff for nonbonded interactions were used. In each simulation, water was initially equilibrated with an immobile peptide molecule for 20 ps followed by equilibration of the whole system for 40 ps to a temperature of 300 K and a pressure of 1 atm before beginning the molecular replacement calculation. A detailed description of the method is given by Hermans et al.[38]

Conformational restraints were used to keep each loop dipeptide within the local minimum of a specific β-turn conformation during molecular replacement simulation. CEDAR confined the conformation by use of a flat-bottom potential, so that the conformational probability distribution was unperturbed locally.[39] The dihedral-angle restraining potential was zero within a 60° range centered at the characteristic value of each main chain dihedral angle for a specific β-turn conformation but increased rapidly beyond this range. For example, for the type-II′ β turn the free energy of the system was not changed by these constraints when $\phi_{L1} = (60 \pm 30)°$, $\psi_{L1} = (-120 \pm 30)°$, $\phi_{L2} = (-80 \pm 30)°$, and $\psi_{L2} = (0 \pm 30)°$.

Estimation of Free-Energy Differences. Molecular replacement simulation[22-24] was used to estimate the free-energy difference (ΔG) between two molecules of different loop dipeptides in the same

(16) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1−109.
(17) Zimmerman, S. S.; Scheraga, H. A. *Biopolymers* **1978**, *17*, 1849−1869.
(18) Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221−232.
(19) Kamatani, Y.; Minakata, H.; Iwashita, T.; Nomoto, K.; In, Y.; Doi, M.; Ishida, T. *FEBS Lett.* **1990**, *276*, 95−97.
(20) Convert, O.; Duplaa, H.; Lavielle, S.; Chassaing, G. *Neuropeptides* **1991**, *19*, 259−270.
(21) Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1988**, *201*, 161−200.
(22) van Gunsteren, W. F. *Protein Eng.* **1988**, *2*, 5−13.
(23) Karplus, M.; Petsko, G. A. *Nature* **1990**, *347*, 631−639.
(24) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 993−1023.
(25) Kollman, P. A.; Merz, K. M. *Acc. Chem. Res.* **1990**, *23*, 246−252.
(26) Hermans, J.; Anderson, A. G.; Yun, R. H. *Biochemistry* **1992**, *31*, 5646−5653.
(27) Tropsha, A.; Hermans, J. *Protein Eng.* **1992**, *5*, 29−33.
(28) Ferguson, D. M.; Radmer, R. J.; Kollman, P. A. *J. Med. Chem.* **1991**, *34*, 2654−2659.
(29) Yan, Y.; Tropsha, A.; Hermans, J.; Erickson, B. W. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7892−7902.
(30) Tropsha, A.; Yan, Y.; Schneider, S. E.; Li, L.; Erickson, B. W. In *Peptides: Chemistry, Structure, and Biology;* Hodges, R. S., Smith, J. A., Eds.; ESCOM Science Publishers: Leiden, The Netherlands, 1994; pp 883−885.
(31) Chandrasekaran, R.; Lakshminarayanan, A. V.; Pandya, U. V.; Ramachandran, G. N. *Biochim. Biophys. Acta* **1973**, *303*, 14−27.
(32) Maigret, B.; Pullman, B. *Theor. Chim. Acta* **1974**, *35*, 113−128.
(33) Carson, M.; Hermans, J. In *Molecular Dynamics and Protein Structure;* Hermans, J., Ed.; Polycrystal Book Service: Western Springs, 1985; pp 165−166.
(34) Hermans, J.; Berendsen, H. J. C.; van Gunsteren, W. F.; Postma, J. P. M. *Biopolymers* **1984**, *23*, 1513−1518.
(35) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces;* Pullman, B., Ed.; Reidel: Dordrecht, 1981; pp 331−342.
(36) Ferro, D. R.; McQueen, J. E.; McCown, J. T.; Hermans, J. *J. Mol. Biol.* **1980**, *136*, 1−18.
(37) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(38) Hermans, J.; Yun, R. H.; Anderson, A. G. *J. Comput. Chem.* **1992**, *13*, 429−442.
(39) Yun, R. H.; Hermans, J. *Protein Eng.* **1991**, *4*, 761−766.

conformation in water. The free-energy difference ΔG between two different molecules in the same conformation (such as GGII and GAII) was computed by describing both molecules in the system. The potential energy of the system, U_p, explicitly depends on the coupling parameter λ.

$$U_p = U_{common} + \lambda U_{molecule1} + (1 - \lambda)U_{molecule2}$$

where U_{common} is the potential energy of the atoms common to the two molecules, and $U_{molecule1}$ and $U_{molecule2}$ represent the potential energy of the nonbonded interactions between the unique atoms of molecule 1 (e.g. GGII) and molecule 2 (e.g. GAII), respectively, with the rest of the system. Thus at $\lambda = 1$ the value of U_p measures the potential energy of the sysetm containing molecule 1 but not molecule 2, whereas at $\lambda = 0$ it measures the potential energy of the system containing molecule 2 but not molecule 1.

CEDAR implements the slow-growth scheme of free-energy simulation,[40] in which the free-energy change for converting molecule 1 into molecule 2 is evaluated as the work done on the system by changing its potential in a quasi-static process.

$$\Delta G = \int_{\lambda=1}^{\lambda=0} \langle \partial U_p / \partial \lambda \rangle \, d\lambda$$

where the barckets "$\langle\ \rangle$" indicate the average over a Boltzmann distribution at each value of λ. In practice, the free-energy change for converting molecule 1 (at $\lambda = 1$) to molecule 2 (at $\lambda = 0$) is estimated as

$$\Delta G = \sum_{\lambda=1}^{\lambda=0} (\partial U / \partial \lambda)\, \delta\lambda \qquad (1)$$

where the value of λ is changed by the same small amount at every step of a molecular dynamics simulation that starts at $\lambda = 1$ and ends at $\lambda = 0$.

A typical replacement cycle involved equilibration of the whole system for 20 ps at $\lambda = 1$, forward replacement for 60 ps as λ was gradually decreased from 1 to 0, equilibration of the whole system for 20 ps at $\lambda = 0$, and reverse replacement for 60 ps as λ was gradually increased from 0 to 1. For each molecular replacement simulation, this replacement cycle was repeated four times starting from different configurations to reduce random noise and to estimate the precision of calculated ΔG values.[38,39] The ΔG value reported for each replacement is the mean $\pm$ root-mean-square deviation (rmsd) of eight independent estimates of ΔG, namely, the ΔG values of the four forward replacements plus the ΔG values of the four reverse replacements with the sign reversed.

Reversibility was measured by examining the superposition of the forward and reverse progress curves and by calculating the hysteresis (the sum of ΔG values for the forward and reverse replacements), which ideally should be zero. The self-consistency of the calculated ΔG values was studied by considering the network of all single-replacement paths between two different loop peptides in the type-II conformation and the symmetry-related network for the type-II′ conformation (Figure 1). These networks contain several closed thermodynamic cycles. The ΔG values between two systems at equilibrium should be independent of the path of transformation, and the ΔG values around a closed cycle involving several systems at equilibrium should be zero. Comparison of the ΔG values for different paths between two loop peptides in the same conformation provided a rigorous test of equilibration of the system and of the self-consistency of the individual ΔG values.

Molecular Symmetry. By definition[3] the four characteristic main chain dihedral angles for the loop residues L1 and L2 of GGII′ ($\phi_{L1} = 60°$, $\psi_{L1} = -120°$, $\phi_{L2} = -80°$, $\psi_{L2} = 0°$) are equal in magnitude but opposite in sign to those of GGII ($\phi_{L1} = -60°$, $\psi_{L1} = 120°$, $\phi_{L2} = 80°$, $\psi_{L2} = 0°$). Thus GGII′, the inverse-glycine (type-II′) β-turn conformer of the achiral loop dipeptide GG, is the mirror image of GGII, the glycine (type-II) β-turn conformer. Because of this symmetry relationship, conformers GGII and GGII′ must have exactly the same free energy. In other words, the free energy for refolding of GGII into GGII′ is zero. Furthermore, since L-alanine and D-alanine are mirror images, the type-II β-turn conformer AGII is the mirror image of the type-II′ β-turn conformer at aGII′, and the type-II β-turn conformer aAII is the mirror image of the type-II′ β-turn conformer AaII′, as illustrated in Figure 2. Thus by symmetry the conformers AGII and aGII′ have the same free energy, and the conformers aAII and AaII′ have the same free energy. By the same argument, seven other mirror-image pairs of loop dipeptides (aGII/AGII′, GAII/GaII′, GaII/GAII′, AAII/aaII′, AaII/aAII′, aAII/AaII′, aaII/AAII′) have the same free energy (Figure 1).

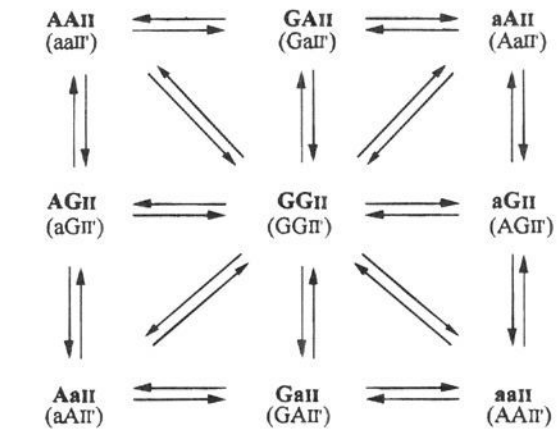


Figure 1. Replacement path graphs for interconverting either the glycine (type-II) β-turn conformers (bold) or the inverse-glycine (type-II′) β-turn conformers (parentheses) of nine loop dipeptides. Calculated estimates of the free-energy change for replacing Gly by L-Ala or D-Ala along each path (arrows) are listed in Table 1.

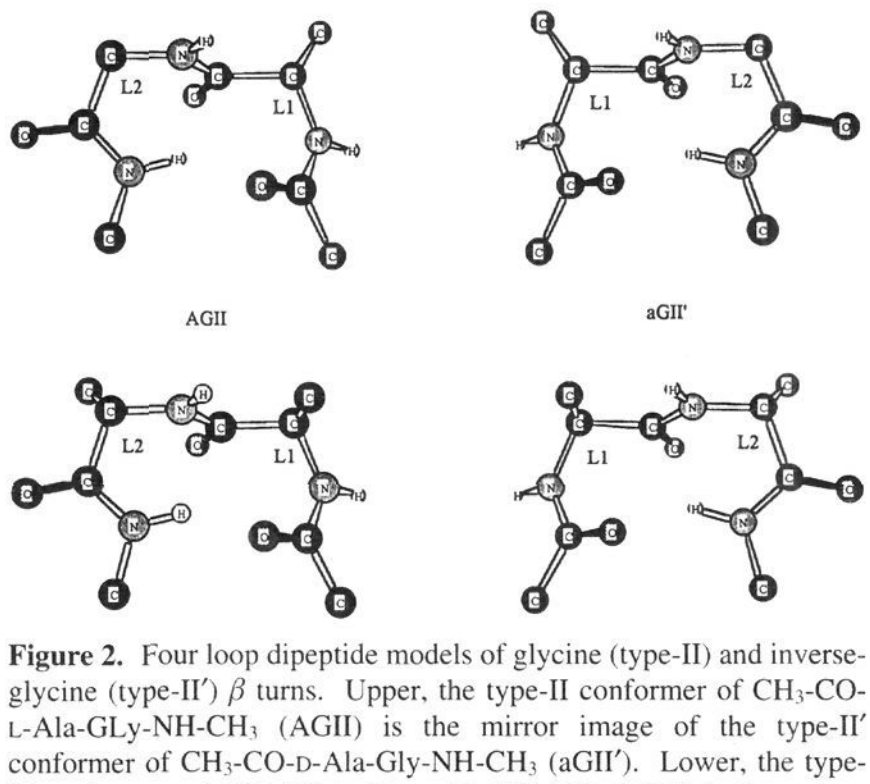
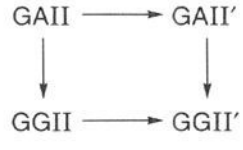


Figure 2. Four loop dipeptide models of glycine (type-II) and inverse-glycine (type-II′) β turns. Upper, the type-II conformer of CH_3-CO-L-Ala-GLy-NH-CH_3 (AGII) is the mirror image of the type-II′ conformer of CH_3-CO-D-Ala-Gly-NH-CH_3 (aGII′). Lower, the type-II conformer of CH_3-CO-D-Ala-L-Ala-NH-CH_3 (aAII) is the mirror image of the type-II′ conformer of CH_3-CO-L-Ala-D-Ala-NH-CH_3 (AaII′).

Thermodynamic Cycles. The free energy for refolding (ΔG_r) and relative free energy for folding ($\Delta\Delta G_f$) of each loop dipeptide were calculated by use of thermodynamic cycles.[41] For example, the following thermodynamic cycle was used to calculate ΔG_r of the GA loop dipeptide from its type-II β-turn conformer into its type-II′ β-turn conformer (GAII → GAII′).



(40) Berendsen, H. J. C ; Postma, J. P. M.; van Gunsteren, W. F. In *Molecular Dynamics and Protein Structure;* Hermans, J., Ed.; Polycrystal Book Service: Western Springs, 1985; pp 43−46.

(41) Tembe, B. L.; McCammon, A. J. *Comput. Chem.* **1984**, *8*, 281−283.

Table 1. Free-Energy Changes for Interconversion of 32 Pairs of Loop Dipeptides in the Glycine (Type-II) or the Inverse-Glycine (Type-II′) β-Turn Conformation

forward replacement path		free-energy change, kcal/mol (mean ± rmsd)			
glycine β-turn conformers	inverse-glycine β-turn conformers	forward ΔG	reverse ΔG	combined ΔG	hysteresis, kcal/mol
GGII → AGII	GGII′ → aGII′	0.96 ± 0.28	−0.86 ± 0.22	0.91 ± 0.25	0.09
GAII → AAII	GaII′ → aaII′	0.78 ± 0.34	−0.96 ± 0.26	0.87 ± 0.30	0.18
GaII → AaII	GAII′ → aAII′	0.94 ± 0.16	−0.71 ± 0.31	0.83 ± 0.26	0.23
GGII → aGII	GGII′ → AGII′	3.71 ± 0.28	−3.46 ± 0.23	3.59 ± 0.28	0.25
GAII → aAII	GaII′ → AaII′	3.79 ± 0.34	−3.81 ± 0.42	3.80 ± 0.35	0.02
GaII → aaII	GAII′ → AAII′	3.91 ± 0.19	−3.83 ± 0.14	3.87 ± 0.16	0.07
GGII → GAII	GGII′ → GaII′	2.81 ± 0.23	−2.68 ± 0.41	2.74 ± 0.33	0.14
AGII → AAII	aGII′ → aaII′	2.93 ± 0.13	−2.52 ± 0.23	2.73 ± 0.28	0.41
aGII → aAII	AGII′ → AaII′	3.10 ± 0.30	−2.86 ± 0.36	2.98 ± 0.33	0.24
GGII → GaII	GGII′ → GAII′	1.06 ± 0.26	−0.82 ± 0.26	0.94 ± 0.28	0.23
AGII → AaII	aGII′ → aAII′	0.80 ± 0.14	−0.64 ± 0.26	0.72 ± 0.21	0.16
aGII → aaII	AGII′ → AAII′	1.39 ± 0.26	−1.26 ± 0.30	1.32 ± 0.24	0.14
GGII → AAII	GGII′ → aaII′	3.96 ± 0.43	−3.42 ± 0.21	3.69 ± 0.42	0.54
GGII → AaII	GGII′ → aAII′	1.58 ± 0.42	−1.88 ± 0.24	1.73 ± 0.36	0.30
GGII → aAII	GGII′ → AaII′	6.49 ± 0.29	−6.45 ± 0.36	6.47 ± 0.30	0.04
GGII → aaII	GGII′ → AAII′	4.73 ± 0.37	−4.62 ± 0.47	4.67 ± 0.40	0.11

The free energy around a thermodynamic cycle is zero, so

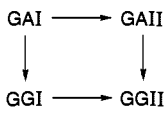
$$\Delta G(\text{GAII}\rightarrow\text{GAII}') - \Delta G(\text{GGII}\rightarrow\text{GGII}') = \Delta G(\text{GAII}\rightarrow\text{GGII}) - \Delta G(\text{GAII}'\rightarrow\text{GGII}')$$

Since $\Delta G(\text{GGII}\rightarrow\text{GGII}')$ is zero by mirror symmetry,

$$\Delta G_r(\text{GAII}\rightarrow\text{GAII}') = \Delta G(\text{GAII}\rightarrow\text{GGII}) - \Delta G(\text{GAII}'\rightarrow\text{GGII}') \quad (2)$$

The latter two values were estimated from molecular replacement simulations. Similar thermodynamic cycles were used to calculate ΔG_r for refolding each of the other loop dipeptides from its type-II β-turn conformation into its type-II′ β-turn conformation.

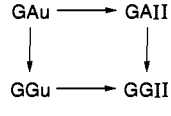
The relative free energy ($\Delta\Delta G_r$) for refolding of a loop dipeptide from the common into the glycine β-turn conformer (relative to the refolding of GG) was calculated similarly. For example, the following thermodynamic cycle was used to calculate $\Delta\Delta G_r$ for refolding of the GA loop dipeptide from its type-I β-turn conformation into its type-II β-turn conformation (GAI → GAII) relative to the same refolding of the GG loop dipeptide (GGI → GGII).



$$\Delta\Delta G_r(\text{GAI}\rightarrow\text{GAII}) = \Delta G(\text{GAI}\rightarrow\text{GAII}) - \Delta G(\text{GGI}\rightarrow\text{GGII}) = \Delta G(\text{GAI}\rightarrow\text{GGI}) - \Delta G(\text{GAII}\rightarrow\text{GGII}) \quad (3)$$

The latter two values were estimated from molecular replacement simulations. Similar thermodynamic cycles were used to calculate $\Delta\Delta G_r$ for refolding of each loop dipeptide from its type-I into its type-II conformation and from its type-I′ into its type-II′ β-turn conformation.

The relative free energy for folding ($\Delta\Delta G_f$) of a loop dipeptide from the unfolded state into a β-turn conformer (relative to the folding of GG) was also calculated by use of a thermodynamic cycle. For example, the following thermodynamic cycle was used to calculate $\Delta\Delta G_f$ for folding of the GA loop dipeptide from its unfolded conformation (u) into its type-II β-turn conformation (GAu→GAII) relative to the same folding of the GG loop dipeptide (GGu→GGII).



$$\Delta\Delta G_f(\text{GAu}\rightarrow\text{GAII}) = \Delta G(\text{GAu}\rightarrow\text{GAII}) - \Delta G(\text{GGu}\rightarrow\text{GGII}) = \Delta G(\text{GAu}\rightarrow\text{GGu}) - \Delta G(\text{GAII}\rightarrow\text{GGII}) \quad (4)$$

The latter two values were estimated from molecular replacement simulations. Similar thermodynamic cycles were used to calculate $\Delta\Delta G_f$ for folding of each loop dipeptide from its unfolded conformation into each of its type-I, type-I′, type-II, and type-II′ β-turn conformations. These calculations require a specific model of the unfolded state. In this paper, we considered two such models employed previously in the literature: the extended conformation,[42] and the random coil.[26] The results for the random coil were obtained earlier by Hermans et al.[26] In this paper, we also performed molecular replacement calculations for the extended conformation (ϵ), in which ϕ_{L1}, ψ_{L1}, ϕ_{L2}, and ψ_{L2} were each kept at (180 ± 30°). As is discussed below, for our systems both models have produced virtually identical results.

Results and Discussion

Free-Energy Differences from Molecular Replacement Simulation. Table 1 lists the calculated ΔG values for interconversion by molecular replacement simulation of 16 pairs of loop dipeptides in either the type-II (glycine) or the type-II′ (inverse-glycine) β-turn conformation. Each value of ΔG for a forward or reverse molecular replacement is the average of four independent simulations. The combined (forward) ΔG value for each interconversion was calculated from eight independent estimates (absolute values of the four forward ΔG and four reverse ΔG). Replacing one or both glycine residues of GGII or GGII′ by alanine increased the free energy of the system by 0.9−6.5 kcal/mol. The reversibility of each interconversion is indicated by the relatively low value of its hysteresis.

Self-Consistency. In principle, only eight molecular replacement calculations are sufficient to evaluate the ΔG values for interconversion of the GG loop dipeptide and the eight alanine-containing loop dipeptides. These interconversions are the eight radial paths that connect the central GG peptide with its eight neighbors at the periphery of the molecular replacement graph (Figure 1). If the molecular replacements are simulated within a well-equilibrated system, the calculated ΔG values between pairs of these thermodynamic states should be independent of the pathway and therefore self-consistent. For example, GGII can be converted into AAII by changing L1 first (GGII → AGI → AAII), by changing L2 first (GGII → GAII → AAII), or by changing both simultaneously (GGII → AAII). From the data in Table 1, the ΔG values for these three independent paths are quite similar (3.64, 3.61, and 3.69 kcal/mol, respectively). The ΔG values for other sets of alternate paths in Figure 1 are

(42) (a) Tidor, B.; Karplus, M. *Biochemistry* **1991**, *30*, 3217−3228. (b) Tobias, D. J.; Sneddon, S. F.; Brooks, C. L.; III. *J. Mol. Biol.* **1990**, *216*, 783−796.

(43) Tobias, D. J.; Sneddon, S. F.; Brooks, C. L., III In *Advances in Biomedical Simulations*; American Institute of Physics Conference Proceedings No. 239, Obernai, France, 1991; pp 174−179.

Table 2. Position ΔG Values for Replacement of Gly by L-Ala or D-Ala in a Loop Dipeptide

loop residue	glycine β-turn conformers[a]	inverse-glycine β-turn conformers[a]	ΔG,[b] kcal/mol (mean ± rmsd)
L1($i+1$)	GXII → AXII	GXII′ → aXII′	0.88 ± 0.27
L1($i+1$)	GXII → aXII	GXII′ → AXII′	3.71 ± 0.27
L2($i+2$)	XGII → XAII	XGII′ → XaII′	2.80 ± 0.31
L2($i+2$)	XGII → XaII	XGII′ → XAII′	0.98 ± 0.25

[a] For X equal to Gly, L-Ala, or D-Ala. [b] Each value is the average of 24 independent estimates of ΔG.

essentially independent of the path taken. These results indicate that the ΔG values of Table 1 are self-consistent and suggest that, by confining each of the main chain dihedral angles to stay within ±30° of its characteristic value, the molecular simulations provided adequate equilibration of the loop dipeptide−water system.

Role of D/L Chirality. The three ΔG values in Table 1 for replacing the Gly residue at the L1 position by L-Ala (0.87 kcal/mol for GAII → AAII, 0.91 kcal/mol for GGII → AGII, and 0.83 kcal/mol for GaII → AaII) are not statistically different and are evidently independent of the nature and chirality of the residue at L2 (L-Ala, Gly, D-Ala). The average of these ΔG values provides an estimate of ΔG(GXII→AXII), the free-energy change for replacing glycine at position L1 by L-alanine for the type-II β-turn conformer of three of the nine loop dipeptides. Other position ΔG values for replacing Gly at position L1 or L2 by L-Ala or D-Ala for the type-II or type-II′ β-turn conformer of a loop dipeptide are calculated similarly (Table 2). Since the four position ΔG values listed in Table 2 are each based on 24 independent simulations, they are statistically more precise than the combined ΔG values shown in Table 1. These results show that for the set of nine loop dipeptides in either the type-II or type-II′ β-turn conformation, changing the D/L chirality at position L1 or L2 is independent of the type of residue (L-Ala, Gly, D-Ala) at the other position.

These position ΔG values clearly indicate that changing the D/L chirality at position L1 or L2 changes the free energy differently for the type-II and type-II′ β-turn conformers. For instance, replacing Gly by D-Ala at position L1 increases ΔG by 3.7 kcal/mol for the type-II β-turn conformer but by only 0.9 kcal/mol for the type-II′ conformer. In contrast, replacing Gly by D-Ala at L2 increases ΔG by only 1.0 kcal/mol for type-II but by 2.8 kcal/mol for type-II′. Inversely, replacing Gly by L-Ala at L1 increases ΔG by only 0.9 kcal/mol for type-II but by 3.7 kcal/mol for type-II′, and replacing Gly by L-Ala at L2 increases ΔG by 2.8 kcal/mol for type-II but by only 1.0 kcal/mol for type-II′. In summary, the type-II β-turn conformer is favored by the loop sequence L-Ala-D-Ala and the type-II′ β-turn conformer by D-Ala-L-Ala. In contrast, the type-I β-turn conformer is favored by the L1-L2 sequence L-Ala-L-Ala and the type-I′ β-turn conformer by D-Ala-D-Ala.[29]

Free-Energy Differences from Position ΔG Values. For each loop dipeptide the ΔG between its type-II (or type-II′) β-turn conformer and GGII (or GGII′) was calculated from the position ΔG values of Table 2. For example, ΔG(AaII) equals the sum of ΔG(GXII→AXII) and ΔG(XGII→XaII). The resulting ΔG values (Table 3) are very similar to the corresponding ΔG values listed in Table 1 but are statistically more precise because they are based on 24 rather than 8 independent simulations. In general, the ΔG values of the type-II β-turn conformers increase as the residue at L1 is changed from Gly to L-Ala to D-Ala and as the residue at L2 changed from Gly to D-Ala to L-Ala. Inversely, the ΔG values of the type-II′ β-turn conformers increase as the residue at L1 is changed from Gly to D-Ala to L-Ala and as the residue at L2 changed from Gly to L-Ala to D-Ala. The most stable alanine-containing conformers are AGII and aGII′. If only the four loop dipeptides lacking the nongenetic residue D-Ala are considered, the most stable L-Ala-containing conformers are AGII and GAII′. This result is consistent with the observation that glycine is the residue most frequently seen at the L2 position of type-II β turns in the crystal structures of natural proteins.[18]

Table 3. Relative Free Eneriges for Two Sets of Loop Dipeptide Conformers

glycine β-turn conformer	inverse-glycine β-turn conformer	ΔG, kcal/mol (mean ± rmsd)
GGII	GGII′	[0]
AGII	aGII′	0.88 ± 0.27
GaII	GAII′	0.98 ± 0.25
AaII	aAII′	1.86 ± 0.37
GAII	GaII′	2.80 ± 0.31
AAII	aaII′	3.68 ± 0.41
aGII	AGII′	3.71 ± 0.27
aaII	AAII′	4.69 ± 0.37
aAII	AaII′	6.51 ± 0.41

Table 4. Relative Free-Energy Changes for Refolding of Nine Loop Dipeptides from the Glycine into the Inverse-Glycine β-turn Conformer

loop dipeptide	refolding path	ΔG_r,[a] kcal/mol (mean ± rmsd)
Aa	AaII → AaII′	4.65 ± 0.55
AG	AGII → AGII′	2.83 ± 0.38
Ga	GaII → GaII′	1.82 ± 0.40
AA	AAII → AAII′	1.01 ± 0.55
GG	GGII → GGII′	0
aa	aaII → aaII′	−1.01 ± 0.55
GA	GAII → GAII′	−1.82 ± 0.40
aG	aGII → aGII′	−2.83 ± 0.38
aA	aAII → aAII′	−4.65 ± 0.55

[a] Based on the position ΔG values in Table 2.

Free-Energy Change for Refolding from the Glycine into Inverse-Glycine β Turns. The free-energy change for refolding of each loop dipeptide from its type-II into its type-II′ β-turn conformation was calculated using eq 2 from the position ΔG values of Table 2. For instance,

$$\Delta G_r(\text{GAII}\rightarrow\text{GAII}') = \Delta G(\text{XGII}'\rightarrow\text{XAII}') - \Delta G(\text{XGII}\rightarrow\text{XAII})$$

The resulting ΔG_r values are listed in Table 4. All three loop dipeptides with D-Ala at L1 are more stable in the type-II′ than in the type-II β-turn conformation. Inversely, all three loop dipeptides with L-Ala at L1 are more stable in the type-II than in the type-II′ β-turn conformation. The aA loop dipeptide most strongly favors the type-II′ β-turn conformation over type-II but the Aa loop dipeptide most strongly favors type-II over the type-II′ conformation. If only the four loop dipeptides lacking the nongenetic residue D-Ala are considered, the GA loop dipeptide most strongly favors the type-II′ β-turn conformation over type-II, whereas the AG loop dipeptide most strongly favors the type-II over the type-II′ conformation.

Relative Free-Energy Changes for Refolding From the Common into Inverse-Common β Turns. For each loop dipeptide $\Delta\Delta G_r$ for refolding from its type-I into its type-II β-turn conformer was calculated from the position ΔG values of Table 2 and those previously published.[29] As seen from

Table 5. Relative Free-Energy Changes for Refolding of Nine Loop Dipeptides from the Common into the Glycine β-Turn Conformation or from the Inverse-Common into the Inverse-Glycine β-Turn Conformation

loop dipeptide	refolding path	loop dipeptide	refolding path	$\Delta\Delta G_r$,[a] kcal/mol (mean ± rmsd)
aA	aAI → aAII	Aa	AaI′ → AaII′	3.41 ± 0.43
GA	GAI → GAII	Ga	GaI′ → GaII′	1.78 ± 0.32
AA	AAI → AAII	aa	aaI′ → aaII′	1.71 ± 0.42
aG	aGI → aGII	AG	AGI′ → AGII′	1.63 ± 0.25
GG	GGI → GGII	GG	GGI′ → GGII′	[0]
AG	AGI → AGII	aG	aGI′ → aGII′	−0.07 ± 0.28
aa	aaI → aaII	AA	AAI′ → AAII′	−0.23 ± 0.40
Ga	GaI → GaII	GA	GAI′ → GAII′	−1.86 ± 0.28
Aa	AaI → AaII	aA	aAI′ → aAII′	−1.93 ± 0.40

[a] Based on position ΔG values in Table 2 and previously published.[29]

Table 6. Free-Energy Changes for Interconverting Four Pairs of Loop Dipeptides in the Extended Conformation

forward replacement path	free energy change, kcal/mol (mean ± rmsd)			hysteresis, kcal/mol
	forward ΔG	reverse ΔG	combined ΔG	
GGϵ → AGϵ	2.19 ± 0.16	−1.87 ± 0.23	2.03 ± 0.25	0.31
GGϵ → GAϵ	2.13 ± 0.34	−1.62 ± 0.17	1.88 ± 0.37	0.51
GGϵ → aGϵ	2.17 ± 0.32	−2.04 ± 0.38	2.11 ± 0.33	0.13
GGϵ → Gaϵ	2.26 ± 0.42	−1.85 ± 0.25	2.05 ± 0.39	0.40

Table 5, all three loop dipeptides with D-Ala at L2 are more stable in the type-II than in the type-I β-turn conformation. Inversely, all three loop dipeptides with L-Ala at L2 are more stable in the type-II′ than in the type-I′ β-turn conformation. The Aa loop dipeptide most strongly favors the type-II β-turn conformation over type-I but the aA loop dipeptide most strongly favors the type-II′ over the type-I′ conformation. If only the four loop dipeptides lacking the nongenetic residue D-Ala are considered, the AG loop dipeptide most strongly favors the type-II β-turn conformation over type-I, whereas the GA loop dipeptide most strongly favors the type-II′ over the type-I′conformation.

Free-Energy Differences for the Unfolded State. We have considered two specific models of the unfolded state: the random coil[26] and the extended (ϵ) conformation.[42] Using the Cedar program and a special protocol based on the analysis of conformational probability distribution, Hermans et al.[26] have reported the free energy of 2 kcal/mol for molecular replacement of Ala for Gly in the random coil state. For each of the four loop dipeptides AG, GA, aG, and Ga, the free-energy difference between its extended conformer and GGϵ was calculated by molecular replacement simulation as before. Starting from the GG loop dipeptide in the extended conformation, the Gly residue at L1 or L2 was replaced by L-Ala or D-Ala and the associated ΔG was calculated from eq 1. The results are listed in Table 6. The mirror-image conformers AGϵ and aGϵ should have the same free energy by molecular symmetry. The average of their very similar combined ΔG values is 2.07 ± 0.29 kcal/mol. Likewise, the mirror-image conformers GAϵ and Gaϵ should have the same free energy by molecular symmetry. The average of their very similar combined ΔG values is 1.97 ± 0.38 kcal/mol. These results indicate that replacing any one of the four α-hydrogen atoms of GGϵ by a methyl group produces essentially the same increase in free energy. Thus in latter calculations an average value for ΔG(GGϵ→AGϵ), ΔG(GGϵ→aGϵ), ΔG(GGϵ→GAϵ), and ΔG(GGϵ→Gaϵ) of 2.02 kcal/mol is used, which is based on 32 independent molecular replacement simulations. Since the position ΔG values of Table 2 are additive independent of the residue (L-Ala, Gly, D-Ala) at the other position, ΔG(GGϵ→AGϵ) and $\Delta\Delta G$(GGϵ→GAϵ) and other appropriate pairs were assumed to be additive. This reasonable assumption produced the same value of 4.04 ± 0.48 kcal/mol for $\Delta\Delta G$(GGϵ→AAϵ), ΔG(GGϵ→aAϵ), ΔG(GGϵ→aAϵ), and ΔG(GGϵ→aaϵ).

Table 7. Relative Free-Energy Changes for Folding of Nine Loop Dipeptides from the Extended Conformation into the Glycine or Inverse-Glycine β-Turn Conformation

folding into the glycine (type-II) β-turn conformer	folding into the inverse-glycine (type-II′) β-turn conformer	$\Delta\Delta G_f$, kcal/mol (mean ± rmsd)
aAϵ → aAII	Aaϵ → AaII′	2.47 ± 0.46
aGϵ → aGII	AGϵ → AGII′	1.69 ± 0.35
GAϵ → GAII	Gaϵ → GaII′	0.78 ± 0.38
aaϵ → aaII	AAϵ → AAII′	0.65 ± 0.48
GGϵ → GGII	GGϵ → GGII′	[0]
AAϵ → AAII	aaϵ → aaII′	−0.36 ± 0.46
Gaϵ → GaII	GAϵ → GAII′	−1.04 ± 0.33
AGϵ → AGII	aGϵ → aGII′	−1.14 ± 0.35
Aaϵ → AaII	aAϵ → aAII′	−2.18 ± 0.48

Table 8. Relative Free-Energy Changes for Folding of Nine Loop Dipeptides from the Extended Conformation into the Common or Inverse-Common β-Turn Conformation

folding into the common (type-I) β-turn conformer	folding into the inverse-common (type-I′) β-turn conformer	$\Delta\Delta G_f$,[a] kcal/mol (mean ± rmsd)
aaϵ → aaI	AAϵ → AAI′	0.88 ± 0.35
Gaϵ → GaI	GAϵ → GAI′	0.82 ± 0.25
aGϵ → aGI	AGϵ → AGI′	0.06 ± 0.24
GGϵ → GGI	GGϵ → GGI′	[0]
Aaϵ → AaI	aAϵ → aAI′	−0.25 ± 0.32
aAϵ → aAI	Aaϵ → AaI′	−0.94 ± 0.34
GAϵ → GAI	Gaϵ → GaI′	−1.00 ± 0.24
AGϵ → AGI	aGϵ → aGI′	−1.07 ± 0.23
AAϵ → AAI	aaϵ → aaI′	−2.07 ± 0.33

[b] Based on ΔG values in Table 6 and published earlier.[29]

The choice of an adequate model for the unfolded state is crucial for accurate determination of the folding free energy. In general, the random coil should be a more rigorous model than the extended conformation. The molecular replacement calculations for the random coil require consideration of the conformational probability distribution for the entire Ramachandran plot,[26] which is computationally very expensive. For our model system, the free energy values for the G → A molecular replacement using the random coil model[26] (2 kcal/mol) and using the extended state (2.02 kcal/mol) were practically the same. Therefore, we have used our results for the extended state (Table 6) to calculate the free energy of folding for model peptides. Whether or not these two models of the unfolded state are equivalent in general remains an open question.

Relative Free-Energy Changes for Folding. The relative free-energy change ($\Delta\Delta G_f$) for folding of a loop dipeptide from the extended state into a β-turn conformation (relative to the folding of the GG loop dipeptide) was calculated by use of eq 4 and data of Tables 2 and 6 and the position ΔG values for the type-I and type-I′ β-turn conformations that we have published previously.[29] Table 7 lists the 18 results for the glycine and inverse-glycine β-turn conformations, and Table 8 provides the 18 results for the common and inverse-common β-turn conformations. The chirality of the amino acid residues at positions L1 and L2 of the loop dipeptides is the major factor that favors one conformer over another. These chiral effects were different for the L1 and L2 positions.

Chiral Patterns for the Glycine and Inverse-Glycine β Turns. The $\Delta\Delta G_f$ values of Table 7 present a chiral pattern and its inverse that are summarized in Table 9. Relative to the GG loop dipeptide, the glycine (type-II) β-turn conformation of each of the alanine-containing loop dipeptides is 1.1 kcal/mol more stable with L-Ala at L1, 1.7 kcal/mol less stable with D-Ala at L1, 1.0 kcal/mol more stable with D-Ala at L2, and

Table 9. Contributions of Alanine Residues to the Relative Free-Energy Changes for Folding of Eight Loop Dipeptides into Four β-Turn Conformations

residue, position	ΔG,[a] kcal/mol common (type-I) β-turn conformation	inverse-common (type-I′) β-turn conformation	glycine (type-II) β-turn conformation	inverse-glycine (type-II′) β-turn conformation
L-Ala at L1	**−1.1**	0.1	**−1.1**	1.7
D-Ala at L1	0.1	**−1.1**	1.7	**−1.1**
-Ala at L2	**−1.0**	0.8	0.8	**−1.0**
D-Ala at L2	0.8	**−1.0**	**−1.0**	0.8

[a] Relative to Gly at L1 and L2; negative contributions (bold) increase the stability of the β-turn conformation.

0.8 kcal/mol less stable with L-Ala at L2, whereas the inverse-glycine (type-II′) β-turn conformation is 1.1 kcal/mol more stable with D-Ala at L1, 1.7 kcal/mol less stable with L-Ala at L1, 1.0 kcal/mol more stable with L-Ala at L2, and 0.8 kcal/mol less stable with D-Ala at L2. In addition, Table 7 shows that all three loop dipeptides with L-Ala at L1 are more stable in the type-II β-turn conformation than in the extended conformation. But all three loop dipeptides with L-Ala at L2 are more stable in the extended conformation than in the type-II′ β-turn conformation. Inversely, all three loop dipeptides with D-Ala at L1 are more stable in the type-II′ β-turn conformation than in the extended conformation. Finally, all three loop dipeptides with D-Ala at L2 are more stable in the extended conformation than in the type-II β-turn conformation.

Chiral Patterns for the Common and Inverse-Common β Turns. The $\Delta\Delta G_f$ values of Table 8 reveal a different chiral pattern and its inverse (Table 9). Relative to the GG loop dipeptide, the common (type-I) β-turn conformation of an alanine-containing loop dipeptide is 1.1 kcal/mol more stable with L-Ala at L1, 0.1 kcal/mol less stable with D-Ala at L1, 1.0 kcal/mol more stable with L-Ala at L2, and 0.8 kcal/mol less stable with D-Ala at L2, whereas the inverse-common (type-I′) β-turn conformation is 1.1 kcal/mol more stable with D-Ala at L1, 0.1 kcal/mol less stable with LhAla at L1, 1.0 kcal/mol more stable with D-Ala at L2, and 0.8 kcal/mol less stable with L-Ala at L2. As shown in Table 8, all five loop dipeptides with L-Ala at L1 and/or L2 are more stable in the type-I β-turn conformation than in the extended conformation. Inversely, all five loop dipeptides with D-Ala at L1 and/or L2 are more stable in the type-I′ β-turn conformation than in the extended conformation.

Predicted Conformational Patterns for the β Turns of Natural Proteins. The four loop dipeptides lacking D-Ala are each most stable in a different β-turn conformation, namely, AAI, GGI (equivalent to GGI′), AGII, and GAII′ (Table 9). But only one of these four loop dipeptides is relatively stable in this β-turn conformation. Thus, AAI is 3.0 kcal/mol more stable than AAI′ (Table 8), 2.7 kcal/mol more stable than AAII′ (Tables 4 and 5), and 1.7 kcal/mol more stable than AAII (Table 5). But GGI′ and GGI are equally stable by symmetry, while AGII is only 0.1 kcal/mol stable than AGI and inversely GAII′ is only 0.1 kcal/mol more stable than GAI′ (Table 5). Thus, having glycine at L1 alone should favor either a type-I′ or type-II′ β turn, at L2 alone should favor either a type-I or type-II β turn, and at both L1 and L2 should favor either a type-I or type-I′ β turn. These predictions provide an explanation for why glycine is found in all four major types of β-turns but is favored at L1 of type-II′ turns, at L2 of type-II turns, and at both L1 and L2 of type-I′ turns.[18] The strongest prediction, however, is that the absence of Gly at both L1 and L2 should favor only the type-I β turn, which in fact is the most common β turn found in natural proteins.

Comparison with the β Turns of Natural Proteins. If the folding of a β turn is primarily guided by local interactions rather than by long-range effects, the β-turn conformers of a given loop dipeptide with lower $\Delta\Delta G_f$ values should occur more often in nature. The values of $\Delta\Delta G_f$ for the loop dipeptide conformations shown in Tables 7 and 8 are consistent with the relative frequency of occurrence of these β-turn conformations in the crystal structures of natural proteins. Four of the nine loop dipeptides explored in this study correspond to dipeptide segments of natural proteins (GG, GA, AG, AA). Tables 7 and 8 indicate that AA favors formation of the type-I β turn and AG favors the type-II β turn. Indeed, glycine is rarely found in common (type-I) β turns but is the most frequent residue at position L2 of glycine (type-II) β turns,[18] which is why the type-II β turn is called the glycine β turn.[2]

Analysis[44] of 101 highly refined nonhomologous proteins[45] in the Brookhaven Protein Data Base[46] has provided the following relative frequencies of occurrence for each β-turn conformation of the dipeptide segments AA, AG, and GA. The type-I/type-I′ subset of loop dipeptide conformers is listed here in decreasing order of stability relative to GGI (Table 8) ($\Delta\Delta G_f$; percent of the AA occurrences): the AAI conformer (−2.1; 10.8%), the AGI conformer (−1.1; 5.2%), the GAI conformer (−1.0; 4.1%), the AGI′ conformer (0.1; 0.6°), the GAI′ conformer (0.8; 0%), and the AAI′ conformer (0.9; 0%). Similarly, the type-II/type-II′ subset of loop dipeptide conformers is listed here in decreasing order of stability relative to GGII (Table 7) ($\Delta\Delta G_f$; percent of the AG occurrences): the AGII conformer (−1.1; 9%), the GAII′ conformer (−1.0; 3.4%), the AAII conformer (−0.4; 0%), the AAII′ conformer (0.6; 0%), the GAII conformer (0.8; 0%), and the AGII′ conformer (1.7; 0%). For both conformer subsets, as the relative conformer stability decreased ($\Delta\Delta G_f$ increased), the relative frequency of occurrence of that conformer in its conformer set decreased. Thus our estimates of the relative stabilities of these β-turn conformers correlate well with their relative frequencies of occurrence in the protein crystallographic database.

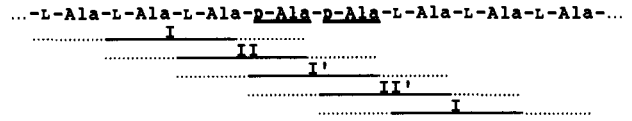
Other Estimates. Tobias et al.[43] have calculated by a conformational forcing method the free-energy change for unfolding of the type-I and type-II β-turn conformers to the extended conformer for a series of blocked dipeptides that included the AG and AA loop dipeptides. They obtained an estimate of the relative free-energy change for folding of AAII relative to AGII of 8.9 kcal/mol, which agrees in sign but not in magnitude with our estimate of 0.8 kcal/mol from the difference between two values in Table 7. Both estimates predict that the AGII conformer should be more stably folded than the AAII conformer. On the other hand, Tobias et al.[43] obtained an estimate of the relative free-energy change for folding of AAI relative to AGI of 2.4 kcal/mol, which disagrees in sign with our estimate of −1.0 kcal/mol from the data in Table 8. Their estimate predicts that the AGI conformer should be more stably folded than the AAI conformer, whereas our

(44) Tropsha, A.; Laiter, S. Unpublished results.
(45) Jones, D. T.; Taylor, W. R.; Thornton, J. M. *Nature* **1992**, *358*, 86−89.
(46) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. D.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Tasumi, M. J. *J. Mol. Biol.* **1977**, *112*, 535.

estimate predicts the opposite. The quantitative differences between our results and those of Tobias et al.[43] could be partially explained by the use of different forcefields and conformational sampling methods. However, the relative frequencies of occurrence[44] in the Brookhaven Protein Database of AGI (5.2% of AG occurrences) and AAI (10.8% of AA occurrences) suggest that the latter conformer is more stable, which supports the negative sign of our estimate.

Chiral Patterns for Placing β Turns in Designed Proteins. Of the 45 loop dipeptide conformers mentioned in Tables 7 and 8, the most stable conformers for each of the four types of β turns are AAI, aaI′, AaII, and aAII′. Specifically, each of these four loop dialanines is calculated to be 2.1 kcal/mol more stable than the corresponding conformer of the GG loop dipeptide (Table 9) and about 1 kcal/mol more stable than the next two most stable conformers of its type (AGI and GAI; aGI′ and GaI′; AGII and GaII; aGII′ and GAII′, respectively). In addition, each of these four loop dialanines is most stable in this β-turn conformation. Thus, AAI is 3.0 kcal/mol more stable than AAI′ (Table 8), 1.7 kcal/mol more stable than AAII (Table V), and 2.7 kcal/mol more stable than AAII′ (Tables 4 and 5). Inversely, aaI′ is 3.0 kcal/mol more stable than aaI, 1.7 kcal/mol more stable than aaII′ and 2.7 kcal/mol more stable than aaII. Similarly, AaII is 4.6 kcal/mol more stable than AaII′ (Table 4), 1.9 kcal/mol more stable than AaI (Table 5), and 1.2 kcal/mol more stable than AaI′ (Tables 4 and 5). Inversely, aAII′ is 4.6 kcal/mol more stable than aAII, 1.9 kcal/mol more stable than aAI′, and 1.2 kcal/mol more stable than aAI. Thus, having D-Ala at L1 alone should favor the type-II′ β turn, at L2 alone should favor the type-II β turn, and at both L1 and L2 should favor the type-I′ β turn.

When designing a synthetic protein, D-Ala should be better than glycine for controlling the locations along the peptide chain of the three less-common types of β turns (types-I′, -II, and -II′). For example, shown below is the energetically most favored type of β turn for each L1-L2 dipeptide of a chain containing two D-Ala residues.



The L-Ala-L-Ala dipeptide favors the common (type-I) β turn, the L-Ala-D-Ala dipeptide favors the glycine (type-II) β turn, the D-Ala-D-Ala dipeptide favors the inverse-common (type-I′) β turn, and the D-Ala-L-Ala dipeptide favors the inverse-glycine (type-II′) β turn. As a practical example, we have used two adjacent D-amino acid residues to favor the inverse-common conformation at all six β turns of recent betabellins, which increased their folding stability and water solubility.[13-15] In general, the favored β turn changes from type-I to type-II to type-I′ on entering a stretch of adjacent D-Ala residues and from type-I′ to type-II′ to type-I on leaving it. Although based only on the chirality of the loop residues L1 and L2 ($i + 1$ and $i + 2$, solid lines), these predictions are probably changed quantitatively but not qualitatively by the chirality of the flanking residues (i and $i + 3$, dotted lines). They may also hold for other D- and L-amino acids.

Conclusions. This study has provided quantitative estimates of the free-energy changes for folding and refolding of nine loop dipeptides that comprise a set of chirally representative β-turn models. Direct determination of the accuracy of our calculated estimates is prevented by the lack of experimental data for the relative free energies for folding of β-turn conformers, which are either unstable in solution[21] or must contain structural constraints to maintain the β-turn conformation.[47] Care was taken during this work to obtain reasonably reliable values for the free-energy changes from the molecular replacement simulations. Earlier molecular simulations with the force field[33,35] used in this work have accurately calculated ligand−receptor binding constants[27] and the α-helical propensities of hydrophobic amino acids.[26] In addition, the present simulations used a time period that gives convergent results.[29] Their reversibility was supported by low hysteresis values and by the self-consistent ΔG values obtained when the same transformation was simulated over different paths. The present simulations avoided the crossing of high-energy barriers, which often leads to large errors, by confining the loop dipeptides to local minima in conformational space and by taking advantage of their molecular symmetry.

One of the most interesting results of this study is that the position ΔG values for folding or refolding of model dipeptides for a given residue in either L1 or L2 appear independent of the amino acid type in the neighboring position. This result raises a question of whether or not the conformational free energy differences between the turn types can be evaluated from the analysis of the Ramachandran-type free-energy surfaces for individual amino acid residues[48] and summing up the individual free energy contributions. This question shall be addressed in future studies.

The set of loop dipeptides described here and in our previous reports[29,30] is the smallest set of peptide models that distinguish the various β-turn conformers. The estimates of the free energy difference between these conformers, obtained in this paper, correlate well with their relative frequency of occurrence in the protein crystallographic database. This correlation suggests that the formation of β-turns in native proteins may be guided primarily by local folding interactions. This peptide model can be used to explore the conformational preferences of any dipeptide through simulation of its relative free energies for folding into various conformations. These molecular simulations are being extended to other types of secondary structures and to a larger set of model dipeptides, which should permit a more quantitative approach to the problem of local protein folding.

Acknowledgment. The authors thank Professor J. Hermans for helpful advice and for providing his CEDAR program and Professor I. Vaisman for helpful discussions. This work was supported in part by a fellowship (to Y.Y.) from Cray Research and the North Carolina Supercomputing Center, by grants (to A.T.) from Tripos Associates, the North Carolina Supercomputing Center, and the UNC Research Council, and by U.S. Public Health Service Grant No. GM 42031 (to B.W.E.) from the National Institute of General Medical Sciences.

JA9403064

(47) Falcomer, C. M.; Meinwald, Y. C.; Choudhary, I.; Talluri, S.; Milburn, P. J.; Clardy, J.; Scheraga, H. A. *J. Am. Chem. Soc.* **1992**, *114*, 4036−4042.

(48) Hermans, J.; Anderson, A. G. In *Theoretical Biochemistry and Molecular Biophysics*; Vol. 2, Proteins; Beveridge, D. L., Lavery, R., Eds.; Adenine Press: Schenectady: New York, 1990; pp 45−61.