# On the evolutionary significance of the size and planarity of the proline ring

**ARTICLE** *in* THE SCIENCE OF NATURE · SEPTEMBER 2012

Impact Factor: 2.1 · DOI: 10.1007/s00114-012-0960-y · Source: PubMed



**4 AUTHORS**, INCLUDING:

Jörn Arnold Behre
Institute of Food Research
**8** PUBLICATIONS **270** CITATIONS
SEE PROFILE

Ingo Althoefer
Friedrich Schiller University Jena
**44** PUBLICATIONS **612** CITATIONS
SEE PROFILE

Stefan Schuster
Friedrich Schiller University Jena
**221** PUBLICATIONS **8,449** CITATIONS
SEE PROFILE

Naturwissenschaften (2012) 99:789–799
DOI 10.1007/s00114-012-0960-y

ORIGINAL PAPER

# On the evolutionary significance of the size and planarity of the proline ring

**Jörn Behre · Roland Voigt · Ingo Althöfer · Stefan Schuster**

Received: 3 May 2012 / Revised: 4 August 2012 / Accepted: 8 August 2012 / Published online: 15 September 2012


**Abstract** Proline is a proteinogenic amino acid in which the side chain forms a ring, the pyrrolidine ring. This is a five-membered ring made up of four carbons and one nitrogen. Here, we study the evolutionary significance of this ring size. It is shown that the size of the pyrrolidine ring has the advantage of being nearly planar and strain-free, based on a general mathematical assertion saying that the angular sum of a polygon is maximum if it is planar and convex. We also provide a sketch of the proof to this assertion. The optimality of the ring size of proline can be derived from a triangle inequality for angles. Quasi-planarity is physiologically significant because it allows an easier and evolutionarily old type of fit into binding grooves of proteins with which proline-rich proteins interact. Finally, we present a comparison with other planar, nearly planar and non-planar biomolecules such as neurotransmitters, hormones and toxins, involving, for example, aromatic rings, cyclopentanone and 1,3-dioxole.

**Keywords** Imino acids · Optimal molecular structure · Planar side chains · Polyproline region · Proline side chain · Ring-shaped amino acids

Communicated by: Sven Thatje

**Electronic supplementary material** The online version of this article (doi:10.1007/s00114-012-0960-y) contains supplementary material, which is available to authorised users.

J. Behre · S. Schuster (✉)
Department of Bioinformatics, Friedrich Schiller University,
Ernst-Abbe-Platz 2,
07743 Jena, Germany
e-mail: stefan.schu@uni-jena.de

I. Althöfer
Department of Mathematics, Friedrich Schiller University,
Ernst-Abbe-Platz 2,
07743 Jena, Germany

R. Voigt
Institute of Mathematics, University of Leipzig,
Johannisgasse 26,
04103 Leipzig, Germany

*Present Address:*
J. Behre
Institute of Food Research,
Colney Lane,
Norwich NR4 7UA, UK

## Introduction

In biological evolution, from an enormous multitude of possibilities, a much smaller (yet in many cases still very large) set of solutions is selected. For example, from all aliphatic amino acids possible, only a few are used in proteins (cf. Grützmann et al. 2011) and from the enormous multitude of conceivable ring-shaped molecules, only a limited fraction occurs in living cells (cf. Koch et al. 2005). In line with Darwinian evolutionary theory, this selection process can be considered as a process of optimisation. Accordingly, many properties of living organisms have been explained, modelled or simulated on the basis of optimality principles. Such principles are extremely useful tools in all fields of biology. For example, properties of metabolic pathways and networks have been analysed from that perspective (cf. Heinrich and Schuster 1996; Banga 2008). Various optimality criteria have been suggested to be relevant for metabolic systems, such as maximisation of pathway fluxes (Kacser and Beeby 1984; Werner et al. 2010) or minimisation of the conversion time from substrate to product (Klipp et al. 2002; Bartl et al. 2010). Examples from other fields of biology are, to mention a few, optimal hematocrit theory (Withers et al. 1991) or the calculation of the cruising velocity of migratory birds (Pennycuick 1997). Also the chemical structure of biologically important

molecules has been studied from an optimality perspective, e.g. the structure of glycogen (Meléndez et al. 1999). It has been argued that the structure of the α-helix is optimal with respect to space-filling properties and avoidance of steric hindrance among side chains (Jean 1994).

The efficient identification of small organic molecules that modulate protein function in vitro and in vivo is crucial in the development of new therapeutic drugs (cf. Koch et al. 2005). Also in that context, optimality considerations are of importance. The chemical space that is relevant to biology has been selected by biological evolution (Koch et al. 2005). Here, we will analyse the structure of one proteinogenic amino acid, proline and consider also other ring-shaped biomolecules. While proline is not a drug, it is here used as an example molecule to demonstrate our computational method. At the end, we will briefly sketch the applicability of the method to neurotransmitters, hormones and drugs.

Proline is the only proteinogenic amino acid in which the side chain is alicyclic. This means that it forms a non-aromatic ring (cf. Lehninger et al. 2000; Berg et al. 2007). Since the nitrogen atom forms bonds with two carbons, proline is, strictly speaking, an imino acid rather than an amino acid. Nevertheless, we will here use the term amino acid both for proline and its analogues with other ring sizes.

The ring structure of proline implies consequences for the properties of those proteins in which it occurs. While both the amino and carboxyl groups of the other amino acids can form hydrogen bonds, the imino group in proline cannot do so when it is integrated in proteins. Therefore, proline does not support α-helices and is, hence, called a helix breaker.

The skeleton of the side-chain ring of proline (called pyrrolidine ring) involves five atoms: four carbons (one of them being the α-carbon atom) and one nitrogen atom and is, therefore, an azacycloalkane (the prefix 'aza' denotes nitrogen, from Greek '*azotikos*' meaning 'without life'). Under the plausible hypothesis that biological evolution has been able to choose from a large number of different options, the question arises why just this structure and size of the ring was chosen rather than, for example, a six-membered ring. Such a six-ring arises relatively easily, for example, by deamination of lysine and ring formation. The corresponding amino acid, piperidine-2-carboxylic acid, called pipecolic acid by trivial name, is indeed a metabolite in the degradation of lysine, for example, in several bacteria (Payton and Chang 1982) and in the mammalian brain (Inoue et al. 2011).

Most ring-shaped biomolecules such as sugars, aromatic amino acids and many hormones involve five-membered or six-membered rings (Koch et al. 2005). However, there are also biomolecules involving larger rings such as the cytochalasins, which are fungal metabolites and usually involve 11- or 14-membered rings (Haidle and Myers 2004). Other examples are provided by the macrolides, which are usually 14-, 15- or 16-membered (Spector et al. 1989). Some macrolides contain 44-membered rings, such as swinholide J produced by the marine sponge *Theonella swinhoei* against grazing by fish (De Marino et al. 2011).

Here, we will analyse the optimal geometric properties of the structure of proline. Non-aromatic ring-shaped molecules usually show the phenomenon of puckering, that is, non-planar conformations (Barnett and Capitani 2006; Kang et al. 2009). Some rings are, however, quasi-planar. The impact of quasi-planarity on binding properties has recently been studied for the case of bile salts (Reschly et al. 2008; Krasowski et al. 2011). We will examine the puckering of azacycloalkanes of various sizes. We do so by considering the angular sums of polygons and of azacycloalkanes. We will interpret the results with respect to the corresponding amino acids and the impact on protein–protein interactions.

## Materials and methods

### Bond angles

We investigate the structural properties of proline mainly by considering the bond angles within its ring. The bond angle between two out of four single bonds at a carbon atom (so-called tetrahedron angle) amounts to 109.5° (cf. Holleman and Wiberg 2001). This angle may somewhat deviate due to mechanical strain imposed by the structure of the whole molecule. In chemistry, this is known as the Baeyer strain (cf. Beyer and Walter 1996). In contrast to carbon (valence four), nitrogen normally has valence three. We should distinguish two cases:

1. When a nitrogen atom is involved in three single bonds, it carries four $sp^3$-hybrid orbitals. Thus, these three orbitals together with the free electron pair approximately form a tetrahedron, as exemplified by ammonia (Fig. 1) (cf. Holleman and Wiberg 2001). Since the orbital with the free electron pair remains closer to the N atom, the tetrahedron geometry is somewhat distorted and the bond angles are about 106.8°, which is still relatively close to the ideal tetrahedron angle.
2. When the N atom is involved in a double bond, it carries one $p$ orbital and three $sp^2$-hybrid orbitals, where the latter form a planar orientation with bond angles of 120° (again with some potential minor variations). Whereas the electron conformation is trigonal–bipyramidal, the resulting molecular shape is trigonal (Fig. 2).

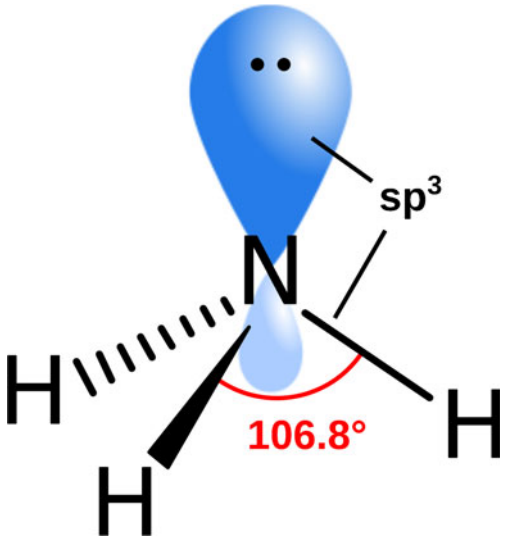


**Fig. 1** 3D structure of ammonia as an example of an N atom with a tetrahedron-shaped bond structure. The three bonds formed by $sp^3$ orbitals overlapping with the 1*s* orbitals of the hydrogens together with the fourth $sp^3$ orbital containing the free electron pair form a tetrahedron. Due to the volume of the latter orbital, the tetrahedron is slightly distorted. Thus, the ideal tetrahedron angle of 109.5° is reduced to 106.8° (Holleman and Wiberg 2001)

The hybridisation of the nitrogen changes when free proline is included in peptide chains. Since the peptide bond in proteins has a partial double-bonded character, the nitrogen is no longer $sp^3$ but $sp^2$ hybridised. In the limiting resonance structure where the N atom is positively charged, no orbital is needed anymore to carry the free electron pair. The N atom then has effectively valence four.

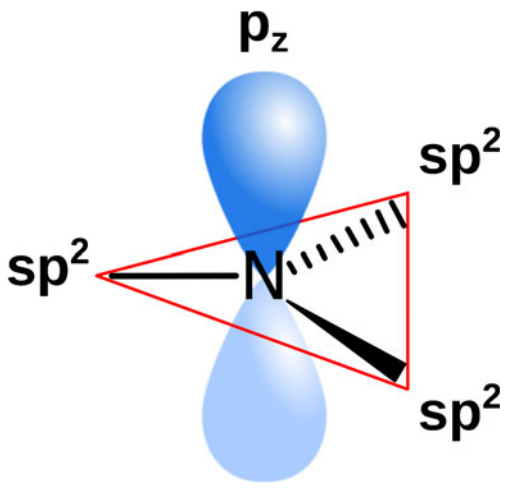


**Fig. 2** $sp^2$-hybridised N atom. The three $sp^2$ orbitals form a planar orientation with bond angles of 120°. The remaining $p_z$ orbital is perpendicular to that plane

## Angular sums and planarity

The sum of the ideal bond angles within the pyrrolidine ring for the two above-mentioned cases is:

1. $106.8° + 4 \times 109.5° = 544.8°$

2. $120° + 4 \times 109.5° = 558°$

Note that the real bond angles can differ from the ideal ones due to distortion by ring formation. The sum of the real bond angles in a ring-shaped molecule is sometimes called endocyclic angle sum (Fritz et al. 1994).

The angle sum in a flat convex pentagon amounts to 540°. The general expression for a planar polygon with $n$ vertices reads $(n-2)\times180°$. This well-established result from geometry can be shown, for example, by the method of induction (see Fig. 3). A triangle has an angular sum of 180°. Note that a digon, that is, a straight line, shows an angular sum of zero, also in agreement with the above expression. Assume that the expression is valid for $n$-gons with a given $n$. Now we introduce an additional vertex $Z$ and, thus, attach a triangle $AZW$ (Fig. 3). In a planar polygon, the angles at any vertex are additive, e.g. $\beta=\beta_1+\beta_2$ at vertex $B$ and similar equations at vertices $A$ and $W$. Due to the induction hypothesis and the angular sum of triangle $AZW$, the angular sum in the $(n+1)$-gon $AB\ldots WZ$ is $(n-1)\times180°$, which completes the proof.

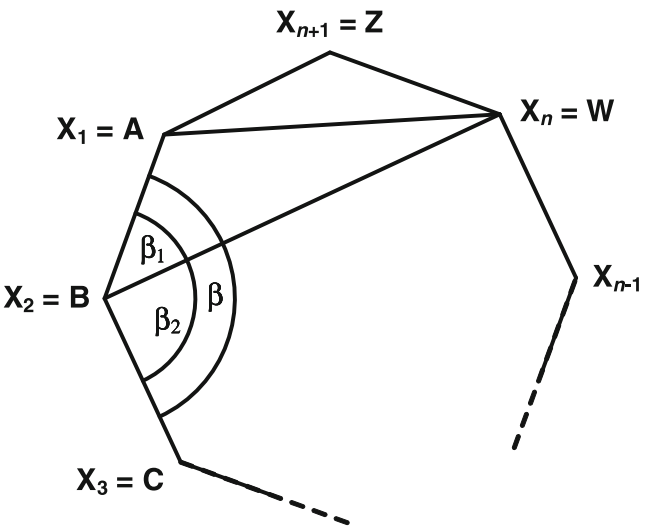


**Fig. 3** Angular sums in planar and skew polygons. The vertices are denoted by $X_1, X_2, \ldots, X_n$ with the order of vertices given by the indices and $X_n$ being connected to $X_1$ again. For simplicity's sake, we write A, B and W for $X_1$, $X_2$ and $X_n$, respectively. For explanations on skew polygons, see 'Electronic supplementary material'

It is important to note that there are two angles at each vertex, one that is larger than 180° and another that is less than 180° (a special case is that both are equal to 180°). Since all bond angles are always less than 180°, we only consider the smaller angle throughout this paper. Thus, we also avoid the problem of distinguishing between interior and exterior angles, which is difficult for non-convex or non-planar polygons. Note that in a concave planar polygon, an exterior angle can be chemically relevant (Fig. 4).

The question arises what the angular sum can be in a non-planar pentagon, alternatively called skew pentagon (Williams 1979). If, for example, only four atoms would be situated in a common plane while the fifth is not, the angular sum would be smaller than in the planar case (Fig. 4).

## Results

### A theorem about the angle sum and planarity

Central to our analysis is the following theorem:

> **Theorem:** *For any given natural number n (n≥2), the sum of the smaller angles in any n-gon in d-dimensional Euclidean space is at most (n−2)×180°. Equality holds if and only if all n vertices lie in a 2-dimensional plane within the d-dimensional space and this planar n-gon is convex. (In this paper, we are only interested in the case where dimension d=3).*

The proof is given in the 'Electronic Supplementary Material'. It works by a triangle inequality for angles and the method of induction.

The optimisation criterion, by which we analyse the structure of proline and its analogues, can be written as the minimisation of the modulus (absolute value) of the difference of the two angle sums in a polygon:

$$\text{Minimise}\left|\text{Sum of ideal bond angles} - (n-2)\times 180^\circ\right|. \quad (1)$$

The second sum refers to the sum of interior angles in convex flat polygons. The idea behind this formula is that the lower the difference between these sums is, the smaller is the deviation from a planar shape. If the difference is zero, the ring-shaped molecule is perfectly planar.

An important condition in the theorem given here is the convexity of the planar polygon. In the case of skew polygons, it is difficult to distinguish between convexity and concavity. One possible definition is by considering the projection into a two-dimensional plane, which is usually done in the graphical representation of the structural formula. For most ring-shaped biomolecules, that representation is convex. This is because their ring size is not large enough to allow a folding of some vertex to the inside of the ring. However, as mentioned in the 'Introduction', there are biomolecules involving larger rings. For simplicity's sake, however, we here restrict our analysis to polygons the projection of which into a 2D plane is convex. In the 'Discussion' section, we will allude to concave molecules.

### Comparing different ring sizes

Biochemical evolution could have led to incorporation of other non-aromatic rings into amino acids. The alicyclic (ring-shaped aliphatic) hydrocarbons with $n-1$ carbon atoms for $n=3, 4, \ldots, 8$ and one nitrogen are shown in Fig. 5. Besides the names given in Fig. 5, there are even other names for the azacycloalkanes, for example dimethylene-imine, trimethylene-imine, tetramethylene-imine, etc. These names are understandable as well, because these compounds involve an imino group and $n-1$ methylene groups ($CH_2$). Many of these substances have trivial names in addition. For example, the name piperidine comes from the fact that this substance was isolated first from pepper (*Piper nigrum*).

In the following, we compare planar polygons with azacycloalkanes containing $sp^2$-hybridised N atoms (as it is the case with proline when it is included in proteins).

*n=3* In azacyclopropane, the angular sum is 180°, while the sum of ideal bond angles in a three-membered ring (with $sp^2$-hybridised N atom) would be $120^\circ+2\times109.5^\circ=339^\circ$, thus considerably more than 180°. Since three points always

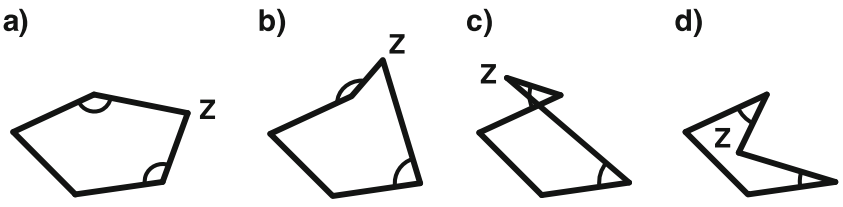


**Fig. 4** Diminution of angular sum in a pentagon upon folding. **a** In a flat convex pentagon, the angular sum is 540°. **b** Two edges (connecting to vertex *Z*) are bent out of the plane. This implies that the two marked angles decrease. **c** The two edges are bent even further, so that the two marked angles decrease further. **d** If vertex *Z* arrives in the plane again, the pentagon is concave. Now the external angle at vertex *Z* is relevant for chemical bonds. Thus, it can easily be seen that the sum of angles is less than 540° in this case

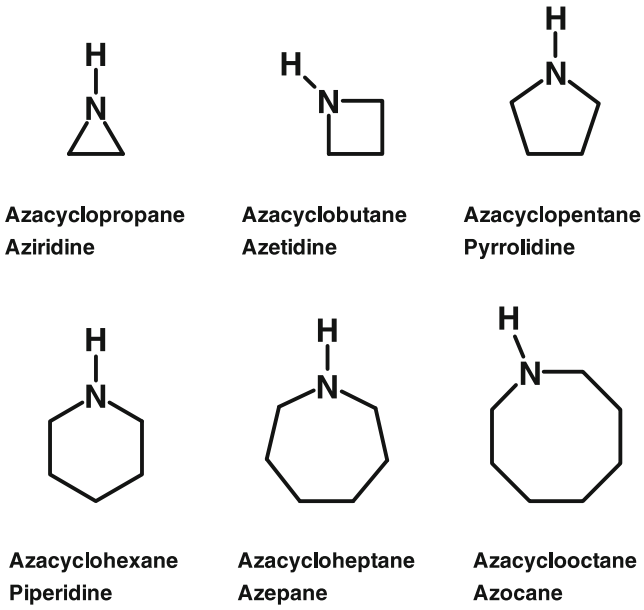


**Fig. 5** Azacycloalkanes with three-membered up to eight-membered rings. Systematic names are given below the chemical structure, followed by trivial names

define a plane, the molecule cannot fold. In azacyclopropane, the bond angles must be considerably smaller than the values given above, which causes strong tensions in the molecule (Baeyer strain). This strain implies that azacyclopropane is very reactive because there is a strong tendency to decrease the ring tension by breaking bonds. In a biochemical context, this implies strong toxicity. To our knowledge, the corresponding amino acid, aziridine-2-carboxylic acid, does not occur in living organisms. However, it has been proposed to be used as an inhibitor of cysteine proteases (Moroder et al. 1992) and as a potential drug against plasmodial pathogens (Schulz et al. 2007).

*n=4* From *n*=4 on, polygons can fold in space, forming skew (non-planar) polygons. In azacyclobutane, the angular sum is 360°, while the sum of the bond angles in a four-membered ring without any strain would be 120°+3×109.5°=448.5°. However, although any folding of the molecule decreases the angular sum and, thus, increases its ring strain, it is slightly folded to reduce the ecliptic conformation of the hydrogens in the ring and thus their van der Waals repulsion (Pitzer strain) (cf. Beyer and Walter 1996). This implies a higher reactivity than in a five-membered ring. Therefore, it is plausible that neither azacyclopropane nor azacyclobutane are used in usual proteinogenic amino acids. Moreover, it is much less likely than in the case of proline that they were formed non-enzymatically in prebiotic evolution. However, the amino acid involving azacyclobutane, azetidine-2-carboxylic acid, does occur in some organisms. It is produced by several plants belonging to the *Liliaceae* (Fowden 1956). Azetidine-2-carboxylic acid is a toxin against grazing by insects and mammals because, upon uptake by an animal, it becomes incorporated into proteins in place of proline (Fowden 1956). The different bond angles compared to proline cause changes in the protein backbone, which results in malfunction. Azetidine-2-carboxylic acid is also produced by *Streptomyces* species and is considered to be an antibiotic (Gross et al. 2008).

*n=5* As shown above, the sum of ideal bond angles in azacyclopentane (pyrrolidine) with an $sp^2$-hybridised N atom is 558°. Although this is still larger than the angular sum of the corresponding planar polygon (540°), the difference is that small that the molecule is nearly planar and strain-free. In proteins, the proline rings exist predominantly in two discrete states, called pucker UP and pucker DOWN (Milner-White et al. 1992; Ho et al. 2005; Song and Kang 2006). Pucker UP (alternatively called $C^{\gamma}$-exo state) refers to the conformation where the $C^{\gamma}$ atom and the carbonyl group are located on opposite sides relative to the plane that is formed by the three atoms $C^{\alpha}$, $C^{\delta}$ and N of the proline ring (Fig. 6; Chakrabarti and Pal 2001; Song and Kang 2006). In the pucker-DOWN conformation ($C^{\gamma}$-endo state), they are on the same side. A further possibility to define these two different conformations is given by the sign of the dihedral angles $\chi_1$ to $\chi_4$ (Ho et al. 2005).

The pucker-DOWN and pucker-UP conformations have also been studied for hydroxyproline (Song and Kang 2006) and pseudoproline residues, in which the $C^{\gamma}H_2$ group of the prolyl ring has been replaced by oxygen or sulphur atoms (Krasowski et al. 2011).

*n>5* The azacycloalkanes with more than five atoms in the ring are strain-free. However, for fulfilling the constraint on the angular sum, they have to be twisted considerably. For example, piperidine-2-carboxylic acid (pipecolic acid) prefers the chair conformation (Budesinsky et al. 2010), which also in cyclohexane has the lowest strain energy of all

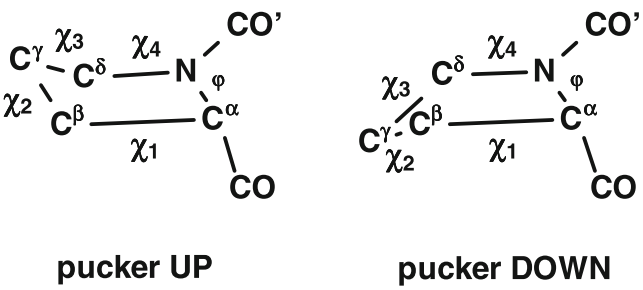


**Fig. 6** The two different conformations, pucker UP and pucker DOWN, of the proline ring. *CO* and *CO'* stand for the carbonyl groups of the proline molecule and of the next amino acid, respectively. $\chi_i$, dihedral angles in the ring

possible conformers. Pipecolic acid occurs as an intermediate in the degradation of lysine (Payton and Chang 1982; Inoue et al. 2011), but not within proteins. While azepane-2-carboxylic acid ($n$=7) is not, to our knowledge, synthesised in living organisms, several of its derivatives are used as antibiotics (Wishka et al. 2011). Eight-membered rings showing some similarity to azocane are formed in some proteins by two adjacent cysteine residues that are linked by disulfide bridges (Creighton et al. 2004). In general, rings with $n$>6 are rare in biology (Koch et al. 2005).

In Fig. 7, the angular sums in planar polygons and the sum of bond angles in azacycloalkanes with a double bond at the N atom (like in proline) are plotted in their dependency on ring size. The intersection point can be calculated by equating the expressions for the two types of angle sums:

$$120^\circ + (x-1) \times 109.5^\circ = (x-2) \times 180^\circ. \quad (2)$$

This gives the value

$$x_{\mathrm{opt}} = 5.255. \quad (3)$$

Only integer values are allowed for the number of atoms. The optimal ring size is given by the integer nearest to $x_{\mathrm{opt}}$:

$$n_{\mathrm{opt}} = 5. \quad (4)$$

This value can be interpreted as the solution to the optimisation principle (1). It is indicative of a special role of pyrrolidine. The sum of ideal bond angles of 558° in the ring of proline is indeed close to the angular sum of 540° in a planar pentagon. Interestingly, in free proline (case 1), which shows an angular sum of 544.8°, the agreement is even better and, thus, free proline is even more planar (Ho et al. 2005). A more planar conformation can also occur within a protein by interaction with amino acids of other proteins (Donnini et al. 2006).

## Functional implications of a planar ring

What physiological advantages result from the quasi-planarity of the ring in proline? It can be hypothesised that this facilitates the binding of proteins involving proline to other proteins. In fact, it is known that polyproline regions within proteins are often responsible for protein–protein interactions (Yu et al. 1994). For example, these regions bind to special binding grooves of SH3 domains, which are indeed relatively narrow (Aitio et al. 2008, 2010). The advantage of a quasi-planar molecule is that it fits more easily in a binding groove than a twisted one (Fig. 8). An interesting analogy is provided by mechanical keys: the blade of many keys is flat because, thus, it slides more easily into the keyway of the lock. A twisted residue, in contrast, would be more difficult to move into a groove because it would require a rotational movement in addition to the translational shift. This often implies an induced fit, that is, a conformational change of the binding groove upon insertion of the ligand, so that the spatial fit is improving.

Although the idea of easy fit of planar ligands is intuitive, biological reality is more complex. Mukherjee et al. (2005) analysed the conformation of the amino group in the side chains of several amino acids. While most earlier studies dealing with hydrogen bonds in protein–DNA complexes model the amino group hydrogen atoms according to the ideal planar configuration, non-planar amino groups improve the geometry of hydrogen bonds (Mukherjee et al. 2005).

Krasowski and coworkers (Reschly et al. 2008; Krasowski et al. 2011) have recently studied the role of quasi-planarity in binding to receptors for the case of bile salts. Bile salts consist of several connected rings. The junction between rings determines whether bile salts have a flat or a bent configuration. Note that for those molecules,

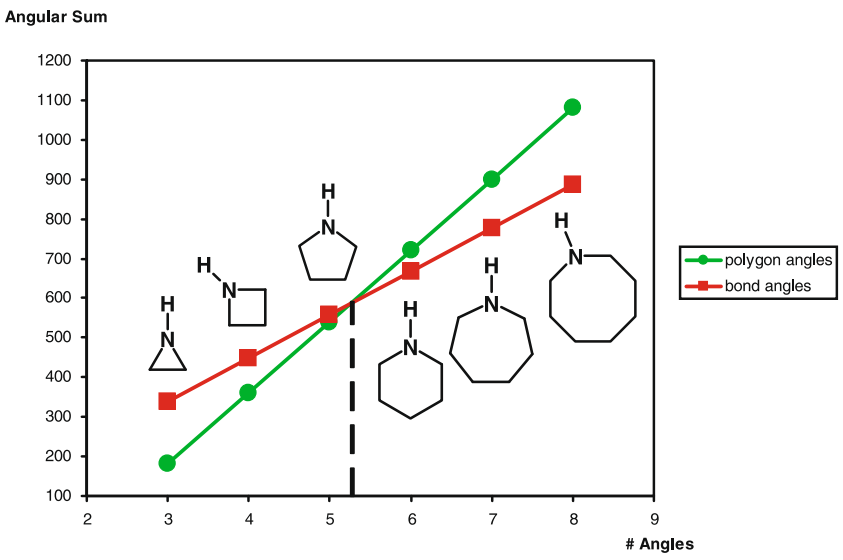


**Fig. 7** Plot of the sum of ideal bond angles in azacycloalkanes with $sp^2$-hybridised N atoms (*squares*) and angular sum of planar polygons (*full circles*) vs. number of angles. The intersection point is situated at an abscissa value slightly above five ($x_{\mathrm{opt}}$=5.255)

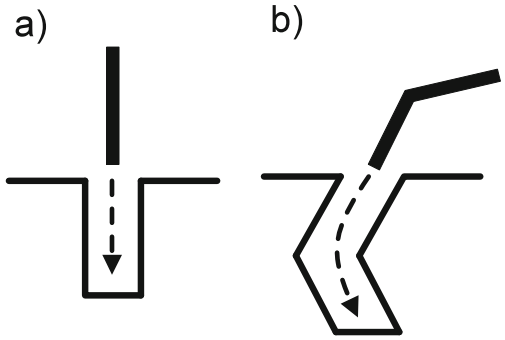


**Fig. 8** Scheme showing two different ways of binding of a ligand to a binding groove. **a** Binding of a planar ligand into a cuboid groove, only requiring a straight movement. **b** Binding of a non-planar ligand into a deep, tilted groove, requiring a torsional movement. This often implies an induced fit

planarity is studied on a larger spatial scale than for proline; the overall configuration can be considered relatively planar even in the case of a zig-zag structure due to non-planarity of the individual rings. An illustrative example is cholesterol, from which the bile acids and salts are derived. It was proposed that quasi-planar bile salts such as 5α-cyprinol-27-sulphate are the likely ancestral ligands while bent bile acids such as chenodeoxycholic acid are evolutionarily younger and require more sophisticated receptor structures for binding (Reschly et al. 2008; Krasowski et al. 2011).

Since amino acids had occurred in prebiotic evolution and also protein–protein interactions are likely to be ancient, it can be assumed that ring-shaped side chains correspond to the ancient type of binding. This is supported by the fact that all the ring-shaped side chains occurring in aromatic amino acids (phenylalanine, tyrosine, tryptophane and histidine) are planar, since aromatic rings have this property. Another advantage of that property of aromatic amino acids is the propensity to form stacking interactions. Also the guanidine group in arginine is planar because the carbon in the centre of that group is involved in one double bond. Thus, the three bond angles are about 120°, so that the four atoms involved in the guanidine group lie in one plane.

As mentioned above, the rings of aziridine and azetidine are planar as well but show intramolecular tensions (Baeyer strain). Larger rings are no longer planar. Besides the more complex binding properties, a further aspect is the synthesis from chain-shaped precursor molecules. In the case of non-planar molecules, it is more difficult to bring the two ends together in the process of ring closure.

### Other ring-shaped biomolecules

Interestingly, there are other molecules involving five-membered rings with the same or similar sums of bond angles as in pyrrolidine. An example is provided by cyclopentanone, that is, cyclopentane with one of the five carbons forming a keto group. This ring is involved in jasmonic acid, a signalling substance in higher plants (Bodnaryk and Yoshihara 1995). The sum of bond angles is 558°. Thus, it shows the same puckering conformations as pyrrolidine. As jasmonic acid must bind to receptor molecules to perform its signalling function, it can be assumed that quasi-planarity is favourable.

Another interesting ring is involved in 1,3-benzodioxole (also known as 1,2-methylenedioxybenzene) (Fig. 9). The benzene part in this double ring is obviously planar. To check the planarity of the 1,3-dioxole ring (involving the methylenedioxy part), we can again calculate the angular sum. At the two carbons shared by both rings, the bond angle is ideally 120° each. The carbon involved in single bonds would ideally contribute the tetrahedron angle of 109.5°. The bond angles at oxygens involved in single bonds are ideally 104.5° (Holleman and Wiberg 2001). The sum is 558.5°, which is again very close to the angular sum of 540° in planar pentagons. Thus, the 1,3-benzodioxole double ring is nearly planar, in line with molecular dynamics simulations (Asiri et al. 2011). It occurs in many natural secondary metabolites such as safrole, which is produced in a wide variety of plants as a natural pesticide (Kamdem and Gage 1995), berberine, which is produced in many plants such as *Berberis* spp. (Birdsall and Kelly 1997), and protopine, an opioid analgesic produced in *Papaveraceae* (poppy plants) (Jiang et al. 2004).

## Discussion

Analyses in terms of optimality principles trace back to very early times of science and philosophy. Famous philosophers such as Aristotle have dealt with optimality questions. Aristotle made a distinction between four different reasons for the structure of things as we observe them: *Causa finalis*, *causa materialis, causa efficiens*, and *causa formalis* (cf. Falcon 2011). 'Why questions' can be answered in different ways. For example, the question 'Why does proline involve a five-membered ring?' could be answered by saying that it is formed from such and such precursors (Aristotle's *causa materialis*) by such and such enzymes (*causa efficiens*), which are encoded in certain genes (*causa formalis*). Usually, however, 'why questions' in the context of optimality

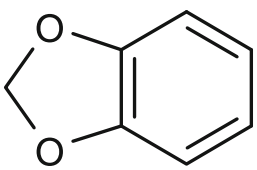


**Fig. 9** Structure of 1,3-benzodioxole. The ring on the left-hand side is called 1,3-dioxole

considerations are asked in the sense of 'to what end?' or 'what is the advantage of…?' (*causa finalis*). This question can be answered in the case of proline by saying that it allows an easy fit into binding grooves of proteins with which the protein containing the proline in question interacts.

It had been proposed that quasi-planar ligands occurred early in evolution (paleomorphic phenotype), while non-planar (bent and twisted) ligands are evolutionarily younger (apomorphic phenotype) (Reschly et al. 2008; Krasowski et al. 2011). It is worth mentioning that an easier binding also implies an easier dissociation of the ligand. Thus, a more intricate spatial fit may imply a tighter binding, even if that is slower. Thus, it depends on the biological function which type of binding is more favourable. For protein–protein interactions, which are usually fast and temporary, an easier fit is likely to be beneficial.

Our main result is the following: We have shown by a relatively simple calculation that the size of the pyrrolidine five-ring involved in proline has, in comparison to other azacycloalkanes, the advantage of being nearly flat and does not imply significant ring strain. Three- and four-membered rings, in contrast, do show ring (Baeyer) strain. Therefore, it is much less likely that they occurred non-enzymatically in prebiotic evolution. Rings larger than five-membered rings, in contrast, require spatially more complex receptor structures for binding. Since the usage of amino acids has started very early in evolution and also protein–protein interactions are likely to be very ancient, it is straightforward to assume that the simpler, paleomorphic mode of binding of ring-shaped side chains enabled by quasi-planarity is relevant for amino acids.

It is clear that proline can only be synthesised if not only the *causa finalis*, but also the other three of Aristotle's *causae* are fulfilled. In humans and many other species, proline biosynthesis starts from glutamate, which involves three carbons in the side chain. Together with the Cα and N atoms, these allow formation of a five-membered ring, yielding pyrroline-5-carboxylate. This is eventually converted into proline (cf. Lehninger et al. 2000). When the smaller precursor aspartate rather than glutamate was used, a four-membered ring would arise. The amino acid harbouring one methylene group more than glutamate is 2-aminoadipate. This is a precursor of lysine. Although glutamate is very abundant in many organisms, aspartate and 2-aminoadipate are available as well. Thus, the material cause is not an obstacle in choosing another ring size during evolution. Also, the efficient and formal causes are unlikely to contradict the occurrence of other ring sizes because azetidine-2-carboxylic acid and pipecolic acid do occur in some organisms, so that evolution could have brought about the necessary enzymes also in other organisms. Consequently, it is most probably the final cause that has determined proline to be chosen in evolution.

Here, we compared amino acids with different ring sizes, not questioning the advantage of the ring. It is an interesting question why a ring-shaped side chain occurred in the first place. Although it is beyond the scope of this paper to answer this question, we here mention some ideas. Proline would have to be compared with, for example, norvaline, which involves an acyclic, unbranched side chain with the same number of carbons (cf. Grützmann et al. 2011). It is clear that such a side chain is mechanically very flexible, as is quantified by a large radius of gyration (Najmanovich et al. 2000). However, collagen (involving much proline) and polyproline regions are stabilised by tight packing of the side chains (cf. Lehninger et al. 2000; Berg et al. 2007). It may be supposed that the secondary structure is stabilised better by the rather stable structure of proline than it would by the flexible side chain of norvaline, in spite of the fact that proline can form only one hydrogen bond rather than two. In fact, hydrogen bonds play a less important role in collagen and polyproline regions than in α-helices.

Our calculation is based on a comparison of the sum of the bond angles and the angular sum of planar convex polygons. In a much less formal way, the sum of the bond angles has been considered earlier by Fritz and coworkers, on studying organosilicon compounds (Fritz et al. 1994). They argued—without formal proof—that a four-membered ring with an angle sum of 344° must be folded. Other authors studied the conformations (puckerings) of cyclic molecules by sophisticated mathematical techniques earlier (Gō and Scheraga 1970; Barnett and Capitani 2006), though without reference to quasi-planarity.

While we have here considered the sum of ideal bond angles and studied the degree of planarity of the ring, a widely used technique is to analyse the planarity of the set of bonds connecting to a nitrogen, based on an argument in terms of bond angles (Schmitt et al. 2011). For example, Song and Kang (2006) considered the sum ($S$) of the three bond angles around the nitrogen of the prolyl ring (i.e., carbonyl C–N–$C^{\gamma}$, $C^{\gamma}$–N–$C^{\alpha}$ and $C^{\alpha}$–N–carbonyl C). If the difference $S-360°$ is about 0, then the imide nitrogen is nearly planar in the $sp^2$ bonding state. As the difference becomes more negative, the degree of planarity (i.e., the pyramidal $sp^3$ character) of the imide nitrogen decreases.

In 'Materials and methods' section, we have compared the pyrrolidine ring with other azacycloalkanes with one $sp^2$-hybridised nitrogen and several carbons. It is also interesting to compare them with other possible structures such as rings with two or three $sp^2$-hybridised nitrogens and several carbons. For a five-membered ring with two or three nitrogens, we would obtain angle sums of $2\times120°+3\times109.5°=568.5°$ and $3\times120°+2\times109.5°=579°$, respectively, which is still relatively close to the angular sum of a planar pentagon.

An important condition in the theorem given here is the convexity of the planar polygon. In future studies, it is worth extending the analysis to concave polygons. Biomolecules involving larger rings such as macrolides could be quasi-planar if they are folded such their projection is concave. The angle sum would then include at least one angle given by 360° minus a bond angle, which gives about 250°. Thus, the angle sum is larger than the sum of bond angles and, thus, could be nearly equal to the angular sum in a planar polygon. However, it can be assumed that the conformation of larger rings can easily and continuously be changed and is, thus, rather unstable. In contrast, five- and six-membered rings usually feature two discrete stable states, which represents a physiological advantage in view of binding into protein pockets.

As for considerations in terms of optimality, it is often not easy to distinguish between cause and effect. Has there been a selection pressure towards planar molecules because of protein–protein interactions, or did the quasi-planar molecule exist first, and only later the protein–protein interactions evolved towards the properties observed today? This is again related to the different causes of Aristotle.

A side remark on teaching: The considerations and calculations on the structure of proline presented here can be very useful in teaching undergraduate courses in biochemistry and molecular biology. Usually, students have to learn by heart the structures of the proteinogenic amino acids, which may be quite uninspiring. As argued earlier (Grützmann et al. 2011), by a formal treatment of amino acids and elucidating certain patterns therein, it should become much easier for students to memorise molecule structures. This is likely to be achieved by teaching them that proline has a quasi-planar ring, how this can be explained, and which functional consequences this has.

Extending the above analysis, it is of great practical interest to study the planarity of ligands such as neurotransmitters, toxins and pharmaceutical drugs. While a thorough analysis is beyond the scope of this paper, it is worth noting that indeed, many of these substances are relatively flat. Prominent examples are provided by derivatives of aromatic amino acids, such as the tyrosine derivatives L -3,4-dihydroxyphenylalanine (L-DOPA), dopamine and adrenaline (epinephrine) (cf. Lehninger et al. 2000; Berg et al. 2007). Well-known tryptophan derivatives are serotonin and melatonin. A further example of a relatively planar substance is aspirin (cf. Lehninger et al. 2000; Berg et al. 2007). Thus, the search for new lead substances in pharmaceutics may be inspired by a special focus on planarity.

Here, we have studied, by way of example, 1,3-benzodioxole. This is important in toxicology and pharmacology because many of the drugs and pharmaceuticals containing this moiety, such as the antidepressant paroxetine and the party drug MDMA, cause metabolic intermediate complex formation with the cytochrome CYP3A4, which is central in the degradation of drugs (Jones et al. 2007). The complex formation, however, inhibits degradation. It is interesting to analyse the role of planarity of the 1,3-benzodioxole moiety in the binding to CYP3A4. In this binding, stacking interactions may play a role.

Many other ligands, however, are non-planar, such as tropane, the core component of cocaine (cf. Vollhardt and Schore 2007). Interestingly, the non-planarity of many toxins such as amanitin and phalloidin, the toxins of the death cap mushroom, *Amanita phalloides* (Benjamin 1995; Hallen et al. 2007), is likely to avoid binding to proteases and, thus, degradation.

Tetrodotoxin is the poison produced, for instance, in *Tetraodontidae* (pufferfish or fugu) (Noguchi et al. 1987; Yotsu et al. 1987) and in the Southern blue-ringed octopus (*Hapalochlaena maculosa*) (Hwang et al. 1989). This is non-planar as well. The basic structure resembles the elementary spatial cell of the diamond crystal (Holleman and Wiberg 2001), which might be important in view of binding to larger receptor pockets. In general, it would be an interesting question to study whether natural product-derived compounds are more likely to be planar than artificial drugs. Such a hypothesis should be readily testable.

Our analysis could also be interesting for designing devices in nanotechnology. For example, when it is advantageous that a ring-shaped device be planar, the angles at the vertices should have such values that the angle sum fulfils the condition mentioned in the above theorem. Planarity is not only interesting for better fit into binding pockets, but also for constructing nanotubes from parallel, planar rings (Horne et al. 2005).

**Acknowledgments** We thank Dmitrij Frishman, Nadine Große, Dirk Hoffmeister, Sabine Hummert, Ralph Lehmann, Jürgen Sühnel and Heiko Stark for stimulating discussions, Ina Weiß for very helpful literature search and two anonymous referees for valuable comments. J.B. was funded by the Bundesministerium für Bildung und Forschung (Germany) within the HepatoSys programme; R.V. was funded by the Deutsche Forschungsgemeinschaft within the RTG 597; S.S. acknowledges financial support by the Deutsche Forschungsgemeinschaft within the Jena School of Microbial Communication.

**Ethical standards** For this publication, no experiment was needed to be conducted.

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

Aitio O, Hellman M, Kesti T, Kleino I, Samuilova O, Pääkkönen K, Tossavainen H, Saksela K, Permi P (2008) Structural basis of PxxDY motif recognition in SH3 binding. J Mol Biol 382:167–178. doi:10.1016/j.jmb.2008.07.008

Aitio O, Hellman M, Kazlauskas A, Vingadassalom DF, Leong JM, Saksela K, Permi P (2010) Recognition of tandem PxxP motifs as a unique Src homology 3-binding mode triggers pathogen-driven actin assembly. Proc Natl Acad Sci U S A 107:21743–21748. doi:10.1073/pnas.1010243107

Asiri AM, Khan SA, Tahir MN (2011) 2-[(1,3-Benzodioxol-5-ylmethylidene)amino]-4,5,6,7-tetrahydro-1-benzothiophene-3-carbonitrile. Acta Cryst E 67:o2162. doi:10.1107/S1600536811029564

Banga JR (2008) Optimization in computational systems biology. BMC Syst Biol 2:47. doi:10.1186/1752-0509-2-47

Barnett MP, Capitani JF (2006) Modular chemical geometry and symbolic calculation. Int J Quant Chem 106:215–227. doi:10.1002/qua.20807

Bartl M, Li P, Schuster S (2010) Modelling the optimal timing in metabolic pathway activation—Use of Pontryagin's Maximum Principle and role of the Golden Section. Biosystems 101:67–77. doi:10.1016/j.biosystems.2010.04.007

Benjamin DR (1995) Mushrooms: poisons and panaceas: a handbook for naturalists, mycologists and physicians. WH Freeman and Company, New York

Berg JM, Tymoczko JL, Stryer L (2007) Biochemistry. WH Freeman and Company, New York

Beyer H, Walter W (1996) Handbook of organic chemistry. Ellis Horwood Ltd, New York

Birdsall TC, Kelly GS (1997) Berberine: therapeutic potential of an alkaloid found in several medicinal plants. Alt Med Rev 2:94–103

Bodnaryk R, Yoshihara T (1995) Structure-activity relationships of cyclopentane analogs of jasmonic acid for induced responses of canola seedlings, *Brassica napus L*. J Chem Ecol 21:1735–1743. doi:10.1007/BF02033673

Budesinsky M, Cisarova I, Podlaha J, Borremans F, Martins JC, Waroquier M, Pauwels E (2010) Structures of cyclic dipeptides: an X-ray and computational study of *cis*- and *trans*-*cyclo*(Pip-Phe), *cyclo*(Pro-Phe) and their *N*-methyl derivatives. Acta Crystallogr B 66:662–677. doi:10.1107/S0108768110040243

Chakrabarti P, Pal D (2001) The interrelationships of side-chain and main-chain conformations in proteins. Prog Biophys Mol Biol 76:1–102. doi:10.1016/S0079-6107(01)00005-0

Creighton CJ, Leo GC, Du Y, Reitz AB (2004) Design, synthesis, and conformational analysis of eight-membered cyclic peptidomimetics prepared using ring closing metathesis. Bioorg Med Chem 12:4375–4385. doi:10.1016/j.bmc.2004.06.015

De Marino S, Festa C, D'Auria MV, Cresteil T, Debitus C, Zampella A (2011) Swinholide J, a potent cytotoxin from the marine sponge *Theonella swinhoei*. Mar Drugs 9:1133–1141. doi:10.3390/md9061133

Donnini S, Groenhof G, Wierenga RK, Juffer AH (2006) The planar conformation of a strained proline ring: a QM/MM study. Proteins 64:700–710. doi:10.1002/prot.21006

Falcon A (2011) Aristotle on Causality. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/. Accessed 20 July 2012

Fowden L (1956) Azetidine-2-carboxylic acid: a new cyclic imino acid occurring in plants. Biochem J 64:323–332

Fritz G, Beetz AC, Matern E, Peters K, Peters EM, von Schnering HG (1994) Bildung siliciumorganischer Verbindungen. 112. Über den Einfluß der Reaktionsbedingungen auf die Umsetzung von $(Cl_3Si)_2CCl_2$ mit Silicium. Die Struktur von 2,2,3,3,5,5,6,6-Octachlor-1,4-bis(trichlorsilyl)-2,3,5,6-tetrasilabicyclo[2.1.1]hexan und 1,1,3,4,6,6-Hexakis(trichlorsilyl)hexatetraen. Z anorg allg Chem 620:136–144. doi: 10.1002/zaac.19946200122

Gō N, Scheraga HA (1970) Ring closure and local conformational deformations of chain molecules. Macromolecules 3:178–187. doi:10.1021/ma60014a012

Gross C, Felsheim R, Wackett LP (2008) Genes and enzymes of azetidine-2-carboxylate metabolism: detoxification and assimilation of an antibiotic. J Bacteriol 190:4859–4864. doi:10.1128/JB.02022-07

Grützmann K, Böcker S, Schuster S (2011) Combinatorics of aliphatic amino acids. Naturwissenschaften 98:79–86. doi:10.1007/s00114-010-0743-2

Haidle AM, Myers AG (2004) An enantioselective, modular, and general route to the cytochalasins: synthesis of L-696,474 and cytochalasin B. Proc Natl Acad Sci U S A 101:12048–12053. doi:10.1073/pnas.0402111101

Hallen HE, Luo H, Scott-Craig JS, Walton JD (2007) Gene family encoding the major toxins of lethal *Amanita* mushrooms. Proc Natl Acad Sci U S A 104:19097–19101. doi:10.1073/pnas.0707340104

Heinrich R, Schuster S (1996) The regulation of cellular systems. Chapman and Hall, New York

Ho BK, Coutsias EA, Seok C, Dill KA (2005) The flexibility in the proline ring couples to the protein backbone. Protein Sci 14:1011–1018. doi:10.1110/ps.041156905

Holleman AF, Wiberg N (2001) Holleman–Wiberg's inorganic chemistry. Academic, Burlington

Horne WS, Ashkenasy N, Ghadiri MR (2005) Modulating charge transfer through cyclic D, L-α-peptide self-assembly. Chem Eur J 11:1137–1144. doi:10.1002/chem.200400923

Hwang DF, Arakawa O, Saito T, Noguchi T, Simidu U, Tsukamoto K, Shida Y, Hashimoto K (1989) Tetrodotoxin-producing bacteria from the blue-ringed octopus *Octopus maculosus*. Mar Biol 100:327–332. doi:10.1007/BF00391147

Inoue H, Sakata Y, Nishio H, Tokumo K, Kojima E, Date Y, Tamura Y, Tsuruta Y (2011) A simple and highly sensitive HPLC method with fluorescent detection for determination of pipecolic acid in mouse brain areas. Biol Pharm Bull 34:287–289. doi:10.1248/bpb.34.287

Jean RJ (1994) Phyllotaxis. A systemic study in plant morphogenesis. Cambridge University Press, Cambridge

Jiang B, Cao K, Wang R (2004) Inhibitory effect of protopine on $K_{ATP}$ channel subunits expressed in HEK-293 cells. Eur J Pharmacol 506:93–100. doi:10.1016/j.ejphar.2004.11.004

Jones DR, Ekins S, Li L, Hall SD (2007) Computational approaches that predict metabolic intermediate complex formation with CYP3A4 (+$b_5$). Drug Metabol Dispos 35:1466–1475. doi:10.1124/dmd.106.014613

Kacser H, Beeby R (1984) Evolution of catalytic proteins: on the origin of enzyme species by means of natural selection. J Mol Evol 20:38–51. doi:10.1007/BF02101984

Kamdem DP, Gage DA (1995) Chemical composition of essential oil from the root bark of *Sassafras albidum*. Planta Med 61:574–575. doi:10.1055/s-2006-959379

Kang YK, Park HS, Byun BJ (2009) Puckering transitions of pseudoproline residues. Biopolymers 91:441–455. doi:10.1002/bip.21150

Klipp E, Heinrich R, Holzhütter HG (2002) Prediction of temporal gene expression. Metabolic optimization by re-distribution of enzyme activities. Eur J Biochem 269:5406–5413. doi:10.1046/j.1432-1033.2002.03223.x

Koch MA, Schuffenhauer A, Scheck M, Wetzel S, Casaulta M, Odermatt A, Ertl P, Waldmann H (2005) Charting biologically relevant chemical space: a structural classification of natural products (SCONP). Proc Natl Acad Sci U S A 102:17272–17277. doi:10.1073/pnas.0503647102

Krasowski MD, Ai N, Hagey LR, Kollitz EM, Kullman SW, Reschly EJ, Ekins S (2011) The evolution of farnesoid X, vitamin D, and pregnane X receptors: insights from the greenspotted pufferfish (*Tetraodon nigriviridis*) and other non-mammalian species. BMC Biochem 12:5. doi:10.1186/1471-2091-12-5

Lehninger AL, Nelson DL, Cox MM (2000) Principles of biochemistry. Worth Publishers Inc, New York

Meléndez R, Meléndez-Hevia E, Canela EI (1999) The fractal structure of glycogen: a clever solution to optimize cell metabolism. Biophys J 77:1327–1332. doi:10.1016/S0006-3495(99)76982-1

Milner-White EJ, Bell LH, Maccallum PH (1992) Pyrrolidine ring puckering in *cis* and *trans*-proline residues in proteins and

polypeptides: different puckers are favoured in certain situations. J Mol Biol 228:725–734. doi:10.1016/0022-2836(92)90859-I

Moroder L, Musiol HJ, Scharf R (1992) Aziridine-2-carboxylic acid. A reactive amino acid unit for a new class of cysteine proteinase inhibitors. FEBS Lett 299:51–53. doi:10.1016/0014-5793(92)80098-2

Mukherjee S, Majumdar S, Bhattacharyya D (2005) Role of hydrogen bonds in protein-DNA recognition: effect of nonplanar amino groups. J Phys Chem B 109:10484–10492. doi:10.1021/jp0446231

Najmanovich R, Kuttner J, Sobolev V, Edelmann M (2000) Side-chain flexibility in proteins upon ligand binding. Proteins 39:261–268. doi:10.1002/(SICI)1097-0134(20000515)39:3<261::AID-PROT90>3.0.CO;2-4

Noguchi T, Hwang DF, Arakawa O, Sugita H, Deguchi Y, Shida Y, Hashimoto K (1987) *Vibrio alginolyticus*, a tetrodotoxin-producing bacterium, in the intestines of the fish *Fugu vermicularis vermicularis*. Mar Biol 94:625–630. doi:10.1007/BF00431409

Payton CW, Chang YF (1982) $\Delta^1$-Piperideine-2-carboxylate reductase of *Pseudomonas putida*. J Bacteriol 149:864–871

Pennycuick CJ (1997) Actual and 'optimum' flight speeds: field data reassessed. J Exp Biol 200:2355–2361

Reschly EJ, Ai N, Ekins S, Welsh WJ, Hagey LR, Hofmann AF, Krasowski MD (2008) Evolution of the bile salt nuclear receptor FXR in vertebrates. J Lipid Res 49:1577–1587. doi:10.1194/jlr.M800138-JLR200

Schmitt V, Schollmeyer D, Detert H (2011) (E, E, E, E)-2,3,5,6-tetrakis {2-[4-(dimethylamino) phenyl]ethenyl}pyrazine. Acta Cryst E 67:o1553. doi:10.1107/S1600536811019714

Schulz F, Gelhaus C, Degel B, Vicik R, Heppner S, Breuning A, Leippe M, Gut J, Rosenthal PJ, Schirmeister T (2007) Screening of protease inhibitors as antiplasmodial agents. Part I: aziridines and epoxides. Chem Med Chem 2:1214–1224. doi:10.1002/cmdc.200700070

Song IK, Kang YK (2006) Conformational preference and *cis-trans* isomerization of 4(*R*)-substituted proline residues. J Phys Chem B 110:1915–1927. doi:10.1021/jp054351h

Spector I, Shochet NR, Blasberger D, Kashman Y (1989) Latrunculins—novel marine macrolides that disrupt microfilament organization and affect cell growth: I. Comparison with cytochalasin D. Cell Motil Cytoskeleton 13:127–144. doi:10.1002/cm.970130302

Vollhardt KPC, Schore NE (2007) Organic chemistry: structure and function. WH Freeman and Company, New York

Werner S, Diekert G, Schuster S (2010) Revisiting the thermodynamic theory of optimal ATP stoichiometries by analysis of various ATP-producing metabolic pathways. J Mol Evol 71:346–355. doi:10.1007/s00239-010-9389-0

Williams R (1979) The geometrical foundation of natural structure: a source book of design. Dover Publications Inc, New York

Wishka DG, Bédard M, Brighty KE, Buzon RA, Farley KA, Fichtner MW, Kauffman GS, Kooistra J, Lewis JG, O'Dowd H, Samardjiev IJ, Samas B, Yalamanchi G, Noe MC (2011) An asymmetric synthesis of (2S,5S)-5-substituted azepane-2-carboxylate derivatives. J Org Chem 76:1937–1940. doi:10.1021/jo102475s

Withers PC, Hillman SS, Hedrick MS, Kimmel PB (1991) Optimal hematocrit theory during activity in the bullfrog (*Rana catesbeiana*). Comp Biochem Physiol A Physiol 99:55–60. doi:10.1016/0300-9629(91)90234-4

Yotsu M, Yamazaki T, Meguro Y, Endo A, Murata M, Naoki H, Yasumoto T (1987) Production of tetrodotoxin and its derivatives by *Pseudomonas* sp. isolated from the skin of a pufferfish. Toxicon 25:225–228. doi:10.1016/0041-0101(87)90245-5

Yu H, Chen JK, Feng S, Dalgarno DC, Brauer AW, Schreiber SL (1994) Structural basis for the binding of proline-rich peptides to SH3 domains. Cell 76:933–945. doi:10.1016/0092-8674(94)90367-0

**Appendix A.1. Sketch of the proof of the theorem**

Consider a (skew or planar) Euclidean $n$-gon ($n \geq 3$) formed by the vertices $X_1$, $X_2$, ..., $X_n$ in 3-(or higher)dimensional Euclidean space where the order of vertices in the $n$-gon is given by the indices and $X_n$ is connected to $X_1$ again. This is illustrated in Fig. 3. Two adjacent edges define one plane and two angles, $\alpha$ and $360°-\alpha$, within this plane. For each vertex, the smaller of the two angles defined by $X_{i-1}$, $X_i$, $X_{i+1}$ is taken, which is then less than, or equal to, 180°. We write A, B, C and W for $X_1$, $X_2$, $X_3$ and $X_n$, respectively, where $X_3$ may coincide with $X_n$. Let $\beta_1$ and $\beta_2$ be the angles formed by the pairs of lines (BA, BW) and (BW, BC). The angle made by the lines BA and BC is denoted $\beta$.

Let W' be the projection of vertex W onto the plane formed by A, B and C (Fig. A.1). We say that the point W' is in the interior of the (smaller) angle(A,B,C) (having magnitude $\beta$) if W' is in the same half-plane as A with respect to line BC and W' is in the same half-plane as C with respect to line BA. A similar definition applies to other angles.

First, we use the following inequality:

$$\beta_1 + \beta_2 \geq \beta \tag{A.1}$$

with equality holding if and only if W is in the same plane as A, B and C (i.e., W = W') and W' is in the interior of angle(A,B,C). The former condition means that the two angles at vertex B are coplanar. Relation (A.1) can be interpreted as a triangle inequality of angles. It is well known in spherical trigonometry (Reid and Szendröi 2005). The proof works as follows. Consider a sphere with a certain radius, $r$, around vertex B (Fig. A.1). A*, W*, and C* are the intersection points of the lines BA, BW and BC, respectively, with this sphere. Spherical distances are measured along great circles, as usual in spherical trigonometry. Each spherical distance between two points defining a certain angle is proportional to that angle. For example: $A^*W^* = \beta_1 r$.

Spherical distances fulfil the conditions on a metric (cf. Reid and Szendröi 2005). One of these conditions is the triangle inequality:

$$A^*W^* + C^*W^* \geq A^*C^* \tag{A.2}$$

Since the distances are proportional to the angles, also the angles fulfil a triangle inequality, that is, relation (A.1).

The theorem given in the Results and Discussion Section can be proved by induction in $n$. The assertion is true for $n = 3$, as the sum of bond angles (abbreviated by SoA) in each triangle is exactly 180°.

Let the assertion be true for some $n \geq 3$. Now we insert a new vertex $X_{n+1} = Z$ in the $n$-gon, without loss of generality between $X_n = W$ and $X_1 = A$ (see Fig. 3). Using the triangle inequality (A.1) twice (at A and at W), we obtain, for the SoA in the ($n$+1)-gon:

$$\text{SoA}[(n+1)\text{-gon}] \leq \text{SoA}[n\text{-gon}] + \text{SoA}[\text{triangle (AWZ)}] = \text{SoA}[n\text{-gon}] + 180° \quad \text{(A.3)}$$

The induction hypothesis says that

$$\text{SoA}[n\text{-gon}] \leq (n-2)*180° \quad \text{(A.4)}$$

Relations (A.3) and (A.4) lead to

$$\text{SoA}[(n+1)\text{-gon}] \leq (n-2)*180° + 180° = (n-1)*180°. \quad \text{(A.5)}$$

Equality holds only when equality holds in both inequalities (A.3) and (A.4). Due to the induction hypothesis, equality in relation (A.4) holds if, and only if, the $n$-gon is planar and convex. In relation (A.3), equality can hold only if Z lies in the same plane as the $n$-gon. Moreover, equality in (A.3) requires that W is in the interior of the (smaller) angle(Z,A,B) and so is A with respect to the angle(Z,W,$X_{n-1}$). This entails that equality in relation (A.4) holds if, and only if, all $n$+1 points lie in one plane and the ($n$+1)-gon is convex, which completes the proof of the theorem.

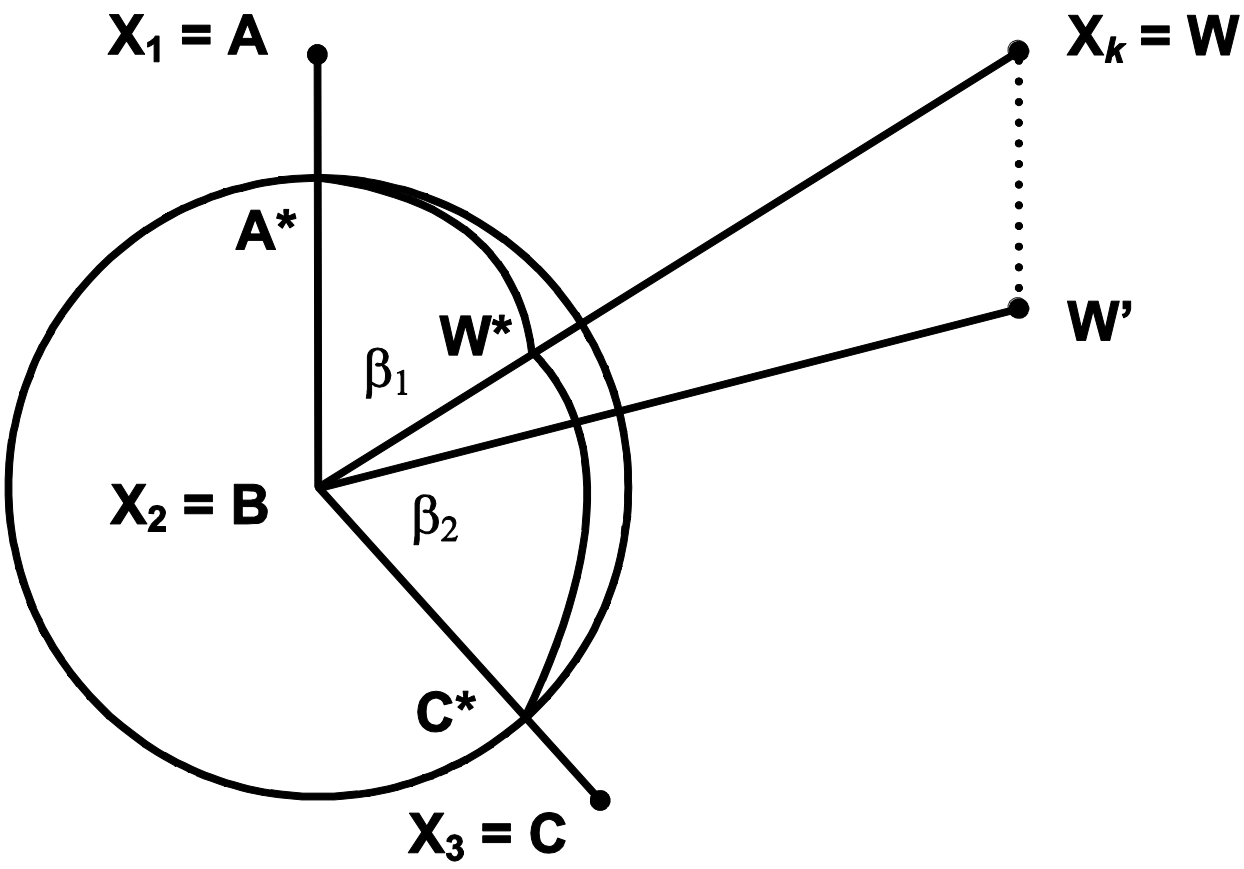


**Fig. A.1.** Schematic picture of the lines and angles considered in the proof of relation (A.1). W' is the projection of vertex W onto the plane formed by A, B and C. Solid circle, sphere around B; A*, W*, C*, intersection points of the lines BA, BW and BC, respectively, with this sphere; solid arcs, segments A*W* and C*W* on great circles on that sphere corresponding to angles $\beta_1$ and $\beta_2$, respectively.

**References**

Reid M, Szendröi B (2005) Geometry and topology. Cambridge University Press, Cambridge