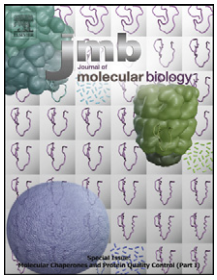
Review




# Prolyl Isomerization and Its Catalysis in Protein Folding and Protein Function

**Philipp A.M. Schmidpeter and Franz X. Schmid**

***Laboratorium für Biochemie und Bayreuther Zentrum für Molekulare Biologie,*** *Universität Bayreuth, 95440 Bayreuth, Germany*

***Correspondence to Franz X. Schmid:*** *Laboratorium für Biochemie, Universität Bayreuth, 95440 Bayreuth, Germany.*
*fx.schmid@uni-bayreuth.de*
http://dx.doi.org/10.1016/j.jmb.2015.01.023
***Edited by J. Buchner***

## Abstract

Prolyl isomerizations are intrinsically slow processes. They determine the rates of many protein folding reactions and control regulatory events in folded proteins. Prolyl isomerases are able to catalyze these isomerizations, and thus, they have the potential to assist protein folding and to modulate protein function. Here, we provide examples for how prolyl isomerizations limit protein folding and are accelerated by prolyl isomerases and how native-state prolyl isomerizations regulate protein functions. The roles of prolines in protein folding and protein function are closely interrelated because both of them depend on the coupling between *cis*/*trans* isomerization and conformational changes that can involve extended regions of a protein.


## Peptidyl-Prolyl Bonds

Peptide bonds are planar, and the two successive $C^{\alpha}$ atoms that flank the peptide bond can be either in *trans* or in *cis* conformation. For peptide bonds before residues other than proline (non-prolyl bonds), the *trans* state is strongly favored, and the equilibrium constant [*trans*]/[*cis*] is usually higher than 100 [1–4]. This is not the case for peptide bonds between proline and its preceding amino acid (Xaa-Pro or prolyl bonds; Fig. 1). Here, the *trans* isomer is only slightly favored over the *cis* isomer unless structural constraints (as in folded proteins) strongly stabilize one isomer relative to the other. In the absence of ordered structure, as in short peptides, the two isomers co-exist in solution, and *cis* contents up to about 30% are observed [5–8].

In proteins with known three-dimensional structure, the conformational state of each peptide bond is usually clearly defined. It is either *cis* or *trans* in every molecule, depending on the structural framework imposed by the folded protein chain. About 7% of all prolyl bonds in folded proteins were found to be in *cis*, compared to only 0.03% of non-prolyl bonds [9,10].

Prolyl *cis*/*trans* isomerizations involve the rotation about a partial double bond (Fig. 1), and they show high activation enthalpies near 80 kJ/mol [5,11]. Therefore, they are intrinsically slow with time constants typically in the range 10–100 s (at 25 °C).

## Prolyl Isomerization in Protein Folding

Garel and Baldwin discovered in 1973 that unfolded ribonuclease A (RNase A) consists of a mixture of fast-folding $U_F$ and slow-folding $U_S$ molecules [12]. A few years earlier, the crystal structure of a variant of RNase A had been solved [13], and it revealed that two prolines (Pro93 and Pro114) adopted the *cis* conformation and two (Pro41 and Pro117) adopted the *trans* conformation in the folded protein. In 1975, Brandts *et al*. coined the proline hypothesis [11] and proposed that the fast-folding $U_F$ molecules show the same prolyl isomers as the native protein (N). Therefore, they are able to complete folding within a few milliseconds. The $U_S$ molecules differ in the *cis*/*trans* isomeric state of one or more Xaa-Pro peptide bonds. They fold slowly because folding must be preceded by the slow isomerization of the incorrect prolines. Subsequently, fast- and slow-folding species were detected in many proteins [14–16].

Of course, non-native prolyl isomers do not block refolding right at its beginning and re-isomerization is

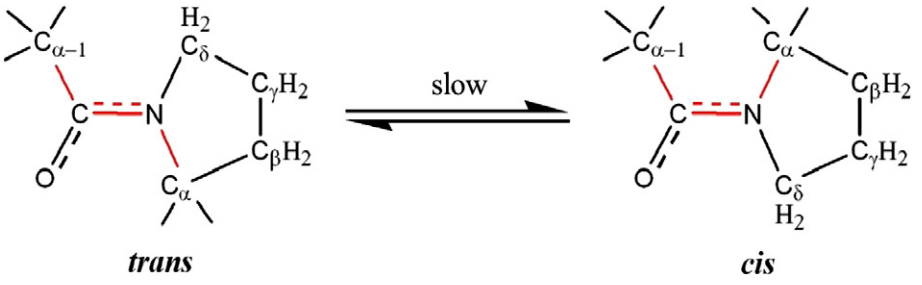



**Fig. 1.** *Trans* and *cis* isomers of a peptidyl-prolyl bond.

not required to be the first step of folding, as suggested initially in the proline hypothesis [11]. Incorrect prolyl isomers can often be accommodated in partially folded intermediates, and in the case of RNase A, we found a folding intermediate with an incorrect proline that showed low catalytic activity already [17–19]. In these folding experiments with RNase A, its four disulfide bonds were left intact, and therefore, the native protein and the partially folded intermediates showed high stability. Thus, in favorable cases, conformational folding can go almost to completion while some prolines are still in their non-native isomeric state. However, the final folding steps require correct prolines, and therefore, late folding reactions are often limited in rate by prolyl isomerizations. These early folding experiments already provided evidence that prolyl isomerizations, in principle, can occur in folded proteins.

The extent of native-like structure that is reached in the course of folding prior to prolyl isomerization depends on the location of the non-native prolyl isomers and on the folding conditions. Generally, incorrect prolyl bonds at the surface of the folded protein or in flexible chain regions will not interfere strongly with conformational folding, and solvent conditions that strongly stabilize folded proteins will also stabilize partially folded structure in intermediates with incorrect isomers. Conformational folding steps and prolyl isomerization are thus mutually interdependent. Incorrect prolyl isomers in a protein chain decelerate its folding, and at the same time, conformational folding prior to the isomerization of a prolyl bond can affect its kinetics and the *cis*/*trans* equilibrium. This close interrelationship between the formation of ordered structure and prolyl isomerization is a key feature of slow-folding steps and, as discussed later, also of reactions in which prolyl isomerizations in folded proteins are used for regulatory purposes.

## Prolyl Isomerization in Folded Proteins

In its early days, crystallography provided us with apparently static structures of small, well-ordered proteins. Consequently, it was assumed that prolyl bonds generally were well ordered as well, occurring in either the *cis* or the *trans* form. This changed when high-resolution NMR spectroscopy became available for determining the structure of proteins in solution.

First evidence for the co-existence of the *cis* and *trans* forms in a folded protein was obtained for staphylococcal nuclease, initially at Pro117 and then at Pro47 [20,21]. At Pro117, the equilibrium could be shifted toward the *cis* conformation by the binding of $Ca^{2+}$ ions or of the inhibitor thymidine 3′5′-phosphate, as well as by the stabilizing H124L substitution. Together, these findings indicated that, in flexible or locally unfolded regions of a protein, the coupling between conformational folding and prolyl isomerization might be too weak to stabilize one isomer strongly over the other. Ligand binding or stabilizing substitutions presumably provide additional folding energy to shift the conformational equilibrium toward the isomer that is best compatible with the fully folded form of the protein. Meanwhile, *cis*/*trans* heterogeneities have been found in many folded protein [22–34].

## Functionally Important Prolyl Isomerizations

The co-existence of *cis* and *trans* isomers in folded proteins suggested that prolyl isomerizations might be used, for example, as slow molecular switches between different functional states of a protein, as already discussed in 1991 [35,36]. The search for such a role gained momentum when the phosphorylation-sensitive prolyl isomerase Ess1 and its human homolog Pin1 were discovered [37–39], which catalyze the isomerization of peptide bonds between phosphoserine or phosphothreonine and proline [40]. Ser/Thr-Pro sequences are targets for many protein kinases, and therefore, it was suggested that prolyl isomerization might be intimately linked with phosphorylation-dependent regulation. Indeed, Pin1 is involved in a large number of regulatory processes ranging from the control of the cell cycle to amyloid formation in the brain [41–44].

Directed mutagenesis is routinely used to examine whether a proline or its neighboring residues are important for the function or regulation of a specific