The flexibility in the proline ring couples to the protein backbone

BOSCO K. HO,[1] EVANGELOS A. COUTSIAS,[2] CHAOK SEOK,[3] AND KEN A. DILL[1]

[1]Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94148, USA
[2]Department of Mathematics and Statistics, University of New Mexico, Albuquerque, New Mexico 87131, USA
[3]School of Chemistry, College of Natural Sciences, Seoul National University, Seoul 151-747, Republic of Korea

(RECEIVED October 4, 2004; FINAL REVISION December 20, 2004; ACCEPTED December 21, 2004)

Abstract

In proteins, the proline ring exists predominantly in two discrete states. However, there is also a small but significant amount of flexibility in the proline ring of high-resolution protein structures. We have found that this side-chain flexibility is coupled to the backbone conformation. To study this coupling, we have developed a model that is simply based on geometric and steric factors and not on energetics. We show that the coupling between φ and χ1 torsions in the proline ring can be described by an analytic equation that was developed by Bricard in 1897, and we describe a computer algorithm that implements the equation. The model predicts the observed coupling very well. The strain in the C^{γ}-C^{δ}-N angle appears to be the principal barrier between the UP and DOWN pucker. This strain is relaxed to allow the proline ring to flatten in the rare PLANAR conformation.

Keywords: proline; pucker; backbone; cyclic ring

We are interested in understanding the variations in the conformations of the proline ring that are observed in the Protein Data Bank (PDB). It is well known that the proline ring exists in two predominant states (Ramachandran et al. 1970; Altona and Sundaralingam 1972). However, a recent study has found that within these states in peptides, there is a significant amount of flexibility (Chakrabarti and Pal 2001). This flexibility is coupled directly to the backbone. What is the nature of this coupling? To answer this question, we have measured proline ring conformations in high-resolution protein structures, and we give a detailed analysis of the degrees of freedom in the proline ring. Our modeling strategy is based on the Bricard equation of the flexible tetrahedral angle (Bricard 1897). It has recently been used to solve the problem of tripeptide loop closure (Coutsias et al. 2004). Here, we apply the Bricard equation to the five-membered ring of proline to generate proline ring conformations. We test our model against the observed structures of the proline ring.

DeTar and Luthra (1977) argued that the proline ring exists in essentially two discrete states, even though proline is a five-membered ring, which has, in principle, a continuum of available conformations (Altona and Sundaralingam 1972). These discrete states are known as the UP and DOWN puckers of the proline ring and have been reproduced in force-field calculations (Ramachandran et al. 1970; DeTar and Luthra 1977). There is also some evidence of a rare PLANAR conformation (EU 3-D Validation Network 1998). However, as these calculations use generic force fields, constraints due to geometry cannot be separated out from constraints due to other energetic factors. Using our analytical approach, we can determine which constraints are due to geometry and which are due to other energetic factors.

Proline is unique among the naturally-occurring amino acids in that the side-chain wraps around to form a covalent bond with the backbone, severely restricting the backbone. Because of the restricted backbone, proline is used in nature in many irregular structures such as β-turns and α-helical capping motifs (MacArthur and Thornton 1991; Chakrabarti

Reprint requests to: Bosco K. Ho, Department of Pharmaceutical Chemistry, University of California San Francisco, 600 16th Street, San Francisco, CA 94148, USA; e-mail: bosco@maxwell.ucsf.edu; fax: (415) 502-4222.

Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.041156905.

and Pal 2001; Bhattacharyya and Chakrabarti 2003), and proline restricts the backbone conformation of neighboring residues (Schimmel and Flory 1968; MacArthur and Thornton 1991). Modeling these structural motifs requires an accurate description of the proline ring. There have been many force-field calculations of the proline ring (Ramachandran et al. 1970; DeTar and Luthra 1977; Summers and Karplus 1990; Némethy et al. 1992). Although the restriction on the φ torsion has been reproduced (Ramachandran et al. 1970; Summers and Karplus 1990), the coupling of the backbone to the proline ring has not. Modeling the coupling and flexibility in the proline ring can be important, for example, in constrained ring peptides (data not shown). Our geometric model of the proline ring captures both these features. It is an efficient algorithm that should be easily implemented in models of structural motifs involving proline.

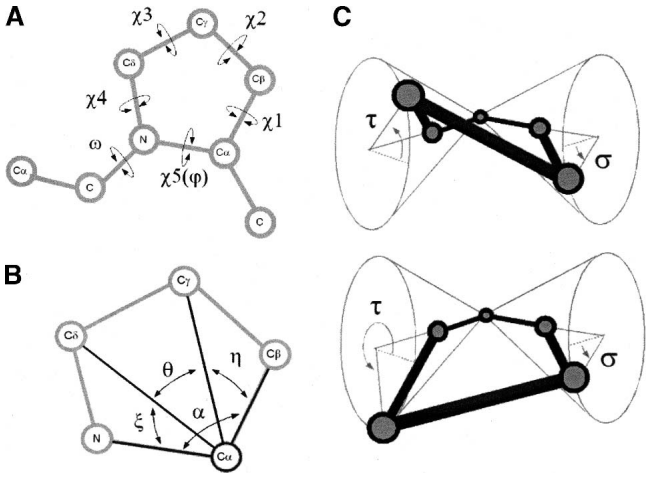

Figure 1. Schematic of the proline ring. (*A*) The torsions in the proline ring. χ5 and φ are related as they measure different torsions around the same central axis. (*B*) The tetrahedral angle in the proline ring has the apex at the C^{α} atom, and the N, C^{β}, C^{γ}, and C^{δ} atoms define the faces of the baseless pyramid that meet at the apex. The apical angles α, η, ξ, and θ are also shown. (*C*) After fixing three of the apical angles (α, η, and ξ), the degrees of freedom consists of the σ and τ torsions, each tracing out a cone. For each pair of values of σ and τ, there is a unique distance between the end points of the two cones. Fixing θ effectively couples σ to τ by fixing the end-to-end distance. If a distance is given, there will be two solutions of τ for σ, and vice versa.

Results

Proline ring conformations in the PDB

In order to determine the conformations of proline, we chose a high-resolution subset of the PDB (Berman et al. 2000) provided by the Richardson laboratory (Lovell et al. 2003) of 500 nonhomologous proteins. These proteins have a resolution of >1.8 Å where all hydrogen atoms have been projected from the backbone and optimized in terms of packing. Following the Richardson method, we eliminate conformations having a B-factor >30, and we only accept proline residues that contain all atoms, including the hydrogens. We define the *trans*-Pro isomer by ω: $90° < \omega < 220°$. Due to the predominance of the *trans*-Pro isomer (4289 counts) over the *cis*-Pro isomer (236 counts), we have focused mainly on the *trans*-Pro isomer.

In the proline ring, there are five endo-cyclic torsions (χ1, χ2, χ3, χ4, and χ5) (Fig. 1A). If we assume planar trigonal bonding at the N atom and tetrahedral bonding at the C^{α} atom, then $\varphi = \chi5 - 60°$. This is approximately satisfied as the observed relationship between the χ5 and φ torsions are relatively linear (Fig. 1B; see also Chakrabarti and Pal 2001). The two discrete states in the proline ring are referred to as the UP and DOWN puckers (Milner-White 1990). UP and DOWN refers to whether the C^{γ} atom is found above or below the average plane of the ring. The four atoms C^{α}, C^{β}, C^{δ}, and N are found close to a planar conformation and can serve as the plane of the proline ring (Chakrabarti and Pal 2001). Another way to characterize the puckers is by the sign of the χ torsions, UP (negative χ1 and χ3, positive χ2 and χ4) and DOWN (positive χ1 and χ3, negative χ2 and χ4). In this study, we follow the method of DeTar and Luthra (1977) in using χ2 to determine the pucker, especially since the observed values of χ2 have the largest magnitude among the χ torsions. However, we also want to include the PLANAR conformation in our analysis. Hence our definition is UP ($\chi2 > 10°$), DOWN ($\chi2 < -10°$), and PLANAR ($-10° < \chi2 < 10°$).

Table 1 lists the parameters of the pyrrolidine ring in the *trans*-Pro isomer—the χ torsions, bond lengths, and bond angles. The bond lengths have little variation; the standard deviation is 0.021 Å. The bond angles, on the other hand, do show some variation. The greatest variation is in the C^{β}-C^{γ}-C^{δ} angle, which has a standard deviation of 2.6°, almost twice that of some of the other angles. This angle is the most flexible because its central atom, the C^{γ} atom, is opposite to the atoms in the C^{α}-N bond, which, in turn, bond to three other heavy atoms. This is in agreement with DeTar and Luthra (1977), who found that most of the mobility in the proline ring observed in crystal structures is found in the C^{γ} and C^{β} atoms and, to a lesser extent, in the C^{δ} atom.

The PDB shows significant correlations between the φ and χ torsions (Table 1). We plot some of these distributions, χ1 versus φ (Fig. 2A), χ3 versus χ2 (Fig. 3A), and χ4 versus χ3 (Fig. 3B). They consist of two lobes of high density with sparse density between the lobes. Although not evident in the correlations, we also found that the χ torsions are coupled to the bond angles (Fig. 4A–C). The strongest coupling is found in C^{β}-C^{γ}-C^{δ} versus χ2 (Fig. 4B), which has the shape of an inverted parabola. In the following sections, we model the observed couplings in the proline ring conformations.

The average χ torsions are near zero, while their standard deviations are large. This is because the proline ring con-

Table 1. *The parameters of the pyrrolidine ring of proline in the* trans*-Pro isomer and correlations with the φ and χ torsions*

	$\langle x \rangle \pm \sigma_n$	φ	χ1	χ2	χ3	χ4	χ5
φ	−64.2 ± 10.0°	1.00	−0.66	0.57	−0.49	0.28	0.81
χ1	−0.1 ± 26.8°		1.00	−0.99	0.96	−0.83	−0.78
χ2	2.1 ± 36.5°			1.00	−0.99	0.91	0.67
χ3	−3.3 ± 32.2°				1.00	−0.95	−0.57
χ4	3.5 ± 17.3°					1.00	0.30
χ5	−2.2 ± 9.8°						1.00
N-C^α	1.464 ± 0.012 Å	0.07	−0.05	−0.08	0.07	−0.07	0.06
C^α-C^β	1.530 ± 0.012 Å	−0.05	−0.04	0.00	0.00	0.01	−0.02
C^α-C^γ	1.501 ± 0.021 Å	−0.03	0.04	−0.01	0.02	−0.03	0.05
C^γ-C^δ	1.513 ± 0.017 Å	−0.06	−0.02	0.04	−0.04	0.03	−0.01
C^δ-N	1.476 ± 0.011 Å	−0.03	−0.04	0.02	−0.02	0.02	−0.02
C^α-C^β-C^γ	103.9 ± 1.8°	−0.03	0.06	0.02	−0.02	0.02	−0.01
C^β-C^γ-C^δ	104.5 ± 2.6°	−0.13	0.01	0.12	−0.11	0.09	−0.06
C^γ-C^δ-N	102.7 ± 1.3°	−0.12	0.01	0.17	−0.18	0.17	−0.16
C^δ-N-C^α	111.5 ± 1.2°	−0.12	0.06	0.11	−0.11	0.11	−0.10
N-C^α-C^β	103.7 ± 1.4°	0.15	−0.02	−0.08	0.07	−0.06	0.04

formations are split into two dominant conformations. We see a double peak in the χ2 frequency distribution (Fig. 4D), which makes the χ2 torsion a good discriminator between the UP and DOWN conformations. The peaks have an asymmetric shape. The φ torsion, on the other hand, is not a good discriminator of the UP and DOWN conformations (Fig. 2C). Table 2 lists the averages of the torsions and bond angles for the two different conformations. Between the UP and DOWN puckers, the bond angles are identical, and the χ2 values have virtually the same magnitude but different signs. The other χ torsions also change sign.

Table 2 lists the averages and standard deviations of the torsions and bond angles of the *cis*-Pro isomer. The bond angles of the UP and DOWN pucker in *cis*-Pro are similar to those of *trans*-Pro. The χ torsions have the same sign, but the magnitude differs by a few degrees. In the *cis*-Pro isomer, the DOWN pucker is massively favored over the UP pucker (see Fig. 2F). Also, for the DOWN pucker, φ has shifted further to the left in the *cis*-Pro isomer (Fig. 2F) compared with the *trans*-Pro isomer (Fig. 2C). This difference is due to a C^α_{i-1}-C steric clash that disfavors conformations of $\varphi > -70°$ and hence favors the DOWN pucker (Pal and Chakrabarti 1999). Another discrepancy appears in the correlation of χ5 versus φ (Fig. 2E), where the observed distribution deviates for the most negative values of φ from the slope that corresponds to ideal trigonal bonding at the N atom and ideal tetrahedral bonding at the C^α atom. Otherwise, we find that the coupling between the internal χ torsions is consistent with those of the *trans*-Pro isomer (data not shown).

The Bricard equation for the tetrahedral angle

According to our PDB statistics, the bond lengths in the proline ring do not vary significantly. However, there is a small amount of variation in some of the bond angles. For the five atoms of the ring, there are $5 \times 3 = 15$ degrees of freedom (DOF). Six of these are due to the absolute position and rotation, which are irrelevant for us. Fixing five of the bond lengths imposes five constraints. Thus the number of DOF for a ring with fixed bond lengths is $15 - 6 - 5 = 4$. If we also fix three of the bond angles, then we will have $4 - 3 = 1$ DOF. We do this below, and we find that modeling proline ring conformations in one dimension is sufficient to understand the observations described in the previous section.

We can identify a tetrahedral angle in the five-membered proline ring. A tetrahedral angle describes the apex of a baseless pyramid. The apex of such a pyramid is formed by the convergence of four planes. In the proline ring, if we place the apex at the C^α atom, then the C^γ, C^δ, C^β and N atoms define the four faces of the baseless pyramid (black lines in Fig. 1B). Each plane that meets at the apex is described by an apical angle: α, η, ξ, and θ (black lines in Fig. 1B). After the apical angles are fixed, there still remains a DOF. This can be described by the planar angles of adjacent side faces. The planar angles are described by the torsions σ and τ (Fig. 1C). We can then make use of the Bricard equation of the flexible tetrahedral angle (Bricard 1897) that relates the two planar torsions to the four apical angles:

$$\cos\theta + \cos\eta\cos\xi\cos\alpha = \sin\alpha(\sin\xi\cos\eta\cos\sigma + \cos\xi\sin\eta\cos\tau) + \sin\xi\sin\eta(\sin\tau\sin\sigma + \cos\alpha\cos\tau\cos\sigma)$$

If we fix the α, η, ξ, θ apical angles of the tetrahedral angle, then the Bricard equation gives the relationship between the σ and τ torsions of the tetrahedral angle (Fig. 1C) and the tetrahedral angle has one DOF. By introducing the projective transformation: $u = \tan\sigma/2$, $v = \tan\tau/2$; the

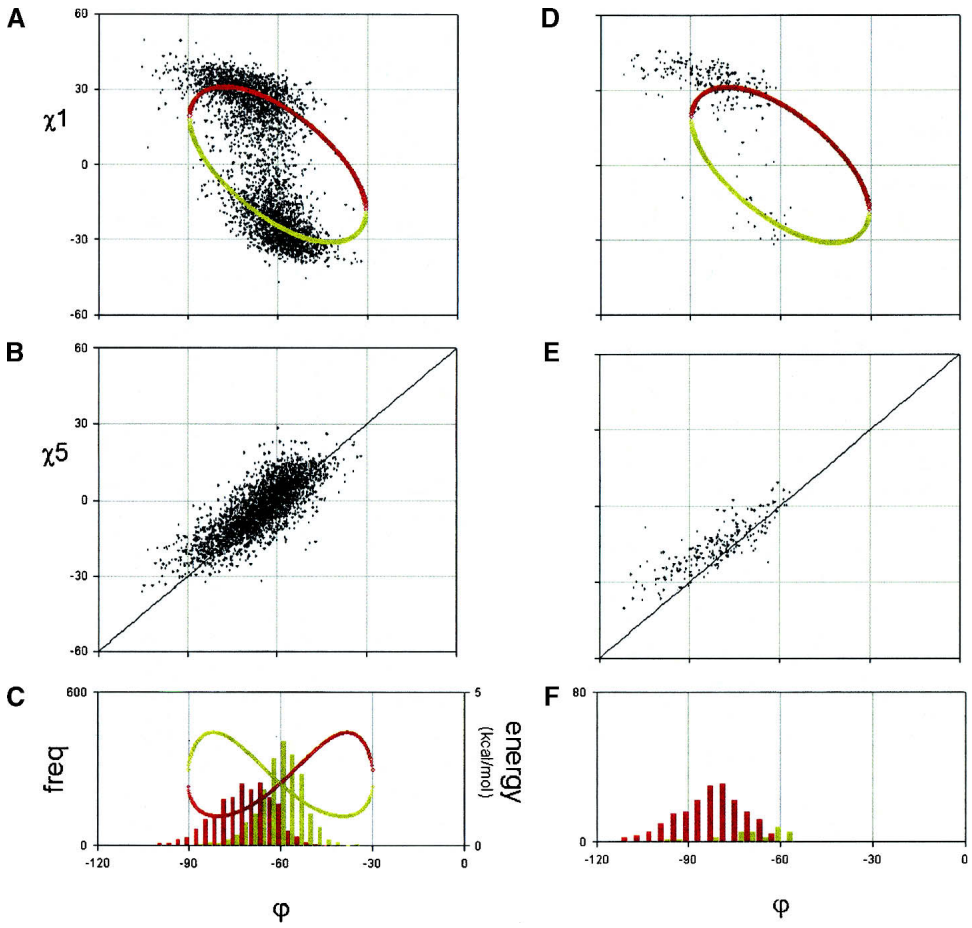

Figure 2. Distributions involving the φ torsion. For the *trans*-Pro isomer: (*A*) plot of χ1 vs. φ, and the curves represent the model based on the flexible tetrahedral angle; (*B*) plot of χ5 vs. φ, where the line corresponds to conformations where there is ideal trigonal planar bonding at N and ideal tetrahedral bonding at C^{α}; and (*C*) the distributions of φ for the DOWN pucker (red) and the UP pucker (yellow). (*D*,*E*,*F*) Corresponding plots for the *cis*-Pro isomer. The colors of the curve represent the UP pucker (red) and the DOWN pucker (yellow). The energy curve in *C* is due to the CHARMM C^{γ}-C^{δ}-N angle term.

Bricard equation becomes a quadratic polynomial in both u and v. Therefore for each value of u (resp. v), there are in general two values of v (resp. u). Thus, there will in general be two solutions when we solve for one of the torsions σ or τ in terms of the other (Coutsias et al. 2004). The full details of the derivation of the Bricard equation can be found in Coutsias et al. (2004).

How can we understand the DOF in the flexible tetrahedral angle? Assume first that the C^{γ}-C^{δ} distance is not fixed. As the other bond lengths are fixed, the triangles containing the α, η, and ξ angles are fixed (Fig. 1B,C). Consequently, the two DOFs are (1) the τ torsion, or the rotation of the C^{γ} atom around the bond C^{α}-C^{β}, which preserves the triangle C^{α}-C^{β}-C^{γ}; and (2) the σ torsion, or the rotation of the C^{δ} atom around the bond C^{α}-N, which preserves the triangle C^{α}-C^{β}-C^{γ} (cones in Fig. 1C). The variation of τ and σ will change the C^{γ}-C^{δ} distance. The conformations of a flexible tetrahedral angle correspond to the coupled values of τ and σ that give the fixed value of the C^{γ}-C^{δ} distance.

Constructing proline ring conformations

We now apply Bricard's equation of the tetrahedral equation to the proline ring. We first fix the four apical angles (Fig. 1B). This effectively fixes three of the five bond angles, where the remaining two bond angles will be coupled. The choice of which bond angles to fix will determine the identity of the σ and τ torsions.

We first place the apex of the tetrahedral angle at the C^{α} atom. We then fix the bond angles centered on the N and C^{α} atoms as these atoms are part of the backbone and are bonded to three other heavy atoms. Of the remaining three angles, the C^{β}-C^{γ}-C^{δ} is the most flexible, so we leave this angle free. Of the two remaining angles, we fix the C^{δ}-N-C^{α} angle as this will the make the σ and τ torsions identical to the χ1 and χ5 torsions of the proline ring (Fig. 1, cf. A and C). As χ5 is related to φ by planarity, we now have an equation that relates φ to χ1. Thus, to construct proline ring conformations:

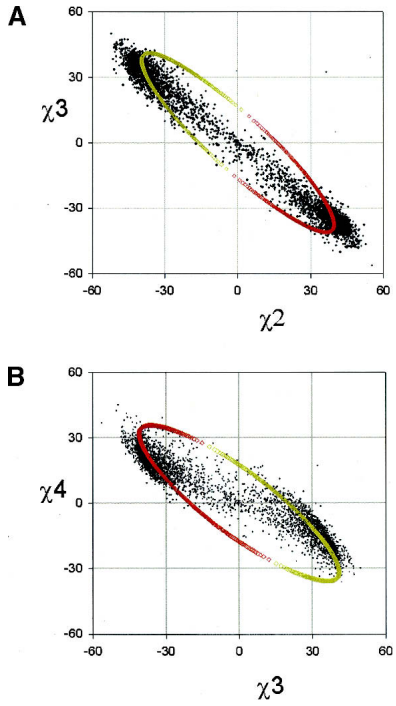

Figure 3. Correlations in the χ torsions. (*A*) χ3 vs. χ2; (*B*) χ4 vs. χ3. The model curve passes through the two lobes of high density, where it is clear that the slope of the two lobes are determined by the model curve and is different to the curve formed by the sparse density of points in between. The colors of the curve represent the UP pucker (red) and the DOWN pucker (yellow).

1. We set the apical angles. For the proline ring, we use the parameters of the average conformation of the UP pucker in Table 2. We set $\alpha = N\text{-}C^{\alpha}\text{-}C^{\beta} = 103.7°$, $\eta = C^{\alpha}\text{-}C^{\beta}\text{-}C^{\gamma} = 103.8°$, and $\xi = C^{\delta}\text{-}N\text{-}C^{\alpha} = 111.3°$. Keeping the bond angles and bond lengths fixed, we use basic trigonometry to calculate the $C^{\alpha}\text{-}C^{\gamma}$ and $C^{\alpha}\text{-}C^{\delta}$ distances. These two distances, combined with the $C^{\gamma}\text{-}C^{\delta}$ bond length, give $\theta = C^{\gamma}\text{-}C^{\alpha}\text{-}C^{\delta} = 36.3°$ (Fig. 1B).

2. We have now obtained the four apical angles (α, η, ξ, and θ) of the Bricard equation. For a given value of φ, we convert φ to $\chi 5 = \varphi - 60°$ and solve the Bricard equation for χ1, which requires the following coefficients:
$A = -\cos\alpha \sin\xi \sin\eta \cos\chi 5 + \sin\alpha \sin\xi \cos\eta$
$B = -\sin\xi \sin\eta \sin\tau$
$C = \cos\theta - \cos\alpha \cos\xi \cos\eta - \sin\alpha \cos\xi \sin\eta \cos\chi_0$.

   Using these coefficients, we have the following condition:

   if $|C/\sqrt{(A^2 + B^2)}| > 1$, then there is no solution for that value of φ.

   Otherwise, we calculate the following:

   $\tau 1 = \arccos[C/\sqrt{(A^2 + B^2)}]$.

   If $B > 0$, then

   $\tau 0 = -\arccos[A/\sqrt{(A^2 + B^2)}]$

   else

   $\tau 0 = +\arccos[A/\sqrt{(A^2 + B^2)}]$.

   For the UP pucker, we set $\chi 1 = \tau 0 - \tau 1$, and for the DOWN pucker, we set $\chi 1 = \tau 0 + \tau 1$. Obviously, there is only one solution if $\tau 1 = 0$, which represents the inflection point between the UP and DOWN puckers.

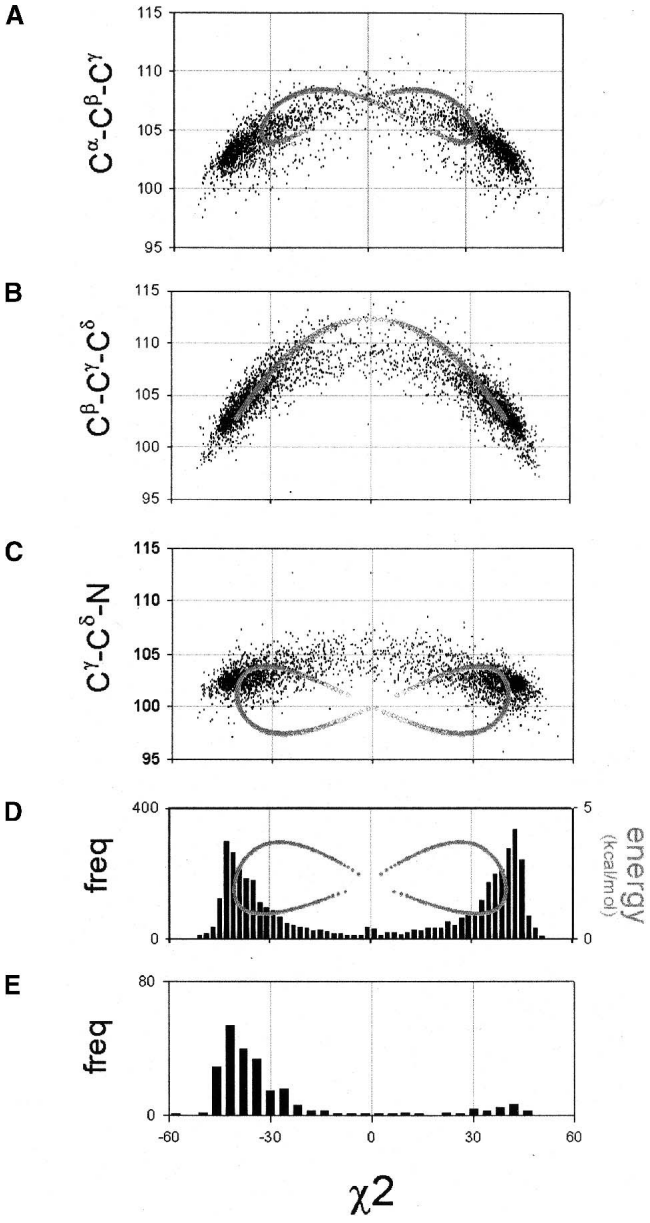

Figure 4. Bond angle strain in the proline ring as a function of χ2 for $C^{\alpha}\text{-}C^{\beta}\text{-}C^{\gamma}$ (*A*), $C^{\beta}\text{-}C^{\gamma}\text{-}C^{\delta}$ (*B*), and $C^{\gamma}\text{-}C^{\delta}\text{-}N$ (*C*). In the disfavored region at $\chi 2 \sim 0°$, the model values of $C^{\gamma}\text{-}C^{\delta}\text{-}N$ are found significantly below the observed values. This is the key strain that separates the puckers. χ2 frequency distributions for the *trans*-Pro isomer (*D*) and the *cis*-Pro isomer (*E*). The energy curve in *D* is due to the CHARMM $C^{\gamma}\text{-}C^{\delta}\text{-}N$ angle term.

Table 2. *Torsions and bond angles [°] for the different puckers*

	PLANAR	UP	DOWN
trans-PRO			
counts	216	2160	1913
ω	180.5 ± 5.3	179.5 ± 3.6	180.5 ± 4.0
φ	−66.8 ± 11.1	−58.9 ± 7.3	−69.8 ± 9.3
χ1	4.6 ± 12.2	−24.9 ± 8.3	27.5 ± 7.5
χ2	−3.1 ± 10.2	36.7 ± 8.6	−36.3 ± 7.7
χ3	0.4 ± 6.0	−33.8 ± 7.3	30.6 ± 7.7
χ4	2.7 ± 7.0	18.6 ± 6.6	−13.6 ± 8.0
χ5	−4.6 ± 11.2	3.7 ± 7.2	−8.6 ± 7.8
C^{α}-C^{β}-C^{γ}	106.7 ± 1.8	103.8 ± 1.7	103.7 ± 1.6
C^{β}-C^{γ}-C^{δ}	108.9 ± 1.8	104.0 ± 2.4	104.4 ± 2.4
C^{γ}-C^{δ}-N	104.7 ± 1.5	102.6 ± 1.3	102.8 ± 1.2
C^{δ}-N-C^{α}	112.4 ± 1.5	111.3 ± 1.1	111.6 ± 1.1
N-C^{α}-C^{β}	105.3 ± 1.7	103.7 ± 1.4	103.4 ± 1.3
cis-PRO			
counts	19	29	188
ω	−3.35 ± 5.35	−1.80 ± 4.42	1.01 ± 5.80
Φ	−88.2 ± 16.1	−68.5 ± 10.2	−82.2 ± 10.3
χ1	25.5 ± 11.7	−18.0 ± 11.2	33.9 ± 5.7
χ2	−18.1 ± 9.2	32.6 ± 11.6	−38.4 ± 6.2
χ3	3.8 ± 4.8	−34.0 ± 8.9	27.6 ± 7.5
χ4	13.3 ± 7.2	23.5 ± 6.3	−6.1 ± 8.3
χ5	−24.5 ± 11.4	−3.7 ± 8.4	−17.3 ± 7.3
C^{α}-C^{β}-C^{γ}	104.7 ± 2.2	104.7 ± 1.6	102.8 ± 1.7
C^{β}-C^{γ}-C^{δ}	107.6 ± 1.7	104.3 ± 2.4	104.3 ± 2.0
C^{γ}-C^{δ}-N	104.9 ± 3.1	102.2 ± 1.8	103.1 ± 1.3
C^{δ}-N-C^{α}	110.5 ± 2.7	111.2 ± 1.2	111.4 ± 1.5
N-C^{α}-C^{β}	104.0 ± 3.1	103.7 ± 0.9	102.7 ± 1.5

3. We now have the χ1 and χ5 torsions. Given the backbone atoms N, C, C^{α} atoms, we use the χ5 torsion, the bond lengths, and angles of the proline ring (Table 1) to place the C^{δ} and C^{β} atoms. Subsequently, we use the χ1 torsion to project the C^{γ} atom from the C^{β} atom.

Modeling the proline ring

Using the algorithm above, we generated the set of allowed proline ring conformations, varying φ from −180° to 0° in steps of 0.1°. From this set of conformations, we extract the model curves. The model curves for the ring angles are cyclic, due to the quadratic nature of the solution (Figs. 2A, 3A,B, 4A,B). The two main lobes of observed density lie along different parts of the cyclic curves with the exception of the region of low density between the two lobes. The fit to the cyclic curves is most evident in the plot of χ4 versus χ3 (Fig. 3B), where the slopes of the two main lobes lie along the cyclic curve, which is different to the slope connecting the two lobes. We conclude that the flexibility within the UP and DOWN pucker is consistent with the flexibility in a five-membered ring with fixed bond lengths and three fixed bond angles. As the χ2 distribution (Fig. 3A) and the φ distribution (Fig. 2A) lie within the limits of the curve, the range of the torsions is determined by the geometry of the five-membered ring.

Although the φ torsion is not a good discriminator between the UP and DOWN pucker, this is an advantage in generating proline conformations. In the graph of χ1 versus φ, the lobes are found along the top and bottom of the cyclic model curve (Fig. 2A). As the Bricard equation gives two solutions of χ1 for every value of φ, the two solutions will automatically correspond to the UP and DOWN pucker.

Some of the properties of the model based on the flexible tetrahedral angle can be anticipated by the pseudo-rotation of cyclic rings (Altona and Sundaralingam 1972). However, there are advantages in our approach compared with the pseudo-rotation approach. Although the pseudo-rotation implicitly contains the twofold degeneracy in the proline ring geometry, our formulation shows this explicitly. Also, the pseudo-rotation angle formulas require two semiempirical parameters. We can derive all necessary parameters directly from the bond lengths and angles of the proline ring.

The strain responsible for puckering

The reason that proline populates two distinct states must be due to some type of strain. Previous calculations have typically explored these conformations by force-field energy minimizations (Ramachandran et al. 1970; DeTar and Luthra 1977; Summers and Karplus 1990; Némethy et al. 1992). However, such studies do not tell us what factors are due to sterics and geometry and what factors are due to other energies. The question is what interaction in the proline ring gives the energy barrier between the UP and DOWN pucker?

There are three types of interactions: Lennard-Jones interaction, bond angle strains, and torsion barriers. However, we found that in energy calculations with CHARMM, the Lennard-Jones interaction and torsion barriers are much weaker than the bond angle strain (see below). Consequently, we focus here on angle strains.

As most of the flexibility in the proline ring lies in the C^{β}, C^{γ}, and C^{δ} atoms, we focus on the variation of the C^{α}-C^{β}-C^{γ}, C^{β}-C^{γ}-C^{δ}, and C^{γ}-C^{δ}-N angles (Fig. 4A–C). We use our model to investigate how the geometry of the closed five-membered ring restricts these angles. In the set of proline conformations generated in the section above, only the C^{β}-C^{γ}-C^{δ} and C^{γ}-C^{δ}-N angles vary as we had fixed the C^{α}-C^{β}-C^{γ} angle. We thus obtain the curves of C^{γ}-C^{δ}-N versus χ2 (Fig. 4C) and C^{β}-C^{γ}-C^{δ} versus χ2 (Fig. 4B). For the C^{α}-C^{β}-C^{γ} variation, we recalculate the cyclic curves where we fix the C^{γ}-C^{δ}-N angle instead of the C^{α}-C^{β}-C^{γ} angle, and place the apex of the tetrahedral angle at the N atom. We thus obtain the curve of C^{α}-C^{β}-C^{γ} versus χ2 (Fig. 4A).

We first note that in all three model cyclic curves, the two lobes of observed density lie on the curves. The differ-

ences are found in the region of sparse density in the PLANAR region ($\chi2 \sim 0°$). In the curves of C^{β}-C^{γ}-C^{δ} versus $\chi2$ (Fig. 4B) and C^{α}-C^{β}-C^{γ} versus $\chi2$ (Fig. 4A), the model curves follow the observed data, even in the sparse region $\chi2 \sim 0°$. In contrast, the model curve of C^{γ}-C^{δ}-N versus $\chi2$ deviates far below the data points near $\chi2 \sim 0°$ but coincides with the data at the UP and DOWN pucker.

All three bond angles tend toward the tetrahedral bonding angle of 109.5°. The geometry of the closed five-membered ring restricts the C^{γ}-C^{δ}-N bond angle in the PLANAR region but relaxes the C^{γ}-C^{δ}-N bond angle at the UP and DOWN pucker. In contrast, the C^{β}-C^{γ}-C^{δ} and C^{α}-C^{β}-C^{γ} angles in both the model curve and the data are quite relaxed in the PLANAR region, as the bond angles are close to the tetrahedral bonding angle. Thus, the ring geometry mostly restricts the C^{γ}-C^{δ}-N bond angle in the PLANAR region. We conclude that the C^{γ}-C^{δ}-N bond angle strain at $\chi2 \sim 0°$ is the main energetic barrier between the UP and DOWN pucker.

Planar conformations of the proline ring

We now analyze the PLANAR conformations of the proline ring. The models presented above fix three of the five bond angles, leaving two angles variable. However, the data show that there are three mobile atoms, the C^{β}, C^{γ}, and C^{δ} atoms (Fig. 1B). All three angles that are centered on these three atoms should be variable. To model this, we generate the family of curves giving the variation of C^{β}-C^{γ}-C^{δ} versus C^{γ}-C^{δ}-N for each value of C^{α}-C^{β}-C^{γ}.

The Bricard equation only has solutions for the range $33° < C^{\alpha}$-C^{β}-$C^{\gamma} < 111°$ (Fig. 5A). For C^{α}-C^{β}-C^{γ} angles approaching the lower limit at C^{α}-C^{β}-$C^{\gamma} = 33°$, the curves get larger. These conformations have completely unrealistic bond angles. The limit at C^{α}-C^{β}-$C^{\gamma} = 111°$ corresponds to a completely flat proline ring. For values of C^{α}-C^{β}-C^{γ} angles approaching the limit at 111°, the curves get smaller, which in the graph of $\chi1$ versus φ (Fig. 5A, cf. with 2A) approaches the point $\varphi = -60°$ and $\chi1 = 0°$. We plot the family of curves for C^{β}-C^{γ}-C^{δ} versus $\chi2$ (Fig. 5B, cf. with 4B) and C^{γ}-C^{δ}-N versus $\chi2$ (Fig. 5C, cf. with 4C). The curve with the smallest cycle has the highest values of C^{γ}-C^{δ}-N. Thus, as the proline ring approaches the region $\chi2 \sim 0°$, the proline ring strains the C^{α}-C^{β}-C^{γ} angle to the maximum in order to relax the C^{γ}-C^{δ}-N steric strain. This flattens the proline ring, which in the $\chi1$ versus φ plot (Fig. 5A), pushes the conformation toward $\varphi = -60°$ and $\chi1 = 0°$. This explains the pattern of sparse density between the two lobes in the PLANAR conformation (Fig. 2A).

Energy calculations of the barrier between the puckers

As we have identified the C^{γ}-C^{δ}-N bond angle strain as the crucial interaction in the closed five-membered ring, we calculate the energy of the ring using a standard C^{γ}-C^{δ}-N bond angle potential. We use $E = 70 * [(108.5 - C^{\gamma}\text{-}C^{\delta}\text{-N}) \pi/180]^2$ kcal/mol (CHARMM22 parameters) (Mackerell et al. 1998). Using the set of allowed ring conformations generated above, we calculate the energy of this term. We also looked at the contribution of Lennard-Jones potentials and the C-C torsion barriers in the proline ring. However these energy terms resulted in energy variations of <0.5 kcal/mol, which are insignificant, compared with the energy variation due to the bond-angle potential of ~3 kcal/mol. The Lennnard-Jones potential contribution is weak because the distance variations are tiny (~0.02 Å). The C-C torsion barrier in proline is $E = 0.16 * [1 + \cos(3 * \chi2)]$ kcal/mol, where the 0.16 kcal/mol energy constant is much weaker than is the bond-angle energy-constant of 70 kcal/mol. As it is beyond the scope of this study to derive force-field parameters, we take the CHARMM parameters as is and focus on the dominant bond-angle potential.

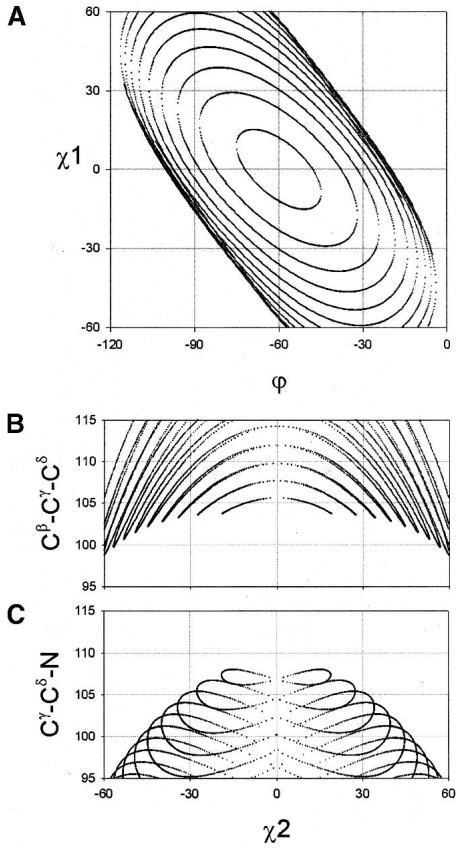

Figure 5. Family of curves for different values of C^{α}-C^{β}-C^{γ} in plot of $\chi1$ vs. φ (*A*), C^{β}-C^{γ}-C^{δ} vs. $\chi2$ (*B*), and C^{γ}-C^{δ}-N vs. $\chi2$ (*C*). The parameterization of each curve is set for C^{α}-C^{β}-C^{γ} in the range $50° < C^{\alpha}$-C^{β}-$C^{\gamma} < 110°$ in 5° steps (*A*) and in 2° steps (*B*,*C*). The curves with large cycles correspond to small C^{α}-C^{β}-C^{γ} angles, whereas the curves with small cycles correspond to large C^{α}-C^{β}-C^{γ} angles. As C^{α}-C^{β}-C^{γ} approaches 111°, the curves in *A* collapse to the point $\chi1 = 0°$, $\varphi = -60°$, in *B* to $\chi2 = 0°$, C^{β}-C^{γ}-$C^{\delta} \sim 104°$, and in *C* to $\chi2 = 0°$, C^{γ}-C^{δ}-N ~ 109°. At this limit, the proline ring is planar.

In Figure 4D, we plot energy versus $\chi 2$ (gray), where the minimum of the energy curve coincides with the observed peaks in the $\chi 2$. We also plotted the energy versus φ for the UP and DOWN puckers (Fig. 2C). The minimum energy at $\varphi \sim -80°$ for the UP pucker is reasonably close to the observed peak for the UP pucker. However, the minimum energy at $\varphi \sim -40°$ for the DOWN pucker is not at the observed distribution of the DOWN pucker. This discrepancy is due to the O_{i-1}-O steric clash in the backbone (Ho et al. 2003) that disfavors values of $\varphi > -50°$ and thus pushes the minimum of the DOWN pucker toward $\varphi \sim -60°$.

Discussion

We have shown that the conformations of the pyrrolidine ring of proline can be understood in terms of the geometry of a five-membered ring. This geometry can be reduced to that of a flexible tetrahedral angle, which we solve using the Bricard equation. Using this equation, we present an algorithm for generating proline conformations from a protein backbone. This algorithm can be easily generalized to other five-membered ring systems.

The Bricard equation for proline gives an analytical relationship between the backbone φ and $\chi 1$ torsions of the proline ring. This relationship captures the coupling of the backbone to the proline ring. For a given backbone conformation, the algorithm generates two symmetric conformations of the proline ring, which correspond to the UP and DOWN puckers. Adding only a C^{γ}-C^{δ}-N bond angle energy term (from CHARMM) is sufficient to explain the barrier between the UP and DOWN puckers. This is an exact algorithm, which only requires knowledge of the bond lengths and angles, and does not require any energy minimization. We show that the method describes well the conformations of proline that are observed in a high-resolution data set from the PDB.

Acknowledgments

We thank Nick Ulyanov for helpful suggestions.

References

Altona, C. and Sundaralingam, M. 1972. Conformational analysis of the sugar ring in nucleosides and nucleotides. A new description using the concept of pseudorotation. *J. Am. Chem. Soc.* **94:** 8205–8212.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. 2000. The Protein Data Bank. *Nucleic Acids Res.* **28:** 235–242.

Bhattacharyya, R. and Chakrabarti, P. 2003. Stereospecific interactions of proline residues in protein structures and complexes. *J. Mol. Biol.* **311:** 925–940.

Bricard, R.J. 1897. Mémoire sur la théorie de l'octaèdre articulé. *J. Math. Pures. Appl.* **3:** 113–150.

Chakrabarti, P. and Pal, D. 2001. The interrelationships of side-chain and main-chain conformations in proteins. *Prog. Biophys. Mol. Biol.* **76:** 1–102.

Coutsias, E.A., Seok, C., Jacobson, M.P., and Dill, K.A. 2004. A kinematic view of loop closure. *J. Comput. Chem.* **25:** 510–528.

DeTar, D.F. and Luthra, N.P. 1977. Conformations of proline. *J. Am. Chem. Soc.* **99:** 1232–1244.

EU 3-D Validation Network. 1998. Who checks the checkers? Four validation tools applied to eight atomic resolution structures. *J. Mol. Biol.* **276:** 417–436.

Ho, B.K., Thomas, A., and Brasseur, R. 2003. Revisiting the Ramachandran plot: Hard-sphere repulsion, electrostatics, and H-bonding in the α-helix. *Protein Sci.* **12:** 2508–2522.

Lovell, S.C., Davis, I.W., Arendall III, W.B., de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. 2003. Structure validation by Cα geometry: ϕ,ψ and Cβ deviation. *Proteins* **50:** 437–450.

MacArthur, M.W. and Thornton, J.M. 1991. Influence of proline residues on protein conformation. *J. Mol. Biol.* **218:** 397–412.

Mackerell Jr., A.D., Bashford, D., Bellott, M., Dunbrack Jr., R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem.* **B102:** 3586–3616.

Milner-White, E.J. 1990. Situations of γ-turns in proteins: Their relation to α-helices, β-sheets and ligand binding sites. *J. Mol. Biol.* **216:** 386–397.

Némethy, G., Gibson, K.D., Palmer, K.A., Yoon, C.N., Paterlini, G., Zagari, A., Rumsey, S., and Scheraga, H.A. 1992. Energy parameters in polypeptides, 10: Improved geometrical parameters and nonbonded interactions for use in the ECEPP/3 algorithm, with application to proline-containing peptides. *J. Phys. Chem.* **96:** 6472–6484.

Pal, D. and Chakrabarti, P. 1999. Cis peptide bonds in proteins: Residues involved, their conformations, interactions and locations. *J. Mol. Biol.* **294:** 271–288.

Ramachandran, G.N., Lakshminarayanan, A.V., Balasubramanian, R., and Tegoni, G. 1970. Studies on the conformation of amino acids, XII: Energy calculations on propyl residue. *Biochim. Biophys. Acta* **221:** 165–181.

Schimmel, P.R. and Flory, P.J. 1968. Conformational energies and configurational statistics of copolypeptides containing L-proline. *J. Mol. Biol.* **34:** 105–120.

Summers, L.N. and Karplus, M. 1990. Modeling of globular proteins: A distance-based data search procedure for the construction of insertion/deletion regions and Pro-non-Pro mutations. *J. Mol. Biol.* **216:** 991–1016.