Electronic Supplementary Information

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

Vladimir Kubyshkin, Patrick Durkin, Nediljko Budisa

Technical University of Berlin, Berlin, Germany

Index

1. General information

Aceturic acid (Ac-Gly), N-acetyl-valine (Ac-Val) and N-acetyl-isoleucine (Ac-Ile), other chemicals and solvents were obtained from commercial sources.

pH was adjusted in aqueous solutions at 23-25°C, and read with HANNA HI221 pH-meter equipped with a glass electrode (for citric and phosphate buffers) and HANNA HI208 pH meter equipped with a gel electrode. pH was calibrated to the standard solutions (pH 4 and 7).

NMR measurement were conducted on Bruker Avance III 500 spectrometer equipped with a BBFO z-gradient probe and Bruker Avance III 700 spectrometer equipped with a TXI z-gradient probe. Variable temperature unit was calibrated with acidified methanol and glycol samples. Infrared spectra were measured at Jasco FT/IR-4100 spectrometer. Mass spectra were recorded at LTQ Orbitrap XL device using electrospray ionization.

2. Chemical synthesis

2.1. Ac-Pro (1)

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).

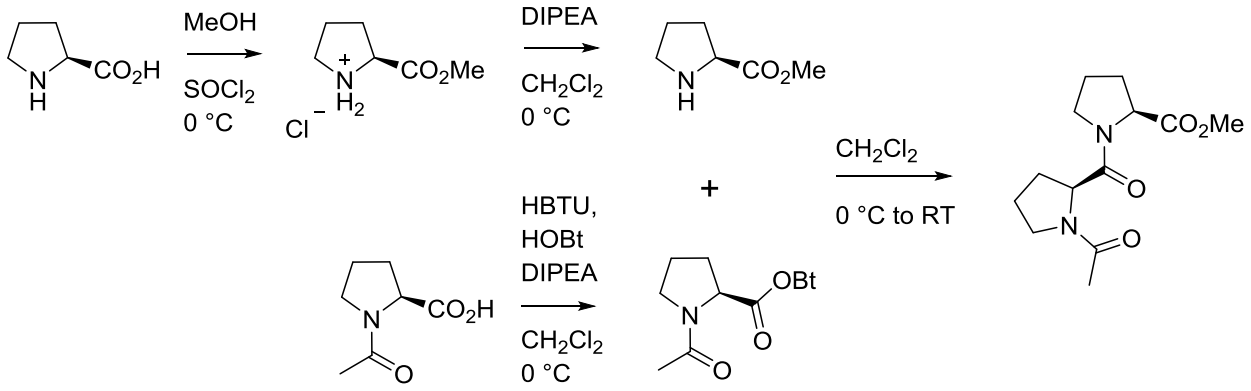
Other Ac-Xaa were synthesized in analogous procedures on smaller scales.

2.2. Ac-Pro-Pro

2.2.1. Solid-phase synthesis

The synthesis was started on L-Pro pre-loaded 2-chlorotrityl resin (1.506 g; 0.55 mmol/g; 0.83 mmol, 1 equiv.). The resin was pre-swollen in dichloromethane (30 min), and washed with DMF. Fmoc-Pro monohydrate (1.2 g; 3.38 mmol, 4 equiv.), HBTU (1.164 g; 3.07 mmol, 3.7 equiv.) and HOBt monohydrate (0.515 g; 3.36 mmol, 4 equiv.) were mixed with DIPEA (1.15 ml; 6.6 mmol, 4 eqiuv.) in 20 ml DMF and added to the resin. Mixing was achieved by nitrogen flow through the mixture. Coupling was continued for 1.5 h, and then repeated with the same portion of the reagents for another 0.5 h. The resin was filtered off, washed with DMF (6 x 20 ml) and with piperidine in DMF (22 % w/w; 3 x 5 ml), then treated with piperidine in DMF solution for the next 20 min. The resin was filtered off, washed with DMF (6 x 20 ml), then pyridine (1.67 ml; 25 equiv.) and acetic anhydride (1.96 ml; 25 equiv.) in DMF (10 ml) were added to the resin and mixing was continued for 1 h. The resin was filtered off, washed with ethanol, dichloromethane and acetonitrile and then dried under vacuum overnight. Peptide (Ac-Pro-Pro) was cleaved from the resin by shaking with hexafluoroisopropanol in dichloromethane (1/4 v/v; 10 ml) for 20 min twice, the resin was washed by dichloromethane (3 x 5 ml). Organic fractions were collected and blown off by nitrogen current for 3 h. Residue was shaken with cold diethylether (25 ml), resulting mixture was allowed to stand at 4 °C, and then centrifuged. Diethylether was decanted and the residue was freeze-dried from water (15 ml). Crude peptide was filtered with activated charcoal (0.9 g) in 30 ml of water and lyophilized to give 148 g (0.58 mmol; 70 % yield) of the peptide as white solid. It was additionally purified on RP-HPLC using Agilent 1260 Infinity system equipped with C18 preparative column. Water-acetonitrile gradient (with 5 mM hydrochloric acid as ion-pairing agent) was used as mobile phase. 72 mg of pure peptide was obtained after lyophilization of the HPLC fractions as white solid. Comparison of the spectra before and after HPLC purification indicated that there was no racemization during synthesis. Mass-spectrum (ESI): 255.1 $[M+1]^+$.

2.2.2. Solution synthesis



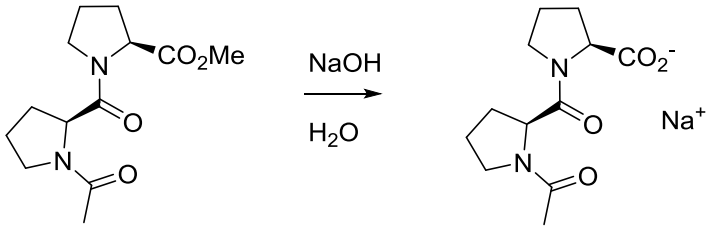
Scheme S1. Synthesis of Ac-Pro-Pro-OMe.

L-proline (3.4 g; 29.5 mmol, 1 equiv.) was dissolved in methanol (HPLC grade; 50 ml) and the solution was cooled down using an ice bath. Thionyl chloride (2.4 ml; 33 mmol, 1.1 equiv.) was added to the stirred reaction mixture within 2 min. After 1 h the cold bath was removed and the mixture was allowed to warm to the ambient temperature overnight (12 h). Solvent was removed under reduced pressure (40 °C in the bath) to give 4.9 g of L-proline methyl ester hydrochloride (HCl*Pro-OMe) as clear oil (quantitative).

^{1}H NMR (CD_2Cl_2, 700 MHz): 9.14 (br s), 4.48 (m, 1H), 3.86 (s, 3 H), 3.56 (m, 1H), 3.49 (m, 1H), 2.43 (m, 1H), 2.19 (m, 1H), 2.14-2.03 (m, 2H).

Ac-Pro (475 mg; 3.02 mmol, 1 equiv.) and HOBt monohydrate (468 mg; 3.06 mmol, 1 equiv.) were mixed with DIPEA (0.63 ml; 3.62 mmol, 1.2 equiv.) in dichloromethane (10 ml) to give a clear solution. The mixture was cooled down in an ice bath under nitrogen atmosphere, and then HBTU (1.114 g; 2.94 mmol, 1 equiv.) was added to the stirred mixture. After 45 min the ice bath was removed and the mixture was stirred additional 15 min before addition to a solution Pro-OMe (prepared as below).

HCl*Pro-OMe (542 mg; 3.27 mmol, 1.1 equiv.) was dissolved in dichloromethane (10 ml), and resulting solution was cooled down in an ice bath. Subsequently DIPEA (0.63 ml; 3.62 mmol, 1.2 equiv.) was added over 3 min, followed by the prepared solution of Ac-Pro-OBt in dichloromethane. Resulting mixture was stirred at 0 °C, after 3 h the ice bath was removed and stirring was continued at ambient temperature for the next 15 min. Reaction was quenched by addition of citric acid solution (10 %; 15 ml). Resulting mixture was filtered to remove the white solid precipitate and the solid was washed with additional dichloromethane (30 ml). Organic fraction (50 ml) was separated from the aqueous layer and washed with citric acid solution (1 x 15 ml), saturated sodium hydrogencarbonate (1 x 15 ml) and brine (1 x 15 ml), dried over sodium sulphate, filtered and concentrated in vacuum. Crude material was purified on a silica gel column using chloroform-methanol (9:1) mixture as mobile phase (R_f = 0.6) to give 578 mg of Ac-Pro-Pro-OMe as clear oil (66 % yield). Mass-spectrum (ESI): 269.1 $[M+1]^+$.



Scheme S2. Saponification of Ac-Pro-Pro-OMe.

Ac-Pro-Pro-OMe (500 mg; 1.86 mmol, 1 equiv.) was dissolved in water (6 ml) and sodium hydroxide solution (1M, 1.85 ml, 1.85 mmol, 1 equiv) was added slowly upon stirring over 1 h. Stirring was continued for another 1 h, during which the reaction progress was monitored with a portion of reaction mixture by ^{1}H NMR (WATERGATE water suppression). Solution was filtered with activated charcoal (0.5 g), and resulting aqueous solution was freeze-dried overnight to give 428 mg of desired Ac-Pro-Pro-ONa as white solid (83 % yield). NMR data showed full consistency of the spectra with the substance obtained in solid-phase synthesis, and indicated no racemization. The sodium salt was later used for pH titrations directly. Mass-spectrum (ESI): 277.1 $[AcPPOH+Na]^{+}$, 255.1 $[AcPPOH+1]^{+}$.

3. Substance characterization

3.1 Ac-Pro (1)

N-acetyl proline

The NMR spectra were interpreted using ^{1}H NOESY, $^{1}H^{13}C$ HSQC spectra and the $^{13}C\{^{1}H\}$ pH series depicted below (Fig. S12).

Ac-Pro acid (pH 1.4)

^{1}H NMR (700 MHz, D_2O):

s-*trans*: 4.33 (dd, 1H, *J* = 8.8 and 4.4 Hz, α-CH), 3.56 and 3.54 (two m, 2H, δ-CH_2), 2.23 and 1.97 (two m, 2H, β-CH_2), 2.02 (s, 3H, CH_3), 1.93 (m, 2H, γ-CH_2);

s-*cis*: 4.56 (dd, 1H, *J* = 9.0 and 2.3 Hz, α-CH), 3.45 and 3.38 (two m, 2H, δ-CH_2), 2.28 and 2.13 (two m, 2H, β-CH_2), 1.93 (s, 3H, CH_3), 1.88 and 1.78 (two m, 2H, γ-CH_2).

$^{13}C\{^{1}H\}$ NMR (126 MHz, D_2O), singlets:

s-*trans*: 176.43 (CO_2H), 173.0 (C(=O)-N), 59.0 (α-CH), 48.5 (δ-CH_2), 29.4 (β-CH_2), 24.2 (γ-CH_2), 21.3 (CH_3);

s-*cis*: 176.35 (CO_2H), 173.3 (C(=O)-N), 60.8 (α-CH), 46.7 (δ-CH_2), 30.8 (β-CH_2), 22.4 (γ-CH_2), 21.2 (CH_3).

Ac-Pro salt (pH 7)

^{1}H NMR (700 MHz, D_2O):

s-*trans*: 4.14 (dd, 1H, *J* = 9 and 4 Hz, α-CH), 3.53 and 3.49 (two m, 2H, δ-CH_2), 2.12 and 1.85 (two m, 2H, β-CH_2), 1.99 (s, 3H, CH_3), 1.85 (m, 2H, γ-CH_2);

s-*cis*: 4.21 (dd, 1H, *J* = 8.8 and 3.6 Hz, α-CH), 3.43 and 3.38 (two m, 2H, δ-CH_2), 2.22 and 1.96 (two m, 2H, β-CH_2), 1.88 (s, 3H, CH_3), 1.79 (m, 2H, γ-CH_2).

$^{13}C\{^{1}H\}$ NMR (126 MHz, D_2O), singlets:

s-*trans*: 179.9 (CO_2^-), 172.3 (C(=O)-N), 61.5 (α-CH), 48.5 (δ-CH_2), 30.0 (β-CH_2), 24.2 (γ-CH_2), 21.4 (CH_3);

s-*cis*: 179.7 (CO_2^-), 173.0 (C(=O)-N), 63.3 (α-CH), 46.9 (δ-CH_2), 31.4 (β-CH_2), 22.8 (γ-CH_2), 21.1 (CH_3).

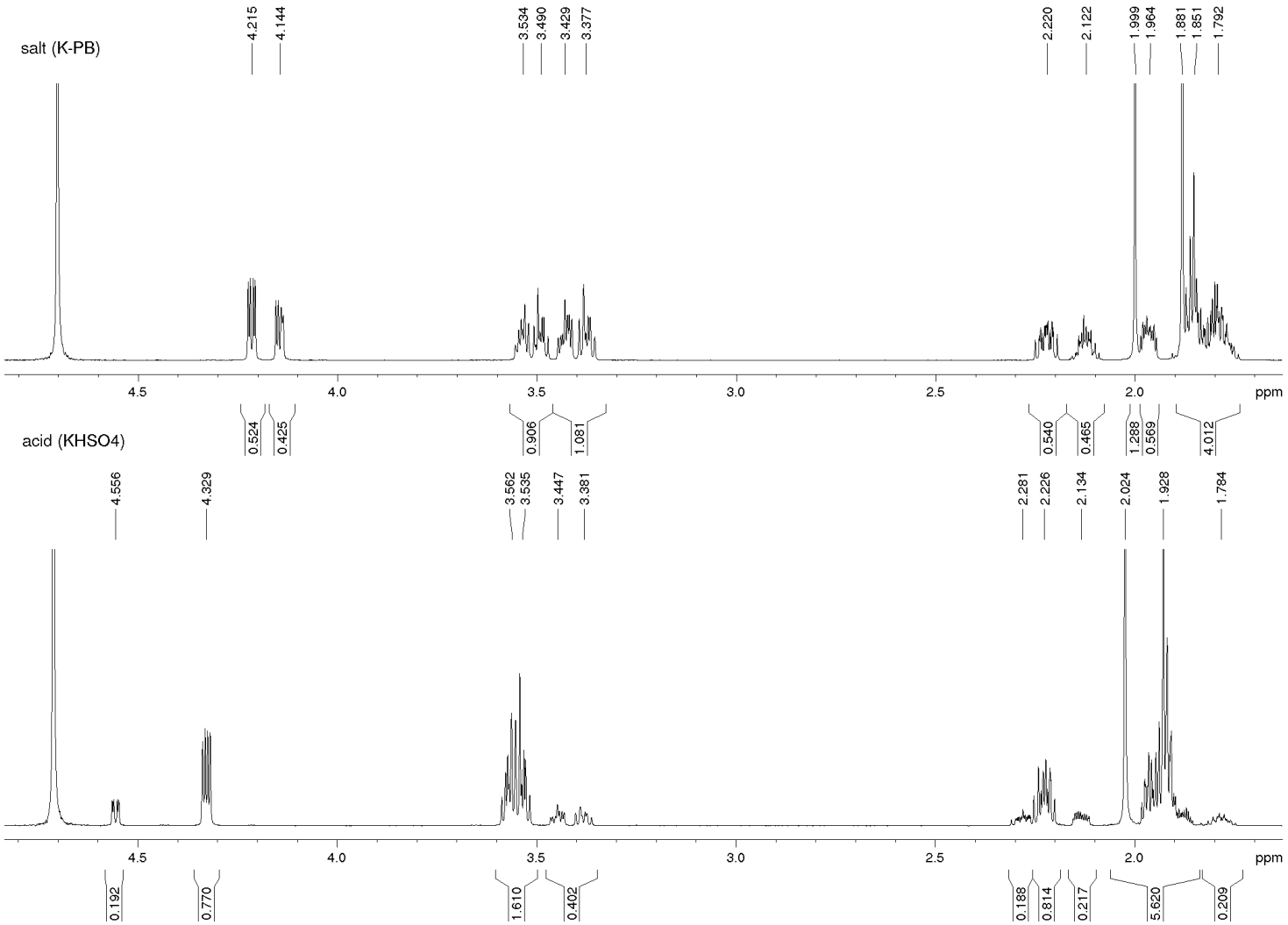


Figure S1. ^{1}H NMR spectra of Ac-Pro in the salt (top trace) and and acid (bottom trace) forms in deuterium oxide at 700 MHz frequency.

3.2 Ac-(4*R*)-Flp (11)

(2*S*,4*R*)-1-acetyl-4-fluoropyrrolidine-2-carboxylic acid

IR bands: 2932, 2795, 2485, 1725, 1590 cm^{-1}. Mass-spectrum (ESI): 176.1 $[M+1]^+$.

The NMR spectra were interpreted using $^{19}F^{1}H$ HOESY (mixing time 500 ms), $^{1}H^{13}C$ HSQC, $^{1}H\{^{19}F\}$ HOHAHA (dipsi2 of 60 ms) and ^{1}H J-resolved assignment experiments. The carbonyl-group assignment was made using $^{1}H^{13}C$ HMBC experiments (cnst = 8 Hz). The proton spectra were taken without and with ^{19}F decoupling. For simplification on Fig. S2 decoupled $^{1}H\{^{19}F\}$ spectra are presented (inverse gated decoupling, no NOE enhancement).

Ac-(4*R*)-Flp acid (pH 1.0)

^{1}H NMR (500 MHz, D_2O):

s-*trans*: 5.35 (dt, J_{HH} = 3.2 Hz, J_{HF} = 52 Hz, 1H, γ-CHF), 4.45 (dd, J_{HH} = 9.7 and 7.9 Hz, 1H, α-CH), 3.93 (ddd, J_{HH} = 13.0 and 2.3 Hz, J_{HF} = 22 Hz, 1H, δ-CH), 3.79 (ddd, J_{HH} = 13.1 and 3.1 Hz, J_{HF} = 38 Hz, 1H, δ-CH), 2.65 (ddddd, J_{HH} = 14.7, 7.7, 2.2 and 1.1 Hz, J_{HF} = 18 Hz, 1H, β-CH), 2.15 (dddd, J_{HH} = 14.7, 10.0 and 3.7 Hz, J_{HF} = 42 Hz, 1H, β-CH), 2.05 (s, 3H, CH_3);

s-*cis*: 5.29 (dt, J_{HH} = 3.4 Hz, J_{HF} = 52 Hz, 1H, γ-CHF), 4.75 (t, J_{HH} = 8.6 Hz, 1H, α-CH), 4.02 (ddd, J_{HH} = 14.0 and 2.5 Hz, J_{HF} = 21 Hz, 1H, δ-CH), 3.45 (ddd, J_{HH} = 14.0 and 3.1 Hz, J_{HF} = 38 Hz, 1H, δ-CH), 2.78 (ddddd, J_{HH} = 15.0, 8.6, 2.5 and 1.3 Hz, J_{HF} = 21 Hz, 1H, β-CH), 2.34 (dddd, J_{HH} = 15.0, 8.7 and 4.1 Hz, J_{HF} = 39 Hz, 1H, β-CH), 1.96 (s, 3H, CH_3).

$^{13}C\{^{1}H\}$ NMR (126 MHz, H_2O+D_2O):

s-*trans*: 175.6 (s, CO_2H), 173.2 (s, C(=O)-N), 92.8 (d, J = 175 Hz, γ-CHF), 57.6 (s, α-CH), 54.2 (d, J = 22 Hz, δ-CH_2), 35.7 (d, J = 22 Hz, β-CH_2), 21.3 (s, CH_3);

s-*cis*: 175.5 (s, CO_2H), 174.0 (s, C(=O)-N), 91.4 (d, J = 176 Hz, γ-CHF), 58.6 (s, α-CH), 52.9 (d, J = 22 Hz, δ-CH_2), 37.3 (d, J = 22 Hz, β-CH_2), 20.8 (s, CH_3).

$^{19}F\{^{1}H\}$ NMR (471 MHz, D_2O):

s-*trans*: -177.90 (s);

s-*cis*: -177.93 (s).

Ac-(4*R*)-Flp salt (pH 7)

^{1}H NMR (500 MHz, D_2O):

s-*trans*: 5.33 (dt, J_{HH} = 3.0 Hz, J_{HF} = 52 Hz, 1H, γ-CHF), 4.27 (dd, J_{HH} = 9.3 and 8.3 Hz, 1H, α-CH), 3.90 (ddd, J_{HH} = 13.2 and 2.2 Hz, J_{HF} = 22 Hz, 1H, δ-CH), 3.77 (ddd, J_{HH} = 13.2 and 3.2 Hz, J_{HF} = 38 Hz, 1H, δ-CH), 2.58 (ddddd, J_{HH} = 14.8, 8.0, 2.2 and 1.1 Hz, J_{HF} = 19 Hz, 1H, β-CH), 2.05 (s, 3H, CH_3), 2.01 (dddd, J_{HH} = 14.8, 9.7, 4.0 Hz, J_{HF} = 42 Hz, 1H, β-CH);

s-*cis*: 5.29 (dt, J_{HH} = 3.1 Hz, J_{HF} = 52 Hz, 1H, γ-CHF), 4.42 (t, J_{HH} = 8.5 Hz, 1H, α-CH), 3.98 (ddd, J_{HH} = 14.0 and 2.6 Hz, J_{HF} = 22 Hz, 1H, δ-CH), 3.49 (ddd, J_{HH} = 14.0 and 3.0 Hz, J_{HF} = 38 Hz, 1H, δ-CH), 2.71 (ddddd, J_{HH} = 15.0, 8.0, 2.5 and 1.3 Hz, J_{HF} = 19 Hz, 1H, β-CH), 2.14 (dddd, J_{HH} = 15.0, 9.2 and 4.0 Hz, J_{HF} = 41 Hz, 1H, β-CH), 1.94 (s, 3H, CH_3).

^{13}C{^{1}H} NMR (126 MHz, H_2O+D_2O):

s-*trans*:178.9 (s, CO_2^-), 172.7 (s, C(=O)-N), 93.1 (d, J = 174 Hz, γ-CHF), 60.0 (s, α-CH), 54.6 (d, J = 22 Hz, δ-CH_2), 36.3 (d, J = 22 Hz, β-CH_2), 21.6 (s, CH_3);

s-cis: 178.8 (s, CO_2^-), 173.8 (s, C(=O)-N), 91.9 (d, J = 173 Hz, γ-CHF), 61.3 (s, α-CH), 53.0 (d, J = 22 Hz, δ-CH_2), 37.9 (d, J = 22 Hz, β-CH_2), 20.9 (s, CH_3).

^{19}F{^{1}H} NMR (471 MHz, D_2O):

s-*trans*: -177.61 (s);

s-*cis*: -177.88 (s).

3.3 Ac-(4*S*)-Flp (12)

(2*S*,4*S*)-1-acetyl-4-fluoropyrrolidine-2-carboxylic acid

IR bands: 2799, 2476, 1704, 1582 cm^{-1}. Mass-spectrum (ESI): 176.1 $[M+1]^+$.

The spectra were interpreted the same way as described above for Ac-(4*R*)-Flp. The $^1H\{^{19}F\}$ inverse gated decoupled spectra are presented on Fig. S3.

Ac-(4*S*)-Flp acid (pH 1.4)

^1H NMR (500 MHz, D_2O):

s-*trans*: 5.36 (dddd, J_{HH} = 5.3, 3.3 and 1.3 Hz, J_{HF} = 52 Hz, 1H, γ-CHF), 4.64 (dd, J_{HH} = 8.9 and 2.7 Hz, 1H α-CH), 3.89 (dd, J_{HH} = 13.2 Hz, J_{HF} = 26 Hz, 1H, δ-CH), 3.82 (ddd, J_{HH} = 13.2 and 3.6 Hz, J_{HF} = 39 Hz, 1H, δ-CH), 2.47 (dm, J_{HF} = 29 Hz, 1H, β-CH), 2.06 (s, 3H, CH_3);

s-*cis*: 5.32 (dt, J_{HH} = 3.4 Hz, J_{HF} = 52 Hz, 1H, γ-CHF), 4.78 (d, J_{HH} = 9.7 Hz, 1H, α-CH), 3.71 (ddd, J_{HH} = 14.4 and 1.4 Hz, J_{HF} = 27 Hz, 1H, δ-CH), 3.64 (ddd, J_{HH} = 14.1 and 3.7 Hz, J_{HF} = 39 Hz, 1H, δ-CH), 2.62 (dd, J_{HH} = 14.9 Hz, J_{HF} = 18 Hz, 1H, β-CH), 2.52 (dddd, J_{HH} = 14.9, 9.7 and 3.3 Hz, J_{HF} = 46 Hz, 1H, β-CH), 2.00 (s, 3H, CH_3).

$^{13}C\{^1H\}$ NMR (126 MHz, D_2O):

s-*trans*: 175.0 (s, CO_2H), 173.4 (s, C(=O)-N), 93.4 (d, J = 172 Hz, γ-CHF), 57.4 (s, α-CH), 54.6 (d, J = 23 Hz, δ-CH_2), 35.7 (d, J = 22 Hz, β-CH_2), 21.3 (s, CH_3);

s-*cis*: 175.4 (s, CO_2H), 173.9 (s, C(=O)-N), 92.3 (d, J = 171 Hz, γ-CHF), 59.0 (s, α-CH), 53.2 (d, J = 24 Hz, δ-CH_2), 37.3 (d, J = 21 Hz, β-CH_2), 21.2 (s, CH_3).

$^{19}F\{^1H\}$ NMR (471 MHz, D_2O):

s-*trans*: -173.05 (s);

s-*cis*: -172.90 (s).

Ac-(4*S*)-Flp salt (pH 7)

^{1}H NMR (500 MHz, D_2O):

s-*trans*: 5.30 (dtt, J_{HH} = 3.7 and 1.0 Hz, J_{HF} = 53 Hz, 1H, γ-CHF), 4.41 (m, 1H, α-CH), 3.85 (dd, J_{HH} = 13.3 Hz, J_{HF} = 27 Hz, 1H, δ-CH), 3.80 (ddd, J_{HH} = 13.3 znd 3.8 Hz, J_{HF} = 37 Hz, 1H, δ-CH), 2.35 (m, J_{HF} = 19 and 43 Hz, 2H, β-CH_2), 2.04 (s, 3H, CH_3);

s-*cis*: 5.26 (dtd, J_{HH} = 3.5 and 1.2 Hz, J_{HF} = 53 Hz, 1H, γ-CHF), 4.41 (m, 1H, α-CH), 3.72 (dd, J_{HH} = 14.6 Hz, J_{HF} = 27 Hz, 1H, δ-CH), 3.62 (ddd, J_{HH} = 14.6 and 4.2 Hz, J_{HF} =39 Hz, 1H, δ-CH), 2.46 (m, 2H, β-CH_2), 1.96 (s, 3H, CH_3).

^{13}C{^{1}H} NMR (126 MHz, H_2O+D_2O):

s-*trans*: 178.39 (s, CO_2^-), 172.7 (s, C(=O)-N), 92.3 (d, *J* = 172, γ-CHF), 59.7 (s, α-CH), 54.6 (d, *J* = 23 Hz, δ-CH_2), 36.5 (d, *J* = 21 Hz, β-CH_2), 21.5 (s, CH_3);

s-*cis*: 178.41 (s, CO_2^-), 173.5 (s, C(=O)-N), 93.4 (d, *J* = 174 Hz, γ-CHF), 61.4 (s, α-CH), 53.2 (d, *J* = 23 Hz, δ-CH_2), 37.8 (d, *J* = 21 Hz, β-CH_2), 21.2 (s, CH_3).

^{19}F{^{1}H} NMR (471 MHz, D_2O):

s-*trans*: -172.23 (s);

s-*cis*: -172.36 (s).

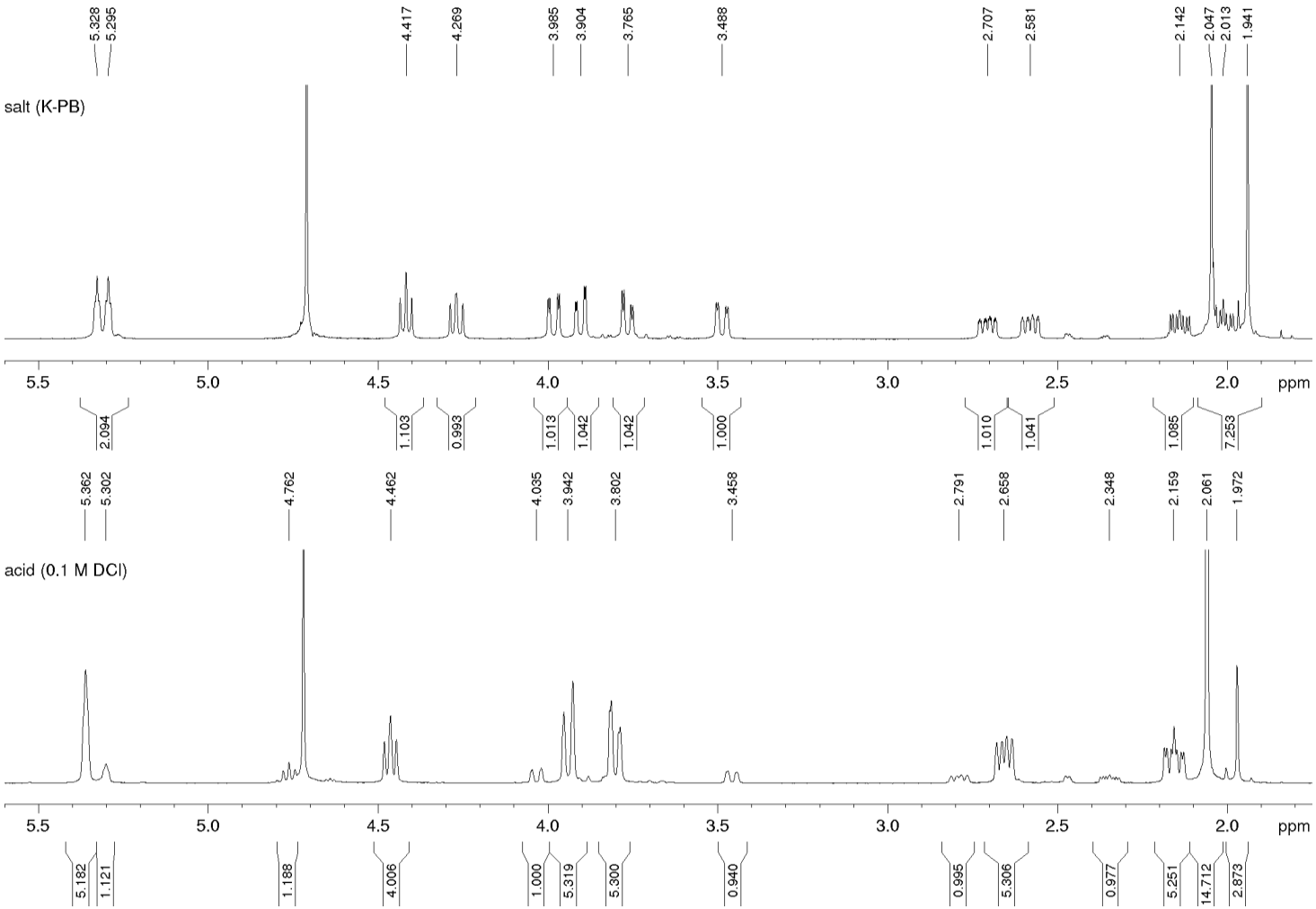


Figure S2. $^{1}H\{^{19}F\}$ NMR spectra of Ac-(4*R*)-Flp in the salt (top trace) and acid (bottom trace) forms in deuterium oxide at 500 MHz proton frequency.

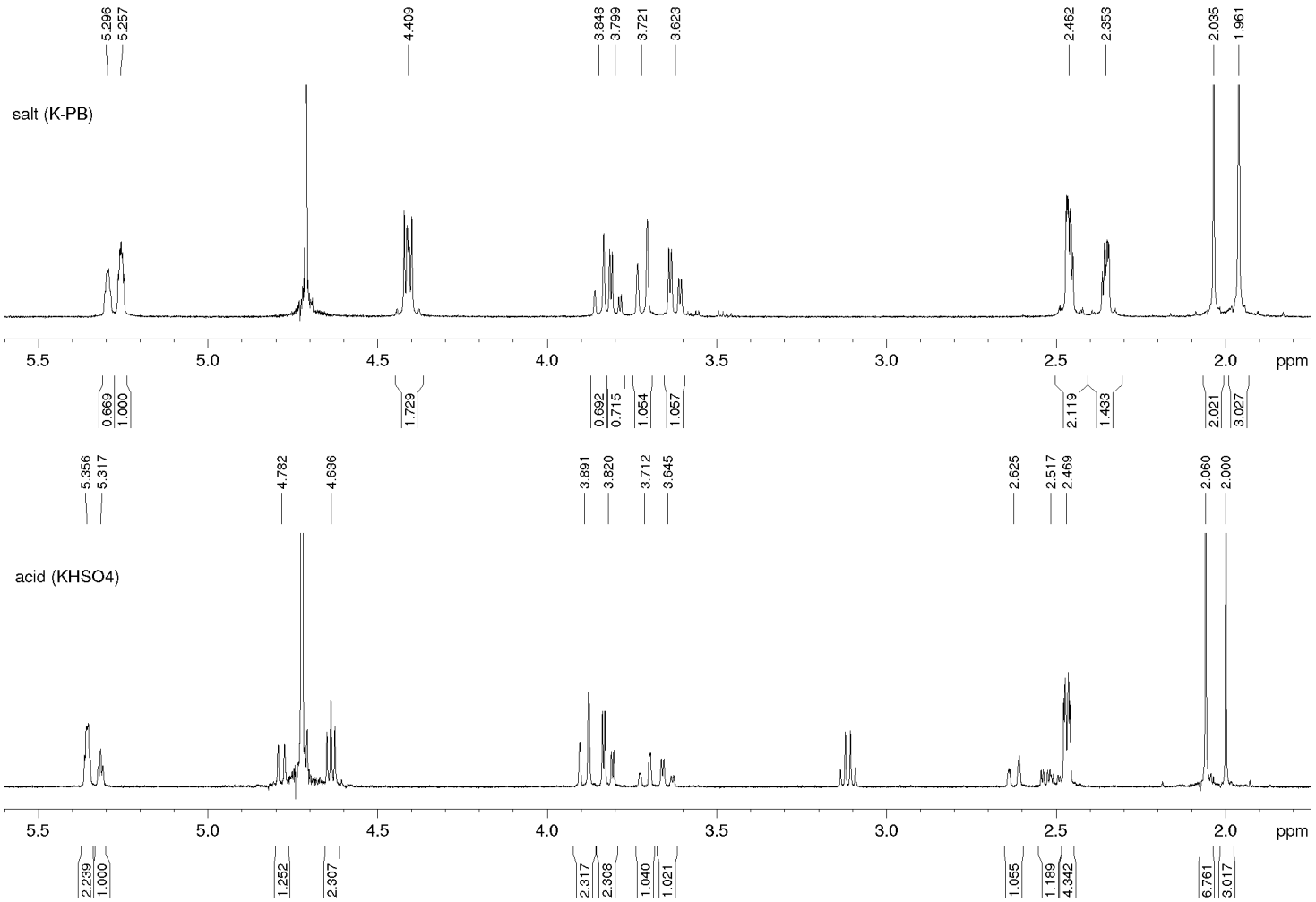


Figure S3. $^{1}H\{^{19}F\}$ NMR spectra of Ac-(4*S*)-Flp in the salt (top trace) and acid (bottom trace) forms in deuterium oxide at 500 MHz proton frequency.

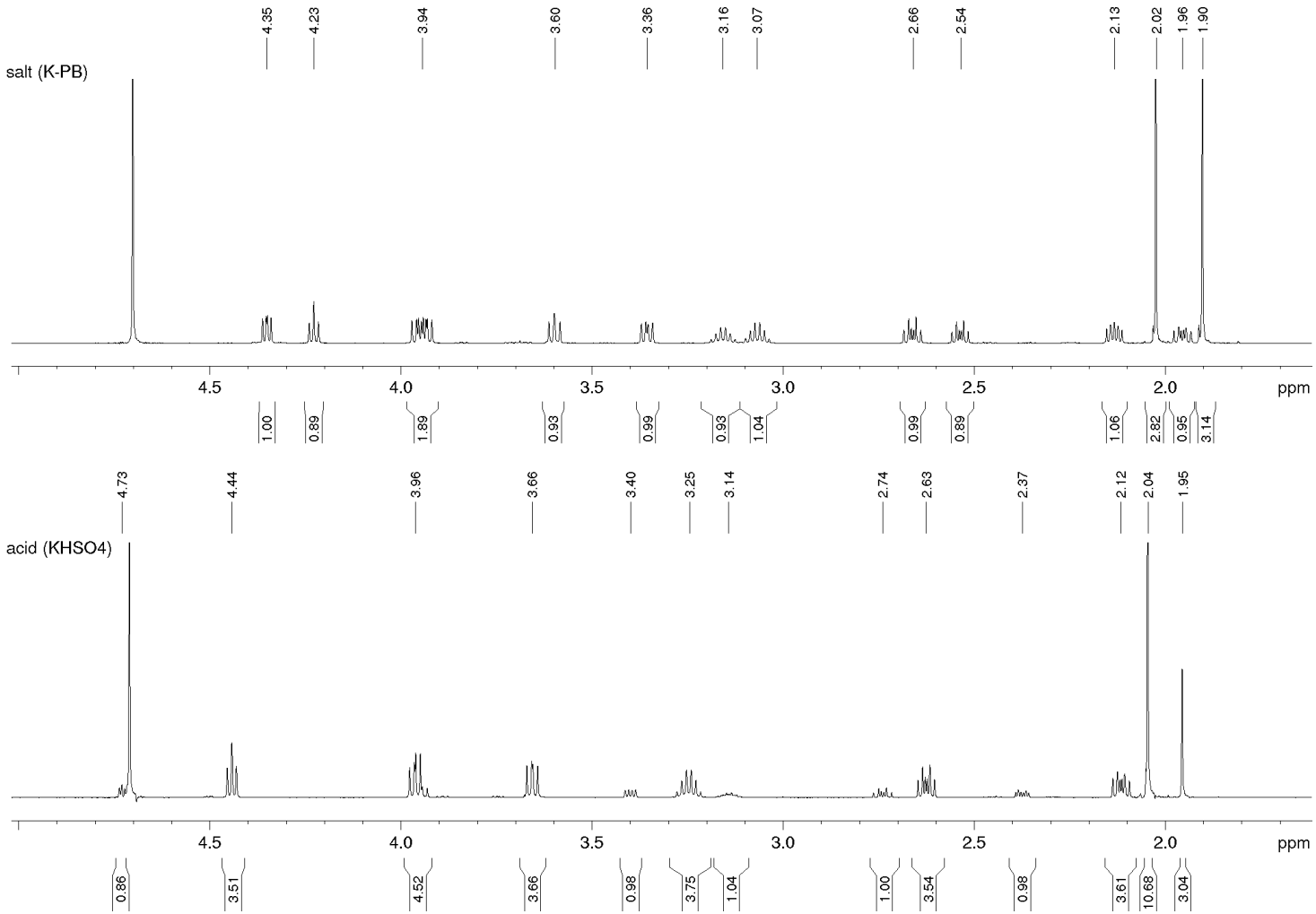


Figure S4. ^{1}H NMR spectra of Ac-(4*S*)-TfmPro in the salt (top trace) and acid (bottom trace) forms in deuterium oxide at 700 MHz frequency.

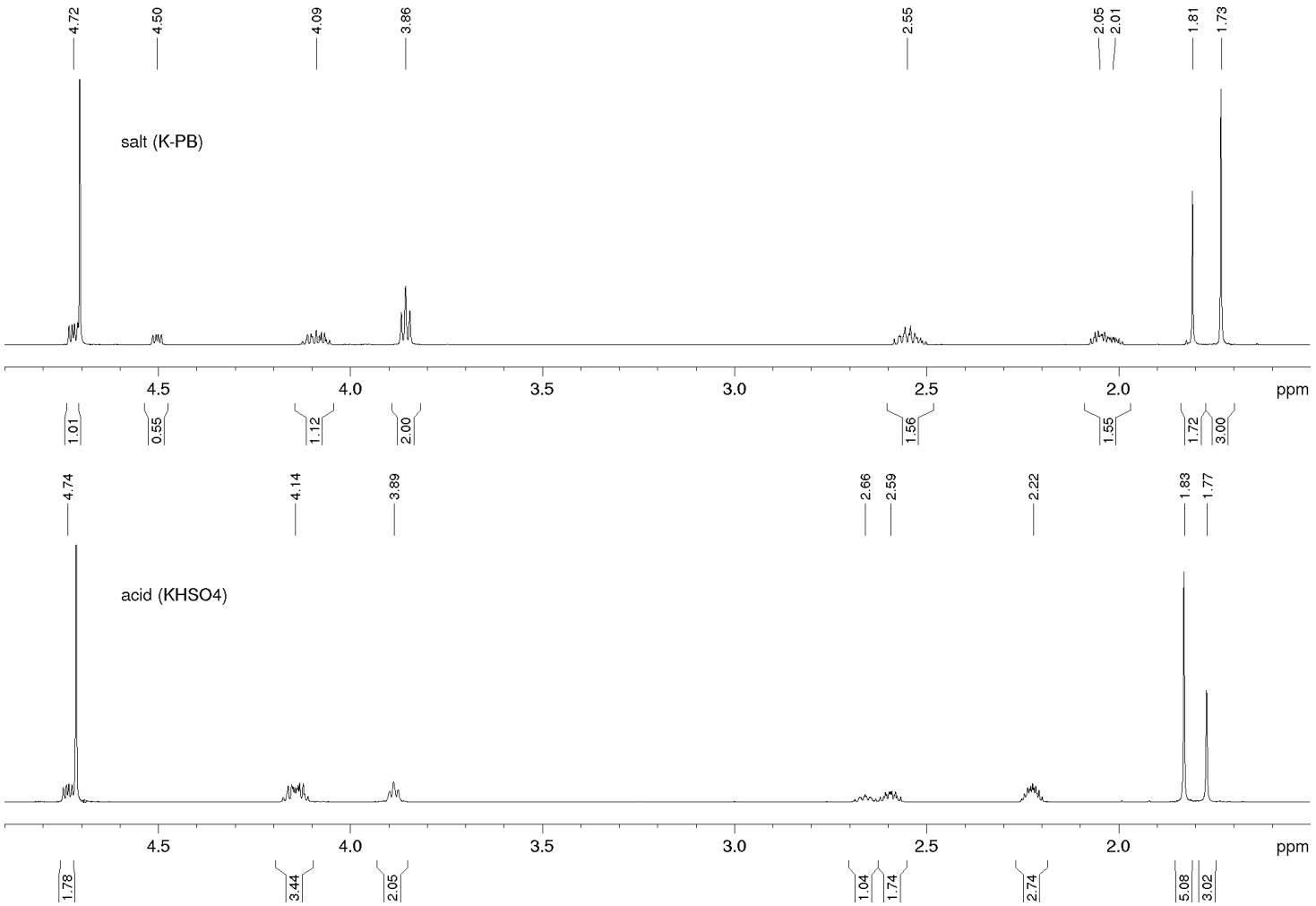


Figure S5. ^{1}H NMR spectra of Ac-Aze in the salt (top trace) and acid (bottom trace) forms in deuterium oxide at 700 MHz frequency.

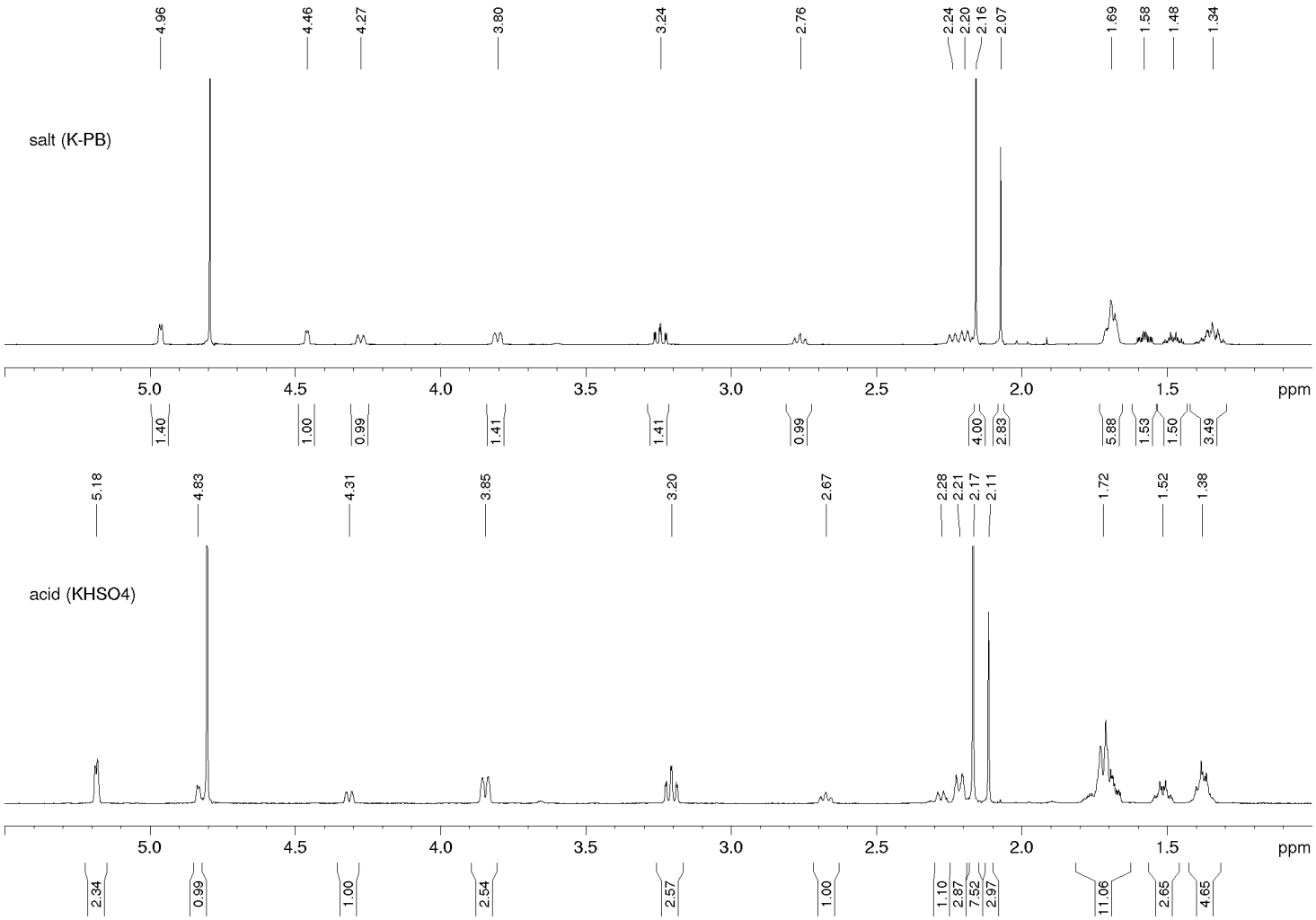


Figure S6. ^{1}H NMR spectra of Ac-Pip in the salt (top trace) and acid (bottom trace) forms in deuterium oxide at 700 MHz frequency.

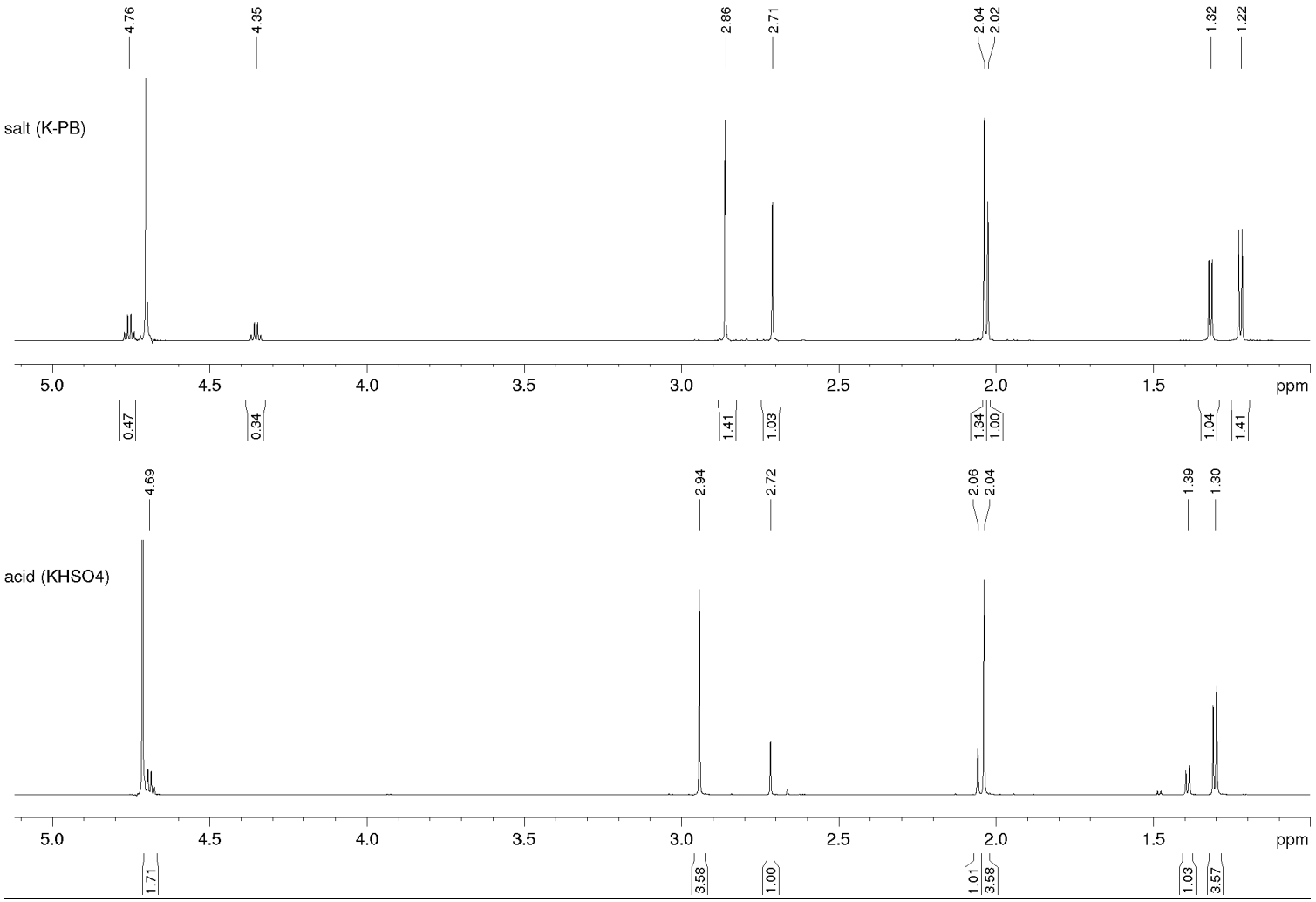


Figure S7. ^{1}H NMR spectra of Ac-MeAla in the salt (top trace) and acid (bottom trace) forms in deuterium oxide at 700 MHz frequency.

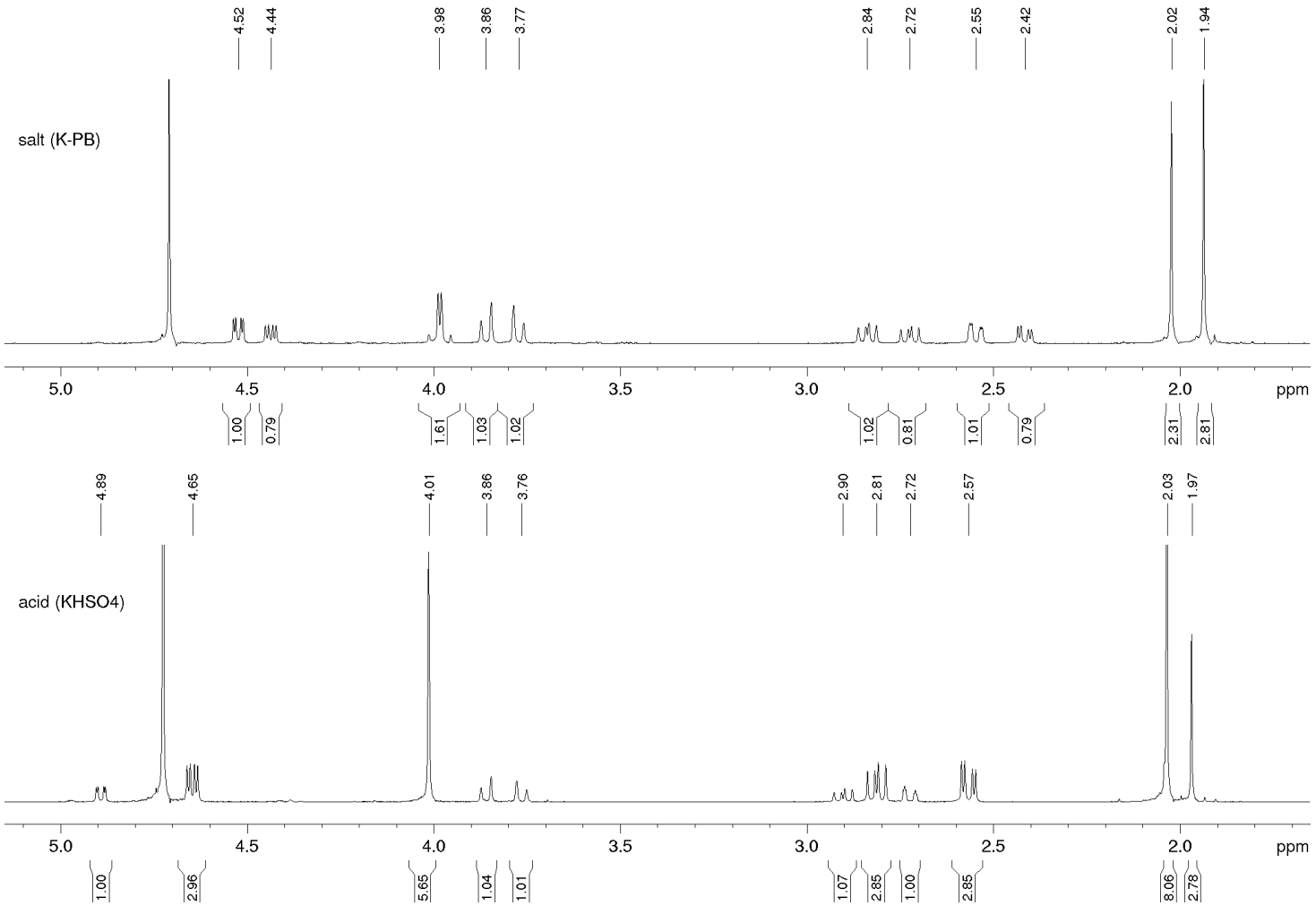


Figure S8. $^{1}H\{^{19}F\}$ NMR spectra of Ac-4,4-F_2Pro in the salt (top trace) and acid (bottom trace) forms in deuterium oxide at 500 MHz proton frequency.

For the rest of the substances the assignment was done based on ^{1}H NOESY, ^{1}H{^{13}C} HSQC and ^{1}H^{13}C HMBC experiments. The ^{1}H spectra are given at Fig S4-S8.

3.4 Ac-(4*S*)-TfmPro (18)

(2*S*,4*S*)-1-acetyl-4-(trifluoromethyl)pyrrolidine-2-carboxylic acid

IR bands: 2893, 2528, 1723, 1593 cm^{-1}. Mass-spectrum (ESI): 226.1 $[M+1]^{+}$.

Ac-(4*S*)-TfmPro acid (pH 1.4)

^{1}H NMR (700 MHz, D_2O):

s-*trans*: 4.44 (t, *J* = 8.3 Hz, 1H, α-CH), 3.96 (dd, *J* = 11.0, 8.6 Hz, 1H, δ-CH), 3.66 (dd, *J* = 11.9, 8.8 Hz, 1H, δ-CH), 3.25 (m, 1H, γ-CH), 2.63 (dt, *J* = 13.4, 8.4 Hz, 1H, β-CH), 2.12 (dt, *J* = 13.4, 8.3 Hz, 1H, β-CH), 2.04 (s, 3H, CH_3);

s-*cis*: 4.73 (dd, *J* = 9.5, 4.6 Hz, 1H, α-CH), 3.95 (m, 1H, δ-CH), 4.30 (dd, *J* = 12.6, 6.5 Hz, 1H, δ-CH), 3.14 (m, 1H, γ-CH), 2.74 (dt, *J* = 14.2, 9.6 Hz, 1H, β-CH), 2.37 (dt, *J* = 14.2, 5.0, 1H, β-CH), 1.95 (s, 3H, CH_3).

^{13}C{^{1}H} NMR (176 MHz, D_2O):

s-*trans*: 174.9 (s, CO_2H), 173.0 (s, C(=O)-N), 126.0 (q, *J* = 278 Hz, CF_3), 58.6 (s, α-CH), 47.2 (s, δ-CH_2), 41.3 (q, J = 29 Hz, γ-CH), 28.3 (s, β-CH_2), 21.2 (s, CH_3);

s-*cis*: 174.9 (s, CO_2H), 173.7 (s, C(=O)-N), 126.6 (d, *J* = 280 Hz, CF_3), 59.7 (s, α-CH), 45.7 (s, δ-CH_2), 39.5 (q, *J* = 29 Hz, γ-CH), 29.5 (s, β-CH_2), 21.0 (s, CH_3).

^{19}F NMR (659 MHz, D_2O):

s-trans: -70.9 (d, J_{FH} = 9 Hz);

s-cis: -71.1 (d, J_{FH} = 10 Hz).

Ac-(4*S*)-TfmPro salt (pH 7)

^{1}H NMR (700 MHz, D_2O):

s-*trans*: 4.23 (t, *J* = 8.4 Hz, 1H, α-CH), 3.93 (dd, *J* = 10.8, 8.4 Hz, 1H, δ-CH), 3.60 (dd, *J* = 10.4, 9.7, 1H, δ-CH), 3.16 (m, 1H, γ-CH), 2.54 (dt, *J* = 13.2, 8.3 Hz, 1H, β-CH), 1.96 (ddd, *J* = 13.2, 9.3, 8.6 Hz, 1H, β-CH), 2.02 (s, 3H, CH_3);

s-*cis*: 4.35 (dd, *J* = 9.0, 6.7 Hz, 1H, α-CH), 3.96 (dd, *J* = 12.2, 8.7 Hz, 1H, δ-CH), 3.36 (dd, *J* = 12.2, 8.6 Hz, 1H, δ-CH), 3.07 (m, 1H, γ-CH), 2.66 (dt, *J* = 13.5, 8.8 Hz, 1H, β-CH), 2.13 (ddd, *J* = 13.5, 7.8, 7.0 Hz, 1H, β-CH),1.90 (s, 3H, CH_3).

^{13}C{^{1}H} NMR (176 MHz, D_2O):

s-*trans*: 178.4 (s, CO_2^-), 172.4 (s, C(=O)-N), 126.2 (q, *J* = 279 Hz, CF_3), 61.0 (s, α-CH), 47.4 (s, δ-CH_2), 41.2 (q, *J* = 29 Hz, γ-CH), 29.0 (q, J = 2 Hz, β-CH_2), 21.5 (s, CH_3);

s-*cis*: 178.2 (s, CO_2^-), 173.4 (s, C(=O)-N), 126.5 (q, *J* = 280 Hz, CF_3), 62.3 (s, α-CH), 45.9 (s, δ-CH_2), 39.8 (q, *J* = 30 Hz, γ-CH), 30.4 (s, β-CH_2), 20.9 (s, CH_3).

^{19}F NMR (659 MHz, D_2O):

s-trans: -70.8 (d, J_{FH} = 9 Hz);

s-cis: -70.7 (d, J_{FH} = 9 Hz).

3.5 Ac-Aze (2)

(*S*)-1-acetylazetidine-2-carboxylic acid

IR bands: 2962, 2892, 2706, 2589, 2500, 1729, 1592 cm^{-1}. Mass-spectrum (ESI): 144.1 $[M+1]^+$.

Ac-Aze acid (pH 1.4)

^1H NMR (700 MHz, D_2O):

s-*trans*: 4.74 (dd, *J* = 10.0, 5.7 Hz, 1H, α-CH), 4.14 (m, 2H, γ-CH_2), 2.59 (m, 1H, β-CH), 2.22 (m, 1H, β-CH), 1.83 (s, 3H, CH_3);

s-*cis*: 5.01 (dd, *J* = 10.1, 5.4 Hz, 1H, α-CH), 3.89 (tm, *J* = 8 Hz, 2H, γ-CH_2), 2.66 (m, 1H, β-CH), 2.22 (m, 1H, β-CH), 1.77 (s, 3H, CH_3).

$^{13}C\{^1H\}$ NMR (176 MHz, D_2O):

s-*trans*: 174.4 (s, CO_2H), 174.0 (s, C(=O)-N), 59.2 (s, α-CH), 49.2 (s, γ-CH_2), 19.4 (s, β-CH_2), 17.6 (s, CH_3);

s-*cis*: 174.63 (s, C(=O)-N), 174.59 (s, CO_2H), 61.7 (s, α-CH), 46.6 (s, γ-CH_2), 19.9 (s, β-CH_2), 18.0 (s, CH_3).

Ac-Aze salt (pH 7)

^1H NMR (700 MHz, D_2O):

s-*trans*: 4.50 (dd, J = 9.9, 5.7 Hz, 1H, α-CH), 4.09 (m, 2H, γ-CH_2), 2.53 (m, 1H, β-CH), 2.02 (m, 1H, β-CH), 1.81 (s, 3H, CH_3);

s-*cis*: 4.72 (dd, J = 10.0, 5.8 Hz, 1H, α-CH), 3.86 (t, J = 7.8 Hz, 2H, γ-CH_2), 2.56 (m, 1H, β-CH), 2.05 (m, 1H, β-CH), 1.73 (s, 3H, CH_3).

$^{13}C\{^1H\}$ NMR (176 MHz, D_2O):

s-*trans*: 178.3 (s, CO_2^-), 173.1 (s, C(=O)-N), 61.8 (s, α-CH), 48.9 (s, γ-CH_2), 20.1 (s, β-CH_2), 17.7 (s, CH_3);

s-*cis*:178.2 (s, CO_2^-), 174.0 (s, C(=O)-N), 64.3 (s, α-CH), 46.2 (s, γ-CH_2), 20.4 (s, β-CH_2), 17.9 (s, CH_3).

3.6 Ac-Pip (3)

(*S*)-1-acetylpiperidine-2-carboxylic acid

IR bands: 2937, 2860, 2523, 1722, 1591 cm^{-1}. Mass-spectrum (ESI): 172.1 $[M+1]^+$.

Ac-Pip acid (pH 1.4)

^1H NMR (700 MHz, D_2O):

s-*trans*: 5.18 (d, *J* = 5.4 Hz, 1H, α-CH), 3.85 (d, *J* = 14.0 Hz, 1H, ε-CH), 3.20 (td, *J* = 13.1, 2.6 Hz, 1H, ε-CH), 2.21 and 1.69 (2m, 2H, β-CH_2), 2.17 (s, 3H, CH_3), 1.72 and 1.51 (2m, 2H, δ-CH_2), 1.72 and 1.38 (2m, 2H, γ-CH_2);

s-*cis*: 4.83 (d, *J* = 5.2 Hz, 1H, α-CH), 4.31 (d, *J* = 13.4 Hz, 1H, ε-CH), 2.67 (td, *J* = 13.3, 2.5 Hz, 1H, ε-CH), 2.28 and 1.76 (2m, 2H, β-CH_2), 2.11 (s, 3H, CH_3), 1.70 and 1.38 (2m, 2H, δ-CH_2), 1.72 and 1.38 (2m, 2H, γ-CH_2);

$^{13}C\{^1H\}$ NMR (176 MHz, D_2O):

s-*trans*: 175.1 (s, CO_2H), 174.0 (s, C(=O)-N), 52.6 (s, α-CH), 44.6 (s, ε-CH_2), 26.0 (s, β-CH_2), 24.3 (s, δ-CH_2), 20.6 (s, CH_3), 20.0 (s, γ-CH_2)

s-*cis*:174.7 (s, CO_2H), 174.2 (s, C(=O)-N), 57.3 (s, α-CH), 40.1 (s, ε-CH_2), 26.5 (s, β-CH_2), 23.9 (s, δ-CH_2), 20.5 (s, CH_3), 20.1 (s, γ-CH_2).

Ac-Pip salt (pH 7)

^1H NMR (700 MHz, D_2O):

s-*trans*: 4.96 (d, *J* = 5.7 Hz, 1H, α-CH), 3.80(dm, *J* = 13.7 Hz, 1H, ε-CH), 3.24 (td, *J* = 13.4, 2.9 Hz, 1H, ε-CH), 2.20 and 1.58 (2m, 2H, β-CH_2), 2.16 (s, 3H, CH_3), 1.70 and 1.48 (2m, 2H, δ-CH_2), 1.68 and 1.33 (2m, 2H, γ-CH_2);

s-*cis*: 4.46 (d, *J* = 5.2 Hz, 1H, α-CH), 4.27 (dm, *J* = 13.6 Hz, 1H, ε-CH), 2.76 (td, *J* = 13.2, 2.5 Hz, 1H, ε-CH), 2.24 and 1.69 (2m, 2H, β-CH_2), 2.07 (s, 3H, CH_3), 1.69 and 1.36 (2m, 2H, δ-CH_2), 1.68 and 1.33 (2m, 2H, γ-CH_2)

$^{13}C\{^1H\}$ NMR (176 MHz, D_2O):

s-*trans*: 178.1 (s, CO_2^-), 172.9 (s, C(=O)-N), 54.7 (s, α-CH), 44.8 (s, ε-CH_2), 27.1 (s, β-CH_2), 24.9 (s, δ-CH_2), 20.7 (s, CH_3), 20.5 (s, γ-CH_2);

s-*cis*: 177.8 (s, CO_2^-), 173.5 (s, C(=O)-N), 59.6 (s, α-CH), 40.0 (s, ε-CH_2), 27.5 (s, β-CH_2), 24.2 (s, δ-CH_2), 20.6 (s, CH_3 and γ-CH_2).

3.7 Ac-MeAla (7)

(*S*)-N-acetyl-N-methyl-alanine

IR bands: 2949, 2504, 1720, 1591 cm^{-1}. Mass-spectrum (ESI): 146.1 $[M+1]^+$.

Ac-MeAla acid (pH 1.4)

^1H NMR (700 MHz, D_2O):

s-*trans*: 4.69 (m, 1H, α-CH), 2.94 (s, 3H, CH_3-N), 2.04 (s, 3H, CH_3-C=O), 1.30 (d, *J* = 7.3 Hz, 3H, β-CH_3);

s-*cis*: 4.69 (m, 1H, α-CH), 2.72 (s, 3H, CH_3-N), 2.06 (s, 3H, CH_3-C=O), 1.39 (d, *J* = 7.2 Hz, 3H, β-CH_3);

$^{13}C\{^1H\}$ NMR (176 MHz, D_2O):

s-*trans*: 175.5 (s, CO_2H), 174.46 (s, C(=O)-N), 54.1 (s, α-CH), 33.5 (s, CH_3-N), 20.9 (s, *C*H_3-C=O), 13.3 (β-CH_3);

s-*cis*: 175.0 (s, CO_2H), 174.55 (s, C(=O)-N), 56.6 (s, α-CH), 29.1 (s, CH_3-N), 20.5 (s, *C*H_3-C=O), 14.3 (β-CH_3).

Ac-MeAla salt (pH 7)

^1H NMR (700 MHz, D_2O):

s-*trans*: 4.76 (q, *J* = 7.4 Hz, 1H, α-CH), 2.86 (s, 3H, CH_3-N), 2.04 (s, 3H, CH_3-C=O), 1.22 (d, *J* = 7.3 Hz, 3H, β-CH_3);

s-*cis*: 4.35 (q, *J* = 7.1 Hz, 1H, α-CH), 2.71 (s, 3H, CH_3-N), 2.02 (s, 3H, CH_3-C=O), 1.32 (d, *J* = 7.1 Hz, 3H, β-CH_3).

$^{13}C\{^1H\}$ NMR (176 MHz, D_2O):

s-*trans*: 179.0 (s, CO_2^-), 174.0 (s, C(=O)-N), 54.9 (s, α-CH), 32.3 (s, CH_3-N), 21.1 (s, *C*H_3-C=O), 19.6 (β-CH_3);

s-*cis*: 178.5 (s, CO_2^-), 174.1 (s, C(=O)-N), 59.1 (s, α-CH), 29.4 (s, CH_3-N), 20.6 (s, *C*H_3-C=O), 15.4 (β-CH_3).

3.8 Ac-4,4-F_2Pro (17)

(*S*)-1-acetyl-4,4-difluoropyrrolidine-2-carboxylic acid

IR bands: 2953, 2789, 2510, 1718, 1589 cm^{-1}. Mass-spectrum (ESI): 194.1 $[M+1]^+$.

Ac-F_2Pro acid (pH 1.4)

$^1H\{^{19}F\}$ NMR (500 MHz, D_2O):

s-*trans*: 4.65 (dd, *J* = 10.0, 4.5 Hz, 1H, α-CH), 4.01 (s, 2H, δ-CH_2), 2.81 (dd, *J* = 14.4, 10.0 Hz, 1H, β-CH), 2.57 (dd, *J* = 14.4, 4.4 Hz, 1H, β-CH), 2.03 (s, 3H, CH_3);

s-*cis*: 4.89 (dd, *J* = 10.1, 2.5 Hz, 1H, α-CH), 3.86 (d, *J* = 13.4 Hz, δ-CH), 3.76 (d, *J* = 13.4 Hz, δ-CH), 2.90 (dd, *J* = 14.3, 10.0 Hz, 1H, β-CH), 2.72 (dt, *J* = 14.4, 1.8 Hz, 1H, β-CH), 1.97 (s, 3H, CH_3).

$^{13}C\{^1H\}$ NMR (176 MHz, D_2O):

s-*trans*: 173.91 (s, CO_2H), 173.3 (s, C(=O)-N), 126.3 (t, *J* = 246 Hz, CF_2), 56.6 (d, *J* = 4 Hz, α-CH), 53.8 (t, *J* = 32 Hz, δ-CH_2), 36.6 (t, *J* = 25 Hz, β-CH_2), 21.1 (s, CH_3);

s-*cis*: 174.00 (s, CO_2H), 173.85 (s, C(=O)-N), 126.2 (t, *J* = 246 Hz, CF_2), 58.5 (d, *J* = 6 Hz, α-CH), 52.6 (t, *J* = 33 Hz, δ-CH_2), 38.1 (t, *J* = 25 Hz, β-CH_2), 20.1 (s, CH_3).

$^{19}F\{^1H\}$ NMR (471 MHz, D_2O):

s-*trans*: -97.8 (d, J_{FF} = 234 Hz, 1F), -102.5 (d, J_{FF} = 234 Hz, 1F);

s-*cis*: -95.2 (d, J_{FF} = 236 Hz, 1F), -104.8 (d, J_{FF} = 236 Hz, 1F).

Ac-F_2Pro salt (pH 7)

$^{1}H\{^{19}F\}$ NMR (500 MHz, D_2O):

s-*trans*: 4.44 (dd, *J* = 9.8, 4.4 Hz, 1H, α-CH), 4.00 (d, *J* = 12.5 Hz, 1H, δ-CH), 3.96 (d, *J* = 12.5 Hz, 1H, δ-CH), 2.72 (dd, *J* = 14.3, 10.0 Hz, 1H, β-CH), 2.42 (dd, *J* = 14.4, 4.4 Hz, 1H, β-CH), 2.02 (s, 3H, CH_3);

s-*cis*: 4.52 (dd, *J* = 9.8, 3.2 Hz, 1H, α-CH), 3.86 (d, *J* = 13.5 Hz, δ-CH), 3.77 (d, *J* = 13.5 Hz, δ-CH), 2.84 (dd, *J* = 14.5, 10.0 Hz, 1H, β-CH), 2.55 (ddd, *J* = 14.3, 3.0, 1.2 Hz, 1H, β-CH), 1.94 (s, 3H, CH_3).

$^{13}C\{^{1}H\}$ NMR (176 MHz, D_2O):

s-*trans*: 177.0 (s, CO_2^-), 172.7 (s, C(=O)-N), 126.8 (t, *J* = 246 Hz, CF_2), 59.1 (d, *J* = 4 Hz, α-CH), 54.1 (t, *J* = 32 Hz, δ-CH_2), 37.5 (t, *J* = 24 Hz, β-CH_2), 21.3 (s, CH_3);

s-*cis*: 177.2 (s, CO_2^-), 173.6 (s, C(=O)-N), 126.4 (t, *J* = 247 Hz, CF_2), 60.9 (d, *J* = 6 Hz, α-CH), 52.7 (t, *J* = 34 Hz, δ-CH_2), 38.8 (t, *J* = 25 Hz, β-CH_2), 20.2 (s, CH_3).

$^{19}F\{^{1}H\}$ NMR (471 MHz, D_2O):

s-*trans*: -97.6 (d, J_{FF} = 232 Hz, 1F), -101.6 (d, J_{FF} = 232 Hz, 1F);

s-*cis*: -96.2 (d, J_{FF} = 233 Hz, 1F), -103.3 (d, J_{FF} = 233 Hz, 1F).

4. Examples of pK_a determination for the carboxyl group in N-acetyl amino acids

4.1 Approach A

Preliminary pK_a determination for Ac-Pro and Ac-(4*S*)-Flp in buffered solutions.

3.85 g of citric acid was dissolved in 200 ml water to make 100 mM citrate concentration. The starting pH was 2.1. It was then titrated with 10 M sodium hydroxide solution; and solutions with different pH values were placed in NMR tubes in 500 µl amounts. For pH 1 500 µl of 0.1 M hydrochloric acid was taken.

2 µl aliquots with Ac-Xaa solutions were added to the NMR tubes to get ≤ 0.5 mM final concentration of the analytes prior to measurements. For Ac-Pro samples 50 µl of 0.3 mg/ml TPS[S1] solution in deuterium oxide was added for referencing and lock. The ^{1}H and ^{19}F NMR spectra were taken on Bruker Avance III 500 spectrometer equipped with a z-gradient BBFO probe. The temperature was set to 298 K. The ^{1}H NMR spectra were recorded in excitation sculpting experiment using W5 pulse tray ("zggpw5") on locked samples. The ^{19}F NMR spectra were recorded in inverse gated ^{1}H decoupled 30-deg pulse experiments ("zgig30").

Following this protocol we were able to determine the acid-base transitions for both rotamers of Ac-Pro since they exhibited characteristic sigmoidal transition in ^{1}H NMR spectra as shown below on Fig. S9.

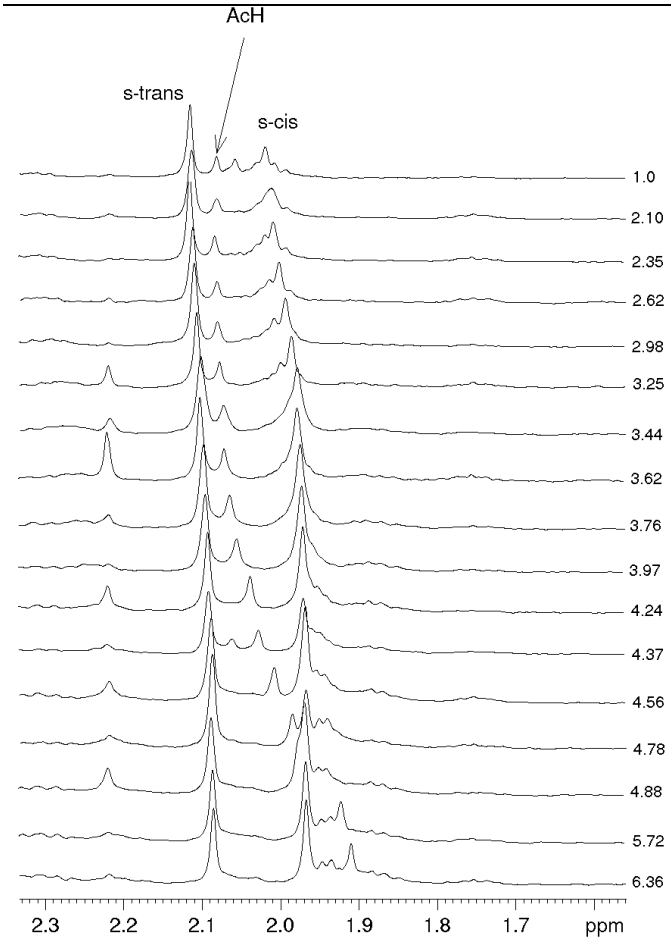


Figure S9. CH_3-range of the ^{1}H NMR spectra of Ac-Pro in Citric Buffer at different pH. The spectra were referenced to the upfield TPS resonance set to zero.
The resonances exhibited a clear sigmoidal shape for both resonances. The marked acetic acid additive resonance had an salt-acid transition at higher pH values. It provided an additional control for the chemical shift and pH.

[S1] Sodium trimethylsilylpropanesulfonate.

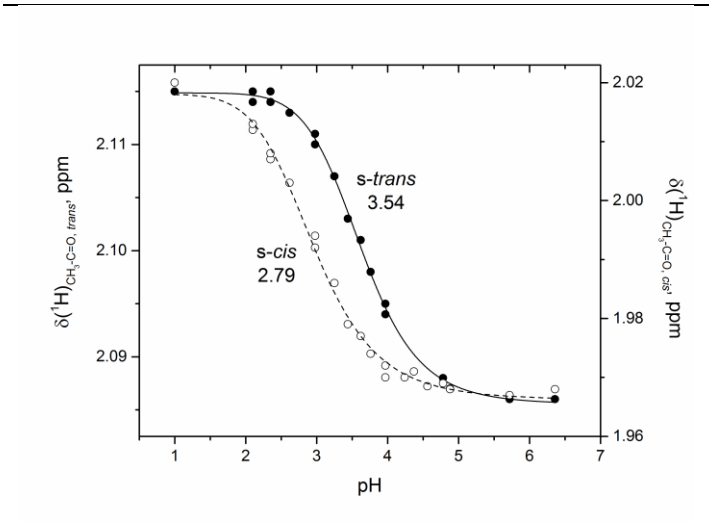


Figure S10. The sigmoidal curves for the CH_3-resonances in Ac-Pro at variable pH. Original data-points are shown with the Logistic fit curves.

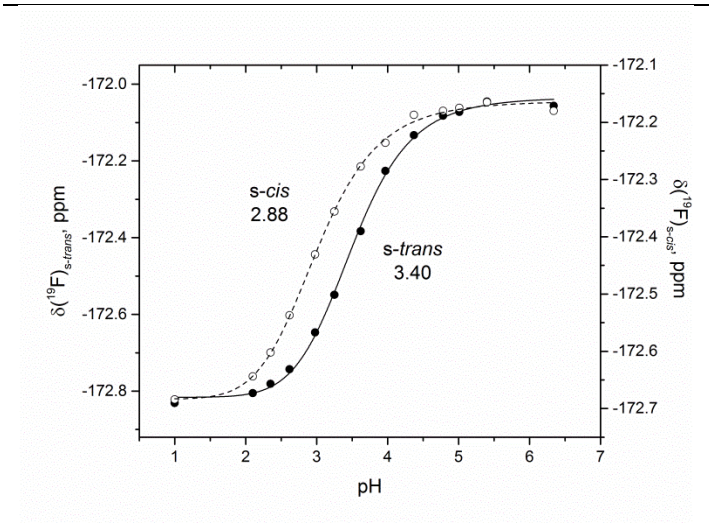


Figure S11. The sigmoidal cirves of the ^{19}F resonances in Ac-(4*S*)-Flp. Original data-points determined in proton-decoupled spectra are shown with the sigmoidal (Logistic) fits to them. No lock was applied to the samples during measurements.

The TPS most upfield resonance was set to 0 ppm. The assignment of the s-*cis* rotameric state at pH 1 was done with a help of 1D selective NOE experiment (mixing time 1 s), where exchange between the s-*cis* and s-*trans* form was detected. Also, the acetic acid additive resonance exhibited the transition well above the pK_a transition of Ac-Pro.

In this way we were able to compute pK_a values and noticed that the population of the rotamers was changing along with the pH changes.

The values of the CH_3 resonances were read out, plotted against pH and fitted according to logistic fit (OriginPro 9.1). The fits were differentiated (1st order) and the derivative maxima were taken as pK_as (Fig. S10).

Derived this way pK_a values were found to be 3.54 for the s-*trans* and 2.79 for the s-*cis* forms of Ac-Pro respectively, as shown on Fig. S10.

However with this method could not enable quantification of the rotameric equilibrium. Also it's drawback was the detection of only one data-point below pH 2.1, therefore the sigmoidal behavior couldn't be concluded certain for the more acidic s-*cis* rotamer.

For Ac-Flp the ^{19}F NMR spectra were recorded without locking the samples. The resonances were read out without referencing and plotted against pH following the same fitting procedure as for Ac-Pro before.

Only for Ac-(4*S*)-Flp we managed to obtain reasonable values as illustrated on Fig. S11. The placement of the data-points indicated that indeed the acid-base transition was complete and the terminal states were detected correctly. Also, no deviation from the sigmoidal behavior was observed in this case.

4.2 Approach B

Direct titration of Ac-Xaa

Solutions containing Ac-Pro, Ac-Flp or Ac-Gly were titrated directly (without buffers) to establish different pH values using potassium hydroxide and hydrochloric acid solutions. 500 µl portions were taken into NMR tubes and then 50 µl of TPS solution (0.3 mg/ml) in deuterium oxide was added to each tube.

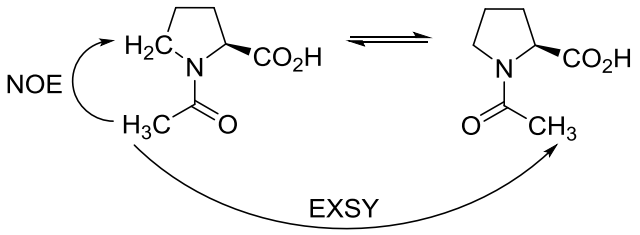
4.2.1 Ac-Pro

The final Ac-Pro concentration in the NMR tubes was 25 mM. For each sample an ^{1}H spectrum (W5 water suppression, "zggpw5"), a $^{13}C\{^{1}H\}$ inverse-gated decoupled ("zgig30") spectrum was collected. In addition, at each pH a selective ^{1}H 1D NOESY ("selnogp") spectra were taken with the frequency offset set on the major CH_3 resonance (s-*trans*). Mixing time was 1 s. These experiments were performed in order to confirm the positioning of the CH_3-resonance in the minor rotamer (by chemical exchange) and confirm the assignment of the major rotamer (by NOE to the δ-CH_2).

Indeed the selective NOESY spectra confirmed a spatial proximity of the CH_3-group (assigned previously as s-*trans*) to the δ-CH_2 with the resonance at ~ 3.6 ppm. Also the exchange with the s-*cis* rotamer appeared only under transition to the acid and was enhanced by decreased pH (Fig. S12, Scheme S3).

Separate appearance of the δ-CH_2 resonances from the two rotamers and invariance of the resonance positions under all pH values enabled us to determine the $K_{trans/cis}$ ratios by integration (Fig. S14). The spectra were processed using Gaussian line broadening (lb = -1.0, gb = 0.05), that enabled the best separation of the multiplets. In addition we applied baseline correction on the region of interest. The $K_{trans/cis}$ values determined this way exhibited a clear sigmoidal behavior (Fig. S16) with the bending point at pH ~ 2.8.

Also the pK_a values determined in this way were 3.56 for the s-*trans* rotamer and 2.79 for the s-*cis* rotamer, respectively (Fig. S15).



Scheme S3. Rotameric transitions in Ac-Pro with indication of the 1D NOESY visible connectivities to the s-*trans* acetyl group, observed as depicted below (Fig. S12).

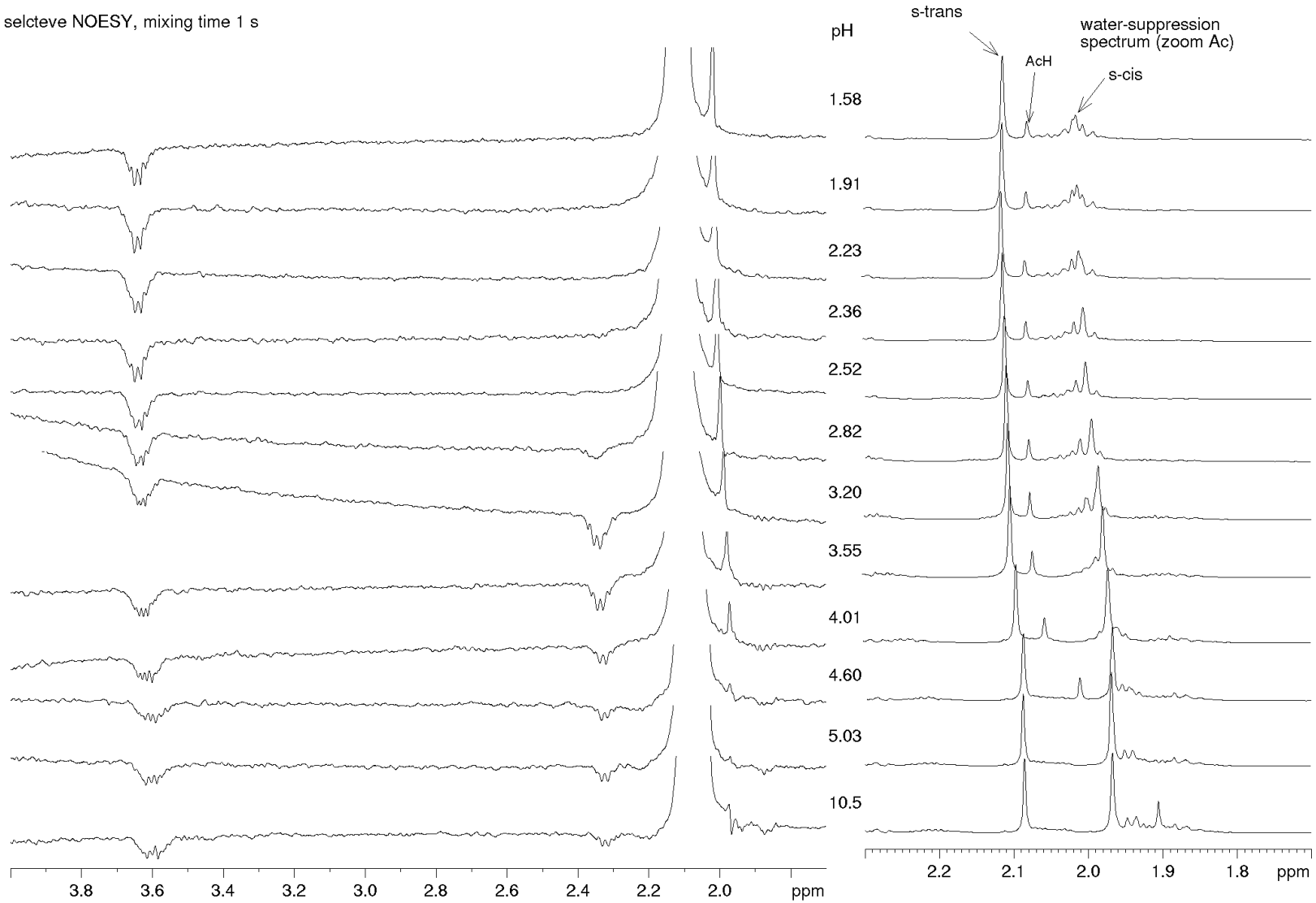


Figure S12. Left: 1D selective ^{1}H NOESY (mixing time 1 s), and Right: CH_3-range of the ^{1}H spectra (W5-water suppression) of Ac-Pro at different pH values. The 1D ^{1}H NOESY spectra exhibited a baseline irregularity since no water-suppression technique was applied within the pulse sequence.

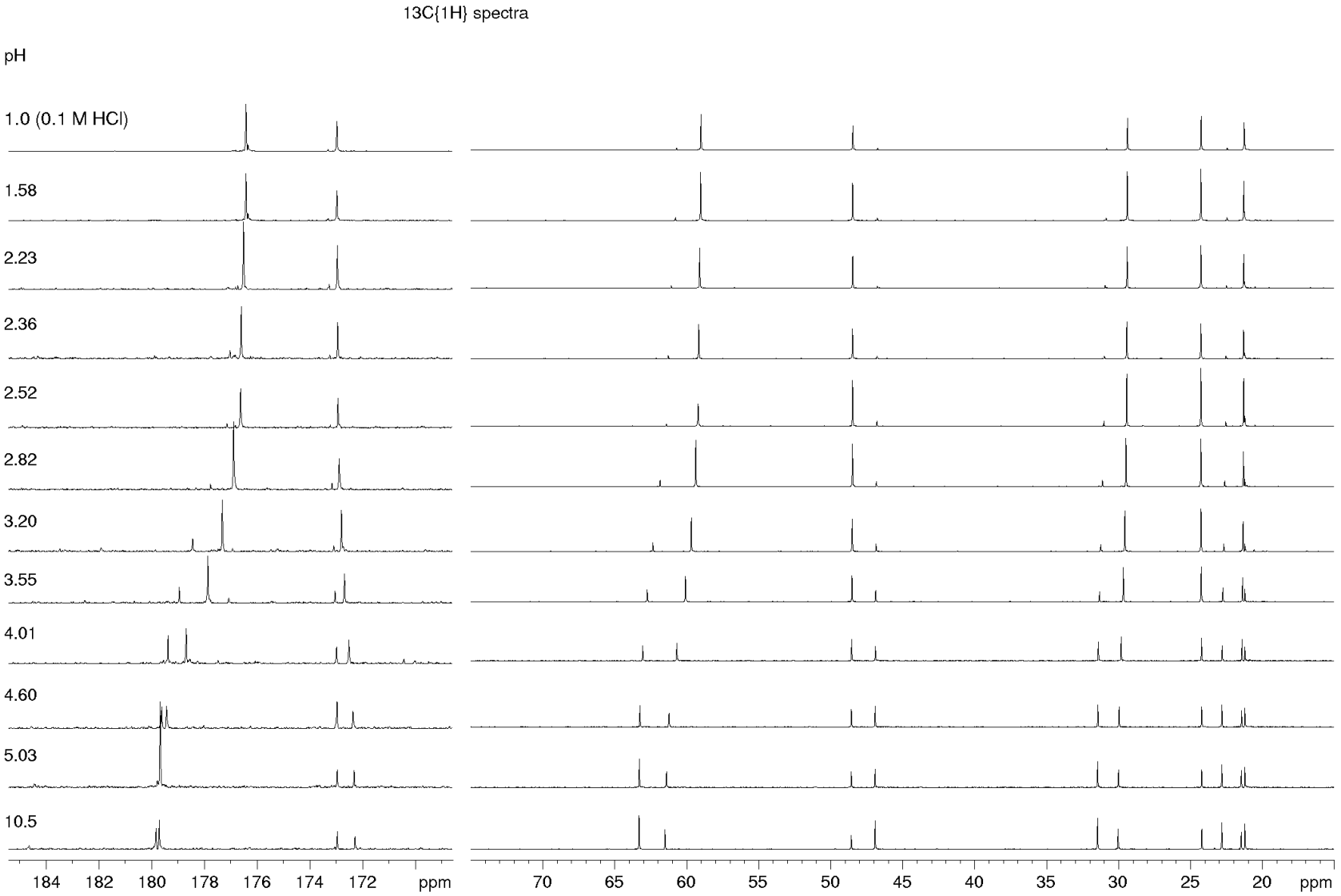


Figure S13. $^{13}C\{^{1}H\}$ NMR spectra of Ac-Pro (25 mM) at different pH. The spectra were taken under inverse gated ^{1}H decoupling with a 30-degree ^{13}C flip angle. Power spectra.

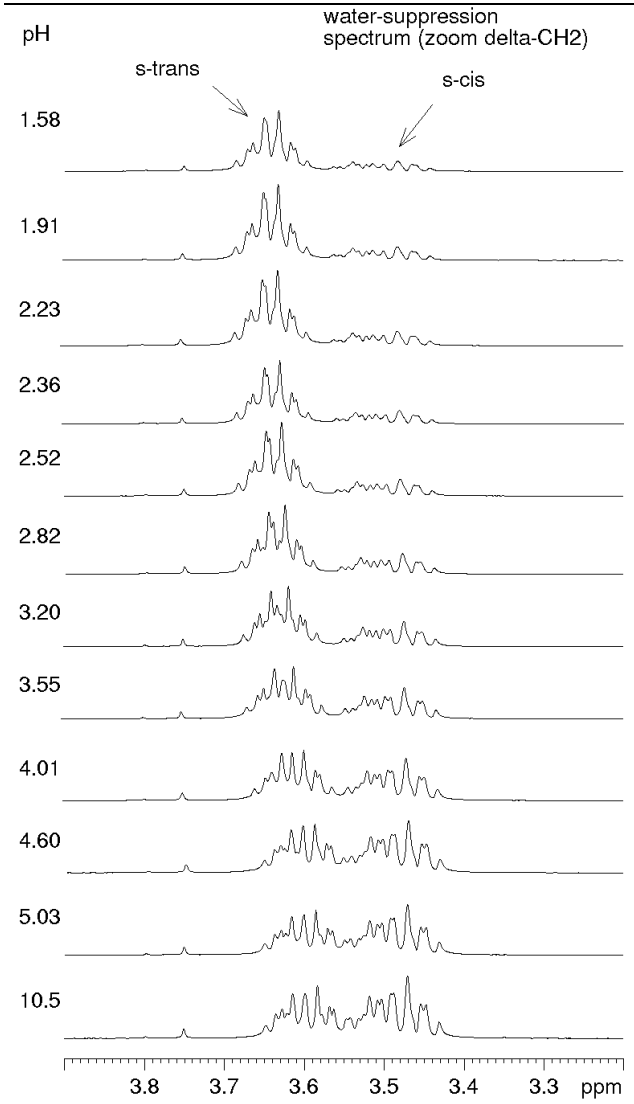


Figure S14. δ-CH_2 region in ^{1}H NMR spectra (W5 water suppression at 4.7 ppm) of Ac-Pro at different pH.

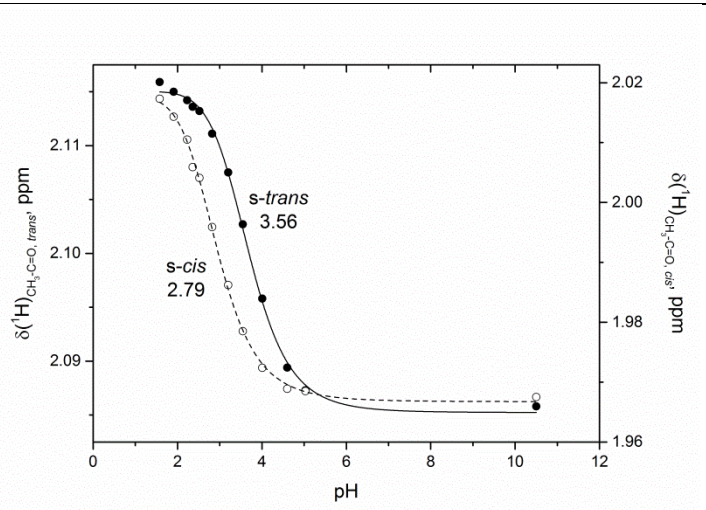


Figure S15. The sigmoidal curves of the CH_3-resonances in ^{1}H NMR of 25 mM Ac-Pro samples prepared by direct titration.

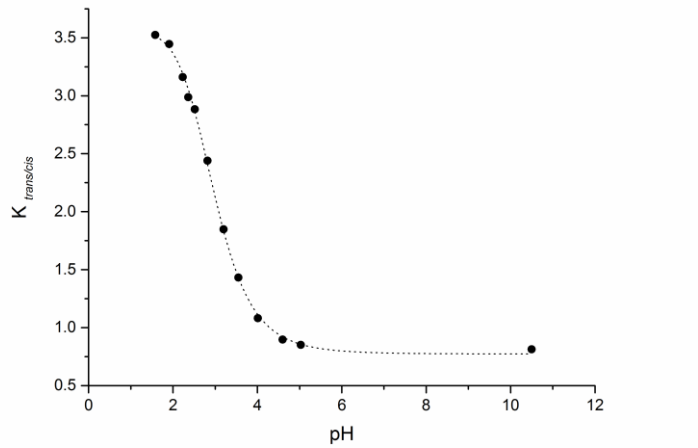


Figure S16. Rotameric ratio ($K_{trans/cis}$) of Ac-Pro dependence from pH.

4.2.2 Ac-Flp

Ac-(4*S*)-Flp and Ac-(4*R*)-Flp were combined in a mixture (4 mM concentration of each analyte) and titrated. 500 µl samples were taken to NMR tube, no deuterium oxide or other components were added thereafter. The $^{19}F\{^{1}H\}$ NMR spectra were measure with inverse gated decoupling ("zgig30") pulse sequence without lock. The chemical shifts were plotted against pH, fitted according to Logistic fits (OriginPro 9.1) and resulting fits were derived (1st order) to get the pK_a values as maxima of the derivative curves. An example of a pH $^{19}F\{^{1}H\}$ NMR spectra series is depicted on Fig. S19.

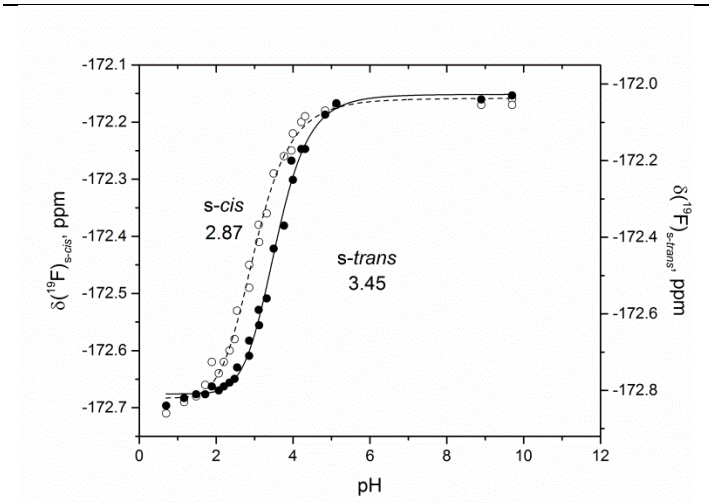


Figure S17. $^{19}F\{^{1}H\}$ NMR resonances of Ac-(4*S*)-Flp under different pH prepared by direct titration. Original data-points and sigmoidal (Logistic) fit curves are depicted.

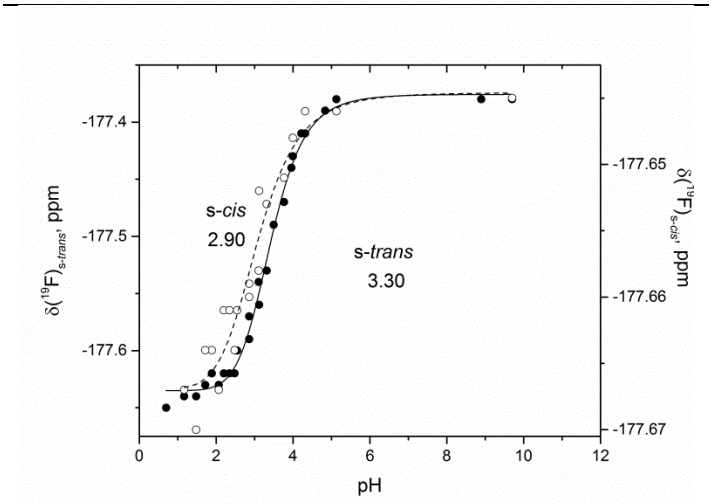


Figure S18. $^{19}F\{^{1}H\}$ NMR resonances of Ac-(4*R*)-Flp under different pH prepared by direct titration. Original data-points and sigmoidal (Logistic) fit curves.

The pK_a values obtained this way for Ac-(4*S*)-Flp (Fig. S17) were consistent with the previously determined in buffered solutions (Fig. S11).

The pK_a values for Ac-(4*R*)-Flp could not be accurately measured in buffered solutions before due to several peculiarities: 1) lack of the data points below pH 2, 2) poor resolution of the rotamer resonances at low pH. By direct titration the values were extracted. Still the s-*cis* resonance position was almost independent from pH and exhibited a quasi-sigmoidal curves delivering a misleading pK_a value of 2.90 (*vide infra*).

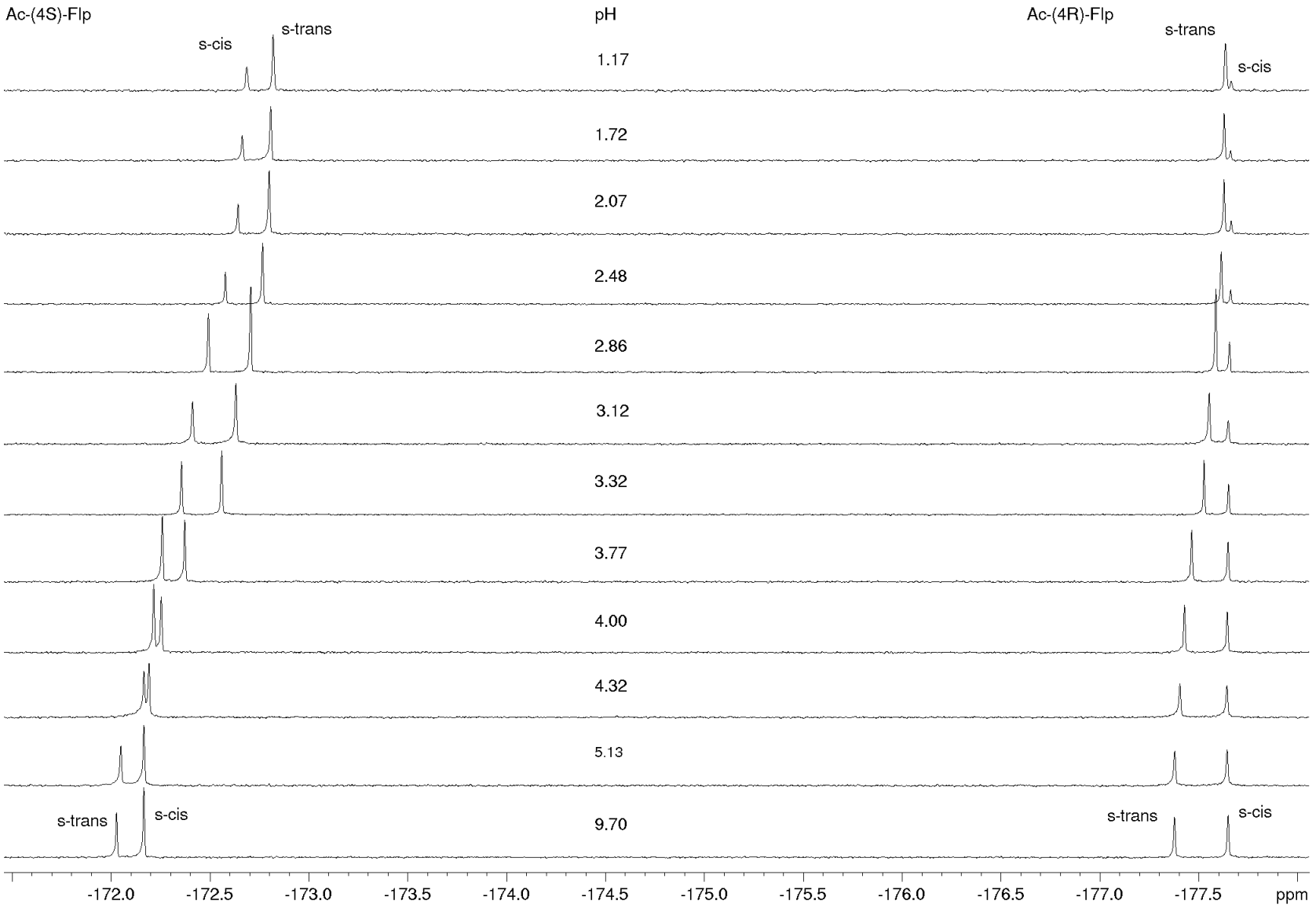


Figure S19. ^{19}F NMR inverse-gated ^{1}H decoupled spectra recorded for Ac-(4*R*)-Flp / Ac-(4*S*)-Flp 1:1 mixture at different pH.

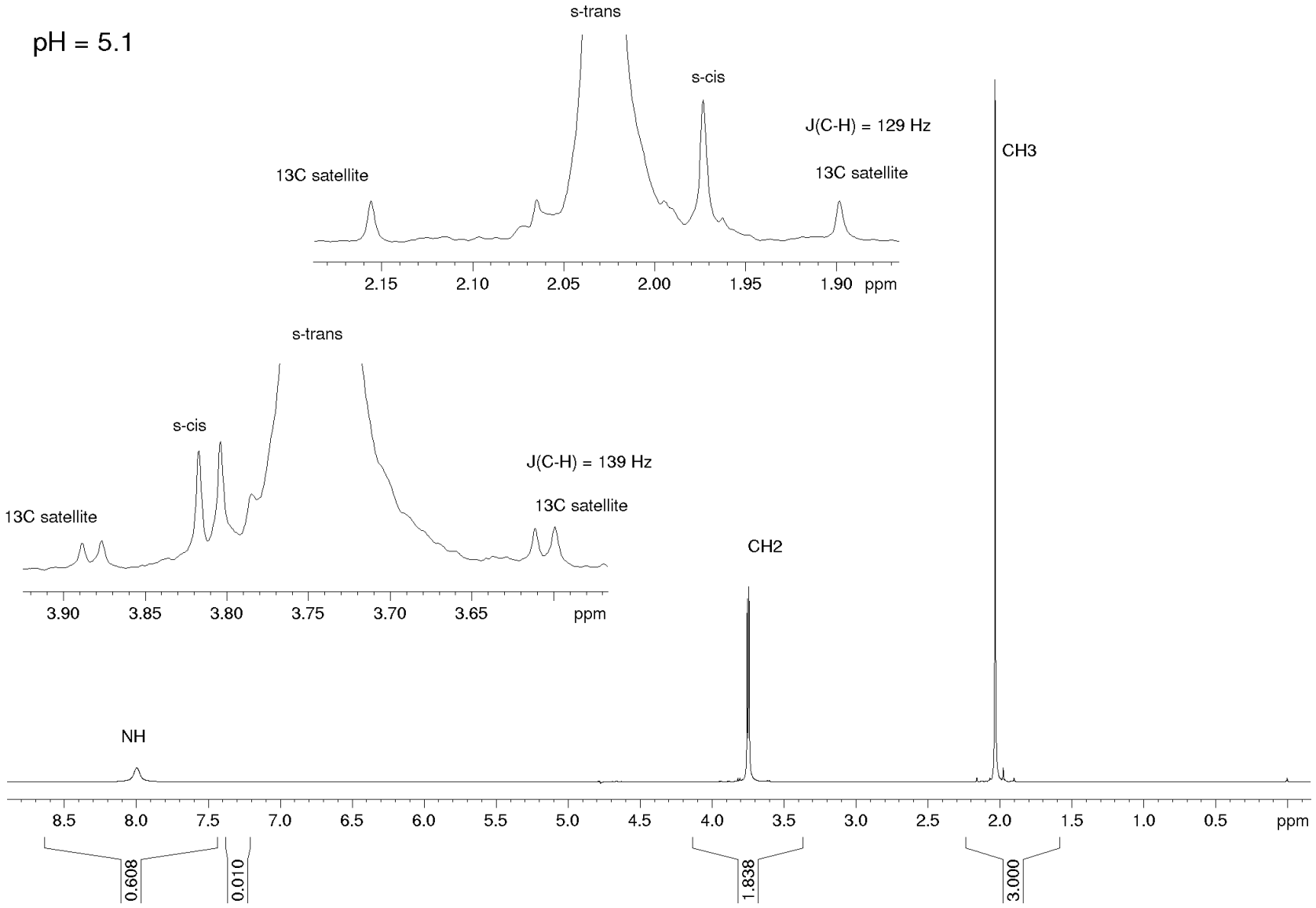


Figure S20. ^{1}H NMR spectrum of Ac-Gly (100 mM) at pH 5.1. Water resonance was suppressed using a W5 pulse tray. The upfield chemical shift of TPS was set as zero.

4.2.3 Aceturic acid (Ac-Gly)

Aceturic acid was titrated at concentration 100 mM. High concentration was needed in order to get the minor rotameric form on a detectable level. 500 µl aliquots were taken into NMR tubes, and 50 µl of TPS deuterium oxide solution (0.3 mg/ml) was added in each one. The W5-water suppression ^{1}H NMR spectra were measured at 500 MHz frequency.

The assignment of the spectrum for basic pH is given on Fig. S20. We noticed sigmoidal behavior for both, the main s-*trans* and the very minor s-*cis* resonances in both cases: α-CH_2 and CH_3. In addition we noticed that the s-*cis* form intensities were decreasing at lower pH. Approximation of the s-cis content was made using NH resonance integration. It was found that indeed the s-*cis* content was estimated being 0.7 % at pH 1.2 and about 1.6 % at pH 5.1 (according to integration of the amide NH region).

Corresponding ^{1}H NMR sigmoidal curves were proceeded as usual (Fig. S21-S22).

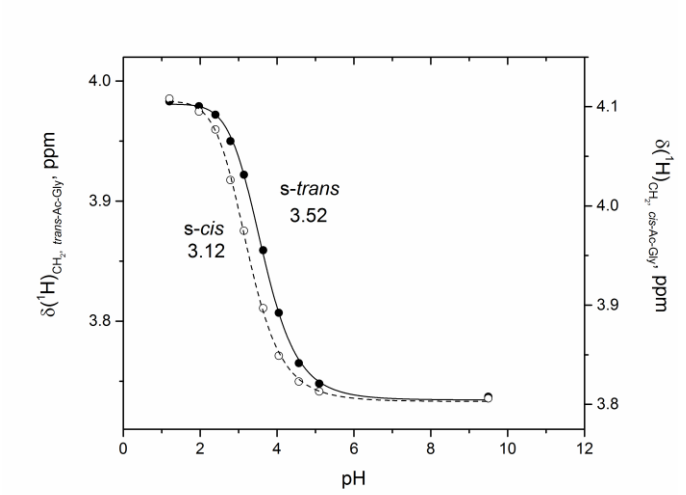


Figure S21. Sigmoidal behavior (original data-points and sigmoidal fit curves) for α-CH_2 ^{1}H resonances of both rotameric forms of aceturic acid.

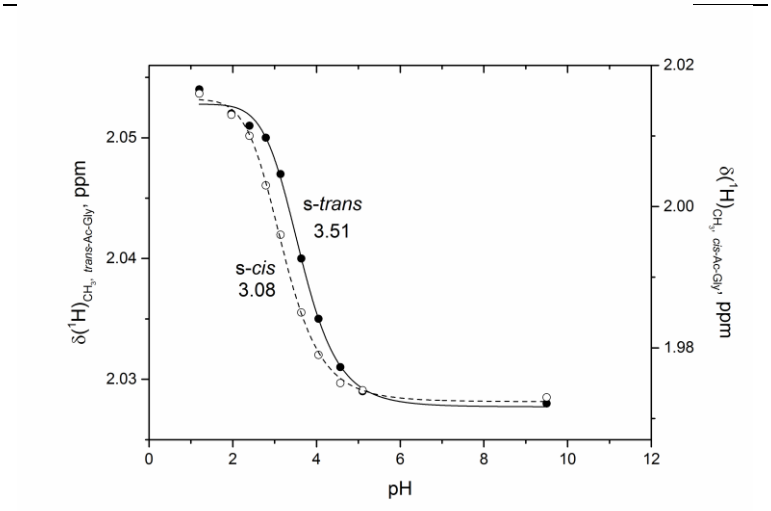


Figure S22. Same for the CH_3 acetyl resonances.

4.3 Approach C

Titration with buffering at high pH.

Ac-Xaa were titrated together with a potassium phosphate buffer. The potassium dihydrogenphosphate was added only above pH 4, to stabilize pH values below 7. Titration was performed by potassium hydroxide and hydrochloric acid. 500 µl aliquots were placed in NMR tubes and 50 µl of 0.3 mg/ml TPS in deuterium oxide solutions was added in each tube. ^{1}H NMR spectra were recorded in W5-water suppressed 1D experiments and referenced to internal TPS downfield resonance afterwards (set as zero).

The concentrations were the following:

Ac-Val	25 mM /	(K-PB 37 mM)
Ac-Ile	25 mM /	(K-PB 37 mM)

α-CH resonance was taken as the most characteristic for plotting. Except of s-*cis*-Ac-Val, for which the acetyl resonance was taken (Fig. S23-S24). The plots were proceeded as before.

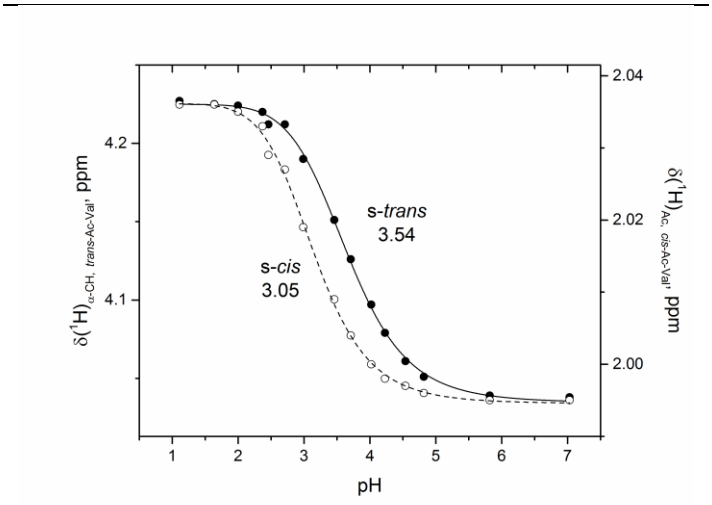


Figure S23. Sigmoidal behavior (original data-points and sigmoidal fit curves) for ^{1}H resonance of s-*trans* Ac-Val (α-CH) and s-*cis* Ac-Val (CH_3-C=O).

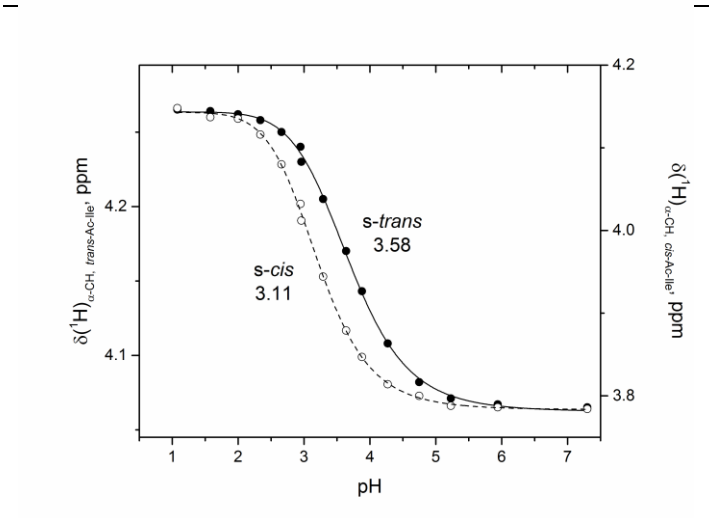


Figure S24. Sigmoidal behavior (original data-points and sigmoidal fit curves) for ^{1}H resonance of rotameric forms of Ac-Ile (α-CH).

4.4 Approach D

Titration in buffered solutions with the references.

The values obtained according to this method are reported in the main body of the paper. Because the pH read out was performed in solutions already contained 10 % deuterium oxide, the pK_a values were corrected thereafter by subtraction of 0.04[S1] and given in the main text. In this chapter of SI we show only the original non-corrected determined pK_as.

Ac-Xaa and potassium dihydrogenphosphate were dissolved 9 ml of water and 1 ml of TPS solution in deuterium oxide was added. The final concentration of the TPS standard was approx. 6 µg/ml (~ 30 µM) and the solvent was 10 % D_2O in H_2O. The final concentrations of the analyte/phosphate buffer was:

Ac-Pro	13 mM / 12 mM
Ac-Gly	43 mM / 12 mM
Ac-(4*S*)-Flp	3 mM / 6 mM
Ac-(4*R*)-Flp	3 mM / 6 mM
Ac-Val	40 mM / 32 mM
Ac-Ile	41 mM / 32 mM

These solutions were directly titrated with potassium hydroxide and hydrochloric acid in order to establish different pH values. 500 µl aliquots of the titrated solutions were taken as ready for measurements.

pH series were measured on Bruker Avance III 700 spectrometer equipped with a z-gradient TXI probe. The ^{1}H NMR spectra were taken in W5-water suppression experiment on 700 MHz frequency, referenced to the upfield TPS resonance (set as zero). Acetyl-group CH_3-resonances were plotted against pH for analysis for all Ac-Pro analytes. For Ac-Gly, Val, Ile we plotted α-CH resonances (CH_2 in the case of Gly), which were the most characteristic and still accessible. The plots were fitted according to Boltzmann fits (OriginPro 9.1), the fit curves were 1st order derived and the maxima of the derivative curves were taken as pK_a values.

Detailed graphical analysis is given on Fig. S25-S38. It should be noted that ^{1}H and ^{19}F resonance plotting for Ac-(4*S*)-Flp exhibited a good consistency, also with the previously determined values. At the same time the ^{1}H NMR results for Ac-(4*R*)-Flp s-*cis* rotamer delivered a very different value to the one found in ^{19}F series in direct titration before (Fig. S18). Indeed the latter value can be considered as a measurement artifact due to very poor response of ^{19}F resonance to the acid-salt transition in this case.

[S1] A. Krężel, W. Bal, *J Inorg. Biochem.* **2004**, *98*, 161-166.

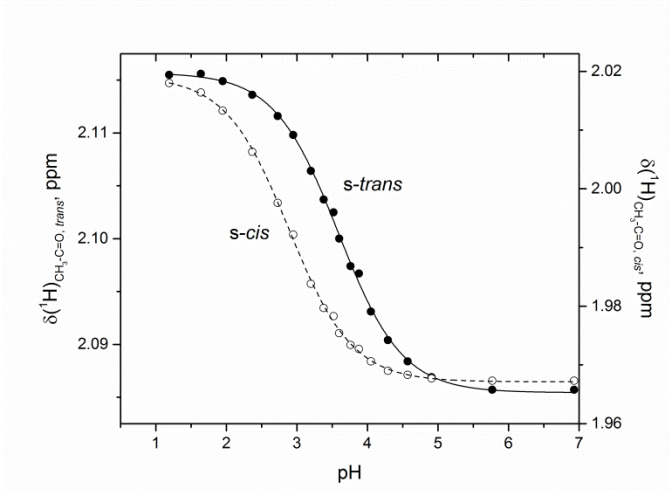


Figure S25. Sigmoidal transitions for Ac-Pro in ^{1}H NMR (CH_3-group).

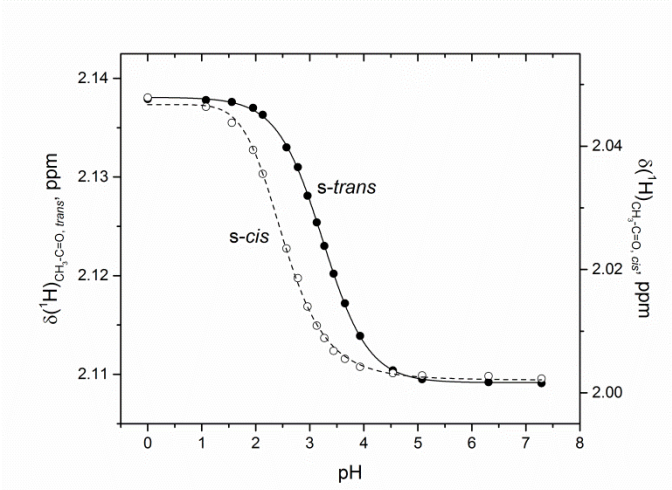


Figure S26. Sigmoidal transitions for Ac-(4*R*)-Flp in ^{1}H NMR (CH_3-group).

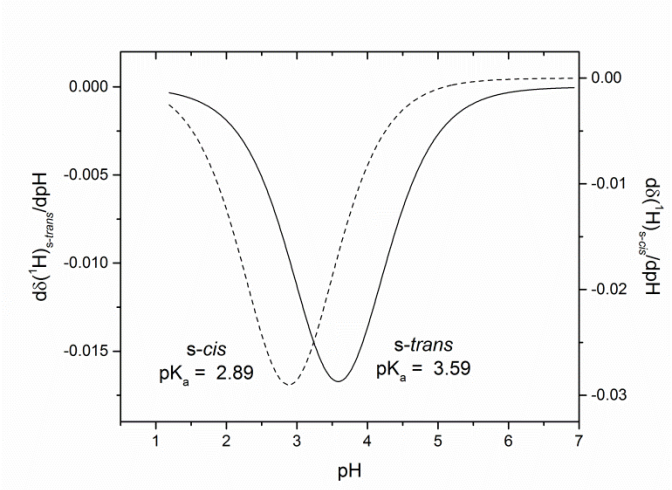


Figure S27. 1st order derivative of the Ac-Pro sigmoidal curves.

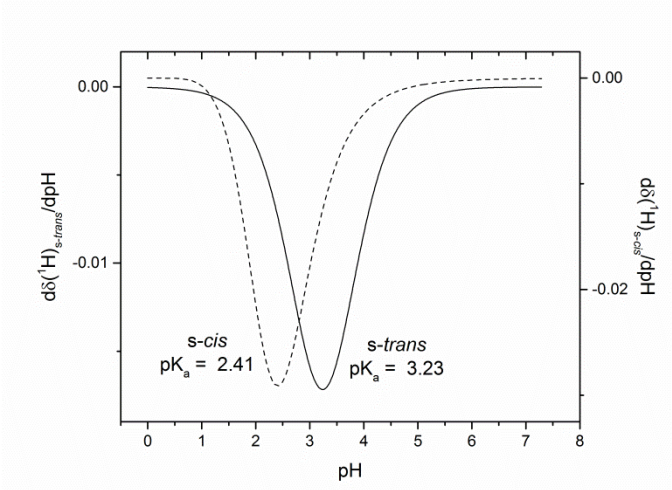


Figure S28. 1st order derivative of the Ac-(4*R*)-Flp sigmoidal curves.

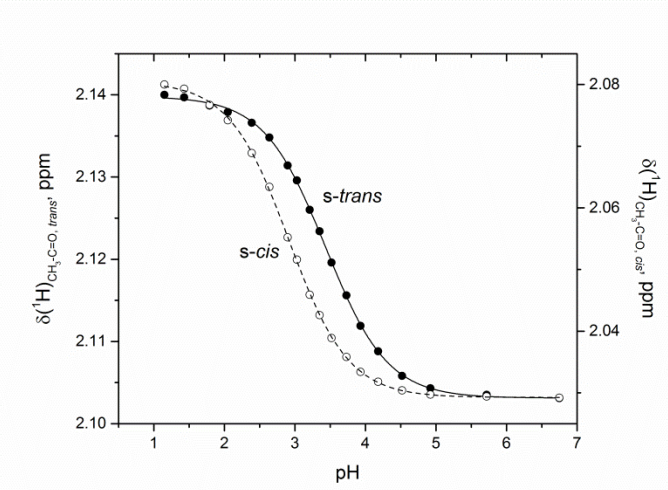


Figure S29. Sigmoidal transitions for Ac-(4*S*)-Flp in ^{1}H NMR (CH_3-group).

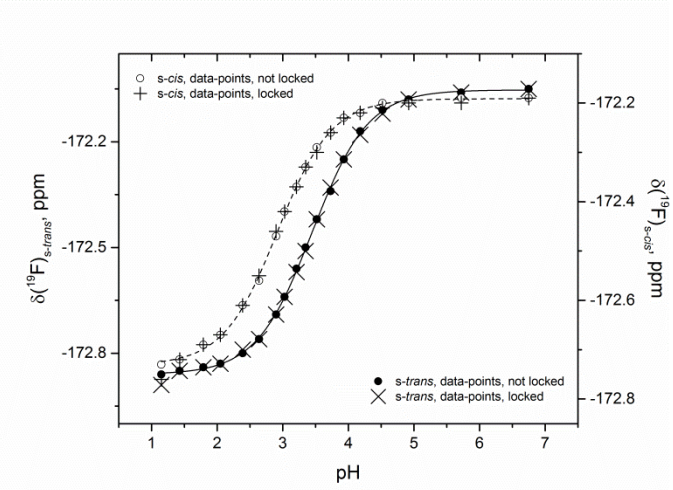


Figure S30. Sigmoidal transitions for Ac-(4*S*)-Flp in ^{19}F NMR with and without deuterium lock applied.

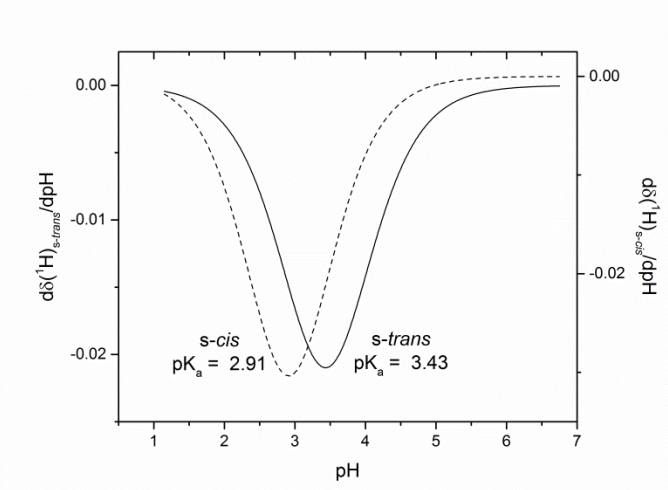


Figure S31. 1st order derivative of the Ac-(4*S*)-Flp sigmoidal curves.

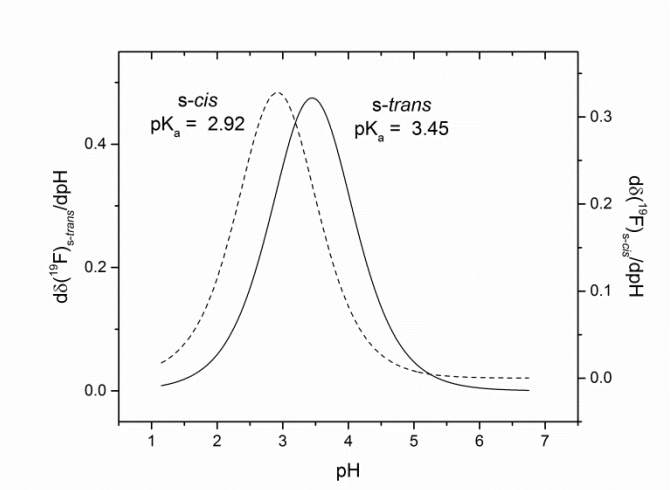


Figure S32. 1st order derivative of the Ac-(4*S*)-Flp sigmoidal curves.

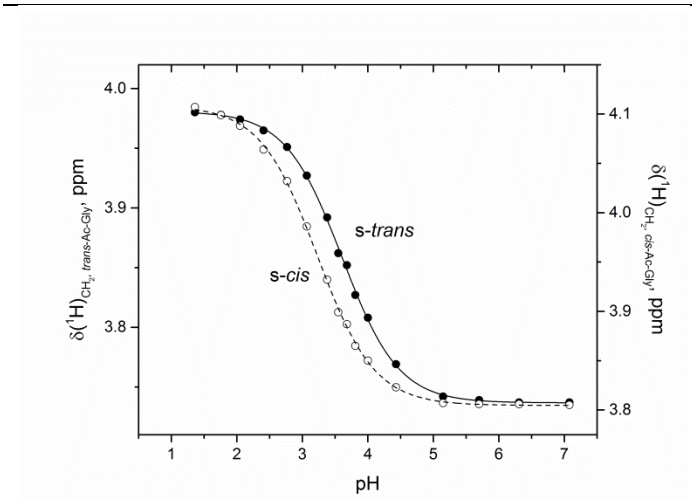


Figure S33. Sigmoidal transitions for Ac-Gly in ^{1}H NMR (CH_2-group).

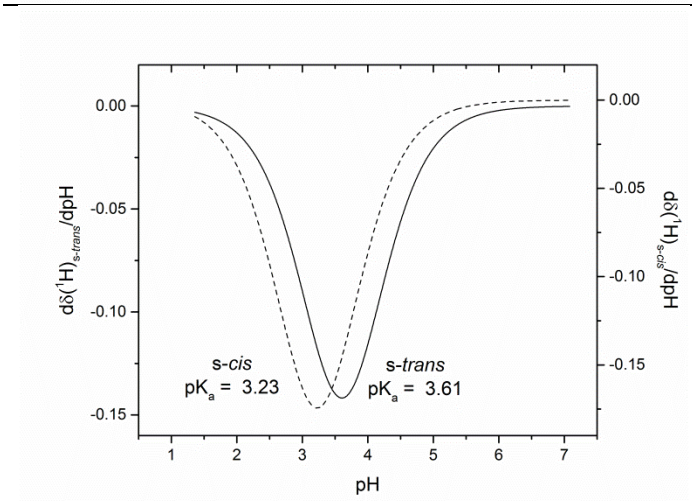


Figure S34. 1st order derivative of the Ac-Gly sigmoidal curves.

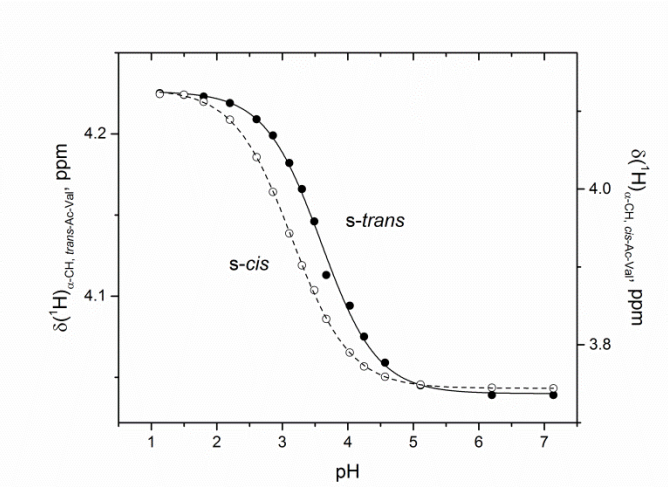


Figure S35. Sigmoidal transitions for Ac-Val in ^{1}H NMR (α-CH).

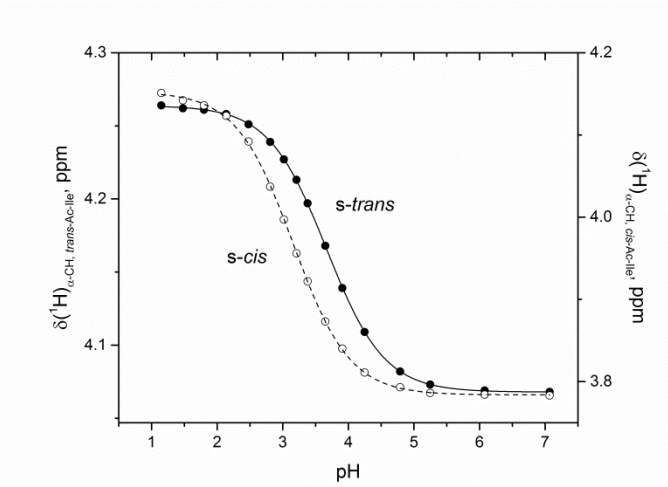


Figure S36. Sigmoidal transitions for Ac-Ile in ^{1}H NMR (α-CH).

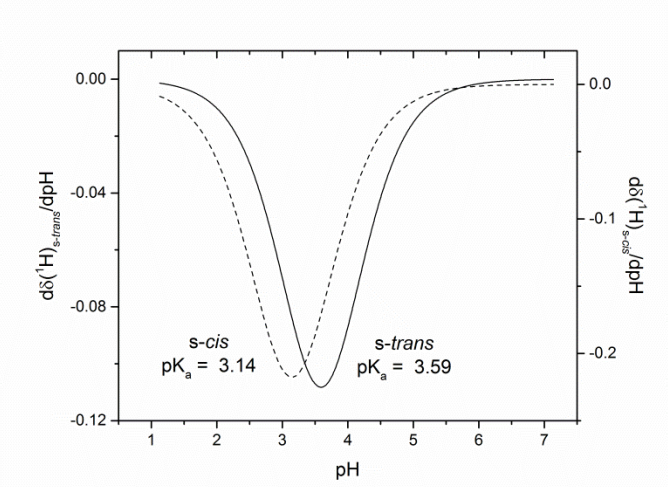


Figure S37. 1st order derivative of the Ac-Val sigmoidal curves.

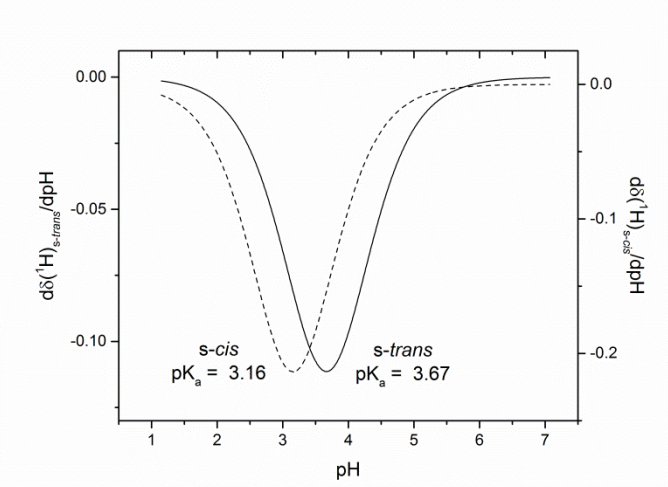


Figure S38. 1st order derivative of the Ac-Ile sigmoidal curves.

5. Thermodynamic parameters of the *trans-cis* isomerization

The *trans/cis* ratio ($K_{trans/cis}$) for Ac-Pro can be determined by integration of the δ-CH_2 resonances in ^{1}H NMR as was illustrated before on Fig. S14. The robust W5-water suppression used to remove the water resonance was set on 4.7 ppm, which was about 1 ppm away from the δ-CH_2 region at 298 K. The δ-CH_2 resonances were not affected by the suppression as was checked by the consistency of the integrals found in spectra recorded at 500 and 700 MHz ^{1}H resonance frequencies (> 500 and > 700 Hz distance from the water position respectively). However, upon increase of the temperature the water resonance position (the suppression position) approaches the resonances of the analyte and the latter started to deviate from the actual values towards decrease of the $K_{trans/cis}$, since the δ-CH_2 of the trans rotamer is more downfield. This was seen as a decrease in the total δ-CH_2 integral compared to the rest of the integrals in the upfield range (see Tab. S1). This was disturbing the initial thermodynamic analysis in heating series, leading to overestimation of both ΔH and ΔS towards more negative values for both ionization states.

Table S1. The initial determination of the $K_{trans/cis}$ of Ac-Pro in ^{1}H NMR spectra recorded using W5-suppression method. Ac-Pro sample was a direct titrated one (25 mM) from the pK_a determination with pH 5.03. The δ-CH_2 downfield signal was about 3.7 ppm.

T, K	δ(H_2O), ppm	integral intensity		'wrong' $K_{trans/cis}$
		δ-CH_2	β-CH	
296.7	4.78	2.00	1.02	0.84
302.1	4.72	2.00	1.03	0.86
307.6	4.67	2.00	1.07	0.81
318.7	4.56	2.00	1.18	0.76
324.3	4.51	2.00	1.26	0.73
329.8	4.46	2.00	1.35	0.68
335.4	4.41	2.00	1.46	0.68

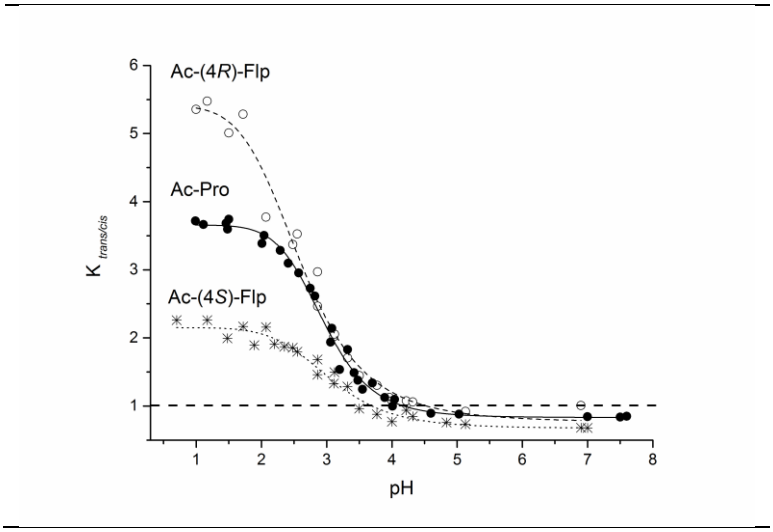


Figure S39. pH dependences of the $K_{trans/cis}$ constants collected over the pK_a determination (298 K).

Therefore, we found the best solution was to measure heating series of ^{1}H spectra in deuterium oxide solution with no suppression method applied. The samples buffered to either acidic (pH 1.4) or basic (pH 7.0±0.2) conditions were prepared in water, lyophilized, then once lyophilized from deuterium oxide and finally taken up into deuterium oxide and measured.

However, for the pH dependences (at 298 K) the ^{1}H NMR W5-suppressed spectra were good enough for the $K_{trans/cis}$ determination. For Ac-Flp $^{19}F\{^{1}H\}$ NMR spectra were integrated (470 MHz). The transmitter offset resonance was set close to the resonances of interest. An integration problem occurred in the case of acidic Ac-(4*R*)-Flp samples, where the resonances from the two rotamers were very close to each other. Gaussian line shape modulation was helpful in resolving these two signals (see Fig. S42). Also, when the concentration was sufficient enough 1-scan spectra were the best resolved. The values for the pH-dependence curves were collected together with the pK_a determination.

Another peculiarity was faced in determination of the $K_{trans/cis}$ constants in the salt forms of Ac-Pro. It turned out that the velocity of the isomerization process was too low (i.e. below accessible for EXSY determination). Therefore the $K_{trans/cis}$ values determined on fresh samples below pH 5 were around 1.0. But after some time (e.g. measured on the next day) or after a heating series the constant equilibrated around 0.81. This is also illustrated on Fig S40, where the fresh Ac-Pro salt sample exhibit hysteresis of the equilibrium values.

Final thermodynamic analysis is given at Fig. S41.

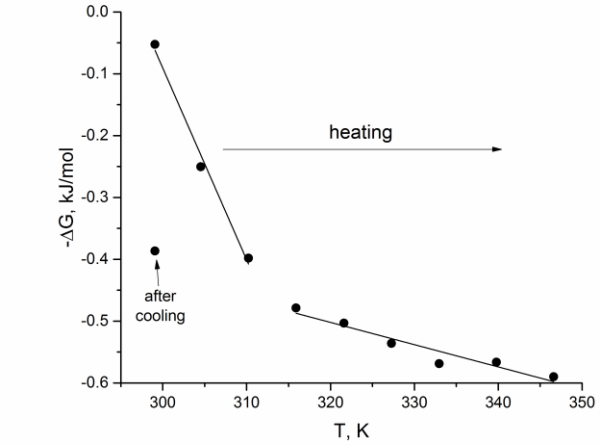


Figure S40. Thermodynamic analysis for Ac-Pro salt sample (pH = 7.0, re-taken in deuterium oxide). The values were obtained in a quick heating temperature series. Whereas the first low-temperature points represent pre-equilibration of the equilibrium ratio, the next points are lying on a "true" straight line. Also, after cooling the free energy value is close to the one from extrapolation of the high temperature line, confirming that the later data points must be true.

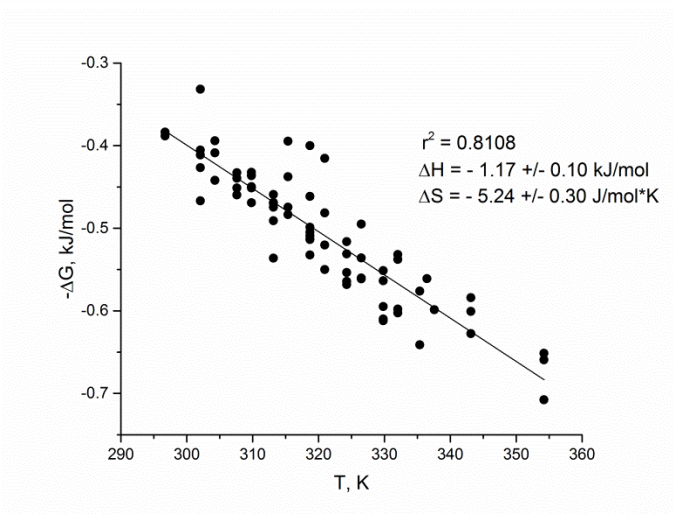


Figure S41. Thermodynamic analysis for the salt form of Ac-Pro (pH = 7.0, re-taken into deuterium oxide).

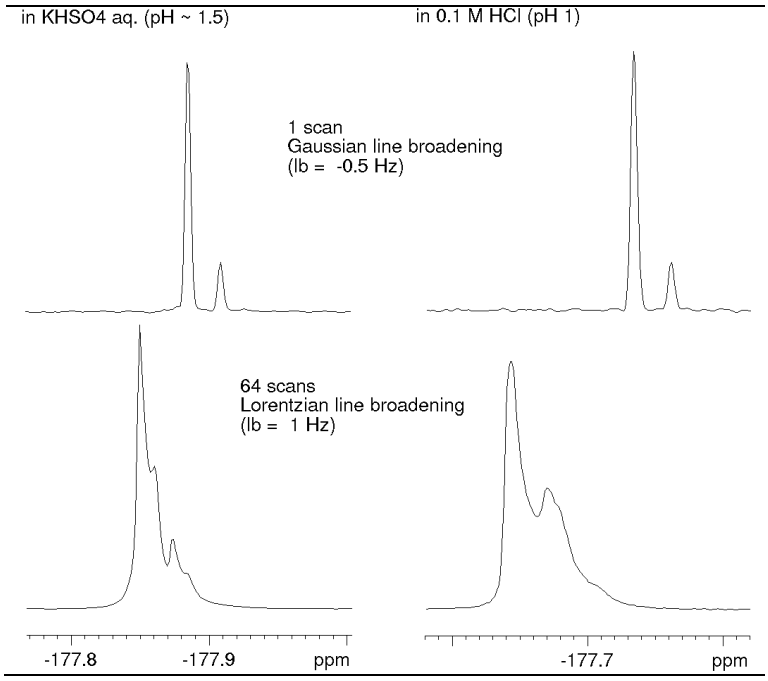


Figure S42. $^{19}F\{^{1}H\}$ NMR spectra of Ac-(4*R*)-Flp in acidic conditions. The resonances from two rotamers were found very close to each other. On the bottom panels multiple-scan spectra with inverse gated decoupling. Due to the warming effect of the decoupling the lines appear blurry. In contrast, one scan spectra processed with Gaussian modulation (top panels) are good resolved for separate integration.

6. Analysis of rotameric populations in Ac-Pro-Pro

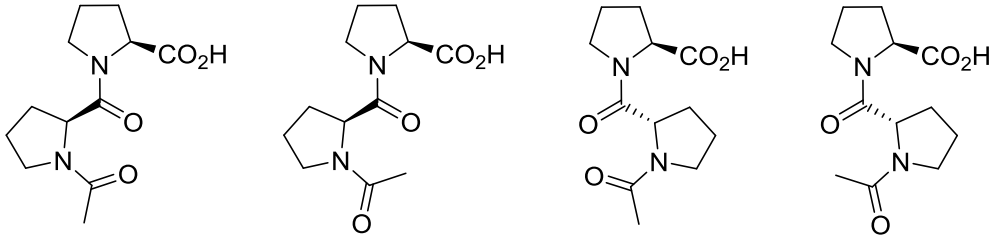


Fig. S43. Rotameric forms of Ac-Pro-Pro.

^{1}H NMR spectra were recorded at 700 MHz resonance frequency. For the reported values we prepared samples containing the following components:

pH 1.5 (acidic form)

Ac-Pro-Pro 75 mM

$KHSO_4$ 140 mM

pH 7.1 (salt form)

Ac-Pro-Pro 75 mM

K-PB 140 mM

The pH was adjusted in aqueous medium, then the samples were freeze-dried, freeze-dried from deuterium oxide (twice) and finally dissolved in amount of deuterium oxide equivalent to the one where pH was adjusted (520 µl sample volume).

6.1 Assignment of the α-CH region

The spectra of the dipeptide were very complicated, since four individual rotameric forms were present (Fig. S43), all containing two inequivalent proline residues. We only assigned the α-CH region, which was then used for integration, correlations with the carbon spectra and the pK_a detection.

Likewise in Ac-Pro, the overall ^{1}H and $^{13}C\{^{1}H\}$ dept45 spectra contained of three characteristic regions: 1) α-CH, 2) δ-CH_2 and 3) the rest (Fig. S44). In the latter group we found characteristic terminal CH_3-group singlets: three in the salt and two in acid, that are marked with asterisks at Fig. S45.

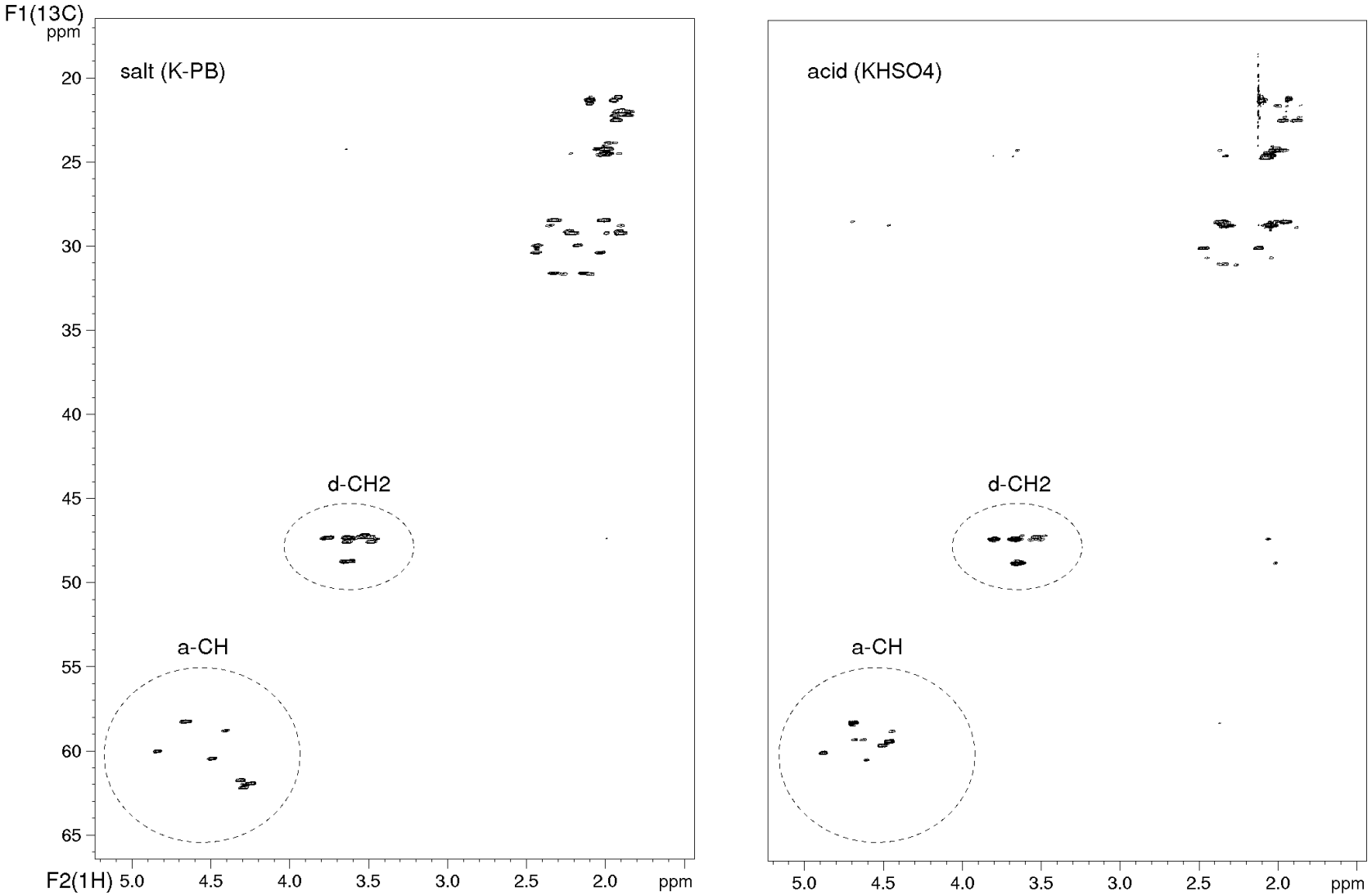


Fig. S44. Slices of the $^{1}H\{^{13}C\}$ HSQC correlations of Ac-Pro-Pro (700&176 MHz frequency) in deuterium oxide solutions.

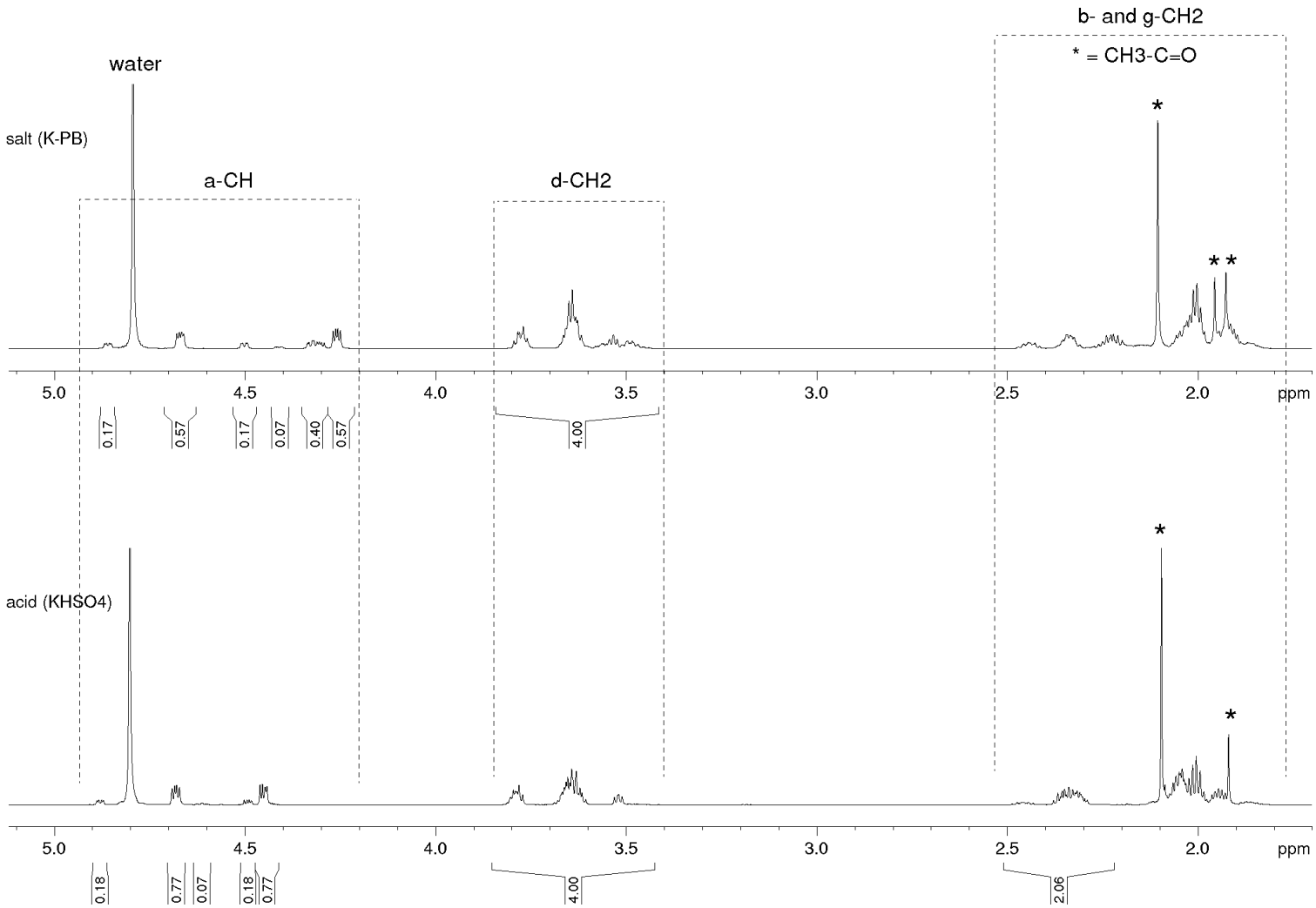


Fig. S45. Overall ^{1}H NMR spectra of Ac-Pro-Pro dipeptide in deuterium oxide solutions at 700 MHz resonance frequency.

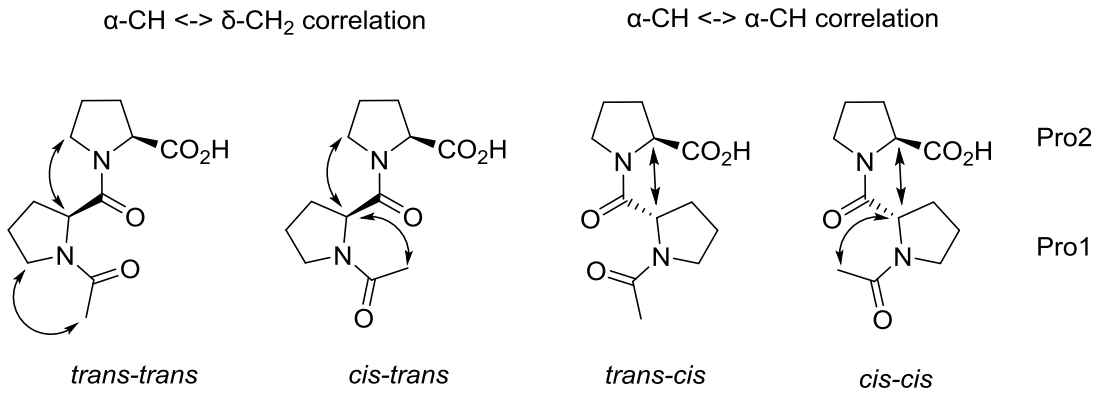


Fig. S46. Expected NOE correlations in Ac-Pro-Pro rotameric forms.

Further analysis was done using NOE-correlations. ^{1}H 2D NOESY experiments were conducted using mixing time 1 s for the acidic form, and 750 ms for the salt. The expected NOE correlations are shown at Fig. S46. The expectations for the rotameric forms were the following:

1) two α-CH of the *trans-trans* rotamer will have the highest absolute intensity, whereas Pro1 α-CH will exhibit NOE with the δ-CH_2 region;
2) in the *cis-trans* forms one of the α-CH (Pro1) will exhibit NOE with δ-CH_2 region and one of the CH_3;
3) in the *trans-cis* form two α-CH (of Pro1 and Pro2) will show NOE with one another, and there will be no NOE with none of CH_3;
4) in remaining *cis-cis* form NOE will be seen between two α-CH (Pro1 and Pro2), and one of them (Pro1) will also exhibit an NOE with a CH_3-group.

The ^{1}H NOESY spectrum for Ac-Pro-Pro in the salt form is shown at Fig. S47. Whereas for *trans-trans* and *cis-cis* forms two α-CH were already assigned to the proline residues (1 or 2), for *trans-cis* form it was only possible to indicate chemical shifts of the α-CH resonances. For *cis-trans* rotamer only one (Pro1) position was immediately found, whereas the other resonance was the only one non-assigned residual signal.

The same analysis afforded ^{1}H NOESY assignment of the Ac-Pro-Pro acidic sample (Fig. S48). The minor forms were in this case of much lower intensity. In addition to NOE, the most intense Pro1 and Pro2 of the *trans-trans* and *cis-trans* rotamers showed presence of the chemical exchange between them (negative cross-peaks).

The best separation of the α-CH chemical shifts was achieved in the ^{1}H{^{13}C} HSQC spectra. Therefore, they were taken for detection of the pK_a transition curves as described in the next chapter. The spectra, that were previously shown at Fig. S44 zoomed to the α-CH region and the transitions are shown at Fig. S49. It is clearly seen that the Pro1 resonances underwent only a little change in both dimensions, whereas the C-terminal Pro2 resonances exhibited the same shift as was seen for Ac-Pro. Namely, upon transition from salt to acid, α-CH resonances were strongly shifted downfield in the ^{1}H dimension and upfield in ^{13}C dimension respectively.

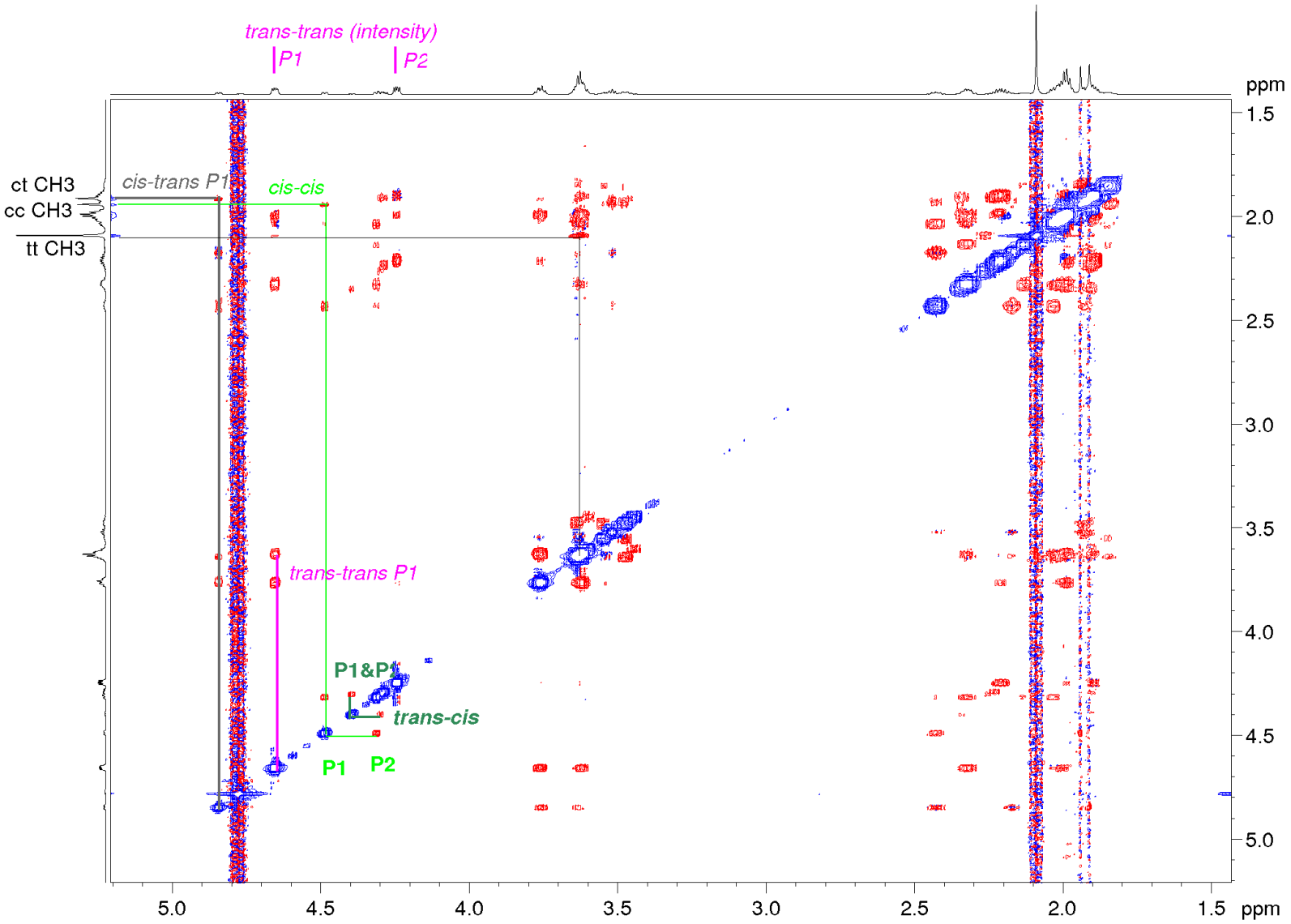


Fig. S47. ^{1}H 2D NOESY assignment for Ac-Pro-Pro in the salt form in deuterium oxide (700 MHz resonance frequency). Mixing time was 750 ms.

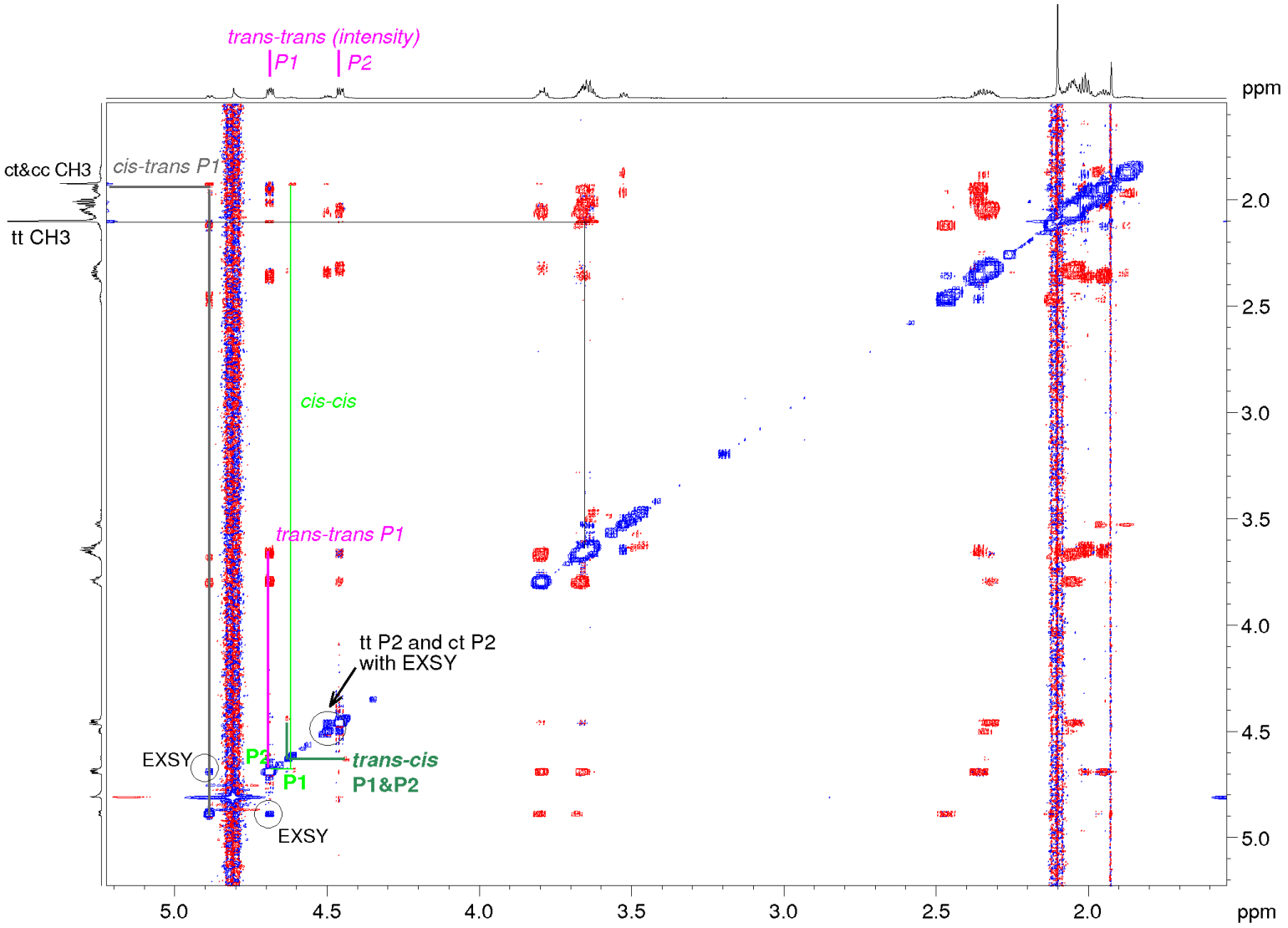


Fig. S48. ^{1}H 2D NOESY assignment for Ac-Pro-Pro in the acid form in deuterium oxide (700 MHz resonance frequency). Mixing time was 1 s.

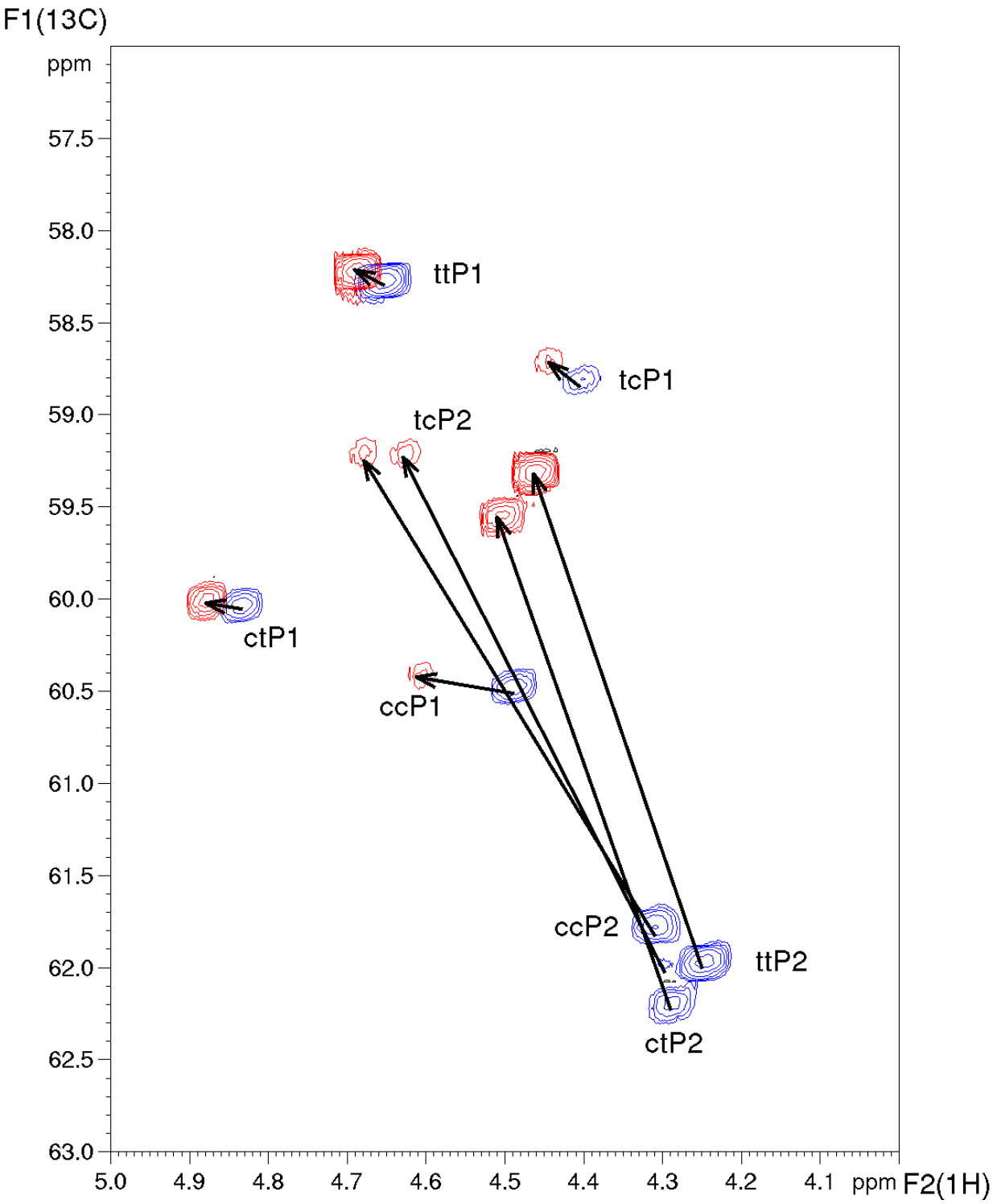


Fig. S49. Overlay of the 1H{13C} HSQC zoomed in the α-CH region for Ac-Pro-Pro salt (blue) and acid (red) samples. Arrows connect same resonances changes upon the salt-acid transition.

6.2 Measurements of the equilibrium populations

^{1}H populations were obtained after integration of several ^{1}H NMR spectra recorded in 1 scan experiments (to avoid relaxation distortions). This is exemplified on Fig. S50. Error was ± 1%.

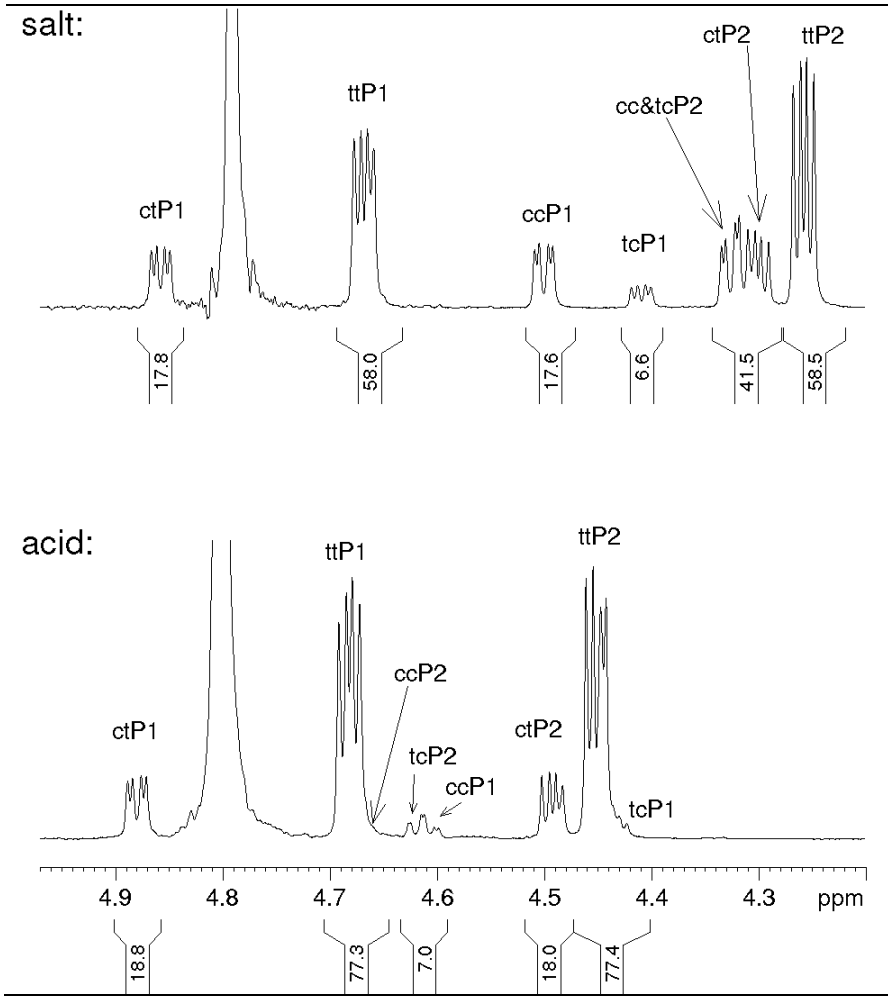


Fig. S50. Integration of the α-CH region in ^{1}H NMR spectra of Ac-Pro-Pro (700 MHz).

For quantitative ^{13}C NMR we adjusted pH in the samples containing the following components:

pH 1.20 (acidic form)

Ac-Pro-Pro 250 mM

K-PB 4 mM

pH 6.90 (salt form)

Ac-Pro-Pro 350 mM

K-PB 4 mM

The pH was adjusted in solutions containing 90 % H_2O and 10 % D_2O. 525 µl of each sample was directly taken for measurements. The $^{13}C\{^1H\}$ NMR spectra were acquired at 126 MHz resonance frequency in 90-degree flip inverse-gated proton decoupling experiments (no NOE enhancement). The transmitter frequency was set on the α-CH range for both ^{13}C and ^1H nuclei. Acquisition time was 1 s, recycling delay 60 and/or 80 s (to ensure full relaxation). Generally 512-768 scans were enough to collect quantitative spectra. Nevertheless, the acidic samples required more scans due to the higher population differences. The spectra were processed with Lorentzian and Gaussian window functions and appropriate baseline correction mode.

The overall integration of the spectra is illustrated on Fig S51. to show overall quantitative character of the experiment. Resonances in α-CH and C=O regions were integrated to achieve quantitative populations of the rotameric forms. Assignment was done via $^1H\{^{13}C\}$ HSQC and $^1H^{13}C$ HMBC spectra. Error was found to be ±1.5 %.

In order to confirm our hypothesis that the domination of the *cis-cis* form over *trans-cis* in the salt form is caused by the ionization state of the terminal carboxyl, and not by the pH alone, we measured all populations in Ac-Pro-Pro-OMe in deuterium oxide solution. Fundamentally the populations obtained for this compound were almost the same as in Ac-Pro-Pro-OH (acidic form), even though, the solution was neutral (Table. S2)

Table S2. Rotameric populations in Ac-Pro-Pro-OMe as found by integration of ^1H NMR spectra (α-CH region, 700 MHz frequency).

	Pro-Pro in s-*trans*		Pro-Pro in s-*cis*	
	trans-trans	*cis-trans*	*trans-cis*	*cis-cis*
	NMR populations, %			
Ac-Pro-Pro-OMe	75	18	4	3

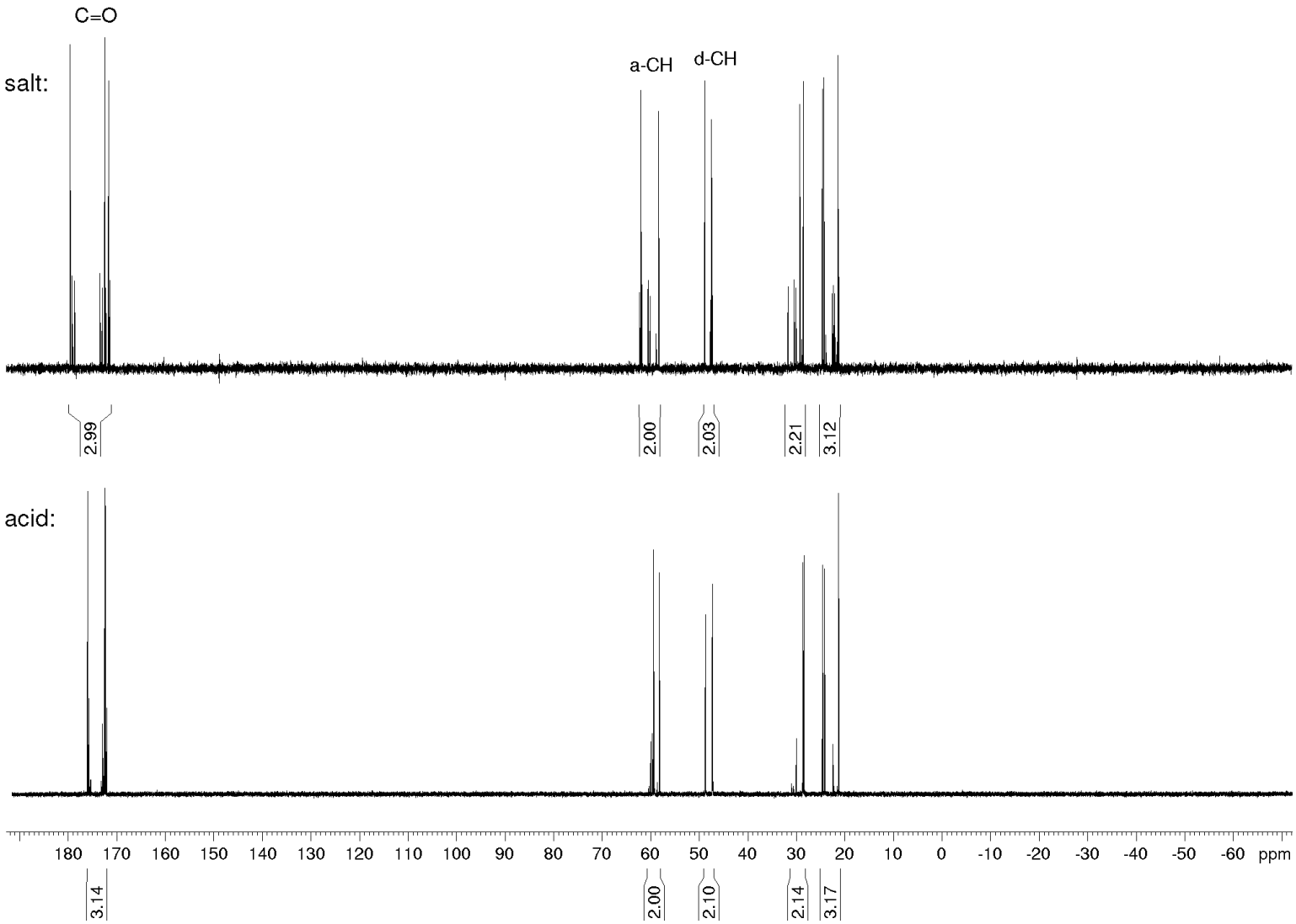


Fig. S51. $^{13}C\{^{1}H\}$ inverse-gated decoupled spectra of Ac-Pro-Pro in different ionization states. Full scale is presented. α-CH region is set to the middle of the spectral window for pulse efficiency.

7. pK_a measurements in Ac-Pro-Pro

Preliminary pK_a measurements were afforded by ^1H experiments and $^{13}C\{^1H\}$ dept45 using approach C (see chapter 4.3) in 40 mM Ac-Pro-Pro solutions buffered with potassium phosphate buffer. Because *trans-trans* and *cis-trans* forms dominated the spectra we obtained pK_a curves from the ^{13}C NMR data for these two forms as depicted on Fig. S52-S53. In addition the pK_a value determined for the *trans-trans* form in ^1H NMR pH series was 3.61 (not shown).

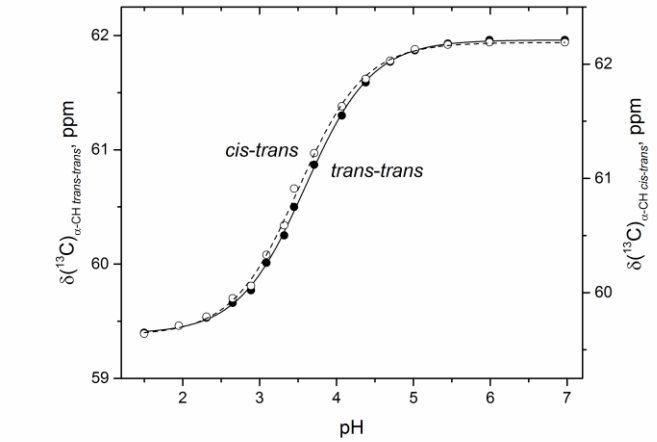


Fig. S52. Sigmoidal transition of Pro2 α-CH ^{13}C resonances in Ac-Pro-Pro *trans-trans* and *cis-trans* rotameric forms.

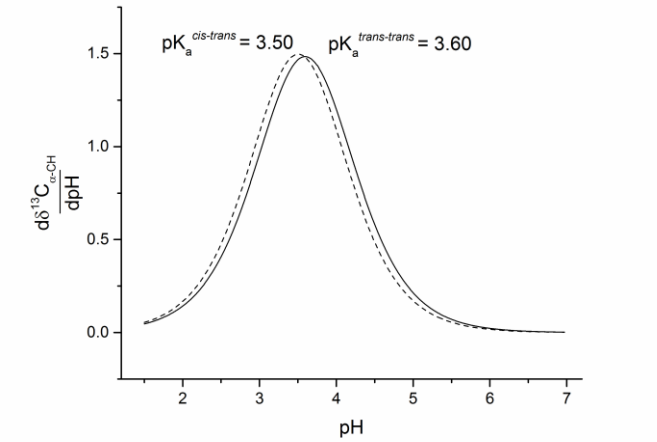


Fig. S53. Corresponding 1st order derivative curves with pK_a values.

A general problem was overlap of several resonances and uncertain assignment of the minor forms upon pK_a transition. In order to obtain all four pK_a values we solved this issue by recording a detailed series of $^1H\{^{13}C\}$ HSQC experiments. Because we were interested only in the α-CH region the spectra were zoomed in both dimensions to the α-CH region. In this way we shortened the acquisition time, whereas the resolution remained good enough (typical FIDRES 3 Hz in ^1H and 19 Hz in ^{13}C dimensions). With HSQC we also were able to measure the samples in H_2O+D_2O 9:1 mixture, whereas the water signal was largely suppressed by the coherence pathway filtering (no flip-back was applied).

25 samples were prepared according to the approach D (chapter S5.4) from a solution containing 100 mM Ac-Pro-Pro and 14 mM K-PB. $^{1}H\{^{13}C\}$ zoomed HSQC correlations were recorded at 700&176 MHz resonance frequency for ^{1}H and ^{13}C respectively. Typical setup was the following:

spectral width: 1.00 x 7.00 ppm (dwell time 714 x 811 μs)

time domain: 512 x 128

acquisition time: 0.37 x 0.052 s (FIDRES 3 x 19 Hz)

recycling delay: 2

number of scans: 2

In several cases (especially for acidic samples) the number of scans or/and time domain in the indirect dimension was increased in order to better detect minor resonances. The values of ^{1}H and ^{13}C Pro2 α-CH chemical shifts were processed as usual (Fig. S54-55). The ^{1}H spectra contained a small additive methanol signal which was used for referencing

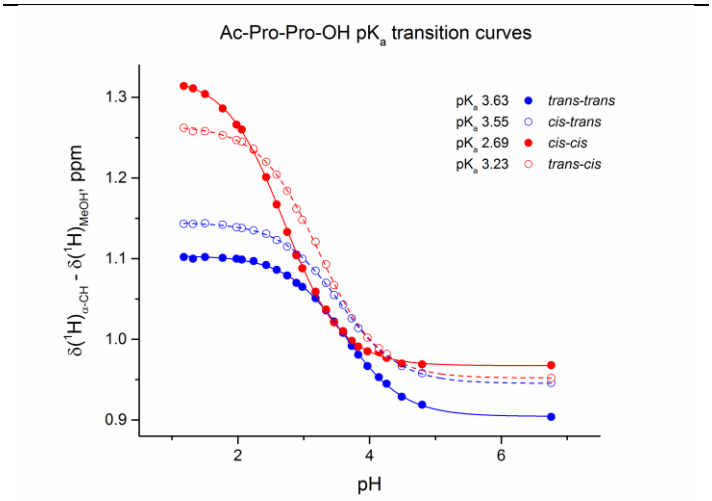


Fig. S54. Sigmoidal transition of Pro2 α-CH ^{1}H resonances in four Ac-Pro-Pro rotameric forms as detected in $^{1}H\{^{13}C\}$ HSQC pH series.

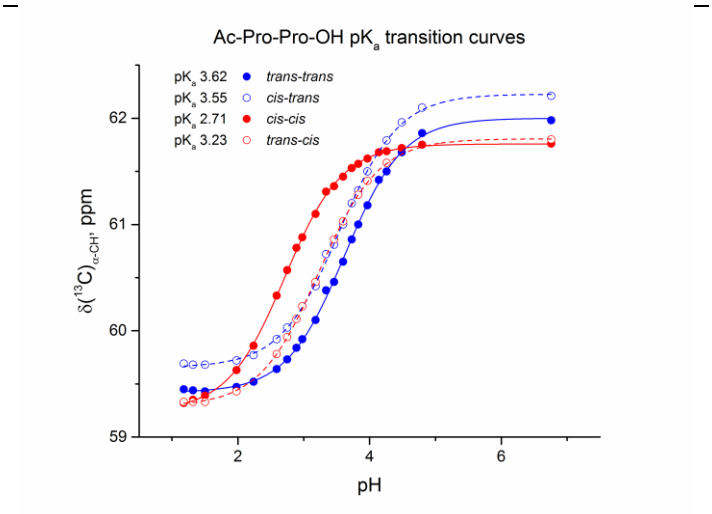
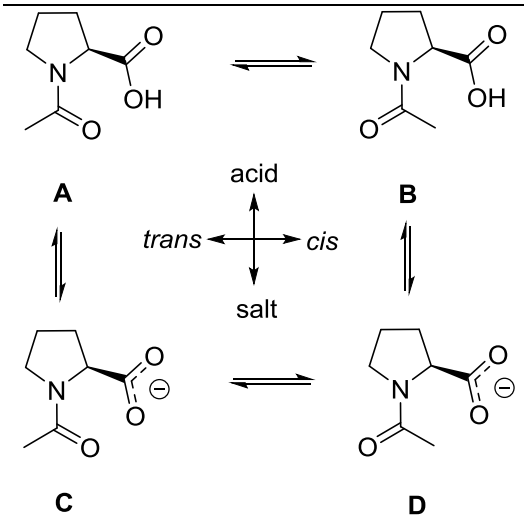


Fig. S55. Sigmoidal transition of Pro2 α-CH ^{13}C resonances in four Ac-Pro-Pro rotameric forms as detected in $^{1}H\{^{13}C\}$ HSQC pH series.

As seen on presented Fig. S54-55 some resonances indeed often overlap in at least one of the dimensions. Though, overlap in both dimensions at the same time was seldom. For final representation we took the mean values from ^{1}H and ^{13}C analysis. Also, because pH was adjusted in a solution containing 10 % of deuterium oxide, we subtracted 0.04 for the final representation.

8. Recalculation of the constants for Ac-Pro



Scheme S4. Equilibrium forms of Ac-Pro in aqueous solution.

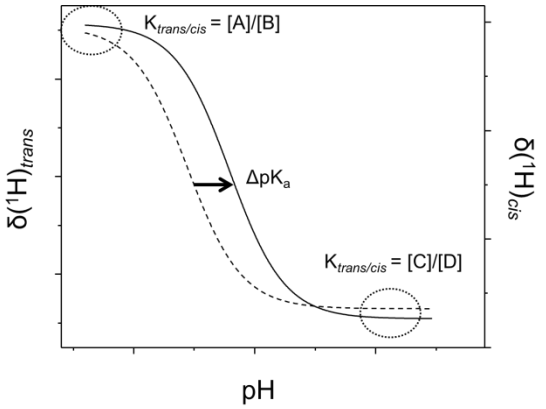


Fig. S56. Schematic representation of the approach.

$$K_a(trans) = \frac{[C][H^+]}{[A]} \qquad K_a(cis) = \frac{[D][H^+]}{[B]} \qquad K_{trans/cis}(acid) = \frac{[A]}{[B]} \qquad K_{trans/cis}(salt) = \frac{[C]}{[D]}$$

K_a defines the free energy difference in the entire pH range, where actual concentration $[H^+]$ defines the ratios [C]/[A] and [D]/[B] for a particular pH value (Scheme S4). $K_{trans/cis}$ constants can be determined in the entire pH range as well as ([A]+[C])/ ([B]+[D]), but only at the terminal pH values, where both forms are acids or when the both forms are salts, this value simplifies down to [A]/[B] or [C]/[D] respectively (Fig. S56).

$\Delta pK_a = pK_a(trans) - pK_a(cis)$

convert ΔpK_a to free energy $\Delta\Delta G$ using $\Delta G = -RT\ln K$

$2.30 \cdot RT\Delta pK_a = \Delta G(K_a trans) - \Delta G(K_a cis) = \Delta\Delta G$
$\Delta G_f(C) + \Delta G_f(H^+) - \Delta G_f(A) - (\Delta G_f(D) + \Delta G_f(H^+) - \Delta G_f(B)) = \Delta\Delta G$

where ΔG_f is a formation energy for substance in aqueous solution

$\Delta G_f(C) + \cancel{\Delta G_f(H^+)} - \Delta G_f(A) - \Delta G_f(D) - \cancel{\Delta G_f(H^+)} + \Delta G_f(B) = \Delta\Delta G$
$\Delta G_f(C) - \Delta G_f(A) - \Delta G_f(D) + \Delta G_f(B) = \Delta\Delta G$
$\Delta G_f(C) - \Delta G_f(D) - \Delta G_f(A) + \Delta G_f(B) = \Delta\Delta G$
$\Delta G_f(C) - \Delta G_f(D) - (\Delta G_f(A) - \Delta G_f(B)) = \Delta\Delta G$

those differences correspond to the free energies for the $K_{trans/cis}$ equilibrium constants

$- RT\ln K_{trans/cis}(salt) - (- RT\ln K_{trans/cis}(acid)) = \Delta\Delta G$

combine expressions for $\Delta\Delta G$

$$-RT \ln \frac{K_{trans/cis}(salt)}{K_{trans/cis}(acid)} = \Delta\Delta G = 2.30RT\Delta pK_a$$

or if to express ΔpK_a

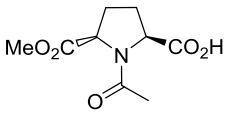
$$\Delta pK_a = pK_a(trans) - pK_a(cis) = \log_{10} \frac{K_{trans/cis}(salt)}{K_{trans/cis}(acid)}$$

These expression is reported in the main text as eq. 2-3.

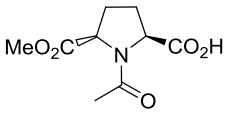
9. Characterization of Ac-(MeO)Pdc (19)

rac-1-acetyl-5-(methoxycarbonyl)pyrrolidine-2-carboxylic acid. Racemic trans-pyrrolidine-2,5-dicarboxylic acid mono methyl ester was provided by Dr. Pavel Mykhailiuk, whereas selective mono-methyl saponification was previously described.[S1] N-acetylation was performed conventionally. The NMR spectra exhibited presence of fully hydrolyzed Ac-Pdc (Fig. S57) due to partial hydrolysis of the substance, that occurred during the sample preparation. Likewise for Ac-Pro-Pro the pK_a transitions were followed by 12 zoomed $^1H\{^{13}C\}$ HSQC at different pH values from 6.9 till 1.3 as illustrated on Fig. S58. For confirmation of the assignment ^1H NOESY and HOHAHA spectra were conducted for the acidic sample. As it is shown on Fig. S59 two downfield resonances exhibited clear NOEs with the CH_3-moieties of the acetyl groups. Whereas in HOHAHA, the correspondence between the α-CH from the carboxyl and carboxymethyl sides were established. The correspondence of the CH_3O resonances was finally established by $^1H^{13}C$ HMBC experiments as shown at Fig. S60.

Ac-(MeO)Pdc (with $KHSO_4$): ^1H NMR (700 MHz, D_2O):

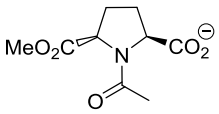


s-*trans*: 4.73 (d, *J* = 9.3 Hz, 1H, C*H*-CO_2H), 4.50 (d, *J* = 8.7 Hz, 1H, C*H*-CO_2Me), 3.66 (s, 3H, CH_3O), 2.34, 2.19 and 2.04 (3m, 4H, CH_2), 1.99 (s, 3H, CH_3-C=O);

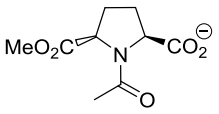


s-*cis*: 4.77 (d, *J* = 8.9 Hz, 1H, C*H*-CO_2Me), 4.47 (d, *J* = 9.5 Hz, 1H, C*H*-CO_2H), 3.72 (s, 3H, CH_3O), 2.34, 2.19 and 2.04 (3m, 4H, CH_2), 1.97 (s, 3H, CH_3-C=O).

Ac-(MeO)Pdc salt (in K-PB): ^1H NMR (700 MHz, D_2O):



s-*trans*: 4.51 (dd, *J* = 9.5, 1.6 Hz, 1H, C*H*-CO_2Me), 4.39 (dd, *J* = 9.1, 1.5 Hz, 1H, C*H*-CO_2^-), 3.67 (s, 3H, CH_3O), 2.30, 2.19, 2.06 and 1.99 (4m, 4H, CH_2), 1.95 (s, 3H, CH_3-C=O).



s-*cis*: 4.71 (d, *J* ~ 7-10 Hz, 1H, C*H*-CO_2Me), 4.29 (d, *J* = 9.4 Hz, 1H, C*H*-CO_2^-), 3.73 (s, 3H, CH_3O), 2.32, 2.12 and 1.91 (3m, 4H, CH_2), 1.94 (s, 3H, CH_3-C=O).

[S1] P. K. Mykhailiuk, S. V. Shishkina, O. V. Shishkin, O. A. Zaporozhets, I. V. Komarov, *Tetrahedron* 2011, **67**, 3091-3097.

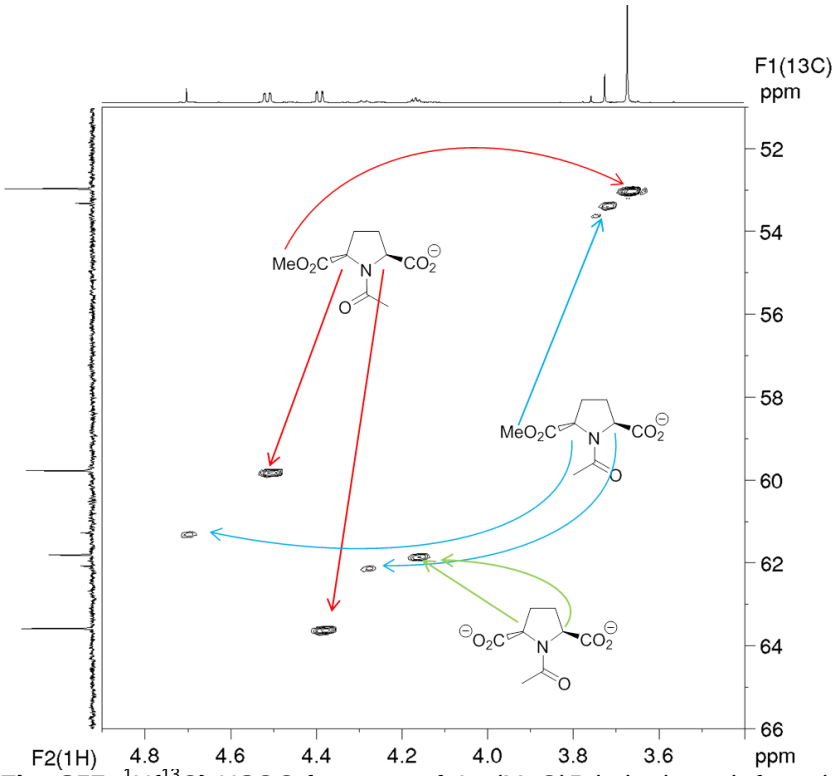


Fig. S57. $^{1}H\{^{13}C\}$ HSQC fragment of Ac-(MeO)Pdc in the salt form (at 700&176 MHz resonance frequency).

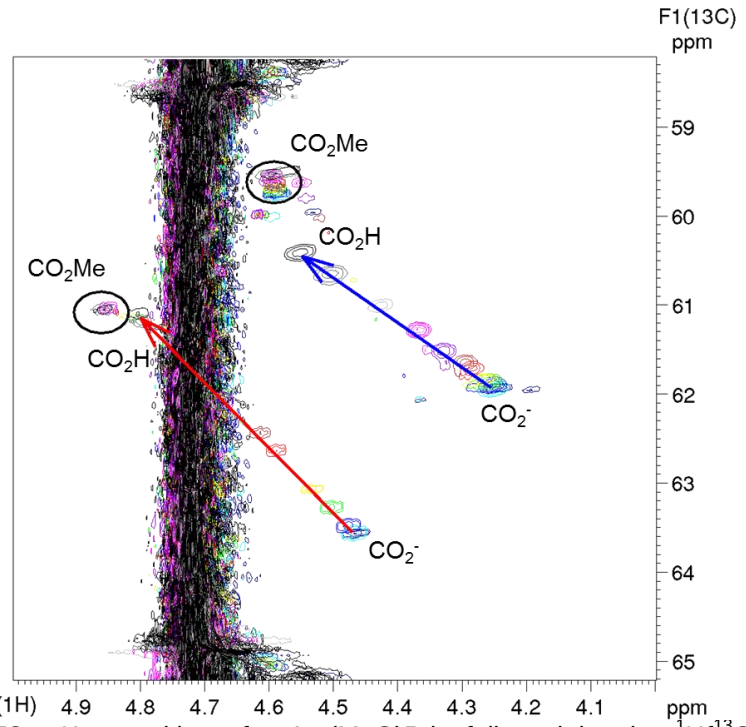


Fig. S58. pK_a transitions for Ac-(MeO)Pdc followed by the $^{1}H\{^{13}C\}$ HSQC spectra zoomed on the α-CH region (700&176 MHz resonance frequency).

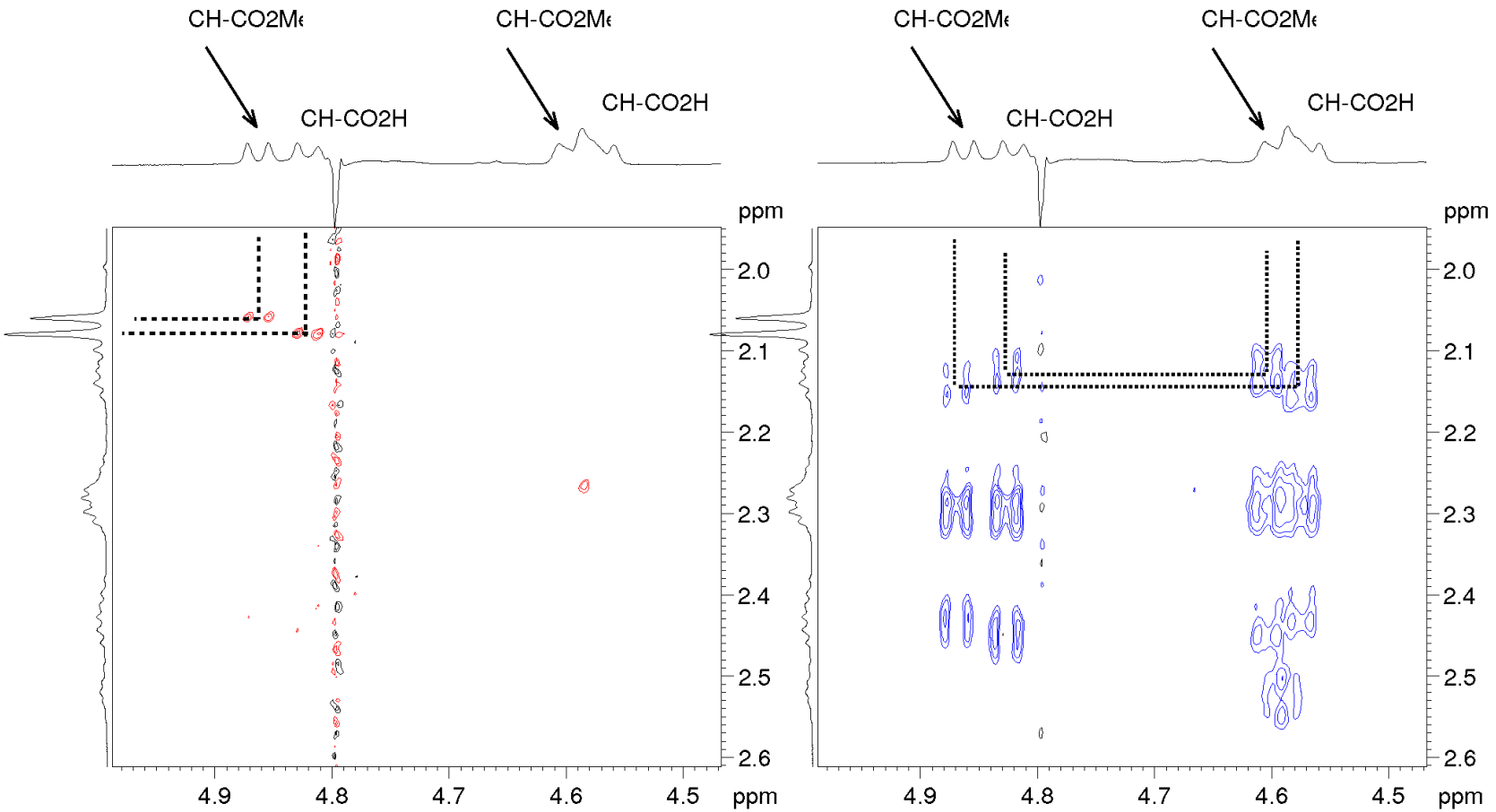


Fig. S59. Fragments of the ^{1}H NOESY (left; mixing time 1 s) and HOHAHA (right; mixing time 60 ms) spectra of Ac-(MeO)Pdc under acidic conditions (pH ≤1.5) at 500 MHz resonance frequency.

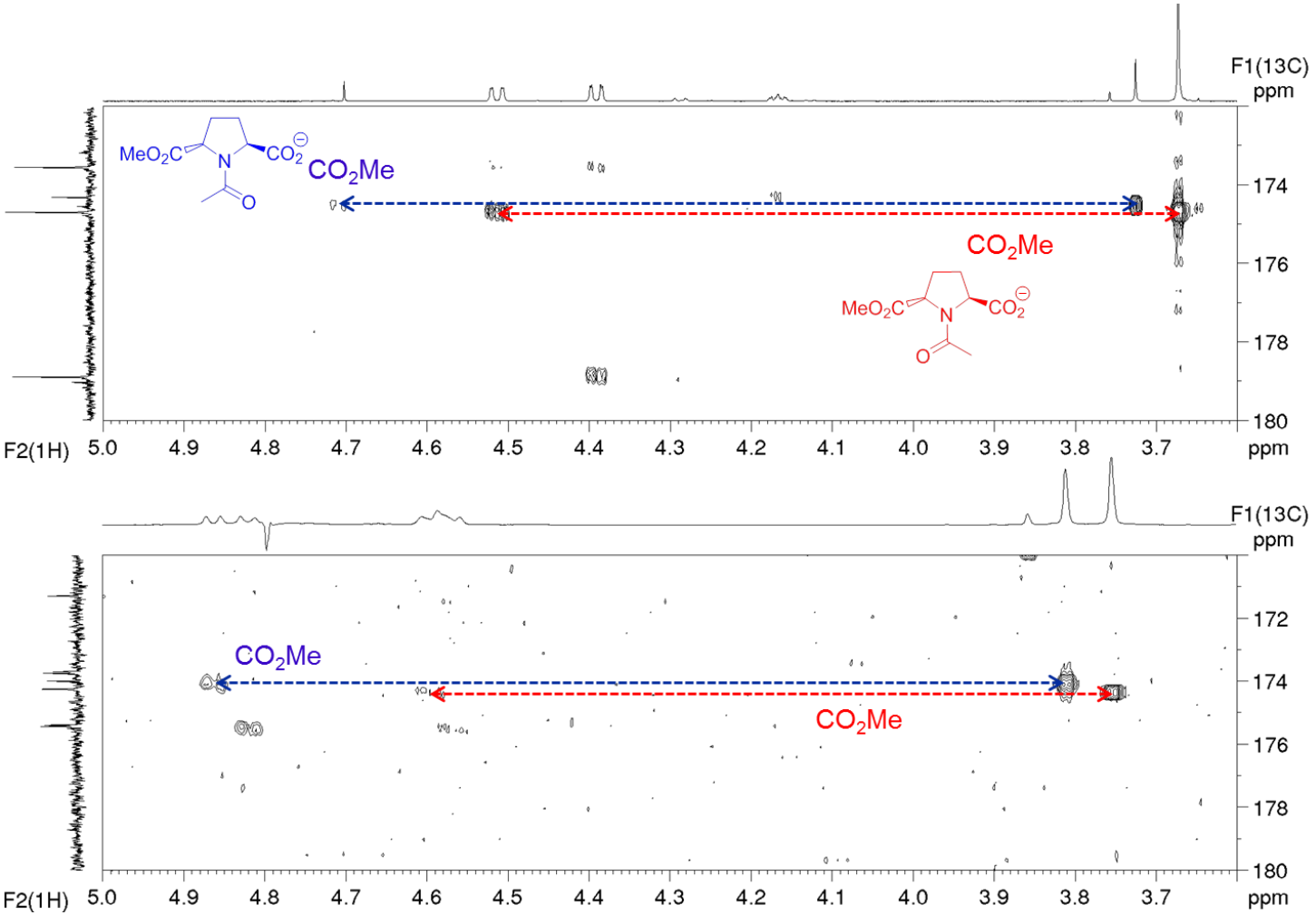


Fig. S60. Fragments of the $^{1}H^{13}C$ HMBC spectra of Ac-(MeO)Pdc as salt (at 700&176 MHz frequency) and acid (at 500&126 MHz) for assignment of the CH_3O resonances.

For the pH dependence of the equilibrium constant we followed the CH_3O signals in the spectra at different pH. CH_3O resonances were chosen, since they didn't interfere with the signals of the fully hydrolyzed substance. The spectra were acquired in ^1H experiments at 700 MHz, whereas the water suppression was afforded by either Watergate W5 or excitation sculpting methods. The former delivered more consistent data (Fig. S61). Concentrations were the following: Ac-(MeO)Pdc 0.9 mM, phosphate 0.9 mM.

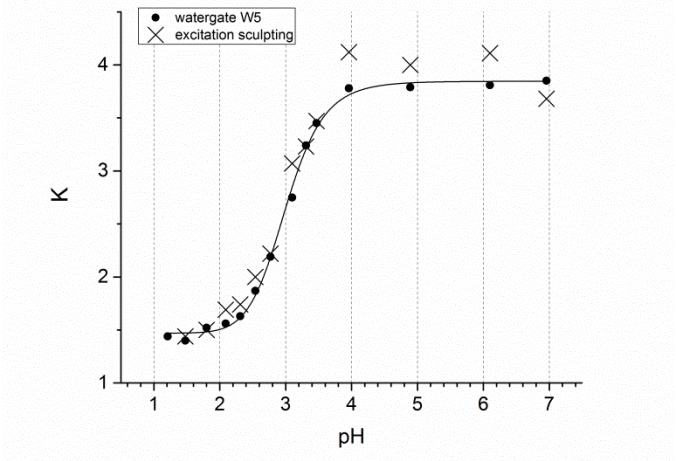


Fig. S61. pH dependence of the equilibrium *trans/cis* ratio in Ac-(MeO)Pdc as followed by ^1H spectra using either W5 or ^1H excitation sculpting water suppression at 700 MHz resonance frequency. The spectra were acquired at 298 K.

In addition, we compared the $K_{trans/cis}$ equilibrium constants of Ac-(MeO)Pdc (determined with the W5 suppression) to the ones predicted from the curve in Fig. 7C in the entire pH range examined. The resulting prediction (Fig. S62) matches the experimental data in the magnitude; however the position of the sigmoidal transition shifts towards more acidic values. This is due to the higher absolute acidity of the Pdc derivative.

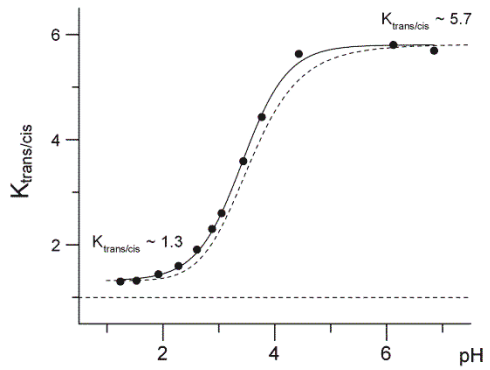


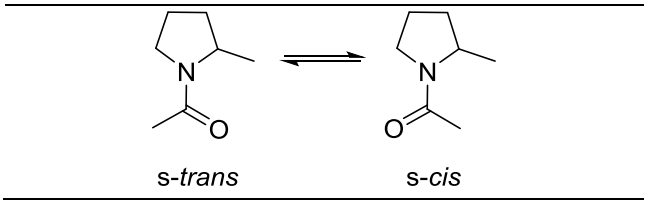
Fig. S62. Equilibrium $K_{trans/cis}$ values for Ac-(MeO)Pdc determined along with the pK_a measurements: solid circles – experimental points, solid line – sigmoidal fit to the experimental points, dashed line – curve predicted from the proline data as $K_{trans/cis}$(Ac-Pro-OMe) divided by $K_{trans/cis}$ fit of Ac-Pro (the one shown on Fig. 7C).

10. Characterization of N-acetyl-2-methylpyrrolidine (21)

Racemic 2-methylpyrrolidine (125 mg; 1.5 mmol, 1 equiv.) was diluted with diethylether (2 ml) and acetic anhydride (152 µl; 1.6 mmol, 1.1 equiv.) was added dropwise. After 15 min the solvent was removed under reduced pressure. Resulting mixture of N-acetyl 2-methylpyrrolidine with acetic acid 1:1 was used thereafter without purification.

The mixture was diluted with deuterium oxide, potassium dihydrogen phosphate was added, and resulting mixture was titrated by potassium hydroxide aqueous solution to the pH mether value 6.9. This sample was then used for NMR measurements and contained the following components: N-acetyl 2-methylpyrrolidine 130 mM, acetate 130 mM, phosphate 72 mM.

The ^{1}H and $^{13}C\{^{1}H\}$ NMR spectra were assigned using ^{1}H NOESY (mixing time 1 s), and $^{1}H\{^{13}C\}$ HSQC experiments. The ^{1}H NMR spectrum is shown at Fig. S63.



Scheme S5. Rotameric states of N-acetyl-2-methylpyrrolidine.

^{1}H NMR (D_2O, 700 MHz):

s-*trans*: 3.99 (m, 1H, 2-CH), 3.49 (ddd, *J* = 3.5, 8.2, 10.6 Hz, 1H, 5-CH), 3.37 (dd, *J* = 7.6, 10.6 Hz, 1H, 5-CH), 1.95 (s, 3H, CH_3-C=O), 1.97-1.80 (m, 3H) and 1.55 (m, 1H, all 3- and 4-CH_2), 1.04 (d, *J* = 6.5 Hz, CH_3-);

s-*cis*: 4.02 (m, 1H, 2-CH), 3.35 (m, 1H, 5-CH), 3.27 (td, *J* = 9.0, 10.8 Hz, 1H, 5-CH), 2.02 (s, 3H, CH_3-C=O), 1.97-1.80 (m, 3H) and 1.63 (m, 1H, all 3- and 4-CH_2), 1.10 (d, *J* = 6.6 Hz, CH_3-).

$^{13}C\{^{1}H\}$ NMR (D_2O, 176 MHz):

s-*trans*: 171.9 (s, C=O), 53.3 (s, 2-CH), 48.0 (s, 5-CH_2), 31.5 and 22.9 (two s, 3- and 4-CH_2), 21.7 (s, *C*H_3-C=O), 18.2 (CH_3-);

s-*cis*: 172.0 (s, C=O), 55.0 (s, 2-CH), 45.7 (s, 5-CH_2), 32.2 and 21.5 (two s, 3- and 4-CH_2), 20.8 (s, *C*H_3-C=O), 19.4 (CH_3-).

Thermodynamic analysis was performed in two temperature series, heating a cooling, in the temperature range 299-342 K. Plotting was performed as described in chapter S5. Results of two series (Fig. S64) were averaged for the final representation in the main text.

Obtained values:

heating series: $\Delta H_{trans/cis}$ = -1.15±0.05 kJ/mol; $\Delta S_{trans/cis}$ = -1.8±0.2 J/mol·K;

cooling series: $\Delta H_{trans/cis}$ = -1.08±0.08 kJ/mol; $\Delta S_{trans/cis}$ = -1.6±0.2 J/mol·K;

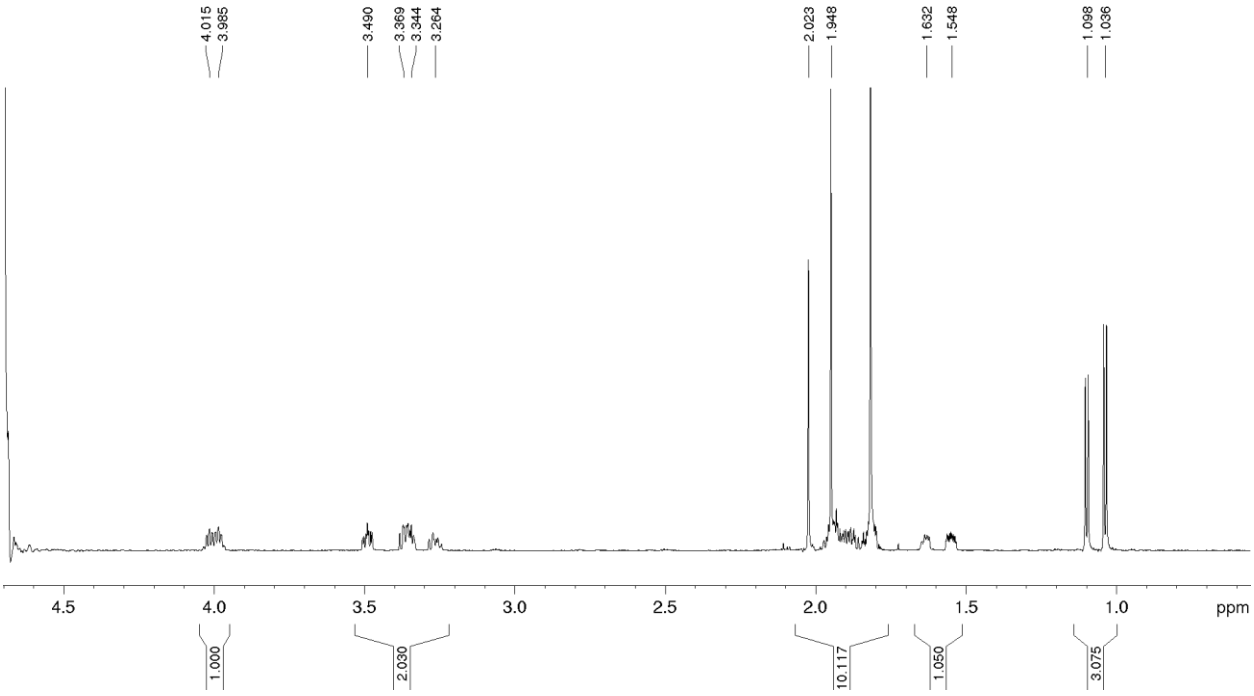


Fig. S63. ^{1}H NMR spectrum of N-acetyl-2-methylpyrrolidine in aqueous acetic and phosphate buffer at 298 K and 700 MHz resonance frequency.

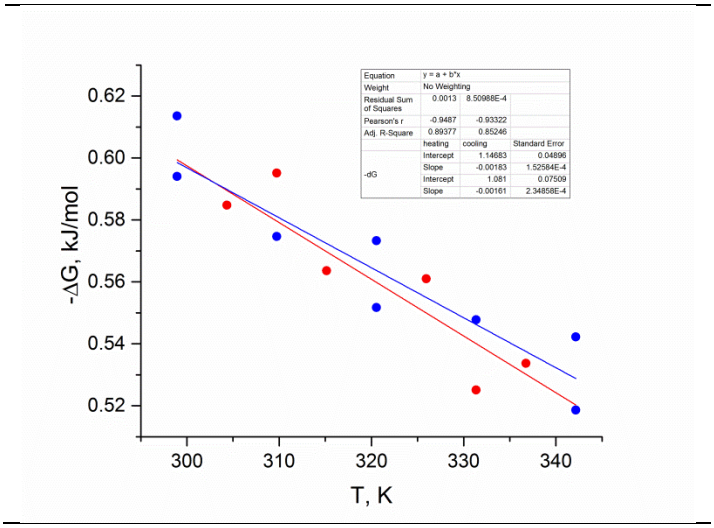
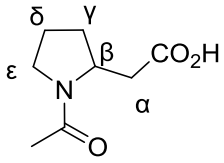


Fig. S64. Thermodynamic analysis for N-acetyl-2-methylpyrrolidine in aqueous buffer. Red: heating series, Blue: cooling series.

11. Characterization of the N-acetyl β-homoproline (20)



The ^{1}H and $^{13}C\{^{1}H\}$ NMR spectra were interpreted using $^{13}C\{^{1}H\}$ dept45, $^{1}H\{^{13}C\}$ HSQC, $^{1}H^{13}C$ HMBC, ^{1}H COSY-45 and ^{1}H NOESY (mixing time 1 s) experiments.

2-(1-acetylpyrrolidin-2-yl)acetic acid

IR bands: 2978, 29.32, 2889, 1703, 1580 cm^{-1}. Mass-spectrum (ESI): 172.1 $[M+1]^{+}$.

Ac-β-HPro salt (pH 7)

^{1}H NMR (D_2O, 700 MHz):

s-*trans*: 4.31 (m, 1H, β-CH), 3.59 and 3.49 (two m, 2H, ε-CH_2), 2.70 (dd, *J* = 14.4, 3.9 Hz, 1H, α-CH), 2.18 (dd, *J* = 14.4, 10.7 Hz, 1H, α-CH), 2.06 (s, 3H, CH_3), 2.06-1.94 (m, 3H, γ-CH_2 and δ-CH), 1.78 (m, 1H, δ-CH);

s-*cis*: 4.30 (m, 1H, β-CH), 3.43 3.49 (m, 2H, ε-CH_2), 2.53 (dd, *J* = 14.5, 4.7 Hz, 1H, α-CH), 2.34 (dd, *J* = 14.5, 10.1 Hz, 1H, α-CH), 2.15 (s, 3H, CH_3), 2.06-1.94 (m, 3H, γ-CH_2 and δ-CH), 1.84 (m, 1H, δ-CH).

$^{13}C\{^{1}H\}$ NMR (D_2O, 176 MHz):

s-*trans*: 180.2 (s, CO_2^-), 172.1 (s, C(=O)-N), 55.2 (s, β-CH), 48.0 (s, ε-CH_2), 40.1 (s, α-CH_2), 29.4 (s, δ-CH_2), 22.8 (s, γ-CH_2), 21.8 (s, CH_3);

s-*trans*: 179.5 (s, CO_2^-), 172.2 (s, C(=O)-N), 57.2 (s, β-CH), 45.7 (s, ε-CH_2), 41.8 (s, α-CH_2), 30.4 (s, δ-CH_2), 21.3 (s, γ-CH_2), 20.9 (s, CH_3).

The ^{1}H NMR spectrum is presented on Fig. S65. The thermodynamic analysis of the *trans-cis* equilibrium was performed as described for other samples. Resulting data were collected from heating and cooling series and averaged. Found values:

$\Delta H_{trans/cis}$ = -0.90±0.09 kJ/mol; $\Delta S_{trans/cis}$ = -2.2±0.3 J/mol·K

Ac-β-HPro acid (pH 1.4)

^{1}H NMR (D_2O, 700 MHz):

s-*trans*: 4.34 (m, 1H, β-CH), 3.60 (ddd, *J* = 11, 8, 4 Hz, 1H, ε-CH) and 3.51 (m, 1H, ε-CH), 2.77 (dd, *J* = 16, 4.5 Hz, 1H, α-CH) and 2.50 (dd, *J* = 16, 9 Hz, 1H, α-CH), 2.07 (s, 3H, CH_3), 2.10-1.94 (m, 3H, γ-CH_2 and δ-CH), 1.81 (m, 1H, δ-CH);

s-*cis*: 4.39 (m, 1H, β-CH), 3.43 (two m, 2H, ε-CH_2), 2.70 (dd, *J* = 16, 4 Hz, 1H, α-CH) and 2.62 (dd, *J* = 16, 10 Hz, 1H, α-CH), 2.15 (s, 3H, CH_3), 2.11-1.94 (m, 3H, γ-CH_2 and δ-CH), 1.90 (m, 1H, δ-CH).

^{13}C{^{1}H} NMR (D_2O, 176 MHz):

s-*trans*: 176.0 (s, CO_2H), 172.6 (s, C(=O)-N), 54.1 (s, β-CH), 48.1 (s, ε-CH_2), 37.2 (s, α-CH_2), 29.8 (σ, δ-CH_2), 22.8 (s, γ-CH_2), 21.6 (s, CH_3);

s-*cis*: 175.3 (s, CO_2H), 172.5 (s, C(=O)-N), 55.9 (s, β-CH), 45.8 (s, ε-CH_2), 38.1 (s, α-CH_2), 30.4 (s, δ-CH_2), 21.1 (s, γ-CH_2), 20.9 (s, CH_3).

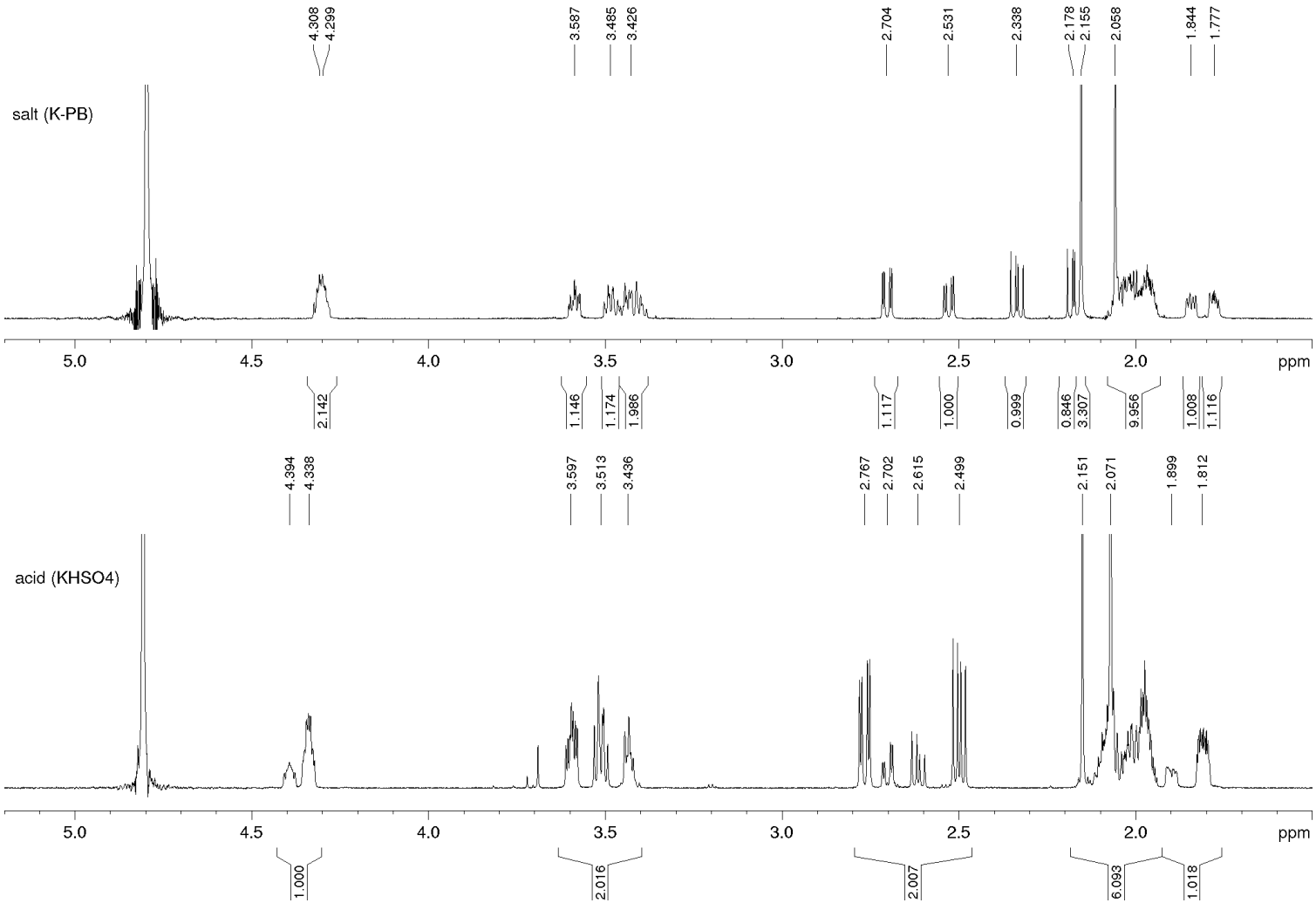


Fig. S65. ^{1}H NMR spectra of Ac-βHPro in the salt (top trace) and acid (bottom trace) forms in deuterium oxide at 700 MHz frequency.

12. Characterization of the N-acetyl 4-methylprolines (15 and 16)

(2*S*,4*R*)-1-acetyl-4-methylpyrrolidine-2-carboxylic acid. ^{1}H NMR spectra at Fig S66.

Ac-(4*R*)-Mep acid (pH 1.4)

^{1}H NMR (700 MHz, D_2O):

s-*trans*: 4.40 (dd, *J* = 9, 3 Hz, 1H, α-CH), 3.71 (dd, *J* = 10, 7 Hz, 1H, δ-CH), 3.10 (dd, *J* = 10, 9 Hz, 1H, δ-CH), 2.36 (m, 1H, γ-CH), 2.09 (ddd, *J* = 13, 6, 3 Hz, 1H, β-CH), 2.01 (s, 3H, CH_3C=O), 1.88 (dt, *J* = 13, 10 Hz, 1H, β-CH), 0.95 (d, *J* = 7 Hz, 3H, γ-CH_3);

s-*cis*: 4.59 (dd, *J* = 9, 1.5 Hz, 1H, α-CH), 3.59 (dd, *J* = 11, 8 Hz, 1H, δ-CH), 2.92 (dd, *J* = 11, 9 Hz, 1H, δ-CH), 2.22 (m, 2H, γ-CH and β-CH), 1.95 (m, 1H, β-CH), 1.91 (s, 3H, CH_3=O), 0.95 (m, 3H, γ-CH_3).

^{13}C{^{1}H} NMR (126 MHz, D_2O):

s-*trans*: 176.2 (s, CO_2H), 173.1 (s, N-C=O), 58.9 (s, α-CH), 54.9 (s, δ-CH_2), 36.6 (s, β-CH_2), 32.1 (s, γ-CH), 21.1 (s, *C*H_3C=O), 16.1 (s, γ-CH_3)

s-*cis*: 176.2 (s, CO_2H), 173.4 (s, N-C=O), 61.0 (s, α-CH), 53.2 (s, δ-CH_2), 38.1 (s, β-CH_2), 30.5 (s, γ-CH), 20.9 (s, *C*H_3C=O), 16.1 (s, γ-CH_3).

Ac-(4*R*)-Mep salt (pH 7)

^{1}H NMR (700 MHz, D_2O):

s-*trans*: 4.21 (dd, *J* = 9, 3 Hz, 1H, α-CH), 3.68 (dd, *J* = 10, 7 Hz, 1H, δ-CH), 3.07 (dd, *J* = 10, 8 Hz, 1H, δ-CH), 2.32 (m, 1H, γ-CH), 1.99 (s, 3H, CH_3C=O), 1.97 (ddd, *J* = 13, 6, 3 Hz, 1H, β-CH), 1.79 (dt, *J* = 13, 9 Hz, 1H, β-CH), 0.94 (d, *J* = 7 Hz, 3H, γ-CH_3);

s-*cis*: 4.26 (dd, *J* = 9, 2.5 Hz, 1H, α-CH), 3.59 (dd, *J* = 11, 7 Hz, 1H, δ-CH), 2.92 (dd, *J* = 11, 9 Hz, 1H, δ-CH), 2.24 (m, 1H, γ-CH), 2.07 (ddd, *J* = 12, 6, 3 Hz, 1H, β-CH), 1.90 (dt, *J* = 13, 9 Hz, 1H, β-CH), 1.88 (s, 3H, CH_3=O), 0.94 (d, *J* = 7 Hz, 3H, γ-CH_3).

^{13}C{^{1}H} NMR (126 MHz, D_2O):

s-*trans*: 179.8 (s, CO_2H), 172.3 (s, N-C=O), 61.5 (s, α-CH), 55.2 (s, δ-CH_2), 37.6 (s, β-CH_2), 31.9 (s, γ-CH), 21.3 (s, *C*H_3C=O), 16.4 (s, γ-CH_3);

s-*cis*: 179.7 (s, CO_2H), 173.0 (s, N-C=O), 63.5 (s, α-CH), 53.4 (s, δ-CH_2), 38.9 (s, β-CH_2), 30.6 (s, γ-CH), 20.9 (s, *C*H_3C=O), 16.4 (s, γ-CH_3).

(2*S*,4*S*)-1-acetyl-4-methylpyrrolidine-2-carboxylic acid. ^{1}H NMR spectra at Fig S67.

Ac-(4*S*)-Mep acid (pH 1.4)

^{1}H NMR (700 MHz, D_2O):

s-*trans*: 4.28 (t, *J* = 8.5 Hz, 1H, α-CH), 3.77 (dd, *J* = 10, 7 Hz, 1H, δ-CH), 3.10 (t, *J* = 19 Hz, 1H, δ-CH), 2.45 (dt, *J* = 13, 7 Hz, 1H, β-CH), 2.32 (m, 1H, γ-CH), 2.03 (s, CH_3C=O), 1.53 (dt, *J* = 12, 10 Hz, 1H, β-CH), 1.00 (d, J = 7 Hz, 3H, γ-CH_3);

s-*cis*: 4.56 (t, *J* = 7.7 Hz, 1H, α-CH), 3.80 (dd, *J* = 12, 7 Hz, 1H, δ-CH), 2.87 (dd, *J* = 12, 9 Hz, 1H, δ-CH), 2.58 (dt, *J* = 13, 8 Hz, 1H, β-CH), 2.23 (m, 1H, γ-CH), 1.91 (s, 3H, CH_3C=O), 1.72 (ddd, *J* = 13, 9, 7.5 Hz, 1H, β-CH), 0.96 (d, *J* = 7 Hz, 3H, γ-CH_3).

^{13}C{^{1}H} NMR (176 MHz, D_2O):

s-*trans*: 176.4 (s, CO_2H), 172.7 (s, N-C=O), 59.5 (s, α-CH), 55.2 (s, δ-CH_2), 37.3 (s, β-CH_2), 33.1 (s, γ-CH), 21.3 (s, *C*H_3C=O), 16.0 (s, γ-CH_3);

s-*cis*:176.5 (s, CO_2H), 173.5 (s, N-C=O), 60.6 (s, α-CH), 53.7 (s, δ-CH_2), 38.5 (s, β-CH_2), 31.4 (s, γ-CH), 20.8 (s, *C*H_3C=O), 16.7 (s, γ-CH3).

Ac-(4*S*)-Mep salt (pH 7)

^{1}H NMR (700 MHz, D_2O):

s-*trans*: 4.10 (t, *J* = 8.5 Hz, 1H, α-CH), 3.73 (dd, *J* = 10, 7 Hz, 1H, δ-CH), 3.08 (t, *J* = 10 Hz, 1H, δ-CH), 2.38 (dt, *J* = 12, 7 Hz, 1H, β-CH), 2.24 (m, 1H, γ-CH), 2.01 (s, 3H, CH_3C=O), 1.41 (dt, *J* = 12, 10 Hz, 1H, β-CH), 0.97 (d, *J* = 7 Hz, 3H, γ-CH_3);

s-*cis*: 4.24 (t, *J* = 8.3 Hz, 1H, α-CH), 3.79 (dd, *J* = 11, 7 Hz, 1H, δ-CH), 2.84 (t, *J* = 10 Hz, 1H, δ-CH), 2.50 (dt, *J* = 13, 7 Hz, 1H, β-CH), 2.16 (m, 1H, γ-CH), 1.88 (s, 3H, CH_3C=O), 1.55 (dt, *J* = 12, 9 Hz, 1H, β-CH), 0.98 (d, *J* = 7 Hz, 3H, γ-CH_3).

^{13}C{^{1}H} NMR (126 MHz, D_2O):

s-*trans*: 180.0 (s, CO_2^-), 172.0 (s, N-C=O), 62.2 (s, α-CH), 55.5 (s, δ-CH_2), 37.9 (s, β-CH_2), 32.9 (s, γ-CH), 21.5 (s, *C*H_3C=O), 16.4 (s, γ-CH_3);

s-*cis*: 180.0 (s, CO_2^-), 173.1 (s, N-C=O), 63.7 (s, α-CH), 53.7 (s, δ-CH_2), 39.4 (s, β-CH_2), 31.5 (s, γ-CH), 20.7 (s, *C*H_3C=O), 16.2 (s, γ-CH_3).

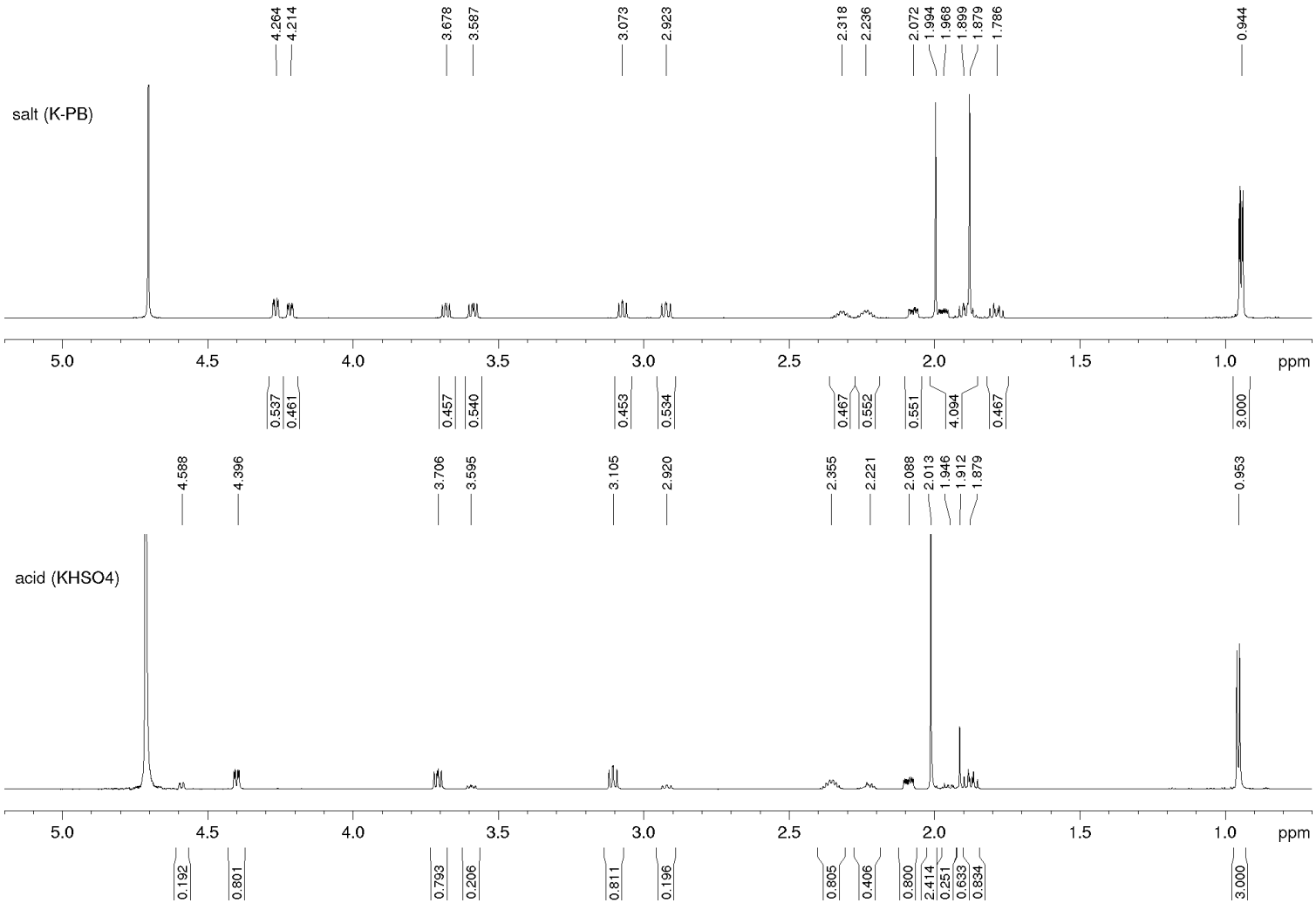


Fig. S66. ^{1}H NMR spectra of Ac-(4*R*)-Mep in the salt (top trace) and acid (bottom trace) forms in deuterium oxide at 700 MHz frequency.

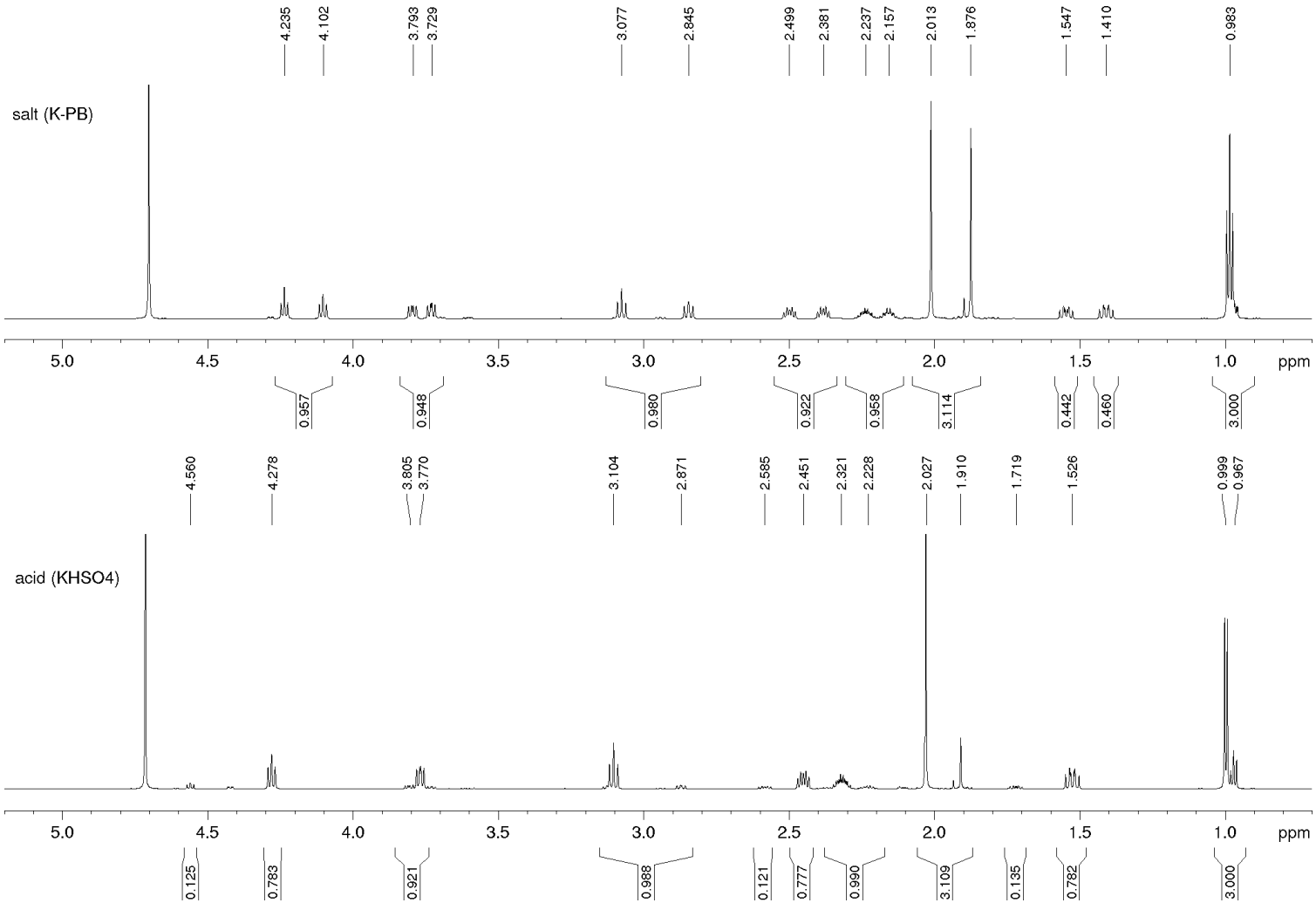


Fig. S67. ^{1}H NMR spectra of Ac-(4*S*)-Mep in the salt (top trace) and acid (bottom trace) forms in deuterium oxide at 700 MHz frequency.

13. Solvent dependence of the *trans/cis* ratio

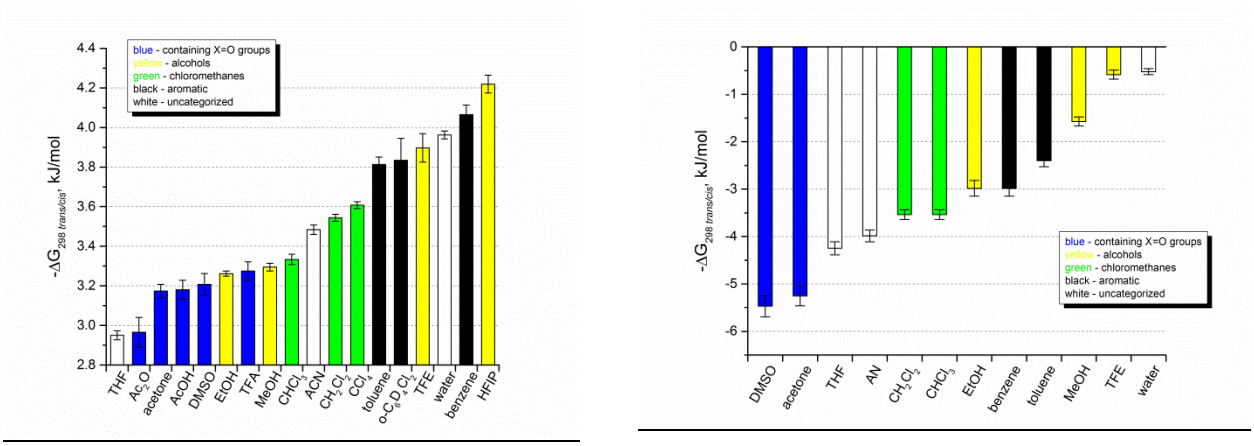


Fig. S68. Free energy $\Delta G_{trans/cis}$ in different solvents as measured at 298 K: left – AcProOMe, right – AcProOK-18-crown-6.