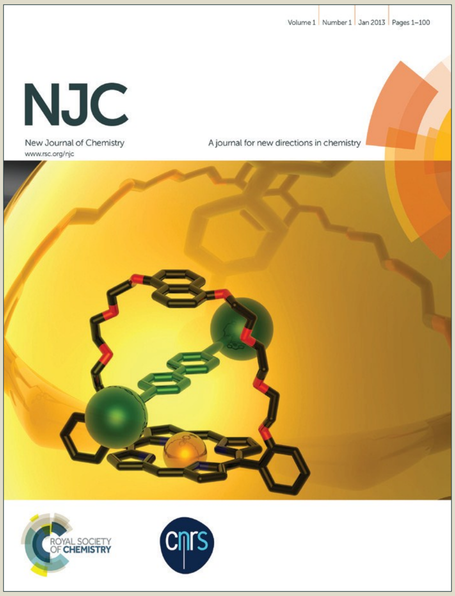
CrossMark
click for updates

View Article Online
View Journal

NJC

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: V. Kubyshkin, P. Durkin and N. Budisa, *New J. Chem.*, 2016, DOI: 10.1039/C5NJ03611A.



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/njc

View Article Online
DOI: 10.1039/C5NJ03611A

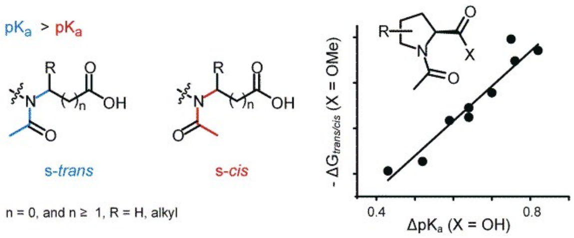
Energetic contribution to both acidity and conformational stability in the peptide models

Vladimir Kubyshkin,* Patrick Durkin and Nediljko Budisa*

Biocatalysis group, Institute of Chemistry, Technical University of Berlin, Müller-Breslau-Str., 10, 10623, Berlin, Germany

E-mails: kubyshkin@win.tu-berlin.de; nediljko.budisa@tu-berlin.de

TOC entry: The acidity difference of the amide rotamers has been revised for a large set of *N*-acetyl amino acids.



Abstract: Acidity of *N*-acyl amino acids is dependent upon the rotameric state of the amide bond. In this work we systematically investigated the acidity difference of the rotamers (ΔpK_a) in the frames of various acetylated amino acids. Our results indicated a mutual interaction of two carbonyl groups of an attractive type. We observed conservative ΔpK_as for acyclic amino acids (2.2-3.0 kJ/mol), whereas in case of alicyclic amino acids, the experimental values revealed strong dependency on the structural

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

context (1.5-4.4 kJ/mol). In homologous amino acids (α-, β-, γ-, etc.), the strength of the attraction decays in an exponential fashion. Furthermore, the interaction can accumulate through a chain of amide bonds in a cascade fashion, as demonstrated by Ac-Pro-Pro dipeptide. As a result, we demonstrate that ΔpK_a is an experimental parameter to estimate increments in the carbonyl-carbonyl alignment, as determined by the amino acid or peptidyl context. This parameter is also important in understanding the roles of amino acids in both protein folding and translation in biological systems as well as their evolutionary appearance in the genetic code.

New Journal of Chemistry Accepted Manuscript

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

1. Introduction

There are a number of non-covalent interactions that govern the folding of polypeptides such as hydrogen bonds, electrostatic attraction (salt bridges) and repulsion, various aromatic stacking interactions (π-π, cation-π, aliphatic-π, aromatic-Pro or aromatic-Met), dipolar interactions and the $n\rightarrow\pi^*$ donation. Some of these, like hydrogen bonding, are very well described in chemistry and biology as their distances and impact are easy to observe by conventional methods such as X-ray crystallography, NMR, CD, IR etc. On the other hand, there are structural motifs, such as the polyproline II helix, where the driving forces behind this organization are more difficult to understand, rationalize and follow experimentally. The left-handed polyproline II (PPII) helix was initially described in 1956[1] and later found to be a ubiquitous structure, not only in fibrillar structures like collagen, but also in many folded and unfolded proteins.[2,3,4] Conversely, the right-handed polyproline I helix (PPI) only has limited biological relevance.[5] Although proline has high PPII propensity, this topology can also be generated in the absence of proline residues, such as in antifreeze glycoproteins.[6]

Clearly, PPII is an essential secondary structure of proteins that participates in major vital biological functions (transcription, signal transduction, ligand-receptor interplay, protein-protein interactions etc.).[7] Thus, the fundamental goal is to understand and (if possible) to quantify the nature and extent of the forces behind its stability. This is particularly relevant for the canonical secondary amino acid proline which has a comparable energy of the s-*trans* and s-*cis* peptide bonds, therefore exhibiting an equilibrium of both peptide bond isomers in solution (Scheme 1).[8] Equally important is

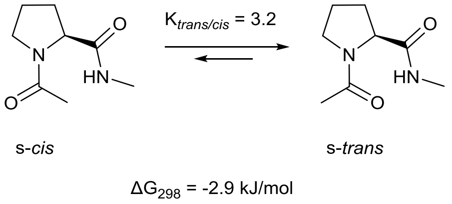
View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

the amount of energy required for the stabilization of this structure, since entropy (at least to a certain extent) should oppose the formation of an ordered structure.

One possible rationalization for this stabilization is the n→π* donation, a donative effect established between the oxygen lone pair of an upstream amide bond to the carbon atom of the downstream carbonyl. Since it represents a form of a nucleophilic attack, it should therefore be enhanced by certain dihedral angles that optimize the angle of donation (Bürgi-Dunitz trajectory). An essential prerequisite for n→π* donation is the s-*trans* conformation (ω = 180 °) of the donating amide (Fig. 1).



Scheme 1. *trans/cis* equilibrium in Ac-Pro-NHMe. The s-*trans* rotamer is preferred over the s-*cis* by approx. - 3 kJ/mol.

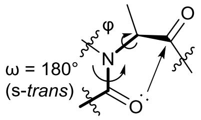


Figure 1. The n→π* interaction in a polypeptide chain.

Once established, the n→π* donation results in 1) an energetic stabilization of particular secondary structural motifs,[9] 2) stabilization of the *trans*-amide bond over the *cis*-isomer,[10] and 3) partial pyramidalization of the carbonyl carbon atom.[11] It has also been argued, however, that a classical dipole-dipole interaction might be mistaken as being an n→π* interaction.[12] Recently, a crystal structure of the PPII helix demonstrated,

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

on the contrary, the existence of a certain degree of pyramidalization of proline carbonyls, which could originate from a specific donation.[13] For a proline residue, indeed, the n→π* interaction can play a significant role, since its ϕ angle (-65±10) is naturally fixed around a value optimal for this donation (-63±6).[9] In proline-rich structures, such as polyproline or collagen, which lack other types of non-covalent interactions, the accumulation of the n→π* donation through a peptide chain may give rise to a significant stabilizing effect due to entropic pre-organization.[14]

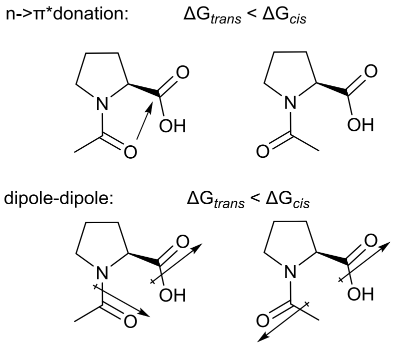


Figure 2. Both n→π* donation and dipole-dipole interaction will result in relative stabilization of the s-*trans* amide (left). Arbitrary orientation of the C-terminal dipole is given for the plane projection here (see also Fig. 14).

A critical question is whether the energetic contribution of this interaction can be of any significance in terms of a single residue. From a chemical standpoint, it appears to be doubtful, that the amide oxygen can donate enough electron density to enable a significant structural stabilization. An alternative explanation might be the dipole-dipole interaction established between the two carbonyl groups. Both effects, n→π* and the dipole-dipole interactions, may potentially result in the relative stabilization of similar

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

dihedral angles in a proline residue (in particular, the *trans*-amide bond), and from this standpoint they cannot be distinguished (Fig. 2).

Further, by using proline analogues containing polar substituents (such as fluorine or hydroxyl-group), additional interacting dipoles will be present which are capable of participating in the overall interaction. Alternatively, in this case the conformational preferences can be explained by the geometric pre-organization of the trajectory between the interacting peptide bonds due to the changes in the pyrrolidine conformational states (*exo*/*endo* pucker equilibrium). As a result, the *trans/cis* equilibrium is known to experience dramatic variations.[15]

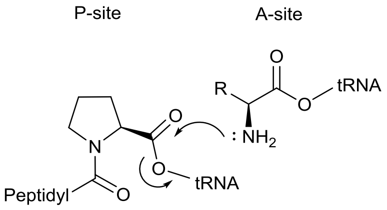


Figure 3. Peptidyl transfer reaction with peptidyl-Pro on the P-site of the ribosome.

The final outcome of this mutual interaction of the polar groups must result in a change of the electronic state of the interacting groups, and hence their chemical reactivities. The latter conclusion is particularly interesting, since the ribosomal translation of proline is also known to be slow and may be as a result of the above mentioned interactions, seen as a donative effect. Furthermore, oligoproline segments in polypeptide sequences slow down the whole translation process, and can induce ribosome stalling,[16] in particular when accommodated on the P-site.[17] Since the C-terminal group in peptidyl-tRNA on the P-site should be attacked by an amino group in

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

the course of a peptidyl transfer reaction on the ribosome (Fig. 3), the interactions that will reduce the electrophilicity of the carbonyl group may also slow down the rate of the reaction. Supporting this suggestion, we previously reported dramatic differences in the translation efficiency in living cells for (4*R*)- and (4*S*)-fluoro- and hydroxyprolines.[18]

To determine a possible logical link between the chemical reactivity and the folding energy, we turned our attention to the acidity of the terminal carboxyl in *N*-acetyl proline and other amino acid residues. Indeed, it has been discovered in the mid-1970^{s} that *N*-acylated proline and sarcosine (Sar) exhibit different acidities of their amide rotameric forms, where the s-*trans* rotamer is less acidic than s-*cis* (Fig. 4).[19, 20] *N*-acylated Pro and Sar were convenient subjects for the studies, since their rotameric populations are of comparable intensity in the NMR spectra, where sensitivity is a general problem. In the times when robust water suppression techniques were still absent,[21] the pK_a for the rotameric states were determined by means of ^{13}C, ^{15}N and ^{17}O NMR, delivering results with a certain degree of discrepancy.[19,22,23]

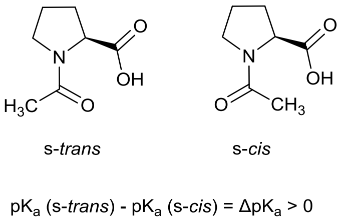


Figure 4. Acidity difference characterized for *N*-acetyl proline (Ac-Pro).

These initial studies revealed, among others, that the difference of the acidity of the two rotamers (ΔpK_a) is defined by the mutual interaction of the amide and carboxyl-groups. The origin of ΔpK_a therefore can be seen in the following way: if the n→π*

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

donation takes place in Ac-Pro, the electrophilicity of the terminal carbonyl group is reduced, but only in the case when the peptide bond is in the s-*trans* conformation.[10b,24] From that, it follows that the resulting acidity of the s-*trans* rotamer should be then lower than in the s-*cis* ($\Delta pK_a > 0$). Consequently, if the dipole-dipole interaction contributes to stabilization of the s-*trans* rotameric form, the effect should be stronger when the terminal group is a carboxylic acid (large dipole) than salt (small dipole). Eventually the effect should result in a reduced acidity of the *trans*-rotamer ($\Delta pK_a > 0$). In a similar fashion to the conformational considerations, both interactions predict the same outcome for the acidity as well.

In this context, we set out to study the acidity differences in various *N*-acetylated amino acids in order to provide an experimental set of parameters for understanding the carbonyl-group alignment in terms of the energy of the interaction. Additionally, we also aimed to provide a solid experimental parameter that would make a logical link between the conformational energy and the chemical reactivity. However, the reader should be fully aware to draw their own conclusion from the present study concerning the molecular origins of the discussed energetic term.

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

2. Experimental Section

2.1. *N*-acylation of the amino acids

Typical procedure for the synthesis of *N*-acetyl proline:

L-proline (2.26 g; 19.7 mmol, 1 equiv.) was stirred with anhydrous dichloromethane (40 ml) and acetic anhydride (1.85 ml; 19.6 mmol, 1 equiv.) until a clear solution was obtained (30 min). Dichloromethane was removed under reduced pressure and the residue was dissolved in water (20 ml) and freeze-dried. Resulting solid contained approx. 1/10 of unreacted proline, therefore it was dissolved in water (25 ml) and filtered through a short ion-exchange column (Dowex® 50WX8, 50-100 mesh). Acidic fractions were collected and freeze-dried to give 2.79 g of *N*-acetyl proline as white solid (91 % yield).

For analytical characterization of the substances **11**, **12**, **15**, **16**, **17**, **18**, **19** and **20** see ESI. For the already known compounds, **1**, **2**, **3**, **7**, and **22** spectral data are also given in the ESI.

2.2. Determination of the carboxyl pK_a in Ac-Xaa

Typical procedure was the following:

Ac-Xaa (5 mg) and potassium dihydrogenphosphate (5-6 mg) were dissolved in 10 ml of water and titrated by potassium hydroxide and hydrochloric acid solutions to different pH values ranging from 7 to 1. 12-15 aliquots with 500 μl volume were taken to NMR

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

tubes, followed by addition of 50 µl TPS (sodium 3-(trimethylsilyl)propane-1-sulfonate) solution in deuterium oxide (0.1 mg/ml) to each tube for lock and referencing. Final concentration in the samples was: Ac-Xaa 2-4 mM, phosphate 7-8 mM, TPS 0.1 mM. ^{1}H NMR spectra were measured at either 500 or 700 MHz frequency in Watergate W5 pulse sequence with the calibrated pulses. For pH < 2 pulses were re-calibrated for each sample due to the increased ionic strength of the solutions. TPS most upfield resonance was set to zero, except of the spectra for the most acidic samples (pH≤1.5), for which TPS resonances deviated due to partial protonation of the sulfonic group. In that case we used any visible additive in the spectrum, for instance, methanol resonance at 3.349 ppm for the referencing. Chemical shifts were read out and plotted against pH. In particular, acetyl CH_3-group resonance was taken with three digits after the decimal in the chemical shift value. In this case s-*trans*-chemical shift was higher than s-*cis*, with the only exception of Ac-MeAla. Chemical shift *vs* pH curves were fitted according to either Boltzman or Logistic fits (OriginPro 9.1), the fits were derivatized and the global maximum of the derivative curves was read as the pK_a value. The derivative curves were checked to exhibit a symmetric shape.

In particular, by this procedure we determined pK_a values for Ac-Sar, Ac-Aze, Ac-Pip. For fluorine containing Ac-Xaa also ^{19}F{^{1}H} NMR spectra were recorded in inverse-gated decoupling experiments (no NOE enhancement).

In addition, for Ac-Pro we determined pK_a values under the following conditions:

Ac-Pro ≤ 0.5 mM / citric buffer 100 mM

pK_a s-*cis* 2.79; s-*trans* 3.54

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

Ac-Pro 25 mM / no buffer

pK_a s-*cis* 2.79; s-*trans* 3.56

Where it was seen, that the pK_a values (in particular, for the s-*trans* form, which is more reliable in detection) are invariant to the offered conditions of the analyte or buffer concentrations.

For Ac-Xaa with the secondary amide (Ac-Gly, Ac-Val, Ac-Ile) we used higher concentrations of the analytes (25, 40 or 100 mM) in order to detect the minor rotameric form. Examples also see in ESI, section 4.

2.3. Equilibrium populations in acids and salts

We found that 140-170 mM solution of potassium hydrogen sulphate solution delivers pH values of 1.4-1.3, also, when some amount of Ac-Pro is present in solution. Therefore the acidic samples were prepared as following. Ac-Xaa (5 mg) and potassium hydrogen sulphate (10-12 mg) were freeze-dried from deuterium oxide (500 µl) and then dissolved in deuterium oxide (525 µl), and this sample was then used for the measurements. Final concentration of Ac-Xaa was therefore 40-70 mM.

Salt samples were prepared as following. Ac-Xaa (10 mg) and potassium dihydrogenphosphate (10-12 mg) were dissolved in water (1 ml) and titrated with potassium hydroxide solution to pH 6.9-7.1. An aliquot of this solution (525 µl) was freeze-dried, freeze-dried from deuterium oxide (2 x 500 µl) and taken into deuterium oxide (525 µl) for the NMR measurements. Final phosphate concentration was therefore

New Journal of Chemistry Accepted Manuscript

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

70-80 mM, Ac-Xaa concentration 40-70 mM. Due to the slow amide rotation is salts, it turned out to be essential that some samples should be pre-equilibrated for at least one day standing at the room temperature prior the measurements.

^{1}H NMR intensities were measured at 700 MHz frequency in 90-pulse experiments in one scan in order to ensure complete pre-relaxation of the nuclei. In particular, the most characteristic α-CH proton resonances characterized with the longest T_1 relaxation times, and s-*trans* resonance relaxed slower than s-*cis*.

2.4. Thermodynamic analysis

Thermodynamic analysis was performed for the salt sample of Ac-Pro and Ac-β-homo-Pro, prepared as described above, and for *N*-acetyl-2-methylpyrrolidine sample (see ESI). Equilibrium populations were measured by ^{1}H NMR at different temperatures. Temperature heating and cooling series were performed in the range 295-355 K. Variable temperature unit was calibrated by standard methanol and glycerol sample measurements.[25]

For the final analysis we collected data points from several series and plotted them in coordinates $\Delta G_{trans/cis}$ versus T, rather than Van ‘t Hoff plot coordinates. Linear fit delivered the values of ΔH and ΔS.

2.5. Rotational barriers in different solvents

Ac-Pro-OK-18-crown-6 complex was prepared as following: Ac-Pro was titrated by potassium hydroxide solution in the presence of a small potassium

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

dihydrogenphosphate additive to pH 7. Followed by addition of equimolar amount of 18-crown-6 in acetonitrile-water the mixture was freeze-dried. Obtained solid contained Ac-Pro-OK / 18-crown-6 / H_2O in 1:1:2 ratio according to the ^1H NMR spectra.

Equilibrium constants for Ac-Pro-OMe and Ac-Pro-OK·18-crown-6·$2H_2O$ complex in perdeuterated solvents were measured by ^1H NMR at 298 K at 700 MHz resonance frequency.

Rotational rates were measured in ^1H 2D cross relaxation experiments with gradients (NOESY-EXSY) with typical mixing time of 1-2 s for exchange detection and 5 ms for referencing. Recycling delay was chosen to be 3-5·T_1 of the nucleus with slowest relaxation (usually, α-CH of the s-*trans* rotamer). The latter was estimated in 1D inversion recovery experiments for each compound/solvent combination. Since relevant spectral windows varied with the solvents, time domain was inset to give 2-4 Hz resolution in the direct and 10-30 Hz resolution in the indirect dimensions, respectively. On the processing step the indirect dimension was zero filled to reach the initial direct dimension resolution. 2D frequency domain spectra were phased, baseline corrected and integrated. Exchange rate matrices were calculated by EXSYCalc freeware (Mestrec). Activation energies were calculated using Eyring equation. In addition to the exchange quantification, the EXSY spectra were interpreted as NOESY for assignment of the rotameric forms.

2.6. Equilibrium populations and pK_a determination in Ac-Pro-Pro

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

Ac-Pro-Pro dipeptide was synthesized both on solid support and in solution synthesis as described in ESI. For the pK_a measurements we followed pH dependences of the most characteristic α-CH resonances. Since the spectra contained eight such resonances (2 proline residues x 4 rotameric forms) in very different intensity close to the water signal, and some of them overlapped, the measurements of all four pK_a values were conducted in a series of 25 $^1H\{^{13}C\}$ HSQC spectra (cnst 150 Hz, recycling delay 2 s) assisted by ^1H Watergate W5 and $^{13}C\{^1H\}$ dept45 experiments on samples with pH between 1 and 7. The spectra windows were zoomed to the α-CH region in both dimensions. The sample concentrations were: 100 mM Ac-Pro-Pro and 14 mM phosphate. Partial overlap of the signals was observed in ^1H and ^{13}C spectra, but simultaneous overlap in both dimensions was seldom. The titration curves and the pK_a values read out were performed as described above for Ac-Xaa (also, see ESI).

For the ^1H NMR measurements of the equilibrium populations the acidic and salt samples in deuterium oxide were prepared as described for Ac-Xaa. The samples contained 75 mM Ac-Pro-Pro and 140 mM potassium hydrogen sulphate or 140 mM potassium phosphate buffer for acid and salt respectively. ^1H NMR spectra were measured at 700 MHz resonance frequency as described for Ac-Xaa. Quantified $^{13}C\{^1H\}$ spectra were acquired in samples containing 250 and 350 mM Ac-Pro-Pro for the acid and salt respectively. The spectra were recorded in 90-pulse inverse gated ^1H decoupled pulse sequence at 126 MHz resonance frequency. Recycling delay was 60-80 s to ensure complete relaxation. 512-768 transients were enough to collect spectra with all four rotamers visible. It is important to note, that the α-CH intensity ratios obtained in quantified $^{13}C\{^1H\}$ NMR spectra were in a good agreement (within ±1 %

View Article Online
DOI: 10.1039/C5NJ03611A

accuracy) with those obtained in $^{13}C\{^{1}H\}$ dept45 spectra with routine setup (recycling delay 2 s, cnst 150 Hz). This observation justifies our approach of the pK_a measurements conducted by HSQC, where the averaged populations of salt/acid upon the pK_a transitions should deliver correct average of the chemical shifts according to their molar ratios.

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

3. Results and Discussion

3.1. *N*-terminal amide *cis*/*trans* rotameric ratio as a function of the C-terminal group

It has been previously demonstrated that *trans*-content in *N*-acyl proline derivatives is dependent from the nature of the C-terminal group. Generally, $K_{trans/cis}$ is higher in the case of the terminal ester compared to the amide due to the differences in the enthalpic term.[26] Further evaluation (Table 1) demonstrates that the preferences towards the s-*trans* conformation gradually decreases in the row X = OMe > OH > NHMe > O^-, i.e. with the decrease in the carbonyl electrophilicity (decrease in the C=O dipole). The same occurs in the case when two C-terminal groups are integrated in the frame of one molecule.

Table 1. Energetic terms of the amide equilibrium for Ac-Pro-X.

Ac-Pro-X	$K_{trans/cis}$ [a)]	ΔG_{298}, kJ/mol	ΔH, kJ/mol	ΔS, J/mol·K	k,[b)] s^{-1}
X = OMe[27]	4.8	- 3.9±0.1	- 5.04 ±0.05	- 3.82 ±0.16	0.84
X = OH	3.8	- 3.3±0.1	-	-	2.4[c)]
X = NHMe[8]	3.2	- 2.9±0.1	- 3.97 ±0.04	- 3.64 ±0.08	0.81
X = O^-	0.81	+ 0.5±0.2	- 1.17 ±0.10	- 5.24 ±0.30	0.075

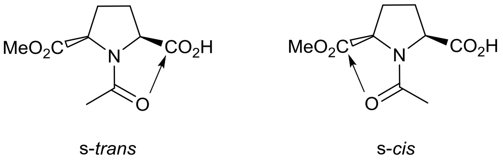
[a)] in aqueous medium at 298 K; [b)] *cis*-to-*trans* rotation rate at 340 K; [c)] the value is the largest because acidic medium catalyzes rotation at high temperature due to the partial amide protonation.

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

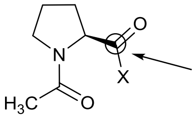
New Journal of Chemistry Accepted Manuscript

In particular, for racemic *N*-acetyl-pyrrolidine-2,5-dicarboxylic acid mono methyl ester (Ac-(MeO)Pdc) the $K_{trans/cis}$ values were predicted from the proline data as $K_{trans/cis}$(Ac-Pro-OH/O$^-$)/$K_{trans/cis}$(Ac-Pro-OMe) are in good agreement with the experimental values (Fig. 5; see also Fig. S61 for comparison in the entire pH range).



$K_{trans/cis}$	experimental	predicted
at pH 1.5 (acid)	0.81	0.78
at pH 7 (salt)	0.18	0.17

Figure 5. In *N*-acetyl-*rac*-(*trans*-)pyrrolidine-2,5-carboxylic acid mono methyl ester, Ac-(MeO)Pdc, two carboxylates compete for the donation from the amide carbonyl. Resulting *trans/cis* equilibrium constant can be predicted from the data on proline.



charge	when X =
+0.25	OMe
+0.24	OH
+0.10	NHMe
+0.02	O^-

Figure 6. Effective charge on the C-terminal carbon atom as dependant from the nature of the C-terminal group.[a)]

[a)] as simulated in MarvinSketch 15.4.27 (ChemAxon Ltd.)

In the terminal case when the C-terminal group is a carboxylate, the effective charge on the carbonyl carbon is close to zero (Fig. 6). Therefore it can be expected that in the salts both n→π* and the dipole-dipole interactions are substantially attenuated, and the *trans/cis* equilibrium is defined by the steric effects. Subtraction of

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

the $\Delta G_{trans/cis}$ values yields the energetic term ΔΔG which accounts for the effect of the peptide bond alignment due to the n→π* and dipole-dipole interactions (eq. 1). Instrumentally this value was experimentally determined by NMR in two different ways as described in the next section.

$$\Delta\Delta G = \Delta G_{trans/cis}(\text{acid}) - \Delta G_{trans/cis}(\text{salt}) \qquad \text{(eq. 1)}$$

3.2. pK_a and rotameric populations in Ac-Xaa

The experimental approach is illustrated in Fig. 7. Because the acid-salt transition is fast on the NMR time scale, the mixture of the two ionization states provides an averaged chemical shift value, whereas the slow amide rotation enables one to observe the two rotamers separately (Fig. 7A). While going from basic to acidic pH values, the chemical shifts exhibit characteristic sigmoidal transitions (Fig. 7B), which is also accompanied by the change in the rotameric populations (Fig. 7C). The difference in the pK_a exhibited by the two rotamers (ΔpK_a) corresponds to the free energy difference ΔΔG = ΔG°(A) - ΔG°(B) - ΔG°(C) + ΔG°(D), which can also be calculated from the equilibrium populations according to eq. 2.

$$\Delta\Delta G = 2.30 \cdot RT\Delta pK_a = -RT \ln \frac{K_{trans/cis}(acid)}{K_{trans/cis}(salt)} \qquad \text{(eq. 2)}$$

Since ΔΔG and ΔpK_a values are equivalent, we can also convert equilibrium constants to the ΔpK_a^* values expressed according to eq. 3.

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

$$\Delta pK_a{}^* = \log_{10}\frac{K_{trans/cis}(acid)}{K_{trans/cis}(salt)} \quad \text{(eq. 3)}$$

Following these considerations, we determined the experimental pK_a and $\Delta pK_a{}^*$ parameters for 19 different *N*-acetyl amino acids with different structures (Fig. 8). In Table 2 we summarize the obtained values. The ΔpK_a and $\Delta pK_a{}^*$ values found were in close agreement, providing a good mutual control. Doubtless, an important observation is that ΔpK_a exhibited positive values (s-*trans* rotamer is less acidic) in all studied cases. In addition, the ΔpK_a and $\Delta pK_a{}^*$ values for the Pdc derivative were in full agreement with the proline data.

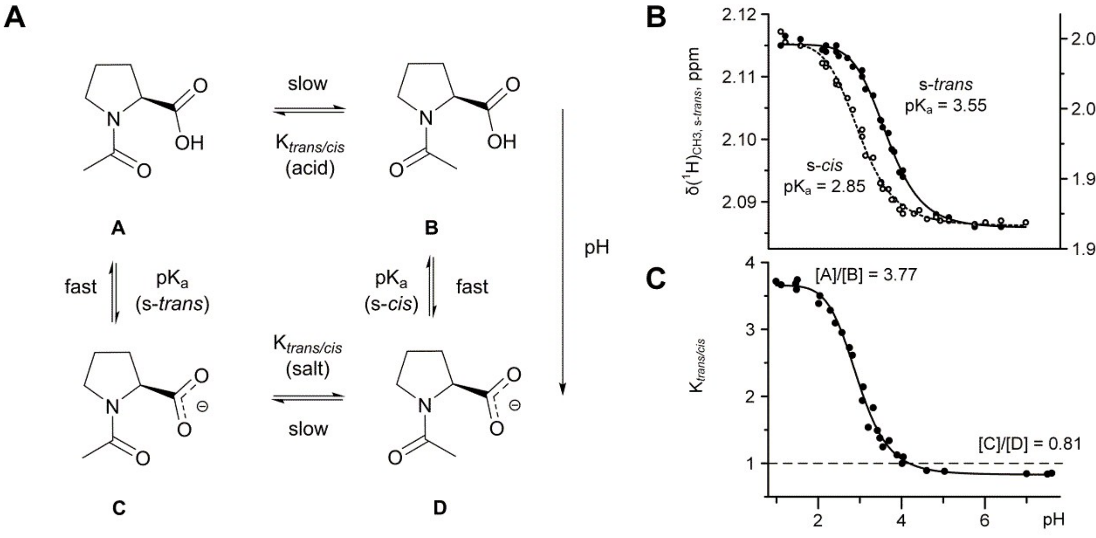


Figure 7. A. Equilibrium forms of Ac-Pro in aqueous medium and their interconversions; **B.** pH dependence of the ^{1}H NMR chemical shifts (CH_3-groups) of the two amide rotamers of Ac-Pro. This plot is equivalent to pH titration curves; **C.** pH dependence of the *trans/cis* equilibrium in Ac-Pro as determined by ^{1}H NMR spectra integration (δ-CH_2 fragment). Observations are made at 298 K.

New Journal of Chemistry Accepted Manuscript

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

Further careful inspection of our experimental data reveals several important observations. Firstly, comparison of the acidity values of the s-*cis* forms and the textbook acidities of the corresponding free amino acids (Pro, Gly, Ala, Phe, Val, Ile) reveals a conservative acidity reduction of 0.8-0.9 units. Considered mechanistically, this is the result of the absence of the compensating ammonium charge in the *N*-acetylated amino acids. At the same time, ΔpK_a appears to be a function of the amino acid context.

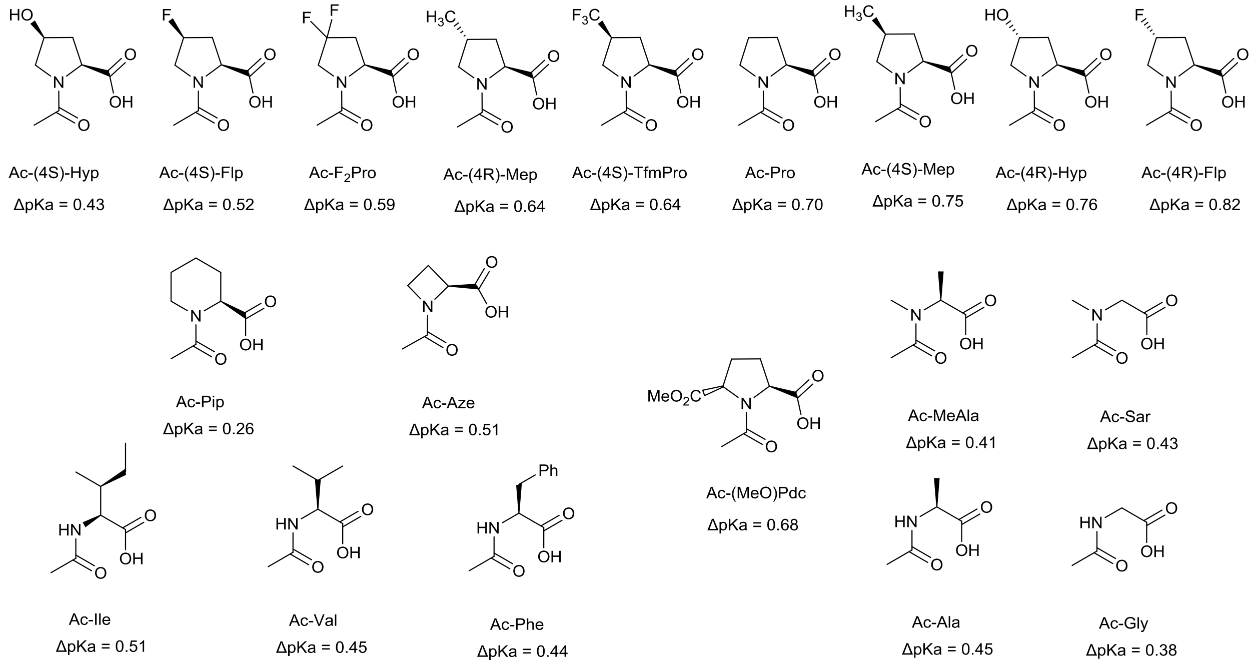


Figure 8. The series of examined *N*-acetyl amino acids with the designated experimentally determined $\Delta pK_a = pK_a$ (s-*trans*) – pK_a (s-*cis*) values.

Secondly, the value of ΔpK_a increases when going from unsubstituted residues (Gly and Sar, ΔpK_a 0.4) to β-branched amino acids (Val and Ile, ΔpK_a 0.5), where the bulky sidechain substituent may reinforce a more stable alignment of the carbonyl-groups. The value was significantly higher for proline (ΔpK_a 0.7), whereas its homologs azetidine-2-carboxylic acid (Aze 0.5) and pipecolic acid (Pip 0.25) exhibit lower ΔpK_a

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

values.[28] Next, incorporation of substituents at position 4 of the proline ring afforded significant changes of ΔpK_a in the range of about 0.4 units.

Table 2. pK_a and amide equilibrium constants determined for the examined *N*-acetylated amino acids at 298 K.

Ac-Xaa	pK_a a)		ΔpK_a a)	$K_{trans/cis}$			$\Delta\Delta G_{298}$, b) kJ/mol	ΔpK_a*
	s-*cis*	s-*trans*		salt	acid	methyl ester		
Ac-Pro, **1**	2.85	3.55	0.70	0.81±0.02	3.77±0.05	4.8[26]	-3.81±0.09	0.67
Ac-Aze, **2**	2.73	3.24	0.51	0.57±0.01	1.67±0.02	n.a.	-2.66±0.08	0.47
Ac-Pip, **3**	3.37	3.63	0.26	1.42±0.01	2.55±0.02	n.a.	-1.45±0.04	0.25
Ac-Gly, **4**	3.19	3.57	0.38	56±1	141±2	n.a.	-2.29±0.08	0.40
Ac-Sar, **5**	3.11	3.54	0.43	1.00±0.01	2.62±0.01	2.8	-2.39±0.03	0.42
Ac-Ala, **6**	3.11	3.56	0.45	62±5	170±13	n.a.	-2.50±0.37	0.44
Ac-MeAla, **7**	3.23	3.64	0.41	1.37±0.03	3.53±0.05	n.a.	-2.34±0.09	0.41
Ac-Phe, **8**	2.98	3.42	0.44	68±1	166±4	n.a.	-2.21±0.10	0.39
Ac-Val, **9**	3.10	3.55	0.45	71±2	239±45	n.a.	-3.01±0.50	0.53
Ac-Ile, **10**	3.12	3.63	0.51	71±2	244±26	n.a.	-3.06±0.32	0.54
Ac-(4*R*)-Flp, **11**	2.37	3.19	0.82	0.91±0.02	5.30±0.20	6.8[26]	-4.37±0.14	0.77
Ac-(4*S*)-Flp, **12**	2.87	3.39	0.52	0.68±0.04	2.27±0.02	2.5[26]	-2.99±0.17	0.52
Ac-(4*R*)-Hyp, **13**	2.39	3.15	0.76	0.95±0.02	5.20±0.09	6.1[15]	-4.21±0.10	0.74
Ac-(4*S*)-Hyp, **14**	3.19	3.62	0.43	0.82±0.03	2.00±0.05	2.4[15]	-2.21±0.15	0.39
Ac-(4S)-Mep, **15**	3.54	2.79	0.75	1.03±0.01	6.24±0.11	7.4[15]	-4.46±0.07	0.78
Ac-(4R)-Mep, **16**	3.48	2.84	0.64	0.85±0.01	3.95±0.10	3.7[15]	-3.81±0.09	0.67
Ac-F_2Pro, **17**	2.34	2.93	0.59	0.79±0.01	2.83±0.03	3.3[26]	-3.16±0.06	0.55
Ac-(4*S*)-TfmPro, **18**	2.57	3.21	0.64	0.90±0.02	3.59±0.09	4.0[29]	-3.43±0.11	0.60
Ac-(MeO)Pdc, **19**	2.68	3.36	0.68	0.18±0.01	0.81±0.02	= 1	-3.74±0.17	0.66

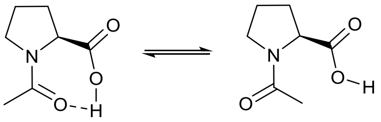
New Journal of Chemistry Accepted Manuscript

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

a) the error of pK_a and ΔpK_a determination is ≤0.05; b) calculated from the equilibrium constants according to eq. 1.



Scheme 2. Hydrogen bond that stabilizes the protonated carboxylate in the s-*trans* conformation in non-polar media.

In the two cases of (4*R*)-Flp and (4*R*)-Hyp, higher ΔpK_a values are accompanied with the higher absolute acidity of both rotameric forms. In contrast, (4*S*)-TfmPro and F_2Pro higher acidity values do not correlate with the increments in ΔpK_a. This is indeed an important observation which contradicts the presence of the hydrogen bond between the carboxyl-group and amide carbonyl that might potentially stabilize the s-*trans* conformation (Scheme 2). Previously in the 1980^{s}, it was reported that "*a hydrogen bond between the terminal carboxyl group and the preceding amide carbonyl could be ruled out in polar media*"[19b] based on considerations of heteronuclear NMR. The fact, that ΔpK_a does not correlate with the absolute acidity further strengthens such conclusions (see also the Pdc example). It further indicates that the effect of donation is not driven by the carboxyl-group electrophilicity alone, but it also clearly accounts for conformational preferences of the proline ring.

Noteworthy, (4*S*)-Hyp, which exhibited low ΔpK_a (as well as $K_{trans/cis}$) turned out to also be the least acidic. It could be that the hydrogen bond between the carboxyl-group and 4-OH group shifts the absolute acidity without affecting ΔpK_a. Another explanation would be the presence of the n→π* donation between these two functionalities. The latter has been reported in simulations in the gas phase.[30] The first explanation seems

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

to be more plausible, since the absolute acidity shift was not observed in (4*S*)-Flp. Indeed, the hydroxyl-group may engage in specific hydrogen bonding with the carboxyl, since a hydroxyl-group is more donative than water, which is not the case for an aliphatic fluorine substituent.[31]

3.3. ΔpK_a and the rotameric populations in esters

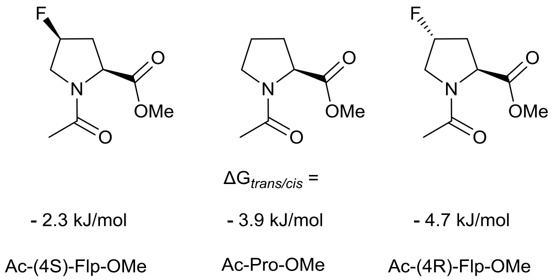


Figure 9. Dominance of the s-*trans* rotamer varies in 4-substituted prolines.[15,27]

Susbequently, we compared the ΔpK_a values found with the already known *trans/cis* preferences in the methyl esters of the *N*-acetylated prolines bearing substitutions in position 4 of the pyrrolidine ring (Fig. 9). We found that the ΔpK_a aligns with their ester population preferences with a decent agreement (Fig. 10). According to the linear free energy correlation this would indicate, that the same phenomenon is underlying the changes in both values. This correlation therefore provides a clear link between ΔpK_a and the electronic properties of the esters.

New Journal of Chemistry Accepted Manuscript

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

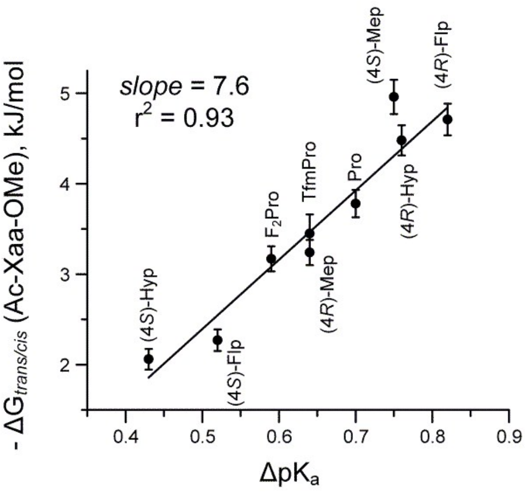
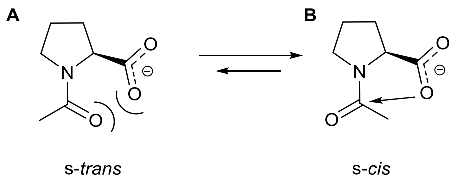


Figure 10. Linear free energy correlation between the Ac-Xaa rotameric acidity differences and Ac-Xaa-OMe equilibrium rotameric populations in 4-substituted proline analogues.

In addition, highly similar ΔpK_a values found for *N*-acetylated canonical amino acids (Gly and Ala) and their *N*-methylated analogues (Sar and MeAla) support the conclusion that ΔΔG originates from geometric alignment of the carbonyl-groups. Thus the ΔΔG values can equally well characterize the peptide bond alignment energy as the values of $\Delta G_{trans/cis}$ for the structurally similar proline analogues. However, the subtraction of the steric effects in ΔΔG values allows comparison of a large variety of the amino acids, and is not restricted to structurally similar systems.

3.4. Effects of the charged carboxylate



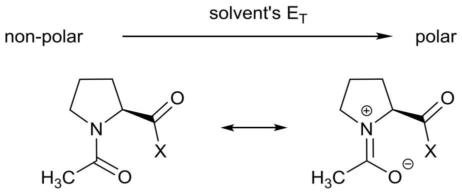
View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

Scheme 3. Alternative explanations for the dominance of s-*cis*-Ac-Pro-O- rotamer in the salt form: **A.** destabilization of s-*trans*-Ac-Pro-O$^-$ due to the repulsion of the polar groups; **B.** stabilization of s-*cis*-Ac-Pro-O$^-$ due to the backward n→π* donation.

It has also been previously reported that the Ac-Pro-O$^-$ salt exhibits dominance of the s-*cis* rotameric form, that was interpreted as the repulsion between the two polar groups, when the terminal carboxyl is charged (Scheme 3A).[32,33] Another hypothetical explanation for the s-*cis* dominance in the salt would be an establishment of a backward n→π* donation from the nucleophilic carboxylate to the carbonyl-group of the amide (Scheme 3B).[34] In the case, when such interactions are present, they will also contribute to the ΔΔG values.



Scheme 4. Rotational barrier increases in more polar solvents with higher ability to separate charges due to a shift of the contributions of their mesomeric structures in the ground states. [35]

In order to test the presence of the carboxylate effects we first studied the dependence of the amide rotational barriers on the charge separation strength of the solvent (E$_T$ values).[36] For the ester form, both activation barriers exhibited the same slope (Fig. 11A) due to the solvent effect on the amide ground state (Scheme 4), irrespective of the s-*trans* and s-*cis* rotameric form.[37]

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

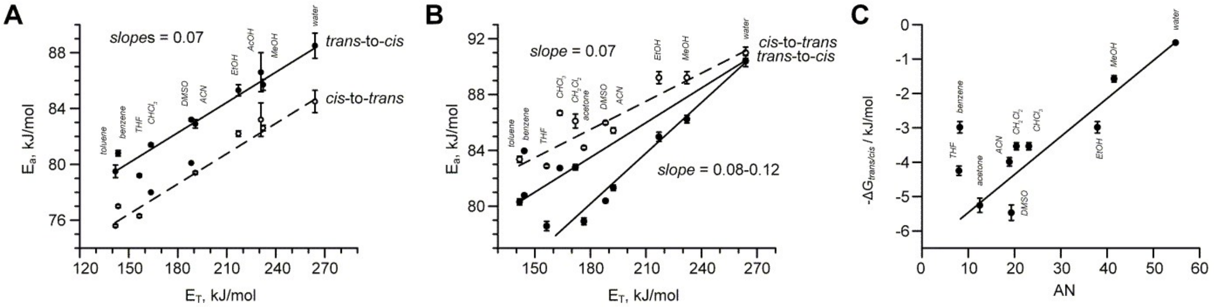


Figure 11. Correlations between the amide rotational barriers and Reichardt-Dimroth solvent parameter (defined as the ability of the solvent to separate charges): A. For 60 mM Ac-Pro-OMe at 298 K; B. For 20 mM Ac-Pro-OK-18-crown-6 complex at 310 K (in dichloromethane at 298 K, in water at 340 K). Straight line, solid circles – activation barriers for the *trans*-to-*cis* isomerization; dashed line, opened circles – for the *cis*-to-*trans* isomerization. C. Correlation of the equilibrium free energy differences in Ac-Pro-O^- and the solvent acceptor number.

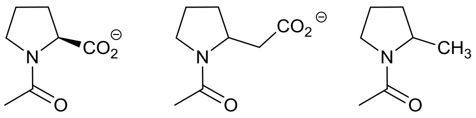
In the case of the salt Ac-Pro-O^- the interaction between the charged carboxylate and the solvent can attenuate the carboxylate interactions shown in Scheme 3. When comparing the salt to the ester, the slopes in the E_T-E_a coordinates will change: 1) for *trans*-*cis* towards an increase in the case of the carboxylate-carbonyl repulsion (Scheme 3A) which is a destabilizing effect, and 2) for *cis*-*trans* towards a decrease in the case of the backward n→π* donation (Scheme 3B) which is a stabilizing effect. Our experimental observations presented in Fig. 11B revealed that the first scenario appeared to be valid, whereas there was no support for the hypothesis that the backward n→π* donation was occurring. As one possible interpretation, the equilibrium was strongly shifted towards the s-*cis* rotamer in the solvents with a low acceptor

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

number (Fig. 11C).[38] Thus, highly polar solvents reduce the carboxylate-carbonyl repulsion. In the terminal case, water seems to quench the repulsion, leading to nearly equilibrated s-*trans* and s-*cis* populations. This is supported by the observation that the salt *trans/cis* ratios found in water (0.81) and water-acetonitrile mixture (0.79) are indistinguishable from the value in trifluoroethanol (0.79) (for comparison of the full solvent set see Fig. S67).



$\Delta H_{trans/cis}$ / kJ/mol =

-1.17 ± 0.10 -0.90 ± 0.09 -1.11 ± 0.06

$K_{trans/cis}$ =

0.81 ± 0.02 1.11 ± 0.01 1.27 ± 0.02

Figure 12. Enthalpic terms of the *trans/cis* equilibrium determined for the salt of *N*-acetyl proline (**1**), *N*-acetyl-β-homoproline (**20**) and in *N*-acetyl-2-methylpyrrolidine (**21**) reveals their close similarity.

The absence of the carboxylate effects is further supported by the finding that $\Delta H_{trans/cis}$ values are nearly identical in substrates similar to Ac-Pro-O$^-$ (Fig. 12). Therefore, it seems to be evident that the ΔΔG values allow quantification of the carbonyl alignment in the esters and acids, whereas the carboxylate effects are opposed by strong water solvation.

3.5. ΔpK_a in non-α-amino acids

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

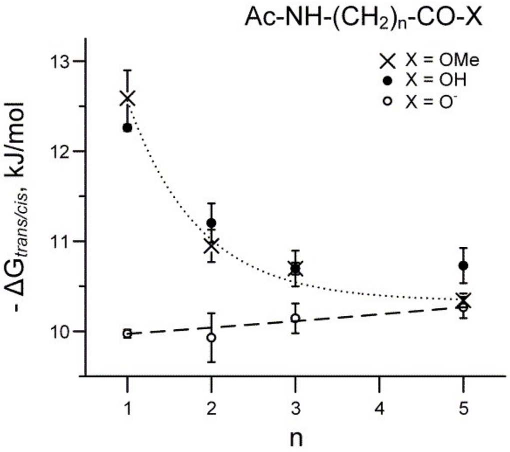


Figure 13. Energy difference between s-*trans* and s-*cis* rotational states in unsubstituted *N*-acetyl amino acids. The attractive effect exponentially decays with increasing chain length.

While in the frame of *N*-acetyl α-amino acids, the ΔpK_a values were found to vary in the range of 0.4-0.8 units, an increase of the chain length between the functionalities should lead to an attenuation of the interaction between the carbonyl groups. Due to the specific nature of the n→π* interaction, we would not expect it to occur with increasing chain length, with this effect already being drastically diminished in the frame of an N-acetyl β-amino acid. The dipole-dipole interaction should also vary dramatically in the homologous *N*-acetyl amino acids, decreasing a stabilizing interaction with increasing chain length. The question, why this should be the case was the addressed through experiment.

Experimentally we found in the case of Ac-βAla both ΔpK_a and ΔpK_a* to be approximately 0.2, which is about a half of the values of Ac-Gly. Next, for Ac-GABA the ΔpK_a* was about 0.1 (Table 3). We observed an exponential decrease of the effect by a factor of 2 for each subsequent CH_2 unit. It is further seen from the Fig. 13, that the

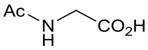
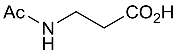
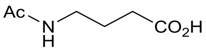
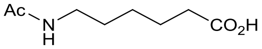
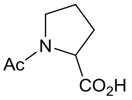
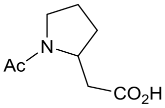
View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

effect originates in the uncharged compounds: acids and esters. Only in these species did we expect a significant C=O dipole (Fig. 6). Then the alignment of this dipole with that of the amide is stronger in the case of the s-*trans* conformation, arising from their closer proximity, as illustrated in Fig. 14.

Table 3. Rotameric acidity differences in the amino acids with increasing chain length.

Xaa	Ac-Xaa structure	pK_a of Xaa	pK_a of Ac-Xaa		ΔpK_a	ΔpK_a*
			s-*cis*	s-*trans*		
Gly (α)	 **22**	2.34	3.19	3.57	0.38	0.40±0.02
βAla (β)	 **23**	3.63	4.15	4.33	0.18	0.22±0.08
GABA (γ)	 **24**	4.08	-	4.65	-	0.10±0.04
ACA (ε)	 **25**	4.37	-	4.69	-	0.08±0.05
Pro (α)	 **1**	2.0	2..85	3.55	0.70	0.67±0.02
βHPro (β)	 **20**	3.58	3.93	4.36	0.43	0.40±0.02

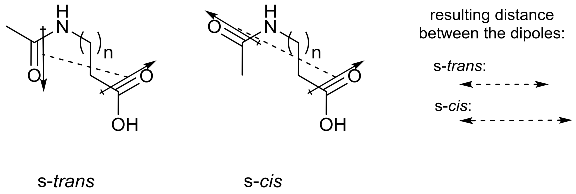


Figure 14. The distance between two interacting dipoles should be closer in the s-*trans* conformation of an *N*-acetyl amino acid.

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

For Ac-Pro and Ac-βHPro, the ΔpK_a values were generally higher, 0.7 and 0.4 units respectively. The higher ΔpK_a for cyclic systems can arise from the reduced degree of conformational freedom present in the ring forms. A cyclic residue necessarily has fewer conformations that it can adopt, leading to a preorganization effect, hence leading to a greater impact on ΔpK_a when compared to the acyclic systems.

3.6. Accumulation of the effect on ΔpK_a in a growing peptide chain

Overall the origin of the ΔpK_a (ΔΔG energetic term) can be seen as donation (of n→π* and/or dipole-dipole origin) that only occurs in the s-*trans* rotamer. Such treatment predicts that the donation effect could accumulate in the all-*trans* peptide chain. In particular, it could provide a motif for PPII folding cooperativity. For this reason we studied both the acidity and the population differences in Ac-Pro-Pro dipeptide. This dipeptide contains four rotameric forms (Fig. 15), and the acidity difference between *trans-trans* and *cis-trans* rotamers would indicate a cascade accumulation of the donative effect (Fig. 16).

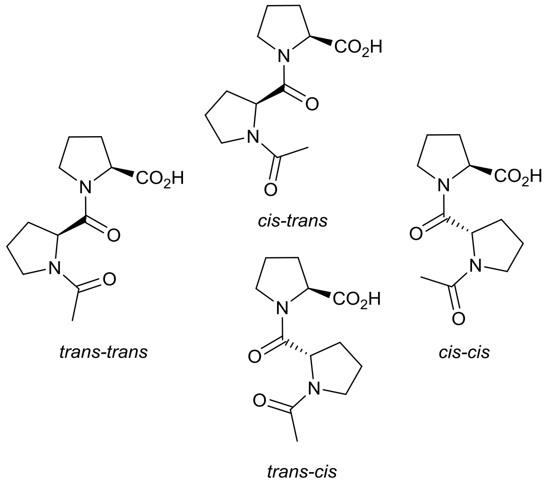


Figure 15. Rotameric forms of Ac-Pro-Pro.

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

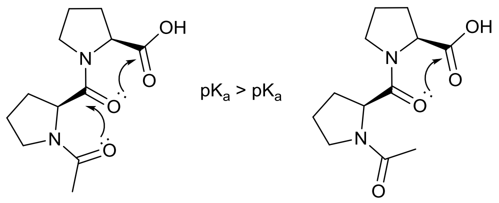


Figure 16. Cascade donation in the chain of s-*trans* amide bonds in Ac-Pro-Pro.

It is well known that Pro-Pro connectivity exhibits the smallest s-*cis* content among all Xaa-Pro fragments (where Xaa is a canonical residue).[39] We also found domination of the s-*trans* Pro-Pro connection in Ac-Pro-Pro which was 93 % in the acidic and 76 % in the salt forms respectively (Table 4). The overall change upon salt-base transition corresponds to the free energy difference $\Delta\Delta G = -RT\ln((93/7)/(76/24)) = -3.6$ kJ/mol (or $\Delta pK_a^* = 0.62$), slightly lower than that which has been observed for Ac-Pro ($\Delta pK_a^* = 0.67$). For the case of the s-*trans* Pro-Pro amide bond rotameric state of the upstream *N*-acetyl moiety exhibited further impact on the C-terminal acidity.. The *trans-trans* form was found to be less acidic than *cis-trans* by 0.08 pK_a units, or 0.11 when calculated from the equilibrium populations (as $\log_{10}((75/18)/(58/18))$). This is a small, but detectable difference which makes about 1/6 of the overall ΔpK_a^*.

Table 4. Thermodynamic populations[a)] and acidities of Ac-Pro-Pro rotameric forms.

	trans-trans	*cis-trans*	*trans-cis*	*cis-cis*
pK_a	3.59	3.51	3.19	2.66
populations, %	Pro-Pro in s-*trans*		Pro-Pro in s-*cis*	
acid (pH 1.4)	75	18	7	
	(75)	(18)	(4)	(3)

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

salt (pH 7.1)	58	18	7	17
	(58)	(17)	(7)	(18)
methyl ester	75	18	4	3

[a] determined by ^{1}H NMR (error ±1 %), in parentheses – ^{13}C NMR intensities (error ±1.5 %).

We suggest therefore, that the cooperativity of the PPII folding is thermodynamically supported by the energy enhancement arising in a cascade fashion, which can be treated as a donation. This consideration can be illustrated in the following way. In a polyproline chain the probability for the amide bond to form a *trans*-amide bond is 93 %. This value is defined by the $\Delta G_{trans/cis}$ of the Pro-Pro connectivity. Statistically, this would mean that for each 14 residues in polyproline, one will have a *cis*-conformation. Taking that the cooperativity provides 1/6 of one peptide bond alignment energy (~ 0.6 kJ/mol), this will immediately decrease the rate of *cis*-bond appearance to one in 18 residues. Further growth of the polypeptide chain may create additional factors, that would enforce stronger alignment.

Noteworthy, we also found dominance of *cis*-*cis* form over *trans*-*cis* in the salt form of Ac-Pro-Pro. This cooperative effect results in the nucleation of the PPI helix (all amide bonds in s-*cis* conformation)[32] already in a context of two residues. Comparison of the acidity values of the *cis*-*trans* (3.19) and *cis*-*cis* (2.66) forms of Ac-Pro-Pro with the acidity of s-*cis*-Ac-Pro (2.85) indicates that this is the effect of destabilization of *trans*-*cis*-Ac-ProPro-O^{-}, rather than stabilization of *cis*-*cis*-Ac-ProPro-O^{-}. We hypothesize that this destabilization could occur in particular due to the close proximity of the *N*-terminal carbonyl oxygen to the C-terminal carboxyl-group in the *trans*-*cis* conformation. However, we didn't study this phenomenon in further detail.

New Journal of Chemistry Accepted Manuscript

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

Finally, similar acidity values found for s-*trans* Ac-Pro (3.55) and *trans*-*trans*-Ac-Pro-Pro (3.59) suppose similar chemical reactivity of the C-terminal carbonyl-group towards nucleophilic reactions. Thus, the oligoproline stalling cannot originate from the donative effects of the carbonyl alignment. This conclusion is supported by our recent kinetic observations of the ribosomal translation rates, which indicate only minor effect of the C-terminal acidity variation to the translational velocity of proline and its analogues.[40]

4. Conclusions

The presented experimental results confirmed that the acidity of *N*-acetylated amino acids is lower when the amide bond is s-*trans* ($\Delta pK_a > 0$) in a general case. Similarly, s-*trans* rotameric populations were always higher for the case of protonated carboxylates ($\Delta pK_a^* > 0$). This phenomenon evidently arises from the mutual interaction of the carbonyl groups in the context of amino acids. Therefore, ΔpK_a, or ΔpK_a^* reflects the magnitude of the peptide bond alignment stabilization energy as evident from comparison of the homologous *N*-acetyl α-, β-, γ- and ε-amino acids. In addition, the energy of this interaction in Ac-Pro-NHMe as compared to Ac-Pro-O^- provides $\Delta\Delta G$ = -3.3 kJ/mol ($\Delta\Delta H$ = -2.8 kJ/mol). This amount accounts for about ½ of one hydrogen bond energy in a β-sheet.[41] Similar value of -2.7 kJ/mol was previously provided by Hinderaker and Raines.[42]

5. Outlook

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

Considerations of the peptide bond alignment energy may also have biological significance, especially in understanding the evolution of the genetic code, which defines evolutionary appearance of the amino acids in translated proteins. Proline is considered to be one of the first amino acids recruited to translation upon transition of the RNA to protein World.[43,44] However, other secondary amino acids are widely abundant in living organisms. For instance, Sar is massively generated by enzymatic methyl-transfer from S-adenosyl-methionine for regulation of cellular C_1-pool.[45] Pipecolic acid, a metabolic product of lysine, is produced in the similar fashion to proline.[46] Aze is a toxic proline substitute which is produced in some plants as a defence mechanism.[47] Despite their widespread metabolic abundance and the fact that they are also found in meteorites,[48,49] Sar and Pip are paradoxically not involved in ribosomal translation. However, nature only recruited proline to the genetic code as the sole secondary amino acid. Among the others, one possible reason for this choice could be the stabilization of the ordered structures with motifs resistant to changing conditions: pH, salt concentrations etc. In other words, the high energy of the peptide bond alignment provided by proline's structure (and may be, α-amino acids in general) might be one of the rationales behind the choice of this amino acid as the basic structural building block for proteins.[50]

View Article Online
DOI: 10.1039/C5NJ03611A

Acknowledgements

This project was granted in the frame of the DFG research group 1805 "Ribosome Dynamics in Regulation of Speed and Accuracy of Translation". The work of P.D. was supported by UniCat cluster TU Berlin. Prof. Igor Komarov (Kyiv, Ukraine) is acknowledged for discussions. We also thank Dr. Pavel K. Mykhailiuk (Kyiv) for provision of a Pdc derivative and discussion.

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

References

---

[1] J. Kurtz, A. Berger, E. Katchalski, *Nature*, 1956, **178**, 1066-1067.

[2] W. A. Elam, T. P. Schrank, A. J. Campagnolo, V. J. Hilser, *Prot. Sci.*, 2013, **22**, 405-417.

[3] A. A. Adzhubei, M. J. E. Sternberg, A. A. Markov, *J. Mol. Biol.*, 2013, **425**, 2100-2132.

[4] It has been described as a preffered conformation in particular: a) polyglutamate, M. L. Tiffany, S. Krimm, *Biopolymers*, 1968, **6**, 1379-1382; b) for polyalanine, M. Mezei, P. J. Fleming, R. Srinivasan, G. D. Rose, *Proteins*, 2004, **15**, 502-507; c) polylysine, A. L. Rucker, T. P. Creamer, *Prot. Sci.*, 2002, **11**, 980-985.

[5] Though, reported in some relelevant examples, for instance: D. Cardinale, G. Guaitoli, D. Tondi, R. Luciani, S. Henrich, O. M. H. Salo-Ahen, S. Ferrari, G. Marverti, D. Guerrieri, A. Ligabue, C. Frassineti, C. Pozzi, S. Mangani, D. Fessas, R. Guerrini, G. Ponterini, R. C. Wade, M. P. Costi, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, E542-E549.

[6] M. M. Harding, P.I. Anderberg, A. D. Haymet, *Eur. J. Biochem.*, 2003, **270**, 1381-1392.

[7] A. A. Adzhubei, M. J. Sternberg, *J. Mol. Biol.*, 1993, **229**, 472-493.

[8] N. W. Owens, C. Braun, J. D. O'Neil, K. Marat, F. Schweizer, *J. Am. Chem. Soc.*, 2007, **129**, 11670-11671.

[9] G. J. Bartlett, A. Choudhary, R. T. Raines, D. N. Woolfson, *Nat. Chem. Biol.*, 2010, **6**, 615-620.

[10] a) J. A. Hodges, R. T. Raines, *Org. Lett.*, 2006, **8**, 4695-4697; b) A. Choudhary, C. G. Fry, K. J. Kamer, R. T. Raines, *Chem. Comm.*, 2013, **49**, 8166-8168.

[11] A. Choudhary, R. W. Newberry, R. T. Raines, *Org. Lett.*, 2014, **16**, 3421-3423.

[12] a) B. Worley, G. Richard, G. S. Harbison, R. Powers, *PLoS ONE*, 2012, **7**, e42075. b) K. J. Kamer, A. Choudhary, R. T. Raines, *J. Org. Chem.*, 2013, **78**, 2099-2103.

[13] P. Wilhelm, B. Lewandowski, N. Trapp, H. Wennemers, *J. Am. Chem. Soc.*, 2014, **136**, 15829-15832.

[14] M. D. Shoulders, K. A. Satyshur, K. T. Forest, R. T. Raines, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 559-564.

[15] M. D. Shoulders, R. T. Raines, *Annu. Rev. Biochem.*, 2009, **78**, 929-958 and referenced therein.

[16] a) S. Ude, J. Lassak, A. L. Starosta, T. Kraxenberger, D. N. Wilson, K.Jung, *Science*, 2013, **339**, 82-85; b) L. K. Doerfel, I. Wohlgemuth, C. Kothe, F. Peske, H. Ulraub, M. V. Rodnina, *ibid.*, 85-88.

[17] a) L. Peil, A. L. Starosta, J. Lassak, G. C. Atkinson, K. Virumäe, M. Spitzer, T. Tenson, K. Jung, J. Remme, D. N. Wilson, *Proc. Natl. Acad. Sci. USA*, 2013, **110**, 15265-15270; b) A. L. Starosta, J. Lassak, L. Peil, G. C. Atkinson, K. Virumäe, T. Tenson, J. Remme, K. Jung, D. N. Wilson, *Nucleic Acid Res.*, 2014, **42**, 10711-10719.

[18] M. Larregola, S. Moore, N. Budisa, *Biochem. Biophys. Res. Commun.*, 2012, **421**, 646-650.

[19] Previously reported values for Ac-Pro: pK_a^D s-*trans* 4.13±0.08 and s-*cis* 3.47±0.08, $\Delta pK_a^D = 0.66\pm0.08$ at 303 K in a) G. R. Bedford, P. J. Sadler, *Biochim. Biophys. Acta*, 1974, **343**, 656-662; and pK_a s-*trans*

New Journal of Chemistry Accepted Manuscript

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

3.36±0.02, s-*cis* 2.79±0.02, ΔpK_a = 0.57±0.02 at 313 K in b) R. N. Hunston, I. P. Gerothanassis, J. Lauterwein, *J. Am. Chem. Soc.*, 1985, **107**, 2654-2661.

[20] C. A. Evans, D. L. Rabenstein, *J. Am. Chem. Soc.*, 1974, **96**, 7312-7317.

[21] Bulk water suppression become available only in 1990^{s} with the advent of pulsed field gradient technique, see: T. Parella, *Magn. Reson. Chem.*, 1998, **36**, 467-495.

[22] pK_a determination using heteronuclear experiments with proton decoupling is risky, due to the differences in both magnetization recycling and NOE enhancement that may occur in the two ionized states, that might produce a distorted average of the chemical shifts.

[23] F. Blomber, H. Rüterjans, K. Lintner, F. Toma, S. Fermandjian, *Org. Magn. Reson.*, 1978, **11**, 598-602.

[24] For instance, the n→π* interaction contributes to the acidity of aspirin: A. Choudhary, K. J. Kamer, R. T. Raines, *J. Org. Chem.*, 2011, **76**, 7933-7937.

[25] a) A. L. Van Geet, *Anal. Chem.*, 1970, **42**, 679-680; b) D. S. Raiford, C. L. Fisk, E. D. Becker, *Anal. Chem.*, 1979, **51**, 2050-2051; c) C. Ammann, P. Meier, A. E. Merbach, *J. Magn. Reson.*, 1982, **46**, 319-321.

[26] C. M. Taylor, R. Hardré, P. J. B. Edwards, J. H. Park, *Org. Lett.*, 2003, **5**, 4413-4416.

[27] C. Renner, S. Alefelder, J. H. Bae, N. Budisa, R. Huber, L. Moroder, *Angew. Chem. Int Ed.*, 2001, **40**, 923-925.

[28] ΔpK_a^* of Ac-Pip calculated from the previously reported values is in a good agreement: 0.25. Whereas for Ac-Aze the salt rotamers have been wrongly assigned: R. E. Galardy, J. R. Alger, M. Liakopoulou-Kyriakides, *Int. J. Pepti. Prot. Res.*, 1982, **19**, 123-132.

[29] V. Kubyshkin, S. Afonin, S. Kara, N. Budisa, P. K. Mykhailiuk, A. S. Ulrich, *Org. Biomol. Chem.*, 2015, **13**, 3171-3181.

[30] A. Lesarri, E. J. Cocinero, J. C. López, J. L. Alonso, *J. Am. Chem. Soc.*, 2005, **127**, 2572-2579.

[31] As supporting this conclusion see also: C. M. Taylor, R. Hardré, P. J. B. Edwards, *J. Org. Chem.*, 2005, **70**, 1306-1315.

[32] M. Kuemin, S. Schweizer, C. Ochsenfeld, H. Wennemers, *J. Am. Chem. Soc.*, 2009, **131**, 15474-15482.

[33] Similar we recently observed in aminoisoxazoles: B. Chalyk, I. Kandaurova, K. Hrebeniuk, O. Manoilenko, I. Kulik, R. Iminov, V. Kubyshkin, A. Tverdokhlebov, P. K. Mykhailiuk, O. Ablialimov, *RSC Adv.* accepted manuscript, DOI: 10.1039/C6RA02365G.

[34] In particular, this effect has been demonstrated along with n)(π repulsion, when proline carboxylate was exchanged to either oxazine or thiazine moieties: D. N. Reddy, R. Thirupathi, S. Tumminakatti, E. N. Prabhakaran, *Tetrahedron Lett.*, 2012, **53**, 4413-4417.

[35] E. S. Eberhardt, S. N. Loh, A. P. Hinck, R. T. Raines, *J. Am. Chem. Soc.*, 1992, **114**, 5437-5439.

[36] K. B. Wiberg, P. R. Rablen, D. J. Rush, T. A. Keith, *J. Am. Chem. Soc.*, 1995, **117**, 4261-4270.

[37] It has also been reported previously, that the solvent effect only has a minor influence on the *trans/cis* ratio: C. B. Braga, L. C. Ducati, C. F. Tormena, R. Rittner, *J. Phys. Chem. A*, 2014, **118**, 1748-1758.

[38] V. Gutmann, *Electrochim. Acta*, 1976, **21**, 661-670.

New Journal of Chemistry Accepted Manuscript

View Article Online
DOI: 10.1039/C5NJ03611A

Published on 31 March 2016. Downloaded by Tulane University on 31/03/2016 14:01:35.

New Journal of Chemistry Accepted Manuscript

---

[39] G. Fischer, *Chem. Soc. Rev.*, 2000, **29**, 119-127.

[40] L. K. Doerfel, I. Wohlgemuth, V. Kubyshkin, A. L. Starosta, D. N. Wilson, N. Budisa, M. V. Rodnina, *J. Am. Chem. Soc.*, 2015, **137**, 12997-13006.

[41] S.-Y. Sheu, D.-Y. Yang, H. L. Selzle, E. W. Schlag, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 12683-12687.

[42] M. P. Hinderaker, R. T. Raines, *Protein Sci.*, 2003, **12**, 1188-1194.

[43] J. T. F. Wong, *Proc. Natl. Acad. Sci. USA*, 1975, **72**, 1909-1912.

[44] H. Hartman, T. F. Smith, *Life (MDPI)*, 2014, **4**, 227-249.

[45] a) M. H. Stipanuk, *Annu. Rev. Nutr.*, 1986, **6**, 179-209; b) A. D. Hanson, S. Roje, *Annu. Rev. Plant Phisiol. Plant. Mol. Biol.*, 2001, **52**, 119-137.

[46] a) Y.-F. Chang, *Biochem. Biophys. Res. Commun.*, 1976, **69**, 174-180, b) H. P. Broquist, *Annu. Rev. Nutr.*, 1991, **11**, 435-448.

[47] a) L. Fowden, *Biochem. J.*, 1956, **64**, 323-332; b) E. Rubenstein, T. McLaughlin, R. C. Winant, A. Sanchez, M. Eckart, K. M. Krasinska, A. Chien, *Phytochemistry*, 2009, **70**, 100-104.

[48] K. A. Kvenvolden, J. G. Lawless, C. Ponnamperuma, *Proc. Natl. Acad. Sci. USA*, 1971, **68**, 486-490.

[49] J. R. Cronin, S. Pizzarello, *Adv. Space Res.*, 1983, **3**, 5-18.

[50] For interesting discussions regarding special role of proline see: 1) B. K. Kay, M. P. Williamson, M. Sudol, *FASEB. J.*, 2000, **14**, 231-241; b) A. C. Forster, *Biotechnol. J.*, 2012, **7**, 835-845.