Cis−Trans Imide Isomerization of the Proline Dipeptide

Stefan Fischer, Roland L. Dunbrack, Jr.,‡ and Martin Karplus*

Contribution from the Department of Chemistry, 12 Oxford Street, Harvard University, Cambridge, Massachusetts 02138

Received April 11, 1994®

Abstract: The *cis−trans* imide isomerization reaction of the proline dipeptide is analyzed. It is shown that the reaction path is complex and involves the imide bond torsion angle ω, the pyramidalization of the imide nitrogen, and the proline backbone torsion angle ψ. A virtual dihedral angle ζ is found to be better suited for describing the progress of the reaction than ω. Adiabatic energy maps are calculated as a function of these coordinates with the empirical CHARMM potential and at the 6-31G*//3-21G *ab initio* level. The gas phase 6-31G* activation barriers for *trans* → *cis* isomerization from the optimized ground state to the transition state are 17.9 and 20.7 kcal/mol for the clockwise (*syn*) and anticlockwise (*anti*) path, respectively. A strong dependence of the activation barrier on ψ is found; its value can change the barrier by as much as 12 kcal/mol. For $\psi \sim 0°$, the C-terminal NH group can interact with either the lone pair of the imide nitrogen (*syn*) or the imide carbonyl oxygen (*anti*); both interactions result in a lowering of the barrier. This "autocatalytic" stabilization of the transition state has implications for the mechanism of catalysis in rotamases.

The *cis−trans* isomerization of peptide bonds is a slow process under normal conditions due to the high barrier resulting from their partial double bond character. Since 10% of peptidyl−proline imide bonds occur as the *cis* isomer in native proteins, *trans* → *cis* isomerization (Figure 1) can be the rate-limiting step in the folding process.[1−3] Two classes of rotamase enzymes, the FK506 binding proteins (FKBPs)[4,5] and the cyclophilins,[6] catalyze *cis−trans* prolyl isomerization by lowering the enthalpy barrier from the aqueous solution value of 19−20[7−9] to 5−6 kcal/mol. They have been shown to act as protein foldases *in vitro*.[6,10−13] The wide distribution of these rotamases in cells suggests that labilization of the imide torsion may be important *in vivo*.[14,15] Moreover, it has been suggested that prolyl isomerization may be involved in regulatory switches.[16]

As part of a study of the catalysis of proline isomerization,[17,18] we analyze the *cis−trans* isomerization reaction path of

cis *trans*

Figure 1. Peptidylprolyl *cis−trans* isomerization.

N-acetylproline methylamide (the proline "dipeptide"), which is the simplest model for a proline containing peptide fragment (see Figure 2,A and C). By use of an empirical potential energy function,[19] we determine the degrees of freedom that are involved in the isomerization reaction. Because the empirical function does not include the changes in electronic structure occurring during the isomerization process, we then calculate adiabatic maps of the *ab initio* energy of the proline dipeptide as a function of the essential coordinates. The utility of *ab initio* calculations has been demonstrated in a detailed study of the simpler case of symmetric peptide bond isomerization in formamide.[20] The adiabatic maps are used here to demonstrate the role of the relevant degrees of freedom in the *cis−trans* prolyl isomerization and their effect on the height of the activation barrier. The essential degrees of freedom are as follows: (1) the imide bond torsion, ω; (2) the out-of-plane deformation of the imide nitrogen (referred to as "pyramidalization"), expressed in terms of an improper dihedral angle, η; and (3) the proline backbone torsion angle, ψ. Because of the coupling between the first two degrees of freedom, the dihedral angle ω is not a good coordinate for describing the isomerization. Instead, a virtual dihedral angle ζ (see Figure 2C) is introduced. It allows a simpler determination of the progress of the *cis−trans* isomerization reaction by effectively removing the amplitude of pyramidalization as an independent degree of freedom. Using ζ as the reaction coordinate, we evaluate the effect of ψ on the activation barrier. Comparison of the results for the proline dipeptide with those for *N*-acetylpyrrolidine (Figure 2B) shows how the C-terminal amide hydrogen of the

‡ Present address: Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143-0446.

® Abstract published in *Advance ACS Abstracts*, November 1, 1994.

(1) Brandts; J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953.
(2) Schmid, F. X.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 4764.
(3) Schmid, F. X.; Baldwin, R. L. *J. Mol. Biol.* **1979**, *133*, 185.
(4) Harding, M. W.; Galat, A.; Uehling, D. E.; Schreiber, S. L. *Nature* **1989**, *341*, 758.
(5) Siekierka, J. J.; Hung, S. H. Y.; Poe, M.; Lin, C. S.; Sigal, N. H. *Nature* **1989**, *341*, 755.
(6) Fischer, G.; Wittmann-Liebold, B.; Lang, K.; Kiehaber, T.; Schmid, F. *Nature* **1989**, *337*, 476.
(7) Drakenberg, T.; Forsen, S. *J. Phys. Chem.* **1970**, *74*, 1.
(8) Roques, B. P.; Garbay-Jaureguiberry, C.; Combrisson, S.; Oberlin, R. *Biopolymers* **1977**, *16*, 937.
(9) Cheng, H. N.; Bovey, F. A. *Biopolymers* **1977**, *16*, 1465.
(10) Bächinger, H. P. *J. Biol. Chem.* **1987**, *262*(*35*), 17144.
(11) Lang, K.; Schmid, F. X.; Fischer, G. *Nature* **1987**, *329*, 268.
(12) Tropschug, M.; Wachter, E.; Mayer, S.; Schönbrunner, E. R.; Schmid, F. X. *Nature* **1990**, *346*, 674.
(13) Schmid, F. X.; Mayr, L. M.; Muecke, M.; Schönbrunner, E. R. *Adv. Protein Chem.* **1993**, *44*, 25.
(14) Gething, M. J.; Sambrook, J. *Nature* **1992**, *355*, 33.
(15) Fischer, G.; Schmid, F. X. *Biochemistry* **1990**, *29*, 2205.
(16) Jayaraman, T.; Brillantes, A.-M.; Timerman, A. P.; Fleischer, S.; Erdjument-Bromage, H.; Tempst, P.; Marks, A. R. *J. Biol. Chem.* **1992**, *267*, 9474.
(17) Fischer, S. (1992) In Curvilinear reaction coordinates of conformational change in macromolecules. Application to rotamase catalysis, Ph.D. Thesis, Harvard University.
(18) Fischer, S.; Michnick, S.; Karplus, M. *Biochemistry* **1993**, *32*, 13820.
(19) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Adv. Chem. Phys.* **1988**, *71*, 1.
(20) Wiberg, K. B.; Laidig, K. E. *J. Am. Chem. Soc.* **1987**, *109*, 5935.

0002-7863/94/1516-11931$04.50/0

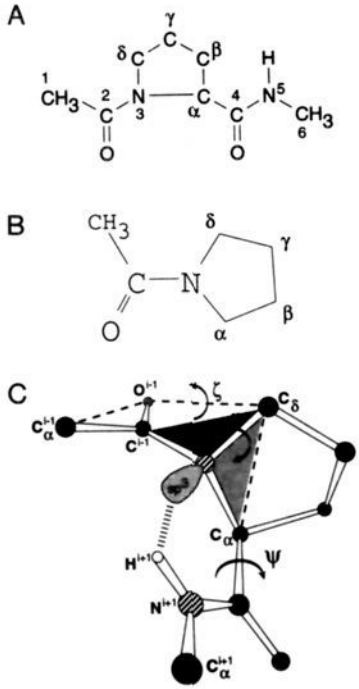

Figure 2. (A) *N*-Acetylproline methylamide (proline "dipeptide"); (B) *N*-acetylpyrrolidine (NAP); (C) proline dipeptide in the *syn/exo* transition state (see text), showing the dihedral angles $\zeta(C_\alpha^{i-1}-O^{i-1}-C_\delta-C_\alpha)$ and ψ. The sp^3 nitrogen lone pair interacts with the C-terminal amide hydrogen. The improper angle η can be defined in several ways. Shown here is the angle between the shaded plane $C_{i-1}N-C_\delta$ and $C_\alpha-N-C_\delta$. Alternatively, it can be the angle between the plane $C_{i-1}-N-C_\alpha$ and $C_\alpha-N-C_\delta$, which was used here.

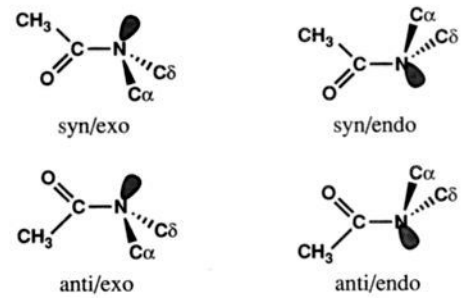

Figure 3. Four transition state structures of the twisted imide bond. The orientation of the lone pair sp^3 orbital is shown in each case. C_α and C_δ are labeled as in Figure 2.

former can interact with the twisted imide group for appropriate values of ψ thereby lowering the imide isomerization barrier. This effect, which can be characterized as "autocatalysis", contributes significantly to the catalytic mechanism of FKBP.[18]

Methods

The molecular mechanics calculations were performed with the CHARMM program,[21] using a developmental version of the all-hydrogen parameter set for proline (Dunbrack, R. L., Joseph-McCarthy, D., Karplus, M., to be submitted for publication). The new parameters, derived from fits to *ab initio* optimized structures and energies of proline derivatives, exhibit improved ring dynamics compared to older extended-carbon models, as demonstrated by NMR data for antamanide.[22] Nonbonded interactions were truncated at 12 Å using the SHIFT function for electrostatics terms and the SWITCH function for the Lennard-Jones terms;[21] a dielectric constant of unity was used. To obtain an overview of the potential energy surface, two-dimensional adiabatic maps of the empirical energy of the proline dipeptide as a function of the dihedral angle pairs (ω,η), (ζ,η), and (ζ,ψ) were computed by minimizing the energy with constraints on the dihedral angles of interest. The angles ω, ζ, and ψ were changed by 10° increments between −180° and +180°; the angle η was changed by 5° increments from +100° to +180° and from −180° to −100°. Complete geometry optimization of the molecule was performed for each set of dihedral angle pairs to obtain an RMS energy gradient smaller than 10^{-3} (kcal mol^{-1})/Å by use of the conjugate gradient minimizer.[21]

Generalized reaction paths between minima on the potential energy surface were then determined with the Conjugate Peak Refinement (CPR) method,[23] implemented in the TRAvel (Trajectory Refinement Algorithm) module of the CHARMM program. CPR starts with the reactant and product conformations to find a connecting minimum energy path, which includes the true saddle point(s). All the degrees of freedom of the molecule can contribute to the reaction path. There is no need to choose a reduced reaction coordinate, such as the internal coordinate ω. If more than one path exists between a given reactant/product pair, each one can be refined by providing an initial guess for one or more intermediate conformers along each path. It has been found from experience that these guesses can be quite crude, as long as they specify the general direction of the path. Given the saddle points from the CPR calculations, adiabatic paths were obtained by controlled steepest-descent minimization with small steps starting at the saddle points; conformers were saved at regular intervals.

Ab initio calculations were performed to determine the energy surface for the isomerization reaction and compare it with the empirical potential results. The GAUSSIAN program[24] at the Hartree−Fock level with the 3-21G and 6-31G* basis sets was used. All internal degrees of freedom were included in optimizing the structures, except for the C−H bond lengths (which were fixed at 1.08 Å). For examining certain portions of the energy surface, geometry optimizations were performed with constraints on ζ and/or ψ.

Results

Two Governing Coordinates. Four transition state conformations are *a priori* possible for the *cis−trans* isomerization of an imide; they are shown schematically in Figure 3. They result from the coupling of two degrees of freedom, (ω,η), each with two possible values at the transition state. One motion is the twist about the imide bond itself, which can be clockwise or counterclockwise. This determines whether carbon C_1 in Figure 2A (C_α^{i-1} in Figure 2C) is on the same side or the opposite site of the proline ring relative to carbon C_4 during the isomerization; the two conformers are referred to as *syn* (clockwise) and *anti* (anticlockwise), respectively. As the imide bond is twisted, the geometry of the imide nitrogen goes from nearly planar sp^2 to essentially tetrahedral sp^3. This motion results in an out-of-plane deformation of the imide nitrogen center (N_3), with carbon C_2 (C^{i-1} in Figure 2C) moving either away from (*exo*) or closer to (*endo*) carbon C_4. Figure 4 shows the adiabatic CHARMM energy map calculated as a function of the dihedral angle ω, which governs imide bond twisting, and the improper dihedral angle η, which governs imide nitrogen pyramidalization; the angle ω is defined in terms of the atoms $C_1-C_2-N_3-C_\alpha$ and the angle η is defined in terms of the atoms $C_2-C_\alpha-N_3-C_\delta$ (see Figure 2A). {Dihedral angles are defined in the notation $X_1-X_2-X_3-X_4$, where the atoms X_i ($i = 1-4$) do not need to be covalently bonded. The dihedral angle (whether normal, improper, or virtual) is defined as the angle

(21) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187.
(22) Schmidt, J. M.; Brüschweiler, R.; Ernst, R. R.; Dunbrack, R. L., Jr.; Joseph, D.; Karplus, M. *J. Am. Chem. Soc.* **1993**, *115*, 8747−8756.
(23) Fischer, S.; Karplus, M. *Chem. Phys. Lett.* **1992**, *194*, 252.
(24) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A.; *Gaussian 92 Revision C*; Gaussian, Inc.: Pittsburgh, PA, 1992.

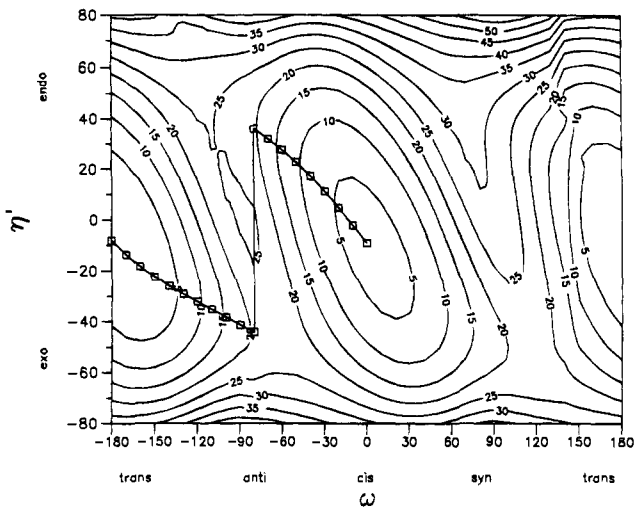

Figure 4. Adiabatic contour map of the CHARMM energy of the proline dipeptide, calculated as a function of the imide bond torsion angle ω and the nitrogen pyramidalization angle η' (see results section for a description of ω and η'). The *trans*, *anti*, *cis*, and *syn* regions are marked along the horizontal axis, the *endo* and *exo* regions along the vertical axis. The reduced adiabatic *trans* → *anti* → *cis* path, calculated as a function of incremental ω, is shown. Energies are in kcal/mol above the *trans* minimum.

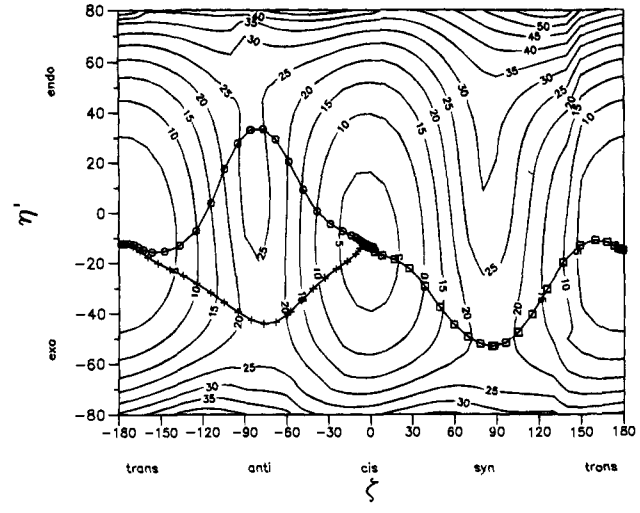

Figure 5. Adiabatic contour map of the CHARMM energy of the proline dipeptide, calculated as a function of the virtual dihedral ζ and the improper dihedral η' (see Figure 2C for a description of ζ). The three generalized adiabatic reaction paths are shown. Energies are in kcal/mol above the *trans* minimum.

between the two planes containing the atoms [X_1, X_2, X_3] and [X_2, X_3, X_4], respectively.} The angle $\eta = 180°$ corresponds to a planar nitrogen and $\eta = \pm 120°$ corresponds to a tetrahedral *endo* or *exo* center. The map is plotted with respect to η', with $|\eta'| = 180° - |\eta|$, which corresponds to the deviation from planarity and is continuous with respect to pyramidalization. For convenience, η' is taken to have a negative sign when the nitrogen is *exo* and a positive sign when it is *endo*. With this convention, $\eta' = -60°$ is tetrahedral in the *exo* form, $\eta' = 0°$ is planar, and $\eta' = 60°$ is tetrahedral in the *endo* form.

There are three saddle regions, located at $\eta' \approx \pm 40°$, indicating that the imide nitrogen is nearly tetrahedral at the transition state. The two *anti* saddle points are located at significantly different values of ω: i.e. *anti/endo* at $(\omega,\eta) = (-109°,33°)$ and *anti/exo* at $(\omega,\eta) = (-63°, -44°)$. The *syn/exo* saddle point is at $(\omega,\eta') = (+118°,-53°)$. There is no *syn/endo* saddle region at $(\omega,\eta') \approx (+60°,+40°)$, because the energy in this region is higher than the intrinsic barrier for *endo* → *exo* inversion of a pyramidal imide nitrogen. On the *anti* ridge, this barrier is about 4.6 kcal/mol for *exo* → *endo* (relative to the *anti/exo* saddle point) or 1.5 kcal/mol for *endo* → *exo* (relative to the *anti/endo* saddle point); the barrier is the local maximum separating the two *anti* saddle points. The increase in energy in the *syn/endo* region is due to the steric repulsion between carbon C_1 and the C-terminus, which does not exist in the *anti/endo* conformation.

Given the results in Figure 4, it is useful to examine whether ω is a satisfactory reaction coordinate. If ω is kept fixed at any value between $-109°$ and $-63°$ or between 80° and 118° there are two local energy minima, one *endo* and the other *exo*, into which the molecule can be optimized. One minimum is close (in conformation and on the map) to a saddle point and high in energy, while the other is distant from the saddle points and has a lower energy. For example, with ω fixed at $-100°$ there is a minimum energy conformer at $\eta' = +41°$, whose energy is 24.2 kcal/mol; it is close to the *anti/endo* saddle point. The other minimum is at $\eta' = -38°$ and has an energy of 15.6 kcal/mol; it is more distant from the *anti/exo* saddle point. This means that specifying only ω does not measure how close the molecule is to a saddle-point conformation. Consequently, the progress of the reaction cannot be monitored with the angle ω alone, but must complemented by specifying η. Qualitatively, this is a consequence of the fact that the contour lines in Figure 4 are skewed with respect to the ω axis, rather than being mostly parallel or perpendicular to it. For the same reason, it is not appropriate to construct the adiabatic *cis–trans* isomerization reaction path as a function of ω. Such a path is plotted for the *anti* transition in Figure 4. As the path climbs out of the *trans* minimum along the *anti/exo* valley, the nitrogen undergoes an *exo* → *endo* umbrella flip before the *anti/exo* saddle point is reached, and the path descends the *anti/endo* valley. This occurs because the steps taken by the minimizer and the steps in ω are not infinitesimal, so that a point on the path can jump from one side of the barrier to the other during the optimization. This creates the illusion that the path "tunnels" through the inversion barrier; no such physical process is involved. For one value of ω the *exo* form is lower in energy and for the next value of ω the *endo* form is lower. The resulting path between the two involves the inversion of the nitrogen and has a barrier of 27.2 kcal/mol. Thus, it requires 4.6 kcal/mol more in energy than the path through the *anti/endo* saddle point with a barrier of 22.6 kcal/mol.

A Coordinate for Imide Twisting. Examination of the structure of the respective saddle points on the energy surface of Figure 4, determined by CPR, revealed an internal reaction coordinate without the limitations of ω. It is the virtual dihedral angle ζ, which is defined in terms of the atoms C_1–O_2–C_δ–C_α (Figure 2A) or C_α^{i-1}–O^{i-1}–C_δ–C_α (Figure 2C). Like ω, $\zeta = 0°$ when the imide group is in the planar *cis* state and $\zeta = 180°$ when it is in the planar *trans* state. Unlike ω, ζ is in the neighborhood of $\pm 90°$ at the saddle points, notwithstanding the nitrogen pyramidalization. The *anti* saddle points on the empirical energy surface are $\zeta = -86°$ in *endo* and $\zeta = -77°$ in *exo*, much closer to each other than the values of ω at these two saddle points ($-109°$ and $-63°$, respectively; see above). The *syn* saddle point is at $\zeta = +87°$ in *exo*. The adiabatic CHARMM energy map of the proline dipeptide as a function of ζ and η' is shown in Figure 5. It resembles Figure 4, but it is significantly less skewed. Because of this, specifying ζ provides an essentially unambiguous measure of how close a conformer is to the saddle-point regions, and the artifactual path found with ω as a reaction coordinate (Figure 4) is not allowed. Once the η' angle is preset to be either positive (*endo*) or

negative (*exo*), it can be optimized for any fixed value of ζ. Although ζ is a virtual dihedral angle, it is defined over atomic centers and can easily be constrained or measured in molecular mechanics and *ab initio* programs. This makes ζ a useful reaction coordinate for describing imide bond twisting in general; it may also be appropriate for other peptide bonds. As shown in the next section, introduction of ζ permits a straightforward exploration of the ψ dependence of the isomerization energy.

The three possible reaction paths corresponding to the saddle points on the map of Figure 4 were calculated with the CPR method by providing a conformer in each saddle region of the map as the initial guess for a path intermediate (see Methods section). The resulting generalized adiabatic paths are projected onto the map shown in Figure 5. To illustrate the qualities of ζ as a reaction coordinate, a good approximation to the *cis* → *syn* → *trans* path, for instance, could be obtained by minimizing the energy at incremental values of ζ. The resulting reduced adiabatic path would encompass the saddle point, and ζ was in effect used as a reaction coordinate in the *ab initio* calculations (see below). However, it should be cautioned that such a one-dimensional adiabatic mapping will successfully determine a reaction path only in well-behaved cases. In general, the approach fails to give a meaningful path, as seen above for the *anti* path adiabatically mapped with ω, even though the angle ω seems to be the "obvious" choice for the reaction coordinate. An early example of the difficulties in the use of an obvious reduced reaction coordinate arose in studies of tyrosine ring flips in the bovine pancreatic trypsin inhibitor.[25] Instead, a method like CPR, which requires no *a priori* selection of reaction coordinates, is recommended for calculating reaction paths, even in small molecules and/or for apparently quite simple reactions. For simple systems, methods such as those developed by Schlegel, which are widely employed in optimizing quantum mechanical surfaces, can also be used effectively.[26]

(ζ,ψ) Empirical Energy Map. In any dipeptide, strong interactions occur between the two polar peptide moieties; corresponding behavior is expected in longer peptides. For the proline dipeptide, the nature of the interactions depends on the orientation of the C-terminal amide, which is essentially determined by the torsion angle ψ (defined in terms of the atoms $N_3-C_\alpha-C_4-N_5$); the angle ϕ is restricted to be near $-60°$ by the proline ring. The role of ψ in the *cis*–*trans* imide isomerization and its influence on the activation barriers was determined from an adiabatic CHARMM energy map as a function of ζ and ψ; see Figure 6. The energy on this surface is continuous everywhere. This is not the case for the (ω,η') or the (ζ,η') adiabatic maps (Figures 4 and 5), where a minor discontinuity in the energy can be observed in the $\omega \sim \zeta \approx 140°$ region, as ψ jumps from a value of about $-20°$ for $\omega \sim \zeta < 140°$ to a value near $+80°$ for $\omega \sim \zeta > 140°$. The fact that the (ζ,ψ) adiabatic energy map is continuous indicates that the degrees of freedom other than ζ and ψ, such as ring puckering or imide pyramidalization, can relax freely for any given value of the (ζ,ψ) pair and do not need to be included as part of the reaction coordinate.

Optimization of the conformers for the map in Figure 6 was started with the imide nitrogen in the *exo* form (i.e. $\eta' < 0°$), which was conserved even though no constraint was applied on η'. In the $-180° < \zeta < 0°$ domain, another (ζ,ψ) adiabatic map exists with the imide nitrogen in the *endo* form, i.e. $\eta' > 0°$ (not shown), which has a slightly higher energy than the

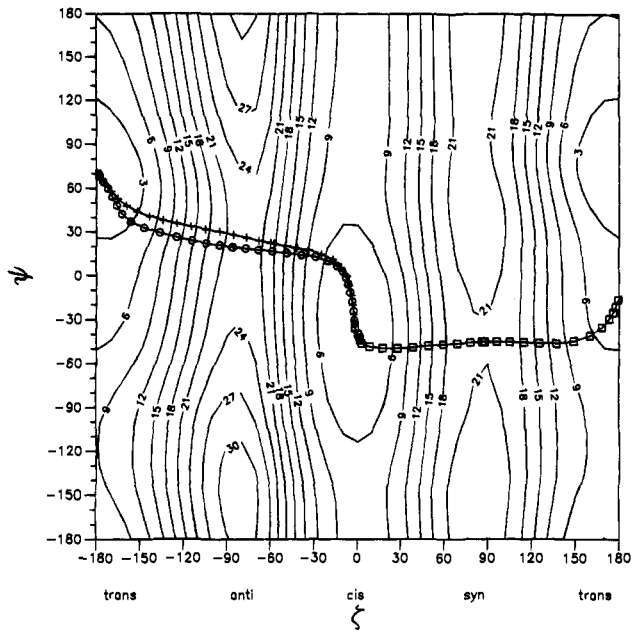

Figure 6. Adiabatic contour map of the CHARMM energy of the proline dipeptide, calculated as a function of the virtual dihedral ζ and the torsion angle ψ. The three generalized adiabatic reaction paths are projected on the map (using the same symbols as in Figure 5). Energies are in kcal/mol above the *trans* minimum.

exo map (as can be deduced from Figure 5). For the $0° < \zeta < 180°$ domain, only the *exo* ($\eta' < 0°$) adiabatic map can be calculated. The *endo* form ($\eta' > 0°$) is unstable in that domain, as indicated by the absence of a *syn*/*endo* saddle point in Figure 5, so that an *endo* (ζ,ψ) map in that domain would require constraining η' to some positive value. The three generalized adiabatic reaction paths shown in Figure 5 are projected onto the map of Figure 6. Because the map is in the *exo* domain, the *exo* paths (indicated by crosses and squares) actually lie on the potential energy surface of the map, but the *endo* path (indicated by circles) does not; it indicates only the (ζ,ψ) variation along the reaction path but not the actual energy.

(ζ,ψ) *Ab Initio* Energy Map. An adiabatic map was calculated at pairs of (ζ,ψ) values taken from the points suggested by the empirical energy surface. These are the *trans* and *cis* valleys and the *anti* and *syn* ridges. Five points, roughly equidistant in ψ (every 75°), were chosen at each of the four corresponding values of ζ. The geometry of these conformers was optimized at the Hartree–Fock level with the 3–21G basis-set (Table 1). The location of the *anti* ridge was estimated by optimizing conformers with ζ fixed at $-75°$, $-80°$, and $-85°$, while allowing ψ to be free (ψ adopted values between 8° and 18°). The energy was highest at $\zeta = -80°$, which was chosen as the constrained value for conformers on the *anti* ridge. The same procedure was used to locate the *syn* ridge; i.e. conformers were optimized with ζ fixed at 80°, 85°, 90°, and 100° (ψ values between 1° and 3° were found) and $\zeta = 85°$ was chosen for the *syn* ridge. Because the *trans* and *cis* structures are stable with respect to ζ, no constraints were applied on ζ for their optimization. In addition, there are certain regions that are stable with respect to ψ; they are *trans* with ψ around 75° (the global energy minimum), *cis* with ψ around 0° (the energy minimum in *cis*), and *anti* and *syn* with ψ around 0° (saddle regions for ζ with positive second energy derivatives with respect to ψ). For these, no constraints were applied to ψ. The energies of the 20 calculated points are listed in Table 2. All geometry optimizations were started with conformers in the *exo* form from the structures obtained with the empirical potential. During *ab initio* optimization, all the *anti* conformers converted into the

(25) Northrup, S. H.; Pear, M. R.; Lee, C.-Y.; McCammon, J. A.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4035–4039.

(26) Schlegel, H. B. *In Modern Electronic Structure Methods*; Yankony, D. R., Ed.; World Scientific: London, 1994.

Table 1. 3-21G Optimized Geometries of Proline Dipeptide (Angles in Degrees)[a]

	anti ($\zeta = -80$)	*cis*		*syn* ($\zeta = 85$)	*trans*	
ψ, deg	η'	η'	ζ	η'	η'	ζ
+150	38.95	6.61	0.17	−25.45	9.14	181.62
+75	33.24	−15.43	−3.51	−47.51	−6.83 ($\psi = 69.96$)[b]	183.88
0	21.33 ($\psi = 11.44$)[b]	−14.45 ($\psi = -1.63$)[b]	2.37	−49.74 ($\psi = 2.14$)[b]	−15.99	178.97
−75	45.46	−10.36	−2.89	−49.97	−14.64	175.67
−150	30.31	4.84	−1.14	−52.55	6.01	173.39

[a] See Results section for an explanation of the nitrogen out-of-plane improper dihedral η' and Figure 2c for the imide torsion angle ζ. Geometry optimization was performed with ζ fixed at −80° for the *anti* and +85° for the *syn* conformers, while fixing ψ at the values indicated (unless otherwise specified). [b] ψ also optimized (value given in parentheses); see text.

Table 2. 3-21G and 6-31G*//3-21G Energies of the Proline Dipeptide[a]

	anti		*cis*		*syn*		*trans*	
ψ, deg	3-21G	6-31G*//3-21G	3-21G	6-31G*//3-21G	3-21G	6-31G*//3-21G	3-21G	6-31G*//3-21G
+150	27.67	25.34	8.54	5.44	26.39	23.88	6.07	2.56
+75	26.94	24.61	11.37	9.00	25.44	22.67	0.00[c]	0.00[c] (0.00)[b,c]
0	21.27	20.72 (20.65)[b]	5.03	2.75 (3.02)[b]	20.64	18.07 (17.91)[b]	6.27	3.31
−75	30.71	25.82	10.76	7.88	27.07	21.75	15.97	11.32
−150	35.18	32.91	10.22	7.87	26.24	21.92	12.28	9.28
$\langle E \rangle_\psi$		20.73		2.78		18.09		0.05

[a] Geometries as given in Table 1 (3-21G optimized), unless otherwise specified. $\langle E \rangle_\psi$ is the Boltzmann average of the energies in that column. [b] Geometry optimized at the 6-31G* level (6-31G* energy given in parentheses). [c] Energies are in kcal/mol above the global *trans* minimum.

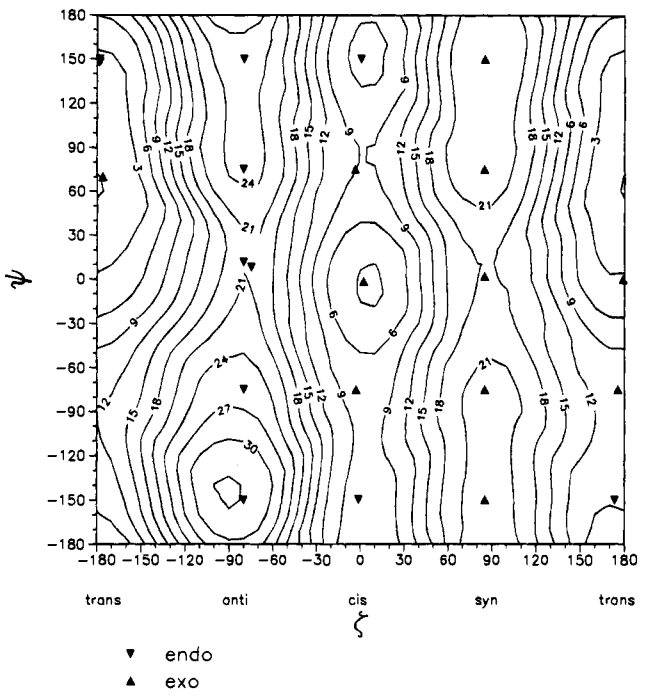

Figure 7. *Ab initio* adiabatic energy map of proline dipeptide, calculated at the 6-31G*//3-21G level over a grid of 20 (ζ,ψ) points, whose locations on the map are marked with triangles (see text and tables). All geometry optimizations were started with the imide nitrogen in the *exo* conformation. Triangles pointing up indicate the points which remained in *exo*, while triangles pointing down indicate the points which were transformed into *endo* during the optimization. Energies are in kcal/mol above the *trans* minimum.

endo form ($\eta' > 0°$), while all *syn* conformers remained *exo* ($\eta' < 0°$) (see Discussion section). For each of these 20 3−21G optimized geometries, the energy was evaluated with the 6-31G* basis set (Table 2) and an adiabatic map (see Figure 7) was constructed from the results by bicubic spline interpolation[27] in terms of (ζ,ψ). The accuracy of the results was checked by performing geometry optimization with the 6−31G* basis set at the four most important points: the *trans* and *cis* minima and the *anti* and *syn* saddle points. Their relative energies, given in Table 2, do not differ significantly from those obtained with the 6−31G*//3−21G calculation. This is in accord with the results of Head-Gordon *et al.*,[28] who have shown that the 3-21G basis set is adequate for geometry optimizations in the alanine dipeptide. The low-resolution *ab initio* adiabatic energy map in Figure 7 displays the same overall features as the corresponding high-resolution empirical map in Figure 6 in terms of the location of valleys, ridges, minima, and saddle regions. This indicates that the points used to calculate the *ab initio* map were chosen appropriately; i.e., the resulting map is expected to display the essential characteristics of the *ab initio* potential energy surface.

Barrier of *N*-Acetylpyrrolidine. To understand the effects of ψ on the activation barrier for *cis–trans* isomerization, the corresponding barrier was calculated for *N*-acetylpyrrolidine (NAP), a molecule which lacks the C-terminal amide unit of the proline dipeptide (see Figure 2B). Due to the symmetry, there is no energetic distinction in the case of NAP between the *syn* and *anti* transitions, nor between the *cis* and *trans* states. The transition state of NAP is quasisymmetric: Disregarding ring puckering, the C_α and C_δ atoms are related through a mirror plane, which contains the heavy atoms of the peptide moiety and is orthogonal to the ring plane. Full geometry optimization at the 6-31G* level of the transition state is then achieved by requiring $\omega = \omega'$ without constraining ω and ω' to specific values (ω' is defined over $C_1-C_2-N-C_\delta$), while maintaining a planar carbonyl carbon C_2 center by fixing the improper dihedral angle defined over $O-C_2-N-C_1$ at 180°. The resulting *ab initio* activation barrier for torsion of the imide bond is $\Delta E^\ddagger = 16.3$ kcal/mol.

(27) Akima, H. A. *ACM Trans. Math. Software* **1978**, *4*(2), 148.
(28) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L., III; Pople, J. *Int. J. Quantum Chem.* **1989**, *16*, 311.

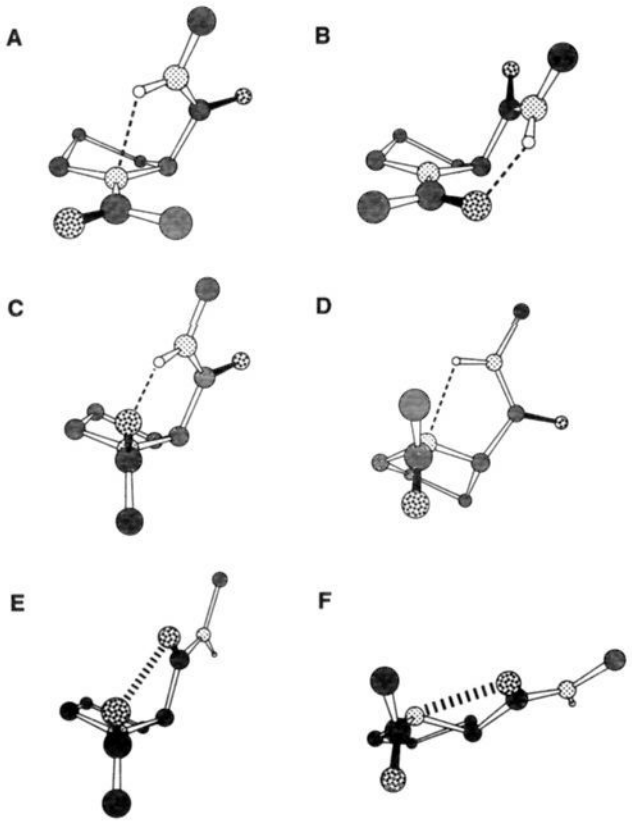

Figure 8. Proline dipeptide conformers. Top: Local minima *cis* (A) and *trans* (B). Middle: Stabilized transition states *anti*,$\psi \cong 0°$ (C) and *syn*,$\psi \cong 0°$ (D). Bottom: Destabilized transition states *anti*,$\psi \cong -150°$ (E) and *syn*,$\psi \cong -150°$ (F). See Table 1 for the geometries of these conformers. Dominant nonbond interactions: attractive (- - -), respulsive (|||). Carbons are gray, nitrogens have small dots, and oxygens have large dots.

Discussion

Empirical vs *ab Initio* Potential. The locations of the *cis* and *trans* minima coincide approximately in the empirical and *ab initio* maps (Figures 6 and 7). The *anti* transition ridge in the *ab initio* surface has a pronounced dip at $\psi \sim 10°$, which is well reproduced by the empirical potential. However, the empirical potential fails to duplicate the depth and the location of the dip in the *syn* ridge at $\psi \sim 0°$ on the *ab initio* map; the dip in the empirical map is at $\psi \sim -45$. These similarities and differences are easily understood. The dip in the *anti* ridge occurs because the carbonyl oxygen O_2 with a partial negative charge interacts with the C-terminal amide group; the interaction is favorable with the positive amide hydrogen H_5 when $\psi \sim 0°$ (Figure 8C) and unfavorable with the negative amide carbonyl oxygen O_4 when ψ is between 150° and 210° (Figure 8E). This effect is reproduced in the CHARMM potential by the Coulomb interactions of the corresponding partial charges on atoms O_2, H_5, and O_4. The dip in the *syn* ridge, on the other hand, results mostly from the inteactions of the lone pair orbital of the sp^3 hybridized imide nitrogen N_3 (see the *syn*/*exo* conformation in Figure 3) with the C-terminal amide. This leads to favorable interactions with the amide hydrogen H_5 when $\psi \sim 0°$ (Figure 8D or Figure 2C) and unfavorable interactions with the amide carbonyl oxygen O_4 when ψ is between 150° and 210° (Figure 8F). Since the empirical potential does not model the electron density of the lone pair, this effect is not reproduced. These interactions explain why the lowest saddle points on the *ab initio* potential energy surface are found around $\psi \sim 0°$, on both the *anti* ridge and the *syn* ridge.

The energy difference of the two maps shown in Figure 6 and 7 yields a correction term,

$$E(\zeta,\psi)_{cor} = E(\zeta,\psi)_{6-31G^*} - E(\zeta,\psi)_{CHARMM}$$

which can be added to the total empirical energy of a protein, so that the imide twisting behavior of the proline residues of interest is correctly reproduced. The derivatives needed for energy minimization or dynamics can be obtained by finite difference calculations. This approach was used in a recent study of the FKBP rotamase, where the quantum correction was applied to the empirical potential of the prolyl substrate and a mechanism of rotamase catalysis consistent with experimental data was proposed.[18]

Nitrogen Pyramidalization. A striking characteristic of the *ab initio* map is that all the points on the *syn* ridge remained in the *exo* form ($\eta' \in [-25°,-53°]$), whereas all the points on the *anti* ridge turned into *endo* ($\eta' \in [21°,45°]$). In their study of the origin of rotation barriers in formamide, Wiberg and Leidig suggested that this is due to the interactions between the dipole created by the lone pair orbital on the sp^3 nitrogen of the twisted amide and the bond dipole of the C_2-O_2 carbonyl.[20] They computed an energy difference of 2.3 kcal/mol between the parallel and antiparallel forms of formamide. Figure 3 shows how these two dipoles are oriented in parallel in the *syn*/*endo* and the *anti*/*exo* conformations, whereas they are oriented antiparallel in the two remaining transition state structures. This leads to the expected destabilization of the former and stabilization of the latter. This effect is not reproduced for the *anti*/*exo* saddle point on the empirical adiabatic energy surface in Figures 4 and 5, due to the lack of a lone pair orbital dipole.

Dipeptide Interactions and Autocatalysis. The interactions between the two peptide moieties affect both the ground state and the transition state. The energy in the two ground-state valleys of the *ab initio* map varies by 6.3 and 11.3 kcal/mol as a function of ψ for *cis* and *trans*, respectively (see Table 2 and Figure 7). In *trans*, the C-terminal H_5 is hydrogen bonded to the carbonyl oxygen O_2, accounting for the greater stability of the *trans* minimum relative to the *cis* minimum (see Figure 8B). In *cis*, the C-terminal H_5 interacts favorably with the imide nitrogen N_3, as indicated by the marked preference for a value of $\psi \sim 0°$ (see Figure 8A). The *cis* minimum is 3 kcal/mol higher in energy than the *trans* minimum. This is significantly larger than the free energy difference in solution, which for proline is in the range 1−1.4 kcal/mol;[9,29] the solvent effect on the *cis*/*trans* free energy difference is estimated by a RISM-HNC to be about 2 kcal/mol stabilizing the *cis* form (H.-A. Yu and M. Karplus, unpublished).

The variation in the *ab initio* energy along the *anti* ridge as a function of ψ is 12.2 kcal/mol. The corresponding variation along the *syn* ridge is 5.8 kcal/mol (see Table 2 and Figure 7). This strong dependence on ψ is the result of stabilizing effects when $\psi \sim 0°$ and destabilizing effects when $\psi \sim 180°$, as described above.

Relative to the global minimum in the planar *trans* configuration ($\psi = 70°$), the lowest *syn* barrier to isomerization is 17.9 kcal/mol and the lowest *anti* barrier is 20.7 kcal/mol (Table 2). Relative to the local minimum in the planar *cis* configuration ($\psi = -2°$), the barriers are 14.9 and 17.7 kcal/mol for the lowest *syn* and *anti* transitions, respectively. Boltzmann averages over ψ for 300 K yield essentially identical results (see Table 2); e.g. $\langle E \rangle_{\psi,trans} = 0.05$ kcal/mol versus the minimum of zero. The calculated *syn* barriers are 2−4 kcal/mol smaller than the

(29) Fersht, A. In *Enzyme Structure and Mechanism*; Freeman: New York, 1985.

activation energy in solution,[30] confirming that water stabilizes the imide bond against twisting. This is consistent with the *ab initio* calculations of Scheiner and Kern,[31] who showed that hydrogen bonding of one or more water molecules to the carbonyl oxygen of N-methylacetamide increases the barrier to *cis–trans* isomerization by up to 2.1 kcal/mol. Since the activation barrier for *N*-acetylpyrrolidine is 16.3 kcal/mol (see above), the stabilizing effect of the C-terminal amide group lowers the *cis* → *syn* barrier by 1.4 kcal/mol.

Such a barrier lowering for proline isomerization corresponds to a form of "autocatalysis". It could occur during the *syn* reaction, with the C-terminal amide of proline dipeptide acting as the "catalyst". Autocatalysis in prolyl *cis–trans* isomerization has been shown to occur in dihydrofolate reductase.[32] During folding of that protein, Pro 66 undergoes a *trans* → *cis* isomerization catalyzed by the positively charged residue Arg 44, which interacts with the imide nitrogen. According to the present study, the interaction is likely to involve the dipole of the nitrogen lone pair orbital. The term "autocatalysis" is applied to both cases because the substrate and the catalyzer are covalently linked. The interaction of a positive charge with the imide nitrogen will always be more favorable in the twisted than in the planar state. Such a charge can be provided by a polar hydrogen or a protonated side chain. In the prolyl–peptide binding site of the cyclophilin rotamase, an arginine and a histidine are in close proximity to the proline of the bound substrate.[33] The presence of a positively charged ammonium ion, when hydrogen bonded to the imide nitrogen, has been shown by *ab initio* calculations to lower the activation barrier, as compared with the uncatalyzed reaction.[38]

Role of Prolines in Proteins. Cis imide bonds, which occur in about 10% of all proline residues in proteins, have a marked preference for turns and bends in the polypeptide chain.[34] They are often found in one of two types of turn structures, referred to as type VIa and type VIb proline turns.[35] The essential difference between these two types of turns is the value of the proline backbone torsion angle ψ. Type VIa turns are characterized by $\psi \sim 0°$, while type VIb turns have ψ in the 180° domain. According to the present study, prolyl imide bonds in type VIa turns are expected to have lower energy barriers than those in type VIb turns. Autocatalysis of the type considered is expected to occur during folding of proteins with native type VIa prolines, but not for type VIb prolines, if the *trans* → *cis* isomerization takes place in the neighborhood of the native conformation. Prolines appear to be critical to a variety of protein functions, as in signal transduction in G proteins[36] and in voltage-gating by GAP junction channels.[37] However, there is no evidence that *cis–trans* imide isomerization is involved.

Rotamase Catalysis. The average height of the *syn* ridge, relative to the *cis* valley, is about 15 kcal/mol, whereas the energy difference between the highest point in the *cis* valley and the *syn* saddle point is only about 9 kcal/mol (Table 2 and Figure 7). An enzyme that catalyzes the *cis* → *syn* → *trans* reaction could take advantage of this by controlling the variation of ψ along the reaction path. In a model for the catalysis of proline isomerization by the rotamase FKBP,[18] the enzyme destabilizes the *cis* ground state by restraining ψ to be about 38°, while ψ is near 14° at the *syn* transition state of the substrate/enzyme complex. This decreases the effective activation barrier by about 3.6 kcal/mol and provides one contribution to the overall enthalpy barrier lowering of 13.5 kcal/mol by the enzyme.[18]

Conclusion

A detailed study of proline isomerization has been made by use of empirical energy functions and *ab initio* calculations. Significant differences between the empirical and *ab initio* energy maps are found. Because imide nitrogen pyramidalization is an essential part of the process, a virtual dihedral angle, ξ, is shown to be a better reaction coordinate for isomerization than the standard peptide angle ω.

The (ξ,ψ) *ab initio* adiabatic map illustrates the importance of the dihedral angle ψ^{Pro} in the *cis–trans* isomerization reaction. An essential element in the transition states for the imide isomerization that distinguishes them from the *cis* and *trans* ground states is the presence of a dipole moment on the imide nitrogen, due to the pyramidalization of the nitrogen and the accompanying shift of its lone pair from a p_z to an sp^3 orbital. Favorable interactions with this orbital dipole moment are expected to be part of the catalytic mechanism of peptidyl–prolyl isomerases like FKBP and cyclophilin. The angle ψ^{Pro} controls the interaction between the proline C-terminal peptide group and the nitrogen lone pair in the transition state. A stabilizing interaction between the amide and the lone pair makes an autocatalytic contribution to catalysis in the FKBP rotamase.

Acknowledgment. The work reported here was supported in part by a grant from the National Science Foundation and a gift from Molecular Simulations, Inc. The calculations were performed on a Cray Y-MP at the National Center for Supercomputing in Illinois and on a Convex C220. We thank Stephen Michnick for valuable discussions and for help with the calculations on NAP.

(30) Stein, R. L. *Adv. Protein Chem.* **1993**, *44*, 1.
(31) Scheiner, S.; Kern, C. W. *J. Am. Chem. Soc.* **1977**, *99*, 7042.
(32) Texter, F. L.; Spencer, D. B.; Rosenstein, R.; Matthews, C. R. *Biochemistry* **1992**, *31*, 5687.
(33) Kallen, J.; Walkinshaw, M. D. *FEBS Lett.* **1992**, *300*, 286.
(34) Stewart, D. E.; Sarkar, A.; Wampler, J. E. *J. Mol. Biol.* **1990**, *214*, 253.
(35) Richardson, J. S.; Richardson, D. C. Principles and patterns of protein conformation. In *Prediction of protein structure and the principles of protein conformation*; Gerald D. Fasman: New York 1989.
(36) Trumpp-Kallmeyer, S.; Hoflack, J.; Bruinvels, A.; Hibert, M. *J. Med. Chem.* **1992**, *35*, 3448.
(37) Suchyna, T. M.; Xu, L. X.; Gao, F.; Fourtner, C. R.; Nicholson, B. J. *Nature* **1993**, *365*, 847.
(38) Danbrack, R. L., Jr. Ph.D. Thesis, Harvard University, 1993.