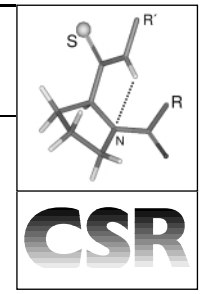
Chemical aspects of peptide bond isomerisation



Gunter Fischer

Max-Planck Research Unit "Enzymology of Protein Folding", Weinbergweg 22, D-06120 Halle/Saale, Germany

Received 31st August 1999

Coplanarity of peptide bonds confers partitioning of peptide chains between two energetically preferred rotational states, *cis* and *trans*. Molecular heterogeneity is particularly pronounced when imino acids like proline form the peptide bond. These conformational substates are prone to isomer-specific biochemical recognition, delayed chain folding and biocatalysis of conformational interconversion attracting broad interest in medicinal chemistry and biotechnology. The present review discusses the structural features of peptide bond conformation in oligopeptides and proteins, and gives an overview of isomer ratios, interconversion rates, and catalysis.

1 Introduction

As a repeating functional group, amide bonds form a considerable part of the backbone of peptides and proteins, thus generating a high molar concentration of amide units in living systems. Due to its multifunctional nature, amide bond reactivity has been established as a topic of experimental and theoretical interest to chemists for a long time. The amide group is characterised by specific properties such as hydroxyimine–amide tautomerism, coplanarity, considerable resistance to nucleophilic attack and acyl group transfer, carbonyl oxygen basicity, reduced nitrogen co-ordination, and favourable hydrogen bond donor/acceptor and acceptor properties for secondary and tertiary amide bonds respectively. Almost all of these peculiarities are of fundamental importance in the structuring and functioning of polypeptide chains, and can be rationalised in terms of amide bond resonance first discussed by Linus Pauling in 1948. Structure analyses of proteins at atomic resolution as well as protein folding experiments revealed exclusive rigidity of the peptide bond among the linkages forming a protein backbone. Hindered peptide bond rotation in conjunction with mobile single bonds on either side of the peptide unit has been found to be the ultimate principle for the formation of the three-dimensional structure of proteins. Since the *cis/trans* isomer-producing rotation about the C–N bond is closely related to amide bond resonance, some molecular aspects of the *cis/trans* isomerisation of secondary amidic and imidic peptide bonds will be discussed. Besides chemical effects related to the theory of bond resonance, shuttling between two different polypeptide conformations in a timed manner is likely to have an impact on biochemical reactivity and cell signalling. Thus, some biological aspects of the *cis/trans* isomerisation of proline-containing polypeptides have been outlined in this review as well.

Gunter Fischer was born in Altenburg, Germany in 1943. After studying Chemistry he received his PhD in 1971 from the Martin-Luther-University of Halle/Saale under the supervision of Alfred Schellenberger. During this time he worked on the carbonyl reactivity of α-ketoacids and α-ketoamides. He followed this up with an education in biochemistry, focusing his research interests on protease chemistry. After a period working in industry in the Germed Research Center of Biotechnology, Berlin, he relocated to Halle in 1988. He was appointed to the chair of Molecular Biochemistry in 1993. He has received the Gottfried-Wilhelm-Leibniz Award of the Deutsche Forschungsgemeinschaft. At present he is a professor and director of the Research Center "Enzymology of Protein Folding" of the Max-Planck-Society in Halle/Saale. In 1984 he discovered peptidyl prolyl cis/trans isomerases, a class of folding helper enzymes. His research interests are molecular recognition, bioactive conformation, and protein folding.

2 Electronic aspects of peptide bond rotation

The classical view of the amide resonance is based on the ability of the nitrogen atom to delocalise its electron lone pair over the whole moiety.[1] The resulting planar framework for the electronic organisation and the consequent partial double bond character of the C–N bond cause restrictions in the number of energy minima in amide bond torsion (dihedral angle ω). Fig. 1*a*

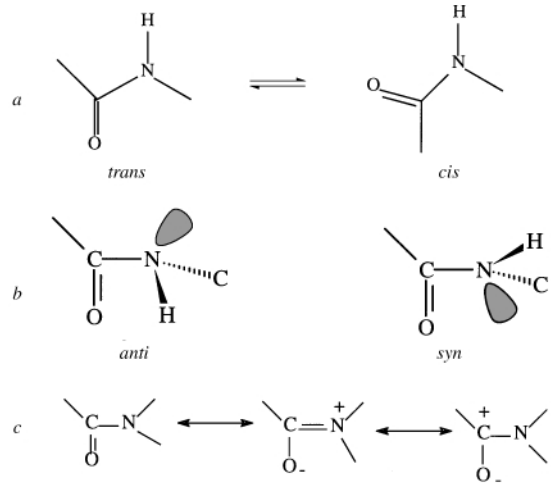


Fig. 1 (*a*) Schematic representation of *cis/trans* isomers of a secondary amide peptide bond; (*b*) alternative transition state configurations for peptide bond rotation; (*c*) canonical structures used to describe the amide group structure.

represents the molecular fragment in just two minimum-energy structures. The geometric isomers at the angles $\omega \approx 0°$ (*cis*) and

$\omega \approx 180°$ (*trans*) are separated by a rotational barrier corresponding to the perpendicular high energy state of $\omega \approx 90°$. The distinction between the *cis* and *trans* isomer of a molecule originates from a geometry-based classification of structures. When the two substituents X and Y are on the same side of the structural unit (Fig. 1*a*) the isomer is *cis*, and it is designated *trans* in the opposite configuration. Ambiguities in the designation of the isomers can be avoided using the precise *E*/*Z* nomenclature. In many cases the *E* conformation is designated *cis* and the *trans* form *Z*, as it is found in peptides and proteins. The semi-rotated, twisted transition state can be described by two energetically different configurations (Fig. 1*b*). In this view atomic motions involved in crossing the internal rotational barrier include C–N bond lengthening, nitrogen and carbon pyramidalisation, solvent cage redistribution and bond rotation. As could be inferred from the resonance model peptide bonds are expected to occur in only two ground state conformations in native proteins. Indeed, the dihedral angles ω are tightly clustered around 0° (*cis*) and 180° (*trans*) in a ratio of about 5 : 95 respectively, when analysing the database of the three-dimensional structures of proteins.

Since the coexistence of *cis/trans* isomers of secondary amides in solution was demonstrated together with the slow C–N bond rotation of dimethylformamide by spectroscopic methods in the mid 1950's, simple amides like formamide, *N*-methylformamide, dimethylformamide and *N*-methylacetamide have often been used to mimic the framework of proteinaceous peptide bonds for theoretical approaches.

Resonance stabilisation of the planar *cis/trans* conformers relative to the twisted state does account for about one-half of the rotational barrier[2] as could be inferred from *ab initio* valence-bond calculations of formamide. Other contributions arise from the lone pair orientation at nitrogen that is arranged perpendicularly to the molecular plane in the low-energy state, even in the absence of conjugation. A stabilising interaction of the lone pair with the adjacent, electron-deficient carbon atom in the planar conformation has been hypothesised. In this model the carbonyl oxygen of the amide group, thought to be tightly involved in bond rotation by charge transfer in classical resonance structures, is only a constantly polarizing element during rotation of the C–N bond.[3] Acetamides are structurally more closely related to peptide units than formamides. Thus, calculations revealed an intrinsic non-equivalence of the rotational barrier heights for the *anti* and *syn* transition states (Fig. 1*b*) that might sensitively depend on the amide structure and the solvent nature. In dimethylacetamide the isomers are separated by a rotational barrier of about 80 kJ mol^{-1} at 60 °C in aqueous solution. Other possible motions in polypeptide chains exhibit much lower rotational barriers rendering the peptide bond isomerisation an uncoupled movement in chain rearrangements. Consistent with the canonical forms of the amide group (Fig. 1*c*) dimethylacetamide showed an increased rotational barrier when moving from a nonpolar to a polar solvent. For the experimental rotational barrier of dimethylformamide a linear relationship with the Dimroth–Reichardt solvent polarity parameter E_T was obtained.[4]

Substituents on either side of the amide unit influence the magnitude of the rotational barrier differently. Electron-donating substituents on the carbonyl side lower the barrier whereas the opposite effect has been obtained on the nitrogen site.[5] However, steric contribution seems to be the most significant determinant on both sides for the ease of crossing the barrier.[6]

3 Secondary amide bond *vs.* imide bond isomerisation

Excepting the side chains of glutamine, asparagine and some C-termini, secondary and tertiary amide bonds are subject to nondegenerate *cis/trans* isomerisations in polypeptides. Nineteen of the twenty gene-coded amino acids form secondary amidic peptide bonds, whereas the *N*-alkylated amino acid proline has the potential to establish an imidic peptide bond at its N-terminal side (Fig. 2a). The term prolyl bond is used for

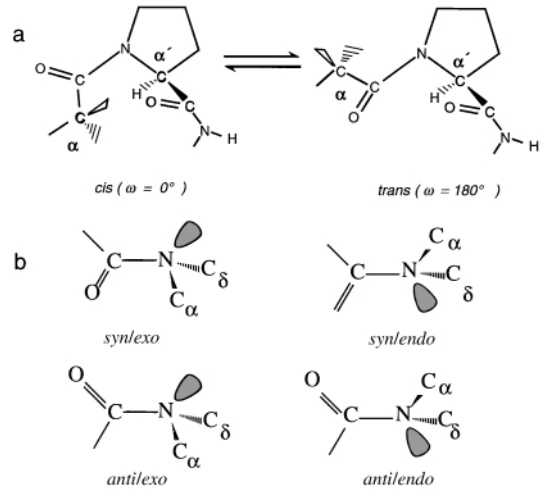
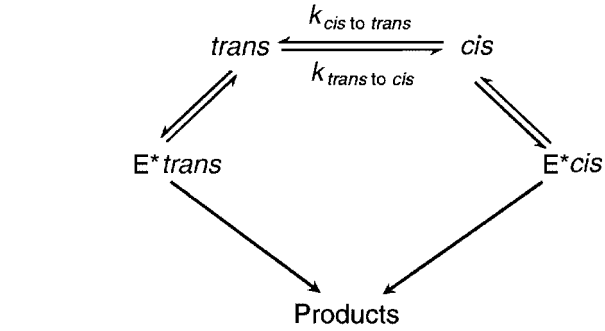


Fig. 2 (*a*) Both isomers of a prolyl bond; (*b*) possible transition state configurations of a rotating prolyl bond.

this linkage throughout the review, and prolyl isomerisation indicates the *cis/trans* isomerisation of the imidic bond preceding proline. Excluding N-terminal proline residues, polypeptides containing *n* prolines can form 2^n prolyl isomers unless structural constraints (such as in folded proteins) stabilise one isomer strongly relative to others. Due to the additional features of having the torsion angle ϕ restricted by fixing the N–alkyl bond in the five-membered ring, and the loss of hydrogen bond donating capability, proline plays a key role in the secondary structures of polypeptides.

The proximity of the C^{α} atoms of alkyl substituents on either side of the amide moiety—in close contact with each other in the *cis* conformation—has been suggested to be important for the low percentage of secondary amide *cis* peptide bonds, because steric strain is released in the *trans* conformer. Typically, for *N*-methylacetamide only 1.4% *cis* isomer has been detected in D_2O.[7] A different situation exists in tertiary amides like prolyl bonds because both conformers meet a comparable steric situation. Many studies of *N*,*N*-dialkylamides have already demonstrated both a rotational barrier of about 80 kJ mol^{-1} and no energetic distinction between *cis* and *trans* populations for symmetric *N*,*N*-disubstitutions. In contrast, distinct aminoacyl substituents attached to the nitrogen atom of proline cause *cis* isomer contents of 5% to 50% dependent on the sequence of the unstructured peptide in which the proline is located. A peptide bond photoisomerises in aqueous solution to the less stable *cis* isomer when excited at the amide absorption at 206 nm. Thus, the ground state population of *cis* isomers can be enhanced at the expense of the respective *trans* isomer by a photomechanism still not known in detail.[8]

Realising prolyl bond rotation opens the possibility of constructing four transition state configurations (Fig. 2b). Calculations show that the conformation of the C–C(O) bond succeeding the proline ring (angle ψ) is critical to the pathway of rotation for *N*-acetylproline methylamide.[9] An intramolecular interaction between this part of the molecule and the proline nitrogen lone pair has been suggested in the transition state which decreases the rotational barrier by about 15 kJ mol^{-1} relative to the rotational pathway lacking the interaction. Even if the rate of prolyl isomerisation is sensitive to the neighbouring residues of proline, it is remarkably resistant to alterations in pH value and buffer composition. Typically, relaxation times from ten to hundreds of seconds can be observed for peptides and proteins at 25 °C. Barriers to rotation of $\Delta G^{\ddagger} = 79$ kJ mol^{-1} (*cis* to *trans*) and $\Delta G^{\ddagger} = 89$ kJ mol^{-1}



$$v = \frac{(k_{trans\ to\ cis} K_m^{trans} k_{cat}^{cis} + k_{cat}^{cis} k_{cat}^{trans} [E]_o + k_{cat}^{trans} K_m^{cis} k_{cis\ to\ trans}) [S]_o [E]_o}{K_m^{cis} (k_{trans\ to\ cis} K_m^{trans} + [E]_o k_{cat}^{trans})}$$

$$[E]_o k_{cat}^{trans} >> k_{cis\ to\ trans} K_m^{cis}$$

$$v = k_{cat}^{cis}/K_m^{cis} [E]_o [S]_o + k_{cis\ to\ trans} [S]_o$$

Scheme 1 The rate of product formation v in a protease (E)-catalysed reaction starting from a mixture of *cis* and *trans* substrate (S). Steady-state approximation has been applied for the evaluation of the rate equation. First-order kinetics independent of protease concentration result, when the specificity constant k_{cat}/K_m for the *cis* isomer of the substrate is low.

(*trans* to *cis*) have been measured for *N*-methylacetamide by ^{1}H NMR at 60 °C in water.[10] Prolyl bonds displayed rotational barriers of the same order of magnitude.

Only recently, ^{1}H NMR studies of oligopeptides containing tyrosine and phenylalanine revealed the presence of *cis* secondary amide peptide bonds in the oligopeptide chain. These were evidenced by signal multiplicity, allowing experimental access to rate constants and isomer ratios.[11] Here, peptide bond rotation appeared to be quite rapid, resulting in relaxation times of $\tau \approx 1$ s at 25 °C. Due to isomer-specific chemical shift differences protons adjacent to the peptide bond gave two separate signals. The *cis* specific upfield shift of about 0.5 ppm for the methyl group signal of alanine flanking the aromatic residue, parallels the isomer-specific shift difference for the C^δ protons of the proline ring. Generally, conformationally dependent signal multiplicity of resonating nuclei in polypeptides often indicates peptide bond isomerisation, because the interconversion dynamics are slow on the NMR time scale at room temperature, and considerable isomer-specific chemical shift differences may exist.

In summary, peptide bonds (prolyl bonds in particular) establish a slowly interconverting conformational multiplicity—in a sequence dependent manner—in peptides and proteins; in accord with the proposed conformational properties of simple imidic bonds.

4 Peptide bond isomers in oligopeptides and proteins

4.1 Measuring isomerisation constants

Due to the dynamic properties of peptide bond isomerisation discussed above, many nuclei offer potential as probes in NMR spectroscopy for determining both interconversion kinetics and ratios of isomers. Using various resonating nuclei, one- and two-dimensional methods such as line-shape analysis, saturation transfer and 2D NOESY (EXSY) experiments can be utilized for the kinetic analysis but require relatively high peptide concentrations.[12] In the case of prolyl bonds, isomer-specific spectral differences exist which can be investigated by a number of spectroscopic methods such as UV–Vis spectra of chemically modified oligopeptides, fluorescence spectroscopy of intramolecularly fluorescence-quenched peptide chains, and far-UV CD spectra. Other analytical methods exhibiting sensitivity for *cis/trans* isomerisation include separation by low-temperature reversed-phase HPLC and capillary electrophoresis, but these were found to be of limited applicability. When the time-course of protein refolding or unfolding is monitored directly by intrinsic fluorescence changes, or other folding probes; or indirectly by double-jump experiments,[13] prolyl isomerisation has frequently been shown to be the rate limiting step. Therefore, folding kinetics may be used as a tool for determining isomerisation rates. However, access to the isomerisation constants of discrete prolyl bonds requires detailed knowledge of the protein structure.

For assaying biocatalysis of isomerisation, the most popular assay was based on isomer-specific proteolysis protocols.[14] According to Scheme 1, the *cis* to *trans* isomerisation of a proline-containing proteolytic substrate becomes rate limiting when relatively high concentrations ($[E]_0$) of an isomer-specific protease were allowed to cleave a very sensitive substrate (large k_{cat} value). Under these reaction conditions, exclusive specificity of the protease for the *trans* substrate leads to rapid depletion of the *trans* substrate whereas the *cis* isomer remains uncleaved. In the case of chymotrypsin or subtilisin for example, proteolysis of peptidyl-Pro-Phe-4-nitroanilides at the anilide bond shows biphasic kinetics with the slower phase identified as *cis* to *trans* isomerisation of the peptidyl–prolyl bond. The time-course of *cis* to *trans* isomerisation can be followed at the 4-nitroaniline absorption by UV–Vis spectroscopy (Fig. 3). The isomer ratio at the ground state of the substrate peptide equals the ratio of amplitudes of the fast and slow kinetic phases when extrapolated to zero time of proteolysis. Obviously, the relative amount of 4-nitroaniline formed by proteolysis of the two critical conformers is dependent on the relative populations of the conformations, because the reaction rate of the cleavage at relatively high protease concentrations (k_{cat}/K_m* $[E]_0$), is much faster than the rates of conformational interconversion. As a precondition, both high conformational discrimination among the isomers and fast first-order reaction rates are required, but

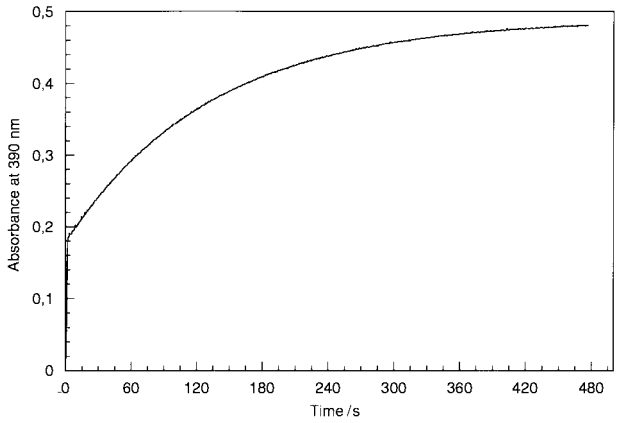


Fig. 3 Time-course of the cleavage of succinyl-Ala-Phe-Pro-Phe-4-nitroanilide by subtilisin type VIII (0.15 mg ml^{-1}) in 0.035 M HEPES buffer pH 7.8 at 13 °C. The reaction was followed by UV–Vis spectroscopy at 390 nm. After a time lag of about 10 s, the kinetic trace represents the *cis* to *trans* isomerisation of the substrate prolyl bond. The reaction was started by injecting into enzyme solution the substrate dissolved in 0.2 M trifluoroethanol–LiCl. About 50% *cis* substrate resulted from the ratio of slow and fast signal amplitudes at zero time, typical of the increased *cis* content in the Li^+-containing nonaqueous stock solution.

these are singular properties of biocatalysis; in particular proteolysis. Thus, limitations of "chemical" conformational analysis as defined in the Curtin–Hammett principle were surmounted in isomer-specific proteolysis assays.

4.2 Ratio of isomers

There are very limited reports on the ratio of isomers of a secondary amide peptide bond in peptide chains. It was determined that in the middle of a linear peptide chain, conformational heterogeneity decreases by favouring *trans* isomers with about 3 kJ mol^{-1} (Table 1).[11] The position where a *cis* peptide bond would be formed in a cyclic peptide cannot yet be predicted. Considering the low *cis* isomer content (Table 1), marked *cis* isomer destabilisation by the actual C- and N-substituents has to be noted for peptides when compared to the prototypic *N*-methylacetamide. Nevertheless, even the low *cis* probability inferred from the data in Table 1 implies the permanent existence of, minimally, a single *cis* peptide bond fluctuating in an unstructured polypeptide chain of 1000 amino acid residues in length.

Table 1 Characteristic constants for the *cis/trans* isomerisation of secondary amide peptide bonds at 25 °C in aqueous solution

Compound	% *cis*[b]	$k^a_{cis\ to\ trans}$ /s^{-1}	$k^a_{trans\ to\ cis}$ /10^{-3} s^{-1}	pH
Ala-Tyr	0.41	0.57	2.4	5.9
Ala-Ala-Tyr[c]	0.21	0.29	0.6	4.4
Ala-Ala-Tyr-Ala[c]	0.16	1.61	2.6	6.3
Ala-Ala-Tyr-Ala-Ala[c]	0.14	2.38	3.3	4.1

[a] From the Arrhenius equation $\ln k = \ln A - E_a/RT$ for 25 °C. [b] From the van't Hoff equation for 25 °C. [c] For the Ala–Tyr peptide bond.

In a recent analysis, 59 non-proline *cis* peptide bonds were found in a database of 747 native proteins; this result probably does not reflect the true level of occurrence. More rigorous calculations will probably result in the discovery of a higher number. Mostly located in β-structures, they often occupy positions critical to biological function. It was hypothesised that *cis* peptide bonds represent high energy structures able to provide stored potential energy to reactions.[15]

Prolyl isomerisation has been studied frequently since it was identified as the rate-limiting step for the refolding of denatured proteins.[16] In short peptides, major determinants of the *cis/trans* ratio were attributable to the local environment, in particular, the nature of the amino acid preceding proline and a negative charge at the proline carboxy group for C-terminal prolines.[17] Investigations of completely unstructured oligopeptides like acetyl-Ala-Xaa-Pro-Ala-LysNH_2 (Xaa is for all gene coded amino acids) shed light on how local effects of the amino acid sequence influence the thermodynamics of prolyl isomerisation.[18] Data depicted in Fig. 4 reveal stabilising proline–Xaa interactions in the *cis* isomer if aromatic amino acids substitute for Xaa. This interaction has been substantiated by the conformational analysis of many other sequences containing the -Xaa-Pro- moiety (Xaa = Phe, Tyr, Trp, His). Proline-rich oligopeptides often show the number of isomers expected from the number of prolines in the chain but the ratio of isomers deviates from those calculated from simple statistics. The reason why still remains unknown but will include stabilising interactions invisible to nuclear Overhauser effects and CD spectroscopy. Under native conditions proteins mostly contain each proline in a conformationally defined state, being either *cis* or *trans*. The energy costs associated with the preferential population of a certain isomer have to be covered by favourable spatial arrangements of interacting groups, although a major contribution to the propensity of an isomer results from local effects. An exception was found with the Pro-Pro moiety that is more often in the *cis* state in proteins than expected from the data of oligopeptides. Recent protein structure database analysis showed a frequency of 4.8% of *cis* conformation for a total number of 8598 prolyl bonds.[18] Unfolding the proteins establishes conformational heterogeneity as typically found by ^{1}H NMR spectroscopy for oligopeptides lacking ordered structures. Rapid refolding of the protein only proceeds to completion with unfolded chains having all prolyl bonds in an already correct isomeric state. Other chains with incorrect

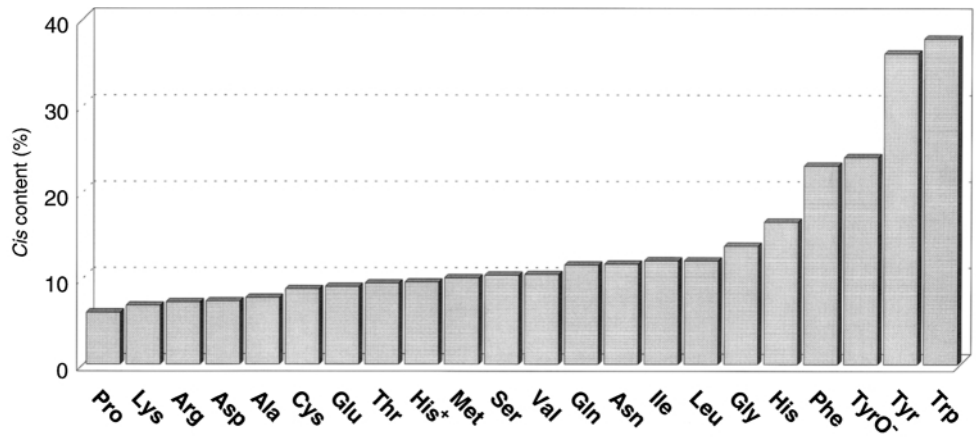


Fig. 4 *Cis* isomer content of the imidic prolyl bond of Ac-Ala-Xaa-Pro-Ala-Lys-NH_2 in aqueous solution pH 5.0 at 4 °C. The data for His^+ and $TyrO^-$ side chain ionisations have been extrapolated from a titration curve.

prolyl isomers collapse into folding intermediates, to be followed by kinetically slow folding phases that have characteristics of either *cis* to *trans* or *trans* to *cis* prolyl isomerisation.[19] However, the occurrence of prolyl isomers has also been reported for an increasing number of proteins in the native state. Moreover, as many as 21 cases of identical or nearly identical proteins have been found in the protein structure database that differ in their isomeric state by a single prolyl bond. Coexisting isomers in solution may present distinct properties to potential reaction partners creating isomer-specific reactions in biochemistry.[20] Unfortunately, isomer-specific chemical properties of the peptide bond isomers such as nucleophilic reactivity, oxygen basicity, hydrogen bonding, metal complexation, and lipophilicity, have not yet been rigorously examined.

4.3 Interconversion kinetics

Obviously, peptide bond isomerisation proceeds by a unimolecular process characterised by the first-order rate constant $k_{obs} = k_{cis\ to\ trans} + k_{trans\ to\ cis}$ for the reversible reaction, whereas either $k_{cis\ to\ trans}$ or $k_{trans\ to\ cis}$ determines the reaction rate if the reverse isomerisation becomes improbable by a coupled, quasi-irreversible reaction. The latter situation dominates in some biochemical events like isomer-specific proteolysis, protein folding, collagen triple helix formation and signal transduction across membranes. For secondary amide peptide bonds the equilibration rate of isomers is almost entirely determined by $k_{cis\ to\ trans}$ (Table 1), whereas in prolyl isomerisations both rate constants usually contribute in a comparable manner; as indicated by the data collected in Fig. 4. A rough comparison can be made of the kinetic constants collected in Table 1 and Fig. 5, taking into account the different substituents on either side of the peptide bond, and different temperatures. Despite these differences, the *trans* to *cis* isomerisation rates seem to be quite similar for secondary amide peptide bonds and prolyl bonds. From these data it becomes obvious that the *cis* ground state of secondary amide peptide bonds is strongly disfavoured relative to the transition state of isomerisation. Virtually identical rate constants became apparent by comparison of secondary amide peptide bonds in oligopeptides (Table 1) with both the forward and reverse isomerisation rates of the Ala38–Tyr39 peptide bond in an RNase T1 variant.[21] It appears that thermodynamic stabilisation of the *cis* secondary amide peptide bond—present in several proteins—has to occur on the level of a decelerated *cis* to *trans* isomerisation rather than an accelerated *trans* to *cis* reaction. Native state prolyl isomerisations, as well as isomerisation during refolding reactions of proteins, exhibit rate constants quite similar to corresponding values of the respective model peptides, except for isomerisations coupled kinetically to other folding events. Decelerated prolyl isomerisations have occasionally been found in folding intermediates, but rate accelerations have also frequently been observed.

Variation of the amino acid preceding proline did not greatly alter prolyl isomerisation rates which were all within one order of magnitude (Fig. 5). To our knowledge, other positions relative to proline did not contribute as much as the Xaa-Pro position. Given this, prolyl isomerisation in unfolded proteins should not depend strongly on the amino acid sequence. However, clustering of *cis* prolyl bonds exemplified by the *cis*/*cis* isomer of the -Pro-Pro- moiety reduces isomerisation rates.[19] Unexpectedly, side chain phosphorylation at -Thr/Ser-Pro- moieties greatly decelerates both *cis* to *trans* and *trans* to *cis* isomerisation rates up to 10 fold. A glutamic acid side chain cannot mimic phosphorylation on Ser/Thr residues in this respect, because the major effect arises in the phosphate dianionic state.[22] Except for proline ring hydroxylation which does not greatly affect prolyl isomerisation[23] other posttranslational protein modifications of residues adjacent to proline have not yet been investigated in detail. The mechanistic picture that emerged for prolyl isomerisation was mainly supported by deuterium isotope effects. Isomerisation rates show no significant solvent deuterium isotope effect, but reveal a large secondary β-deuterium isotope effect of 1.05 for a Cα-Gly double-deuterated Gly-Pro moiety.[24] In conjunction with the enthalpy control of prolyl isomerisation already known from Eyring plots, the transition state is characterised by considerable bond rotation, no significant solvent reorganisation and a lack of nucleophilic or electrophilic solvent participation.

In a protein, prolyl isomerisation most affects the spatial arrangement of the residue preceding proline by shortening of the Cα atom distance during *trans* to *cis* interconversion. However, the geometric expansion or contraction is propagated through the chain to remote sites. Probably, a whole segment of the chain following proline is pulled toward the isomerising bond (U. Reimer and G. Fischer, submitted). Taken together, peptide bond isomerisations are reminiscent of a mechanical on/off switch because two stable positions are separated by a switching barrier. Such a kind of periodic main chain movement might provide a basis for utilising proline as an internal clock in the control of protein functionality.

4.4 Environmental effects

Both isomer ratio and rate constants remain constant over a broad pH range unless dissociable groups are present near to the peptide bond.[11,24] However, some reaction parameters change considerably in nonaqueous environments. For acetyl-Gly-Pro-

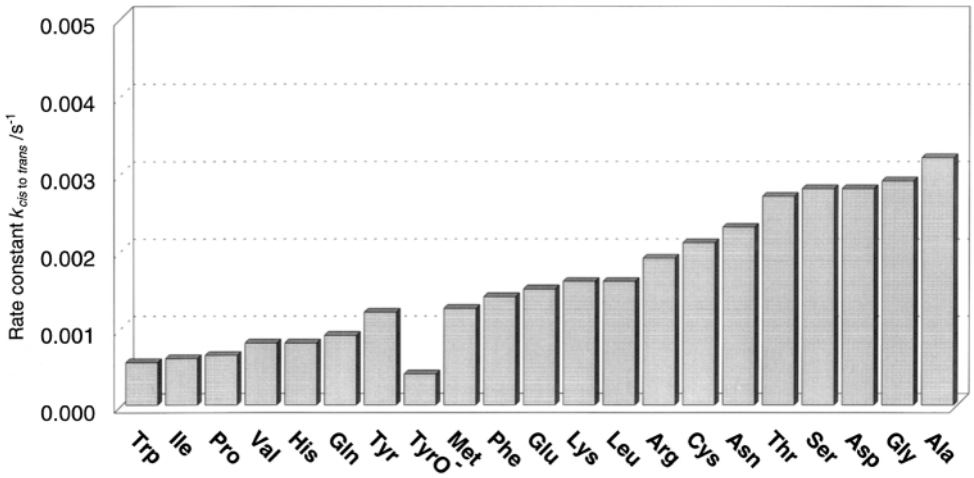


Fig. 5 Rate constants for the *cis* to *trans* isomerisation of the imidic prolyl bond of Ac-Ala-Xaa-Pro-Ala-Lys-NH_2 in aqueous solution pH 5.0 at 4 °C.

OMe the free energy of activation $\Delta G^{\ddagger}$ was linearly correlated to the amide I vibrational mode of acetyl-Pro-OMe when both parameters were measured in various solvents; allowing the conclusion of a direct proportionality between the hydrogen bond donating capability of the solvent and the ease of isomerisation.[25] In this respect, water and trifluoroethanol cause the highest rotational barriers with an approximately 50 fold rate deceleration compared to toluene. Apparently, lipophilic solvents contribute much to stabilisation of the nonpolar, carbonylamine-like transition state of isomerisation, whereas water stabilises the C–N bond against twisting by preferentially solvating the ground state carbonyl group. Using the same prolyl ester, solvent effects on the isomer ratio proved to be small. However, the presence of the Lewis acid Li^+ (0.2 M) in water-free trifluoroethanol or tetrahydrofuran strongly enhances the amount of *cis* isomer in the presence of an -Ala-Pro- motif,[26] but overall applicability of this Li^+ effect is limited to certain amino acids preceding proline. Generally, metal ions did not influence isomerisation constants of simple amides or protected peptides in water.

Surprisingly, aqueous surfactants and phosphatidyl choline vesicles considerably affect prolyl isomerisation.[27] When monitoring first-order isomerisation kinetics of prolyl isomerisation of oligopeptides with respect to detergent concentration, saturation kinetics appeared with a maximal rate enhancement well above the critical micelle concentration of the detergent (Fig. 6). In fact, micelles were generally able to

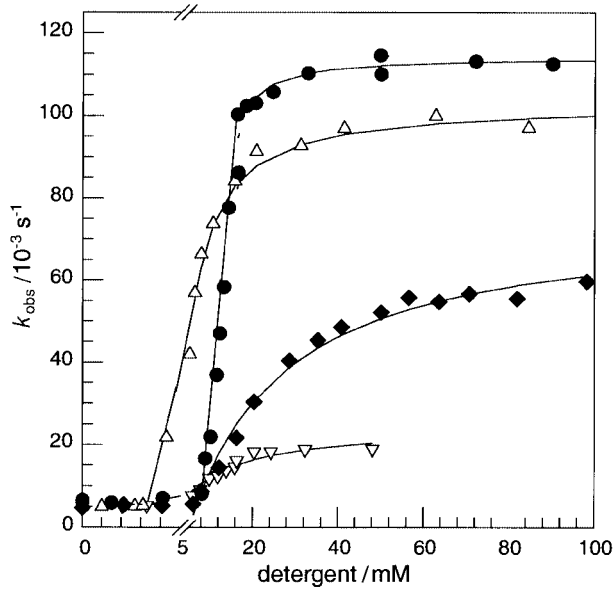


Fig. 6 Dependence of the first-order rate constants k_{obs} of isomerisation on the concentration of different detergents: dodecyltrimethylammonium bromide (●), *N*-dodecyl-*N*,*N*-dimethyl-3-ammoniopropane-1-sulfonate (SB12) (Δ), 3-[(3-cholamidopropyl)dimethylammonio]propane-1-sulfonate (◆), sodium dodecyl sulfate (▽).

accelerate prolyl isomerisations. An up to 23 fold rate enhancement was observed for a combination of the zwitterionic detergent SB12 with a hydrophobic tetrapeptide, and the micelle-induced decrease of the isomerisation barrier was entirely enthalpy-driven. The rate constants k_{mic} determined under saturating conditions did not vary greatly with the charge of the micelles, indicating close contact of the micelle-buried prolyl bond with the hydrocarbon chains of the detergents. Concomitant to the micellar rate effects, the *cis* conformation is generally destabilised relative to the *trans* form. Micelles formed by a variety of detergents average 2.8 ± 0.2 kJ mol^{-1} for the difference in the free energy $\Delta\Delta G^0$ in favour of *trans* isomers. Fig. 7 presents an impressive picture of a concentration jump that allows the simultaneous recording of the decreased isomer ratio and the accelerated isomerisation when a prolyl peptide is sequestered in micelles.

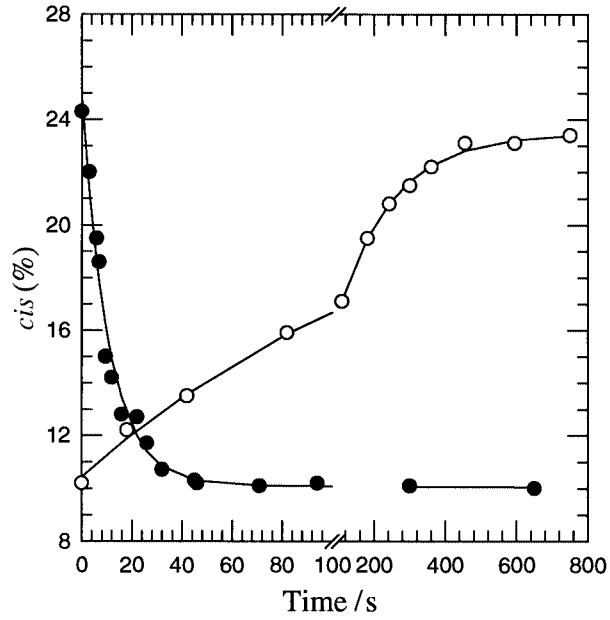


Fig. 7 Dependence of the percentage of the total *cis* isomer content of succinyl-Ala-Phe-Pro-Phe-4-nitroanilide on incubation time at the specified conditions recorded after micellar jumps with *N*-dodecyl-*N*,*N*-dimethyl-3-ammoniopropane-1-sulfonate (SB12). Jumping into the micellar phase at 16 mM detergent concentration (●); jumping out of micelles was performed through eightfold dilution of a solution of 16 mM detergent (o).

The occurrence of a *cis* amidic peptide bond at the Ile1-Lys2 moiety of the 17 mer peptide amide bombolitin dissolved in aqueous sodium dodecyl sulfate, underlines by the dramatic shift of 20 kJ mol^{-1} for ΔG^0, the impact of micelles (and, by analogy, of membranes) on the peptide bond conformation,[28] but indicates also the unpredictability of micellar effects.

5 Catalysis of peptide bond isomerisation

5.1 Chemical catalysis

Formally possible mechanisms for the catalytic acceleration of peptide bond isomerisations are collected in Fig. 8. Among them, specific acid catalysis by the rare *N*-protonation (pK_a about −7) can be easily obtained at very high proton concentration, but this is of course not relevant to biology. The failure of intermolecular general acid/base effects (Fig. 8, route 2) to catalyse peptide bond isomerisation was established by observing the independence of rate constants of buffer composition over a pH range between 3 and 10. Nucleophilic catalysis, which is quite usual for carbon–carbon double bond *cis/trans* isomerisations, has not yet been observed in the nonenzymatic case for peptides (Fig. 8, route 1). The accelerated rate of *cis/trans* isomerisation of amides at high pH values might be interpreted as an indication of the existence of a hemiorthoamide anion intermediate. Catalysis by distortion, which proceeds *via* a stabilised twisted transition state (Fig. 8, route 3), may be caused by prolyl bond desolvation, electrostatic destabilisation or steric strain applied to the ground state of the peptide bond. However, only a moderate increase in the rate of isomerisation was obtained by using specially constructed antibodies that were complementary to a structure with a twisted peptide bond. Haptens were chosen in imitation of the α-dicarbonyl moiety of the macrolide FK506 thought to mimic the perpendicular arrangement in the transition state of isomerisation[29] (Fig. 9). The k_{cat} values of the catalytic antibodies produced are rather similar to those already found for the micellar catalysis described above. Removal of the Cα proton of the amino acid preceding proline will lead to a

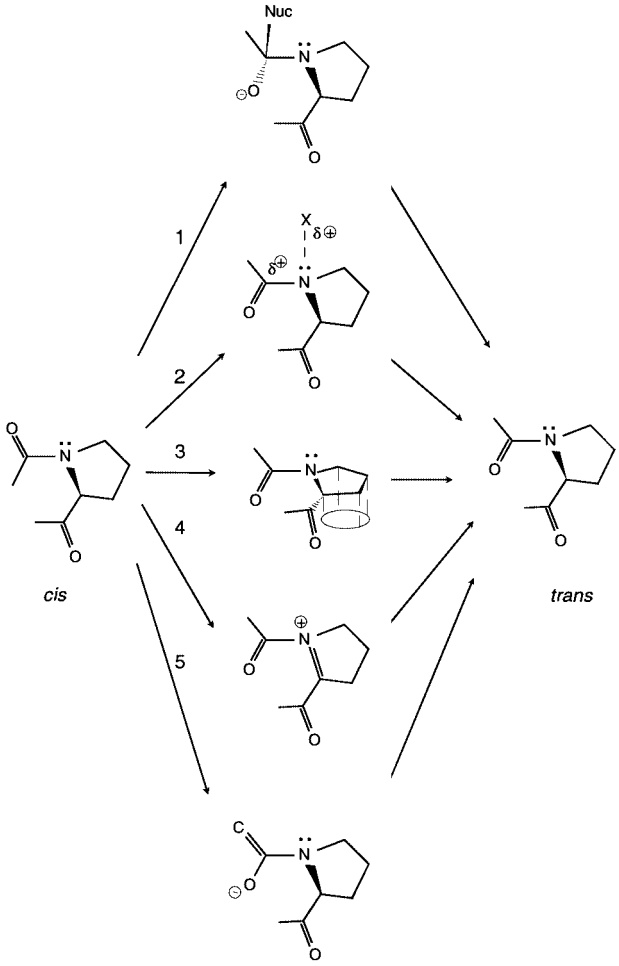


Fig. 8 Hypothetical reaction mechanisms for prolyl *cis* to *trans* isomerisation. They all function by forcing electrons to leave the C–N bond as it is realised in the reaction intermediates or transition states.

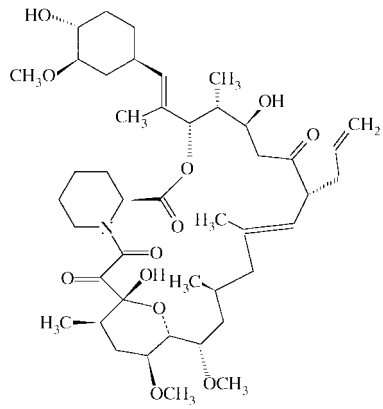


Fig. 9 The peptidomacrolide FK506.

decreased rotational barrier of the peptide bond following the carbanion (Fig. 8, route 4). Proline racemase provides the microenvironment for proton removal through a pair of properly positioned cysteine residues.[38] Catalysis *via* formation of a carbonium ion intermediate (Fig. 8, route 5) appears to be improbable energetically.

Thus, under cellular conditions, intermolecularly acting, nonenzymatic catalysis is difficult to obtain. Rather than intermolecular catalysis, intramolecular assistance of prolyl isomerisation might be operative in polypeptides because of the potential of manifold catalytic groups to approach the reaction center by folding. In principle, all five catalytic routes depicted in Fig. 8 can be utilised but only electrophilic assistance has been detected to date. An intramolecular –NH··$\text{N}_{\text{proline}}$ hydrogen bond has already been proposed to assist prolyl isomerisation by general acid catalysis, utilising the weak proton donating ability of the secondary amide peptide bond succeeding proline in oligopeptides (Fig. 10). *Ab initio* calculations of *N*-

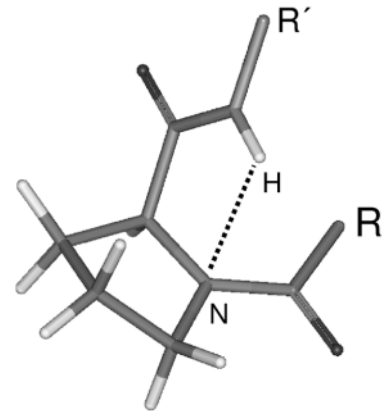


Fig. 10 Intramolecular hydrogen bond thought to stabilise the rotated state of a prolyl bond.

acetylproline methylamide reveal a transition state stabilisation by the –NH··$\text{N}_{\text{proline}}$ hydrogen bond of about 6 kJ mol^{-1}.[9] Due to the competition for the hydrogen bond by the solvent this catalytic route should not contribute much to transition state stabilisation in water. In contrast, the imidazolium side chain of a histidine residue in the context of the -His-Pro- moiety shows a general acid catalysed kinetic term of up to 10 fold of the uncatalysed reaction rate, even in water.[31] In support of the transition state stabilisation by the imidazolium proton, a deuterium solvent isotope effect of the *cis* to *trans* isomerisation rate of $k_{H_2O}/k_{D_2O} = 2.0 \pm 0.1$ for the -His-Pro- isomerisation at pH 3.5 in angiotensin III has been observed. Obviously, these acceleration effects were not caused by constraints applied by secondary or tertiary structures but were determined solely by the amino acid sequence. Thus, it has characteristics of a general effect operating at near physiological pH values in all proteins containing the His-Pro moiety. The stabilised transition state proposed in Fig. 11 only becomes operative with a preceding

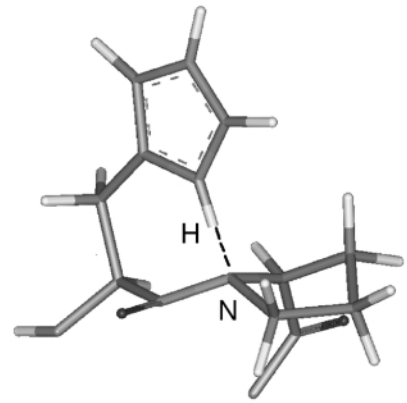


Fig. 11 The intramolecular hydrogen bond plays a critical role in the prolyl isomerisation of protonated -His-Pro- moieties.

histidine, whereas the same residue in the context of the -Pro-His- segment proved to be ineffective. Moreover, a properly positioned arginine side chain, not adjacent in the sequence to proline, was shown to accelerate prolyl isomerisation in native dihydrofolate reductase due to proximal positioning of the interacting bonds.[32] In contrast isomerisation rates appear to be

unaffected by arginine when positioned on either side of proline.

5.2 Enzyme catalysis

Despite the lack of powerful intermolecularly acting chemical catalysts an enzyme catalysing peptide bond isomerisation was discovered in 1984 in pig kidney.[14] It proved to be inactive toward both non-proline *N*-alkyl amino acid moieties and secondary amide peptide bonds but was highly active toward a variety of proline-containing oligopeptides. This new enzyme class was termed peptidyl prolyl *cis/trans* isomerases (abbreviated PPIases; EC number 5.2.1.8), representing the sole example of a biocatalyst directed to a conformational interconversion known to date. Enzymes catalysing prolyl isomerisations evolved early during the course of evolution, and evolution of three different enzyme families took place soon after the progenitor had appeared. According to our current knowledge, these families comprise cyclophilins (abbreviated: Cyp), FK506 binding proteins (abbreviated: FKBPs) and parvulins (abbreviated: Par) that are unrelated in their amino acid sequences to each other, have distinct substrate specificities and prove to be sensitive to different types of inhibitors. They are highly abundant in most tissues and cells, and are found ubiquitously distributed across all organisms. An additional subfamily, the ribosome-bound prokaryotic trigger factors, exhibits weak sequence similarity to FK506-binding proteins but lacks any FK506-binding ability typical of other FKBPs. Currently, more than 300 different amino acid sequences homologous to the three PPIase families have been recorded in the protein sequence databases. Usually these enzymes are composed of a catalytic core, which is also found as a single domain in the prototypic PPIases human Cyp18, human FKBP12 and *E. coli* Par10, and a different number of additional domains or short motifs covering other functional properties. The minimal peptide chain length of substrates for catalysis has been shown to be a tripeptide, and productive binding of peptide chains extends up to polypeptide length. On the other hand, many PPIases essentially require the location of the reactive bond of the substrate in the context of secondary binding sites or a specific spatial organisation of the polypeptide chain, thus creating features of stereo- and regiospecificity. Members of the PPIase families resemble perfectly evolved enzymes regarding catalytic efficiency. For oligopeptide substrates both the Michaelis constants K_m and the turnover numbers k_{cat} turned out to be high ($> 10^4$ s^{-1}), yielding specificity constants $k_{cat}/K_m > 10^7$ M^{-1} s^{-1}. This magnitude of the bimolecular rate constant already approaches diffusion controlled limits for enzyme reactions indicating a highly evolved catalytic machinery. Obviously PPIases applied in catalytic amounts are not able to alter the *cis/trans* equilibrium of the substrates due to the absence of an energy-providing reaction cycle. Unlike chaperones, PPIases are classical enzymes. In a refolding batch reaction suffering from misfolding and protein aggregation, an increase of the yield of native protein was not obtained by supplementing the refolding buffer with the catalytic core domain of PPIases. Additional domains, complementing the catalytic core and conferring chaperone-like properties to PPIases, must be present to obtain a higher yield of correctly refolded protein. Therefore, all prototypic PPIases belong to the category of folding helper enzymes as protein disulfide isomerases do.[33]

Apart from their putative role in accelerating slow folding steps, compartmentalised PPIases are likely to have pleiotropic effects within cells. A promising approach to the elucidation of these effects is based on the identification of proteins intracellularly associated with PPIases. Copurification, affinity chromatography, immunoprecipitation and the yeast two-hybrid screen provide valuable tools for characterising PPIase–protein heterooligomeric complexes. A number of proteinaceous ligands have been identified in this way. Along with other biological properties of PPIases they are collected in review articles.[20,34]

5.3 Mechanism of enzyme catalysis

All reaction routes depicted in Fig. 8 might play a role in enzymatic catalysis because the enzyme protein can provide the catalytic functions required for substrate activation and transition state stabilisation. Many high resolution crystal structures of PPIases complexed to oligopeptide substrates, tight binding reversible inhibitors and binding proteins have now been collected in the Brookhaven database, and have been investigated regarding the enzyme mechanism.[35,36] A major drawback to this approach to uncovering a reliable enzyme mechanism is that substrate-containing crystals may generate dead-end complexes making mechanistic conclusions unreliable. Accordingly, from the structural studies a clear picture of the mechanism of catalysis has not emerged. Moreover, present data do not support a common mechanism for all PPIase families. Of particular interest, in view of the reaction routes collected in Fig. 8, is that the parvulin-like PPIase human Pin1 shows clear evidence for a nucleophilic mechanism involving a cysteine SH group-assisted rehybridisation toward sp^3 character of the prolyl bond carbonyl group in the Pin1–Ala-Pro complex.[37] On the other hand cyclophilins may benefit from general acid catalysis because site-directed mutagenesis of Arg55Ala inactivates the human Cyp18 enzyme, and secondary β-deuterium effects of enzymatic rate constants did not support a covalent intermediate (Fig. 8, route 1) on the catalytic pathway. In addition, crystal structures revealed a sandwich-like arrangement of the prolyl ring between the Arg55 side chain and His126. Both residues are within hydrogen bonding distance of the prolyl bond nitrogen of a tetrapeptide substrate, which makes this residue a suitable candidate for electrophilic assistance.

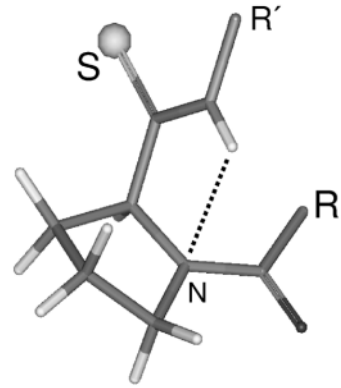


Fig. 12 Intramolecular hydrogen bond of the -Xaa-Pro-ψ[CS-NH]- moiety.

Transition-state stabilisation of the enzyme-bound substrate by the intramolecular hydrogen bond shown in Fig. 10 has been suggested to enhance enzymatic isomerisation rates because *ab initio* calculations in conjunction with molecular mechanics result in a type VIa β-turn conformation of a substrate chain bound to human FKBP12.[30] The tendency to form the $NH\cdot\cdot N_{proline}$ hydrogen bond is intrinsic to that type of VIa β-turn conformation. However, the acidification of the relevant NH group in the thioxopeptide derivative shown in Fig. 12, which should lead to an enhanced transition state stabilisation by H-bond strengthening, could increase isomerisation rates neither in the uncatalysed reaction in water nor in FKBP12 catalysis.

Current models of PPIase catalysis suggest a stepwise release of catalytic power, including a minor contribution from the binding of the desolvated transition state to the hydrophobic cavity formed by the active site of cyclophilins and FKBPs. PPIase variants produced by site-directed mutagenesis often have a small enzyme activity of 0.1–2% of wild type enzyme which is related to this part of the catalytic power. This contribution might be mimicked by catalytic antibody and micellar catalysis. However, the major part of catalysis may arise either from electrophilic assistance or covalent intermediates tetrahedrally ligated to the enzyme protein depending on the PPIase family considered. Only specific substrates correctly recruiting all secondary binding sites provided by the enzyme may be able to make the whole catalytic machinery active. In fact oligopeptides containing side-chain phosphorylated Ser/Thr-Pro moieties produce a certain pH profile at high k_{cat}/K_m values with the parvulin-like Pin1 PPIase whereas the same enzyme showed a perturbed pH profile at a much lower level of catalysis with the unspecific -Glu-Pro- or -Ala-Pro peptides.[22]

6 Outlook

Two-state behaviour of isomerising peptide bonds provides a switch to a polypeptide chain that has the potential to control slow interconversion dynamics of the protein backbone by means of amino acid sequence, three-dimensional structure or biocatalysis. Actuating the switch causes backbone rearrangements around the isomerising bond. Regarding *cis/trans* isomerisations prolyl bonds are of special importance because of the high propensity for the *cis* state in unfolded polypeptides and native proteins.

For oligopeptides some reliable methods can now be used to determine the isomer ratio and the isomerisation rates as well as the isomer-specificity of biochemical processes *in vivo* and *in vitro*. However, attempts to detect and to understand native state isomerisations of proteins are still at a preliminary stage. Much development in this field has been initiated by the discovery of three families of *cis/trans* isomerases all of which are known to catalyse specifically prolyl isomerisations. It remains an open question whether nature evolved additional peptide bond isomerases, but prolyl substrate libraries might offer the opportunity to discover novel peptidyl prolyl *cis/trans* isomerases. Low molecular weight compounds that can discriminate between *cis* and *trans* isomers of polypeptides have not been developed but might be useful for disrupting proline-mediated interactions of protein domains like the WW and SH3 domain in cell signalling. Similarly, small molecules effective in catalysing peptide bond isomerisations under physiological conditions may greatly improve our understanding of the role of this process in biology but are still lacking.

7 Acknowledgements

I would like to acknowledge Dr Ulf Reimer, Dr Cordelia Schiene and Jason Smith for fruitful discussions, critical reading of the manuscript and preparing figures.

8 References

1 L. Pauling, *The Nature of the Chemical Bond*, Cornell University Press, Cornell, 1948.
2 D. Lauvergnat and P. C. Hiberty, *J. Am. Chem. Soc.*, 1997, **119**, 9478.
3 N. J. Bennet, Q.-P. Wang, H. Slebocka-Tilk, Y. Somayaij, R. S. Brown and D. Santarsiero, *J. Am. Chem. Soc.*, 1990, **112**, 6383.
4 K. B. Wiberg, P. R. Rablen, D. J. Rush and T. A. Keith, *J. Am. Chem. Soc.*, 1995, **117**, 4261.
5 A. Rauk and S. A. Glover, *J. Org. Chem.*, 1996, **61**, 2337.
6 C. H. Yoder and R. D. Gardner, *J. Org. Chem.*, 1981, **46**, 64.
7 A. Radzicka, L. Pedersen and R. Wolfenden, *Biochemistry*, 1988, **27**, 4538.
8 P. S. Li, X. G. Chen, E. Shulin, S. and A. Asher, *J. Am. Chem. Soc.*, 1997, **119**, 1116.
9 S. Fischer, R. L. Dunbrack and M. Karplus, *J. Am. Chem. Soc.*, 1994, **116**, 11931.
10 T. Drakenberg and S. Forsen, *J. Chem. Soc., Chem. Commun.*, 1971, 1404.
11 G. Scherer, M. L. Kramer, M. Schutkowski, U. Reimer and G. Fischer, *J. Am. Chem. Soc.*, 1998, **120**, 5568.
12 D. Kern, G. Kern, G. Scherer, G. Fischer and T. Drakenberg, *Biochemistry*, 1995, **34**, 13594.
13 F. X. Schmid, R. Grafl, A. Wrba and J. J. Beintema, *Proc. Natl. Acad. Sci. USA*, 1986, **83**, 872.
14 G. Fischer, H. Bang and C. Mech, *Biomed. Biochim. Acta*, 1984, **43**, 1101.
15 B. L. Stoddard and S. Pietrokovski, *Nat. Struct. Biol.*, 1998, **5**, 3.
16 J. F. Brandts, H. R. Halvorson and M. Brennan, *Biochemistry*, 1975, **14**, 4953.
17 C. Grathwohl and K. Wuthrich, *Biopolymers*, 1976, **15**, 2043.
18 U. Reimer, G. Scherer, M. Drewello, S. Kruber, M. Schutkowski and G. Fischer, *J. Mol. Biol.*, 1998, **279**, 449.
19 F. X. Schmid, "Protein folding" in: *Protein: A Comprehensive Treatise*, G. Allen, Ed., JAI Press Inc., Greenwich, CT, 1999, Vol. 2, pp. 153–215.
20 G. Fischer, *Angew. Chem., Int. Ed. Engl.*, 1994, **33**, 1415.
21 C. Odefey, L. M. Mayr and F. X. Schmid, *J. Mol. Biol.*, 1995, **245**, 69.
22 M. Schutkowski, A. Bernhardt, X. Z. Zhou, M. H. Shen, U. Reimer, J. U. Rahfeld, K. P. Lu and G. Fischer, *Biochemistry*, 1998, **37**, 5566.
23 E. S. Eberhardt, N. Panasik and R. T. Raines, *J. Am. Chem. Soc.*, 1996, **118**, 12261.
24 R. K. Harrison and R. L. Stein, *J. Am. Chem. Soc.*, 1992, **114**, 3464.
25 E. S. Eberhardt, S. N. Loh, A. P. Hinck and R. T. Raines, *J. Am. Chem. Soc.*, 1992, **114**, 5437.
26 J. L. Kofron, P. Kuzmic, V. Kishore, E. Colon-Bonilla and D. H. Rich, *Biochemistry*, 1991, **30**, 6127.
27 M. L. Kramer and G. Fischer, *Biopolymers*, 1997, **42**, 49.
28 E. Bairaktari, D. F. Mierke, S. Mammi and E. Peggion, *J. Am. Chem. Soc.*, 1990, **112**, 5383.
29 J. T. Yli-Kauhaluoma, J. A. Ashley, C. H. L. Lo, J. Coakley, P. Wirsching and D. Janda, *J. Am. Chem. Soc.*, 1996, **118**, 5496.
30 S. Fischer, S. Michnick and M. Karplus, *Biochemistry*, 1993, **32**, 13 830.
31 U. Reimer, N. Elmokdad, M. Schutkowski and G. Fischer, *Biochemistry*, 1997, **36**, 13802.
32 F. L. Texter, D. B. Spencer, R. Rosenstein and C. R. Matthews, *Biochemistry*, 1992, **31**, 5687.
33 M. J. Gething and J. Sambrook, *Nature*, 1992, **355**, 33.
34 A. Galat and S. M. Metcalfe, *Prog. Biophys. Mol. Biol.*, 1995, **63**, 67.
35 Y. D. Zhao and H. M. Ke, *Biochemistry*, 1996, **35**, 7362.
36 G. D. Van Duyne, R. F. Standaert, P. A. Karplus, S. L. Schreiber and J. Clardy, *J. Mol. Biol.*, 1993, **229**, 105.
37 R. Ranganathan, K. P. Lu, T. Hunter and J. P. Noel, *Cell*, 1997, **89**, 875.
38 W. J. Albery and J. R. Knowles, *Biochemistry*, 1986, **25**, 2572.

Review a803742f